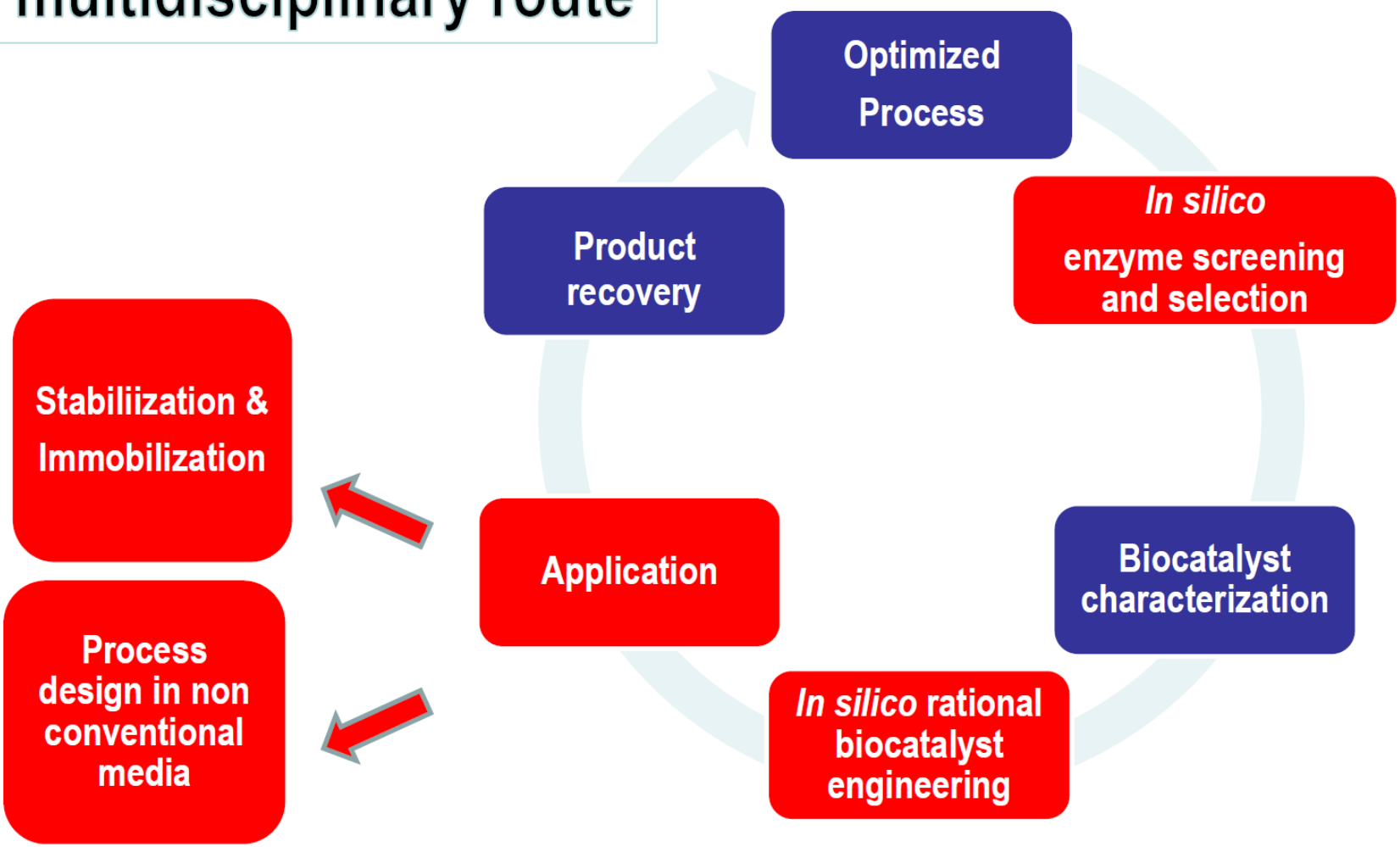
Biocatalysts in chemical reactions and organic synthesis

Strategies for planning rationally biocatalyzed reactions

Developing a biocatalyzed process: a multidisciplinary route

Contribution from chemists?



The reaction: identifying the most suitable enzyme and the reaction conditions to carry out the biotransformation at lab-scale (proof of concept)

The protein: Trying to predict enzyme properties (stability, ractivity, selectivity) on a rational basis in order to preserve such features as long as possible throughout the biotrasformation

The process: how making the enzyme and the biotransformayion effective at large (industrial) scale, the process ecomically viable and environmentally sustainable

Aim:
making enzymes working under unconventional «desperate» conditions while keeping most of their catalytic activity

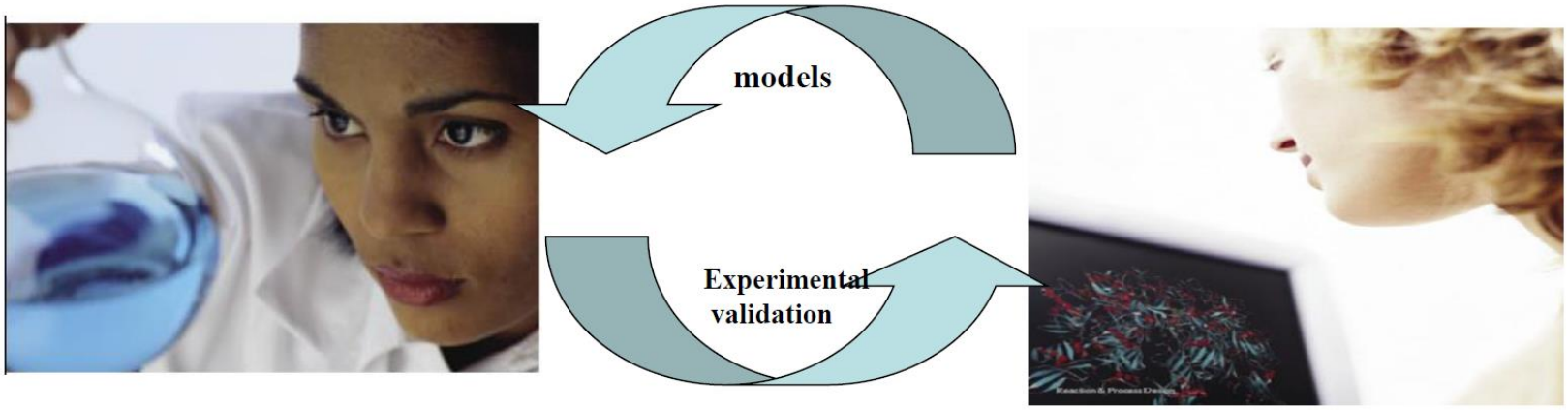
Solid to solid:

Organic solvent

Bulk –solventless - transformation

UNIVERSITÀ DEGLI STUDI DI TRIESTE
Dipartimento di
Scienze Chimiche e Farmaceutiche

Integrating experimental and computational methods



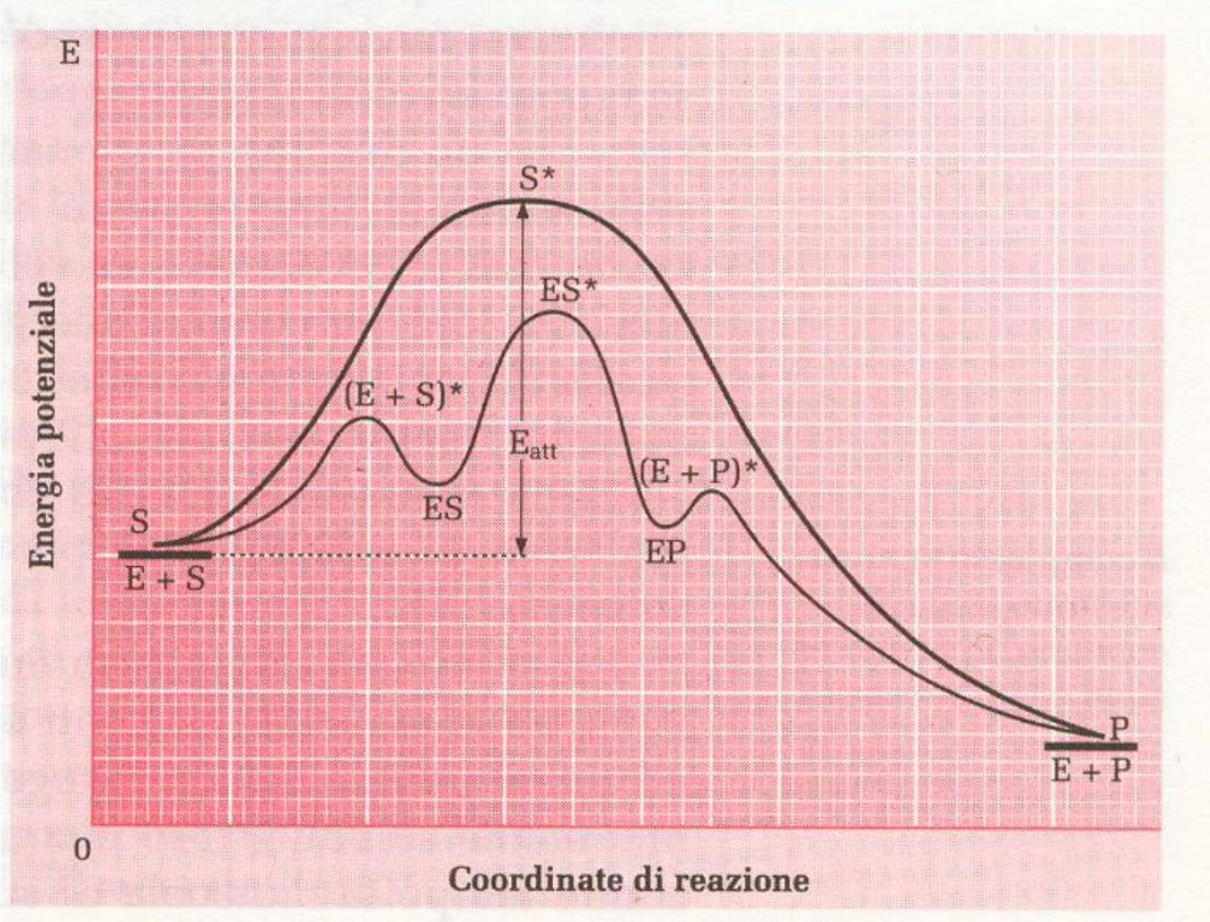
Computational simulations and predictions must be competitive with experimental approaches & time scale

Identifying a suitable enzyme and predict its ability to transform a certain substrate with the desidered chemo-regio-stereo-selectivity (when required)

How can we study and predict enzyme properties?

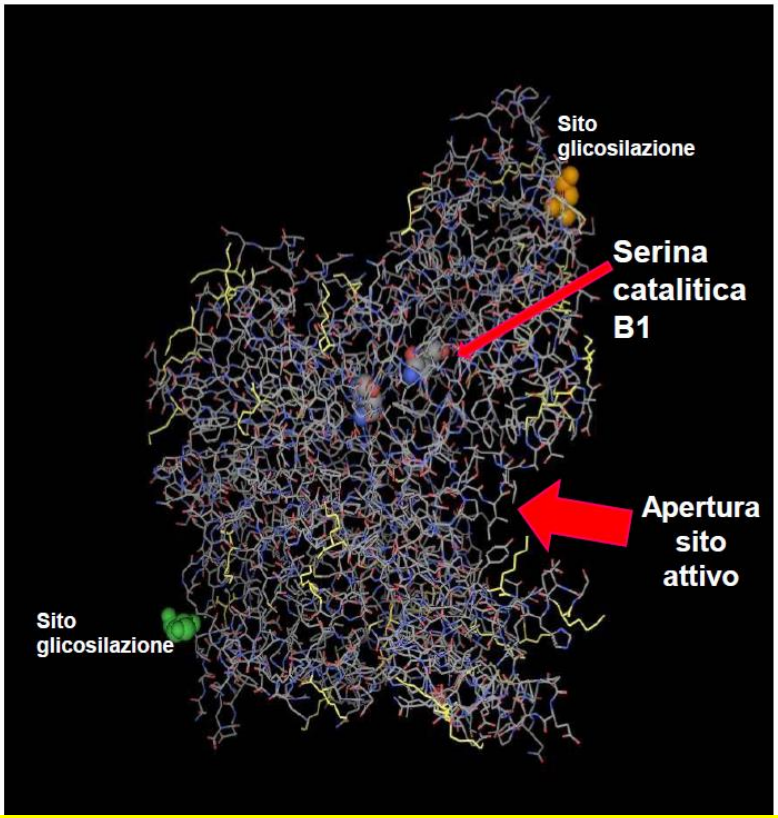
The theory:

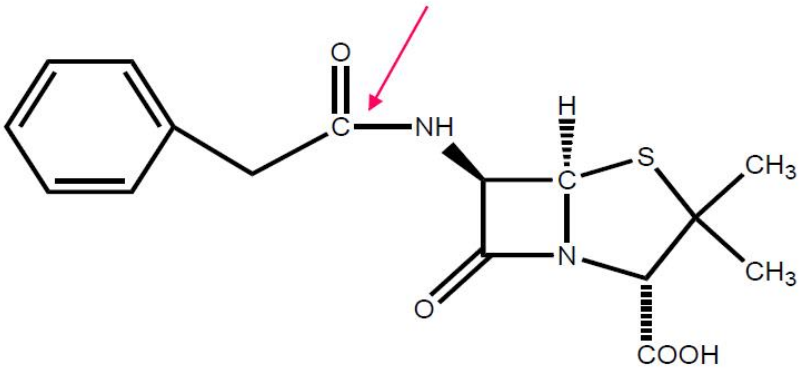
1) **Enzyme-substrate recognition**
2) **Catalysis and Transformation**



$$E + S \rightleftharpoons ES \rightarrow ES^* \rightarrow EP \rightleftharpoons E + P$$

Molecular simulation for the study of enzyme-substrate recognition: penicillin G inside penicillin G amidase (PGA), a serine hydrolase (dimer)





Penicillin G

Molecular models: a computational simulation of tridimentional structures

MODEL:

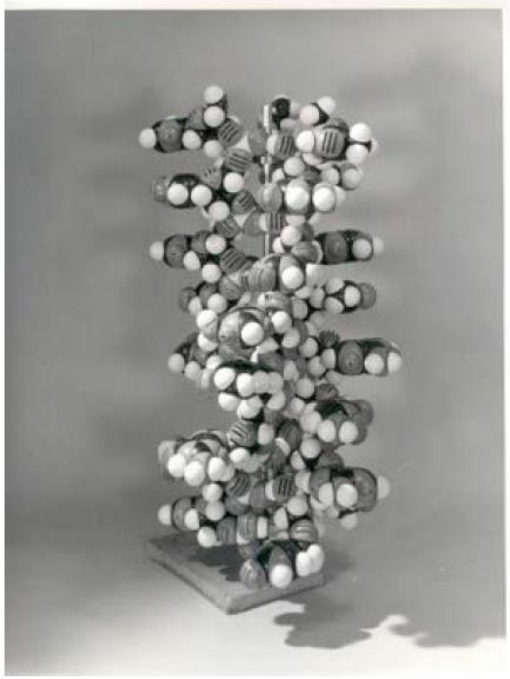
A simplified or idealized description or conception of a particular system, situation, or process, often in mathematical terms, that is put forward as a basis for theoretical or empirical understanding, or for calculations, predictions, etc.; a conceptual or mental representation of something

OR

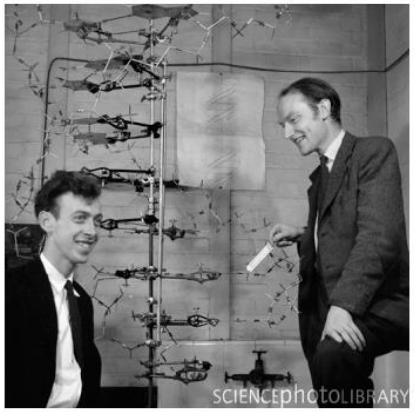
A description of structure

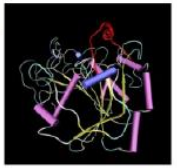
Molecular modelling therefore deals with models of molecules, or more generally, molecular phenomena

Molecules models

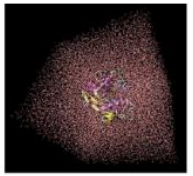


Classical/traditional models



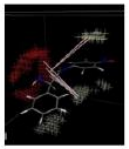


STEERED MD



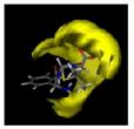
CLASSICAL MD

New methods:
Computer simulations

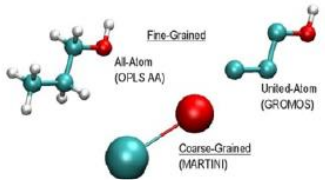


GRIND DESCRIPTORS

GRID DESCRIPTORS

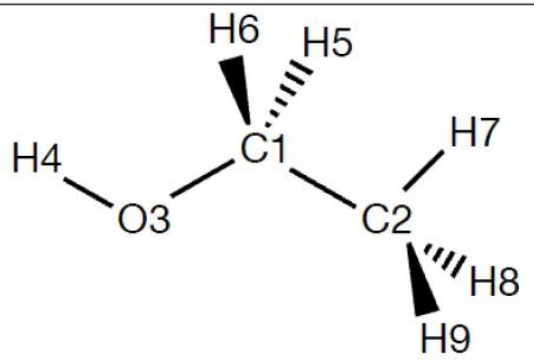


MARTINI MD



Three-dimensional structure: coordinates

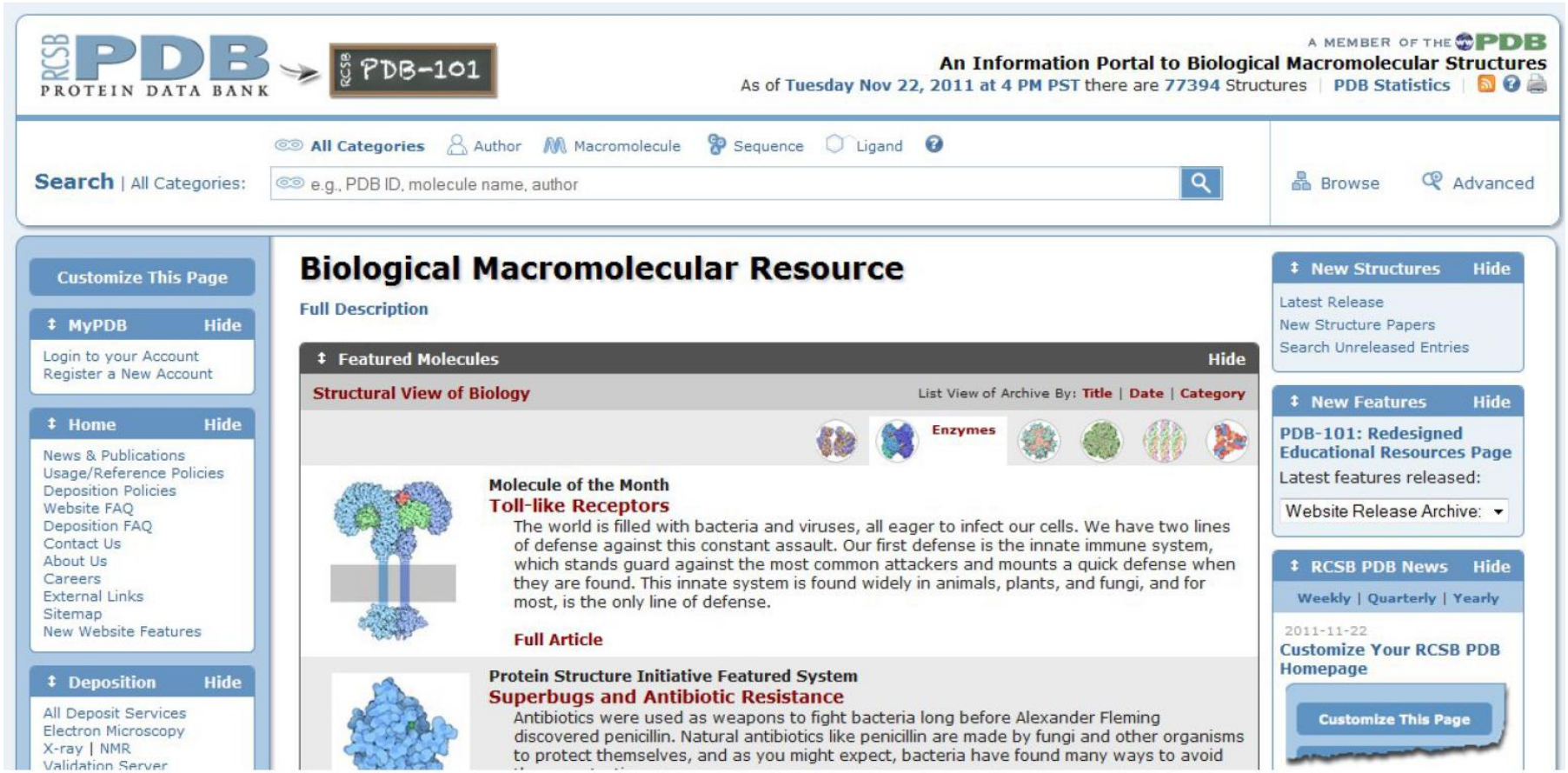
A primary requirement for handling molecules and molecular structures on the computer is obviously to be able to specify the location of the atoms in 3D space



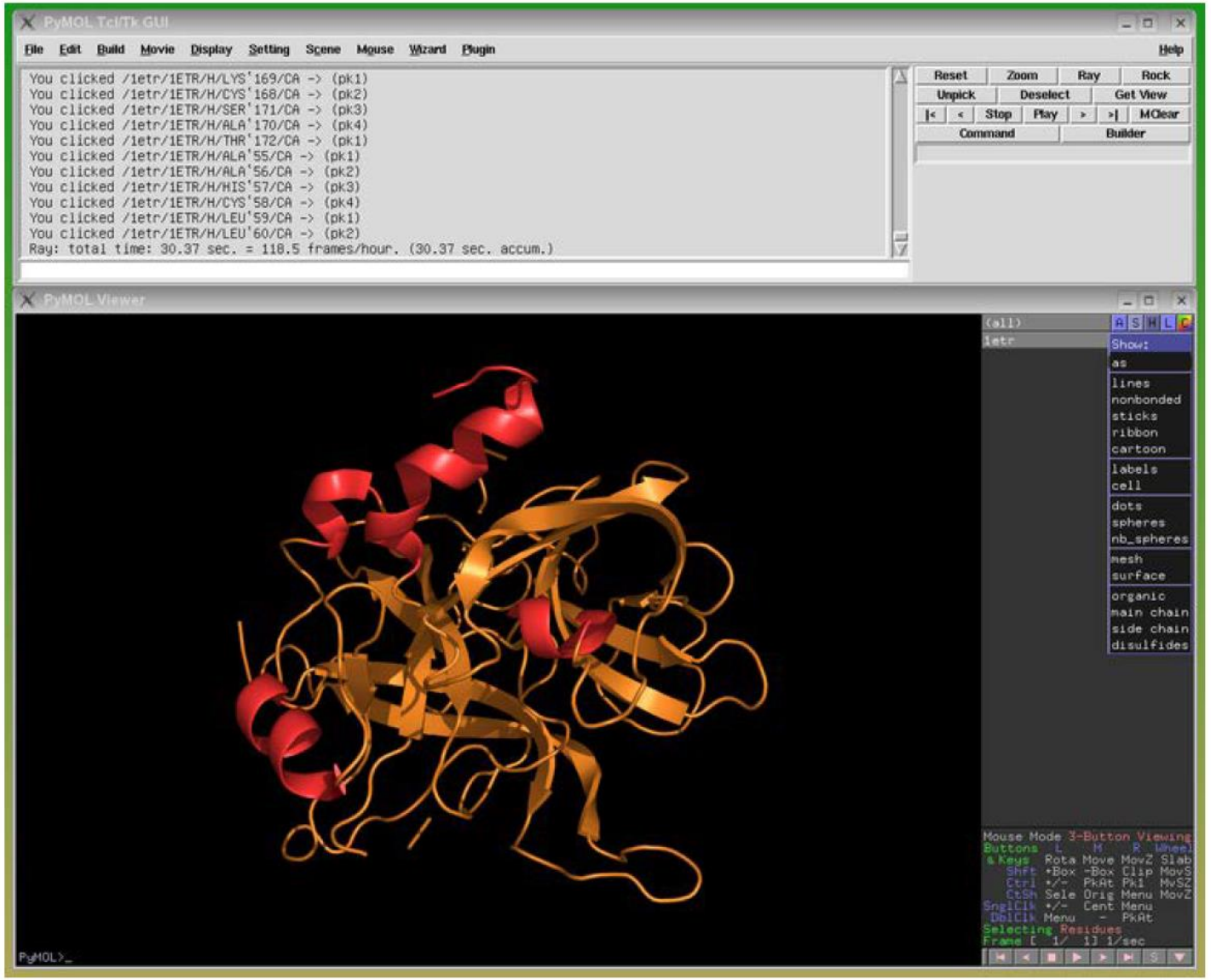
9				
1	C	0.000000	0.000000	0.000000
2	C	0.000000	0.000000	1.450000
3	O	1.319933	0.000000	-0.466667
4	H	1.319933	0.000000	-1.416667
5	H	-0.513360	0.889165	-0.363000
6	H	-0.513360	-0.889165	-0.363000
7	H	-1.026719	0.000000	1.813000
8	H	0.513360	0.889165	1.813000
9	H	0.513360	-0.889165	1.813000

coordinates

Protein structures:
www.pdb.org

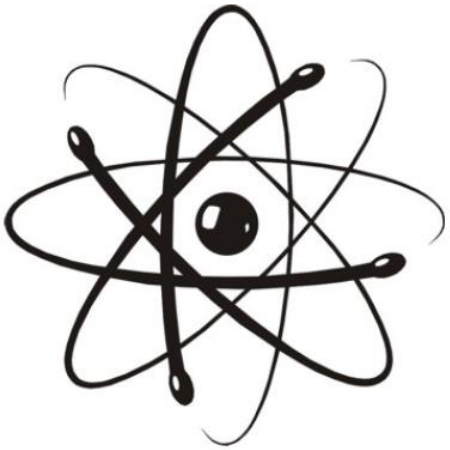


Structure visualization and manipulation: PyMOL



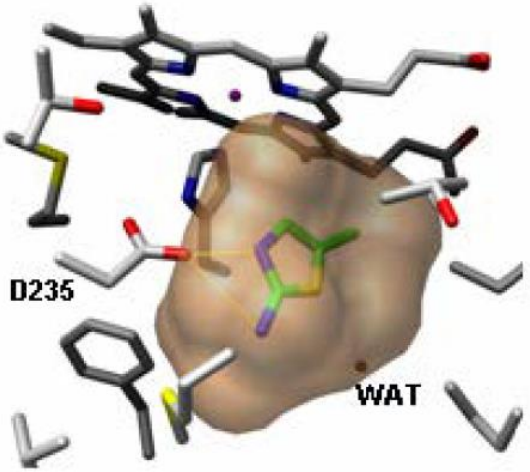
MOLECULAR MODELLING

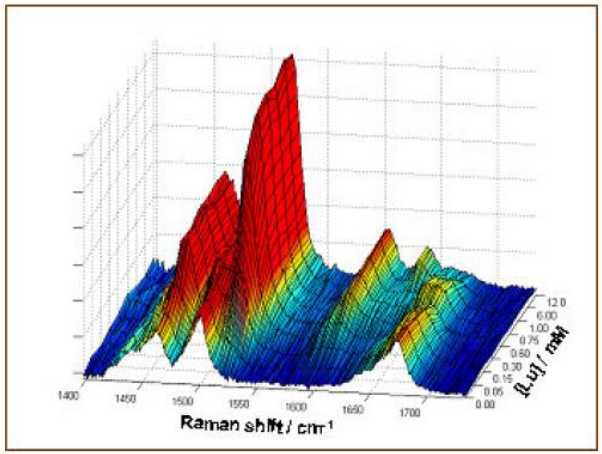
The term molecular modelling indicates a series of different techniques able to describe, to represent and to manipulate 3D-structures



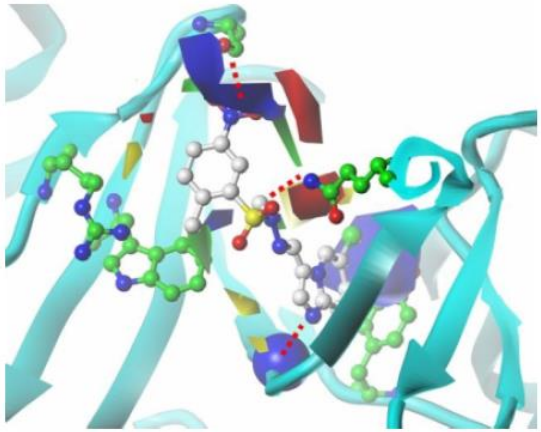
Quantum Mechanic

Molecular Docking





Molecular Dynamic



Statistic (3D-QSAR)

MOLECULAR DOCKING

Molecular docking is a method that predicts how molecules approach and interacts each other

Based on conformational search and the analysis of steric and electrostatic interactions between molecules

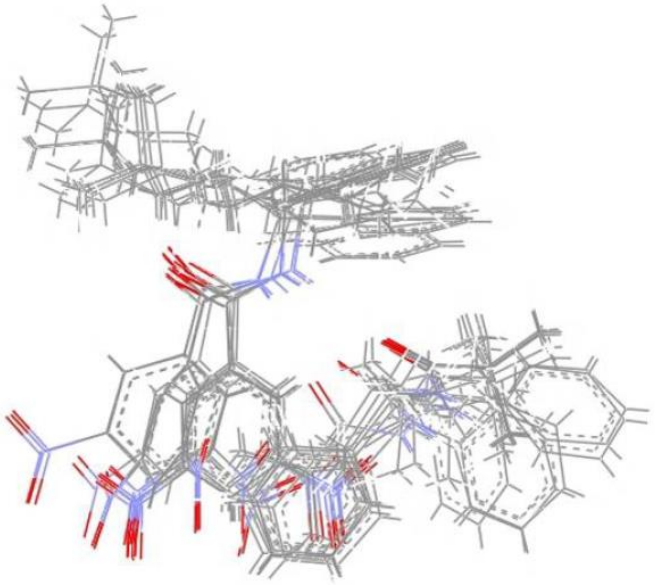
The method can be used to predict the strength of association or binding affinity between two molecules

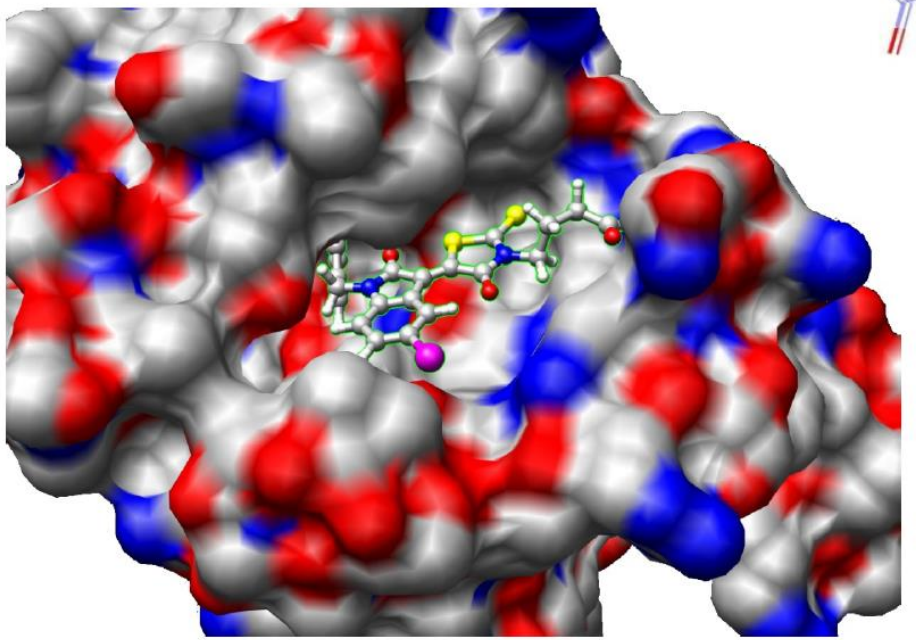
Docking applications concern different research fields:

- **Interactions between macromolecular receptor and ligand with low molecular weight (i.e. enzyme-substrate)**
- **Interactions between macromolecular receptor and macromolecular ligand (i.e. protein-protein, DNA-protein, DNA-DNA)**
- **Interaction between low molecular weight receptor and low molecular weight ligand (i.e. inclusions)**

MOLECULAR DOCKING

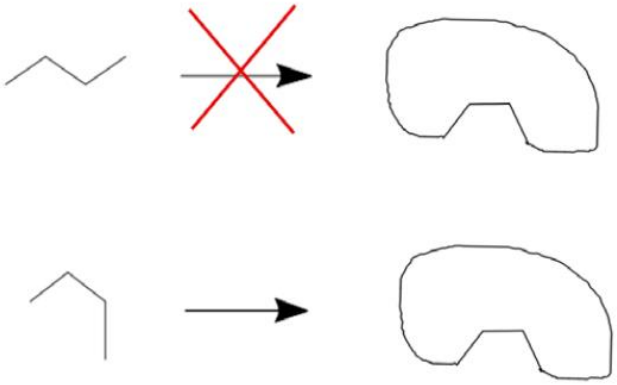
1- Conformational search



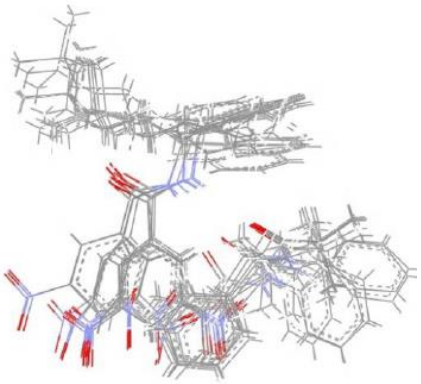


2- Placement

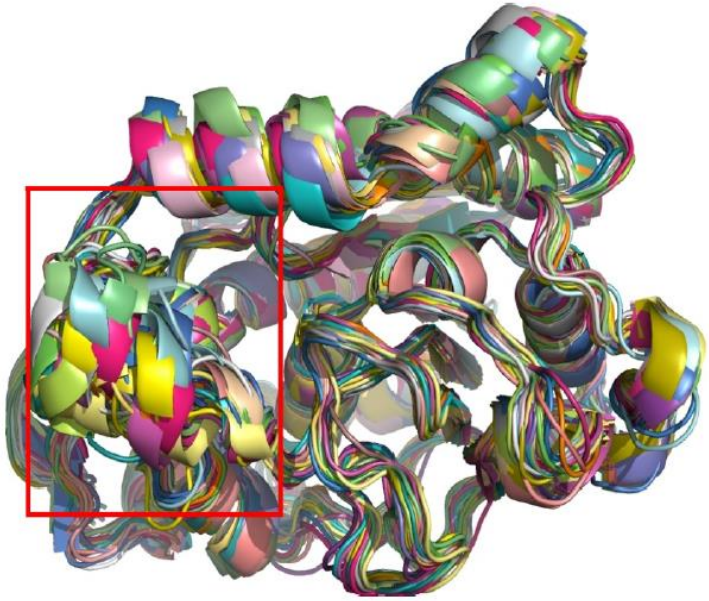
CONFORMATIONAL SEARCH



The same molecule can not be docked in all its possible conformations



Not only the small molecules like ligands can assume several conformations. Big molecules, like proteins have the same characteristic



METHODS:

1. **Systematic search**
2. **Model building**
3. **Simulation-based**
4. **Genetic algorithms**

PLACEMENT

Approach molecules each other and generate docking poses classified by a scoring function

Rigid Docking

Molecules are kept rigid and the best conformation of each molecule is determined by a previous conformational search and energy minimization

Flexible Docking

Molecules are free to move: several conformations of the small molecule are placed and than the generated poses are refined by a new conformational search (usually big molecule are kept rigid)

<u>Scoring functions</u> calculate the potential energy of each docking pose (steric and electrostatic interactions)

AutoDock 4.2.6 Download Page

AutoDock is now distributed freely under the GNU GPL for all to use.

If you plan to use AutoDock for commercial purposes we encourage donations to the Olson Laboratory to help support further development of the AutoDock suite of programs.

Please make out any donations checks to:

The Scripps Research Institute
c/o Prof. Arthur J. Olson

and send them to:

Prof. Arthur J. Olson
Department of Molecular Biology, MB-5
The Scripps Research Institute
La Jolla, CA 92037
USA

Many thanks!

Select the platform and/or source code.
Release 4.2.6 Notes.

```
AutoDock4.2.6 features improved input checking and an output format suitable for automated analysis.
 Multiple search methods can be used in a single AutoDock4.2.6 job.
 AutoDock 4.2.6 is available for more platforms.
 The process of compiling new atomic parameter tables into AutoDock and AutoGrid is documented in the README file.
```

	▪ AutoDock 4.2.6 User Guide (PDF, 2.2 MB) ▪ Examples (11MB) md5sum 44148341b0b7f97c894a19b45bbb4210

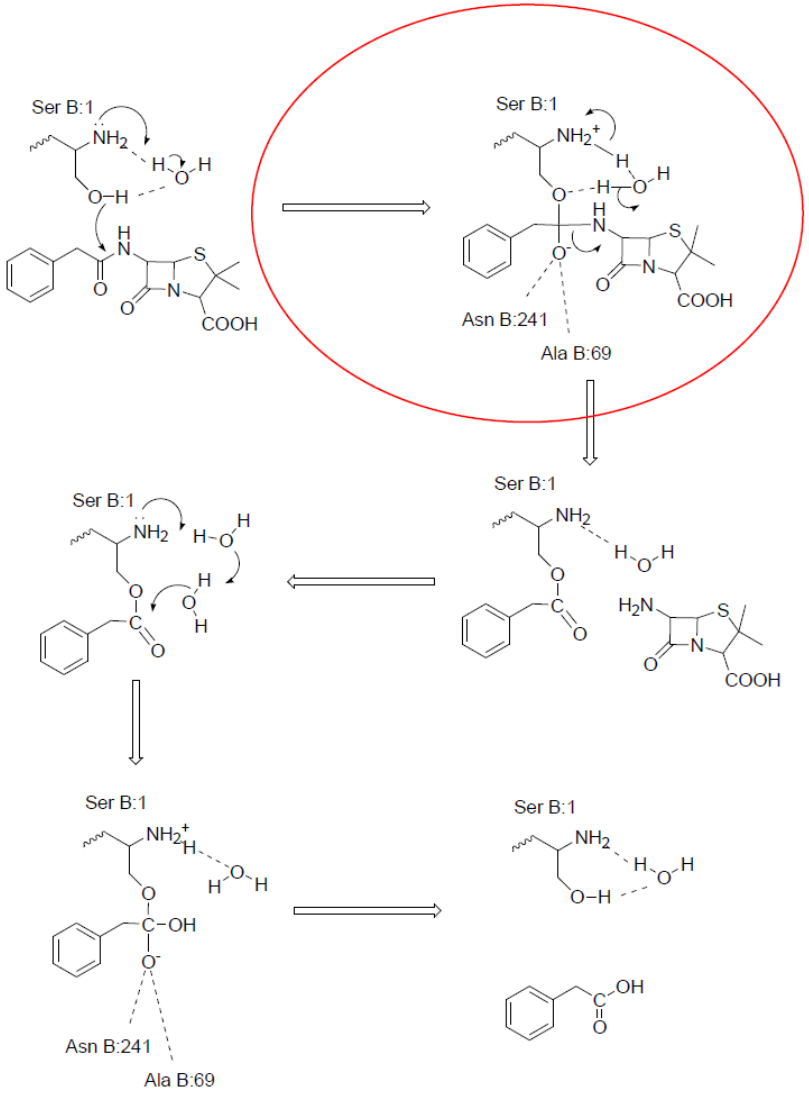
	▪ Linux: Intel (32-bit) (667K) md5sum e3b18a7f399525c6edbea4b05f26e850
	▪ Linux: Intel (64-bit) based on command 'uname -r' output:
	  ▪ 2 - Linux: Intel (64-bit) (743K) md5sum 8c175d4f7b9b1529fdf8d3abf9c90772
	  ▪ 3 - Linux: Intel (64-bit) (764K) md5sum 0ff500576d03abd97c8e543af6e99dd2

	▪ Mac OS X including 10.5 (Leopard), 10.6 (Snow Leopard), 10.7 (Lion), 10.8 (Mountain Lion), 10.9 (Mavericks) that works on both 32-bit and 64-bit Power PC (PPC) and Intel (2.3MB) md5sum ce1333e17b53d1c7bd9734adf81d7ca0
	▪ Windows (577K) md5sum 8dacd08691ec206060ba7e84d32a1cfa
	▪ Sun Solaris (Sparc) (808K) md5sum b4662f5023ee68e4da4e6afabe9af7fe
	▪ Source code (35MB) md5sum f4942c8e8c47aca7f3a2ae8794259067

APPLICATIONS

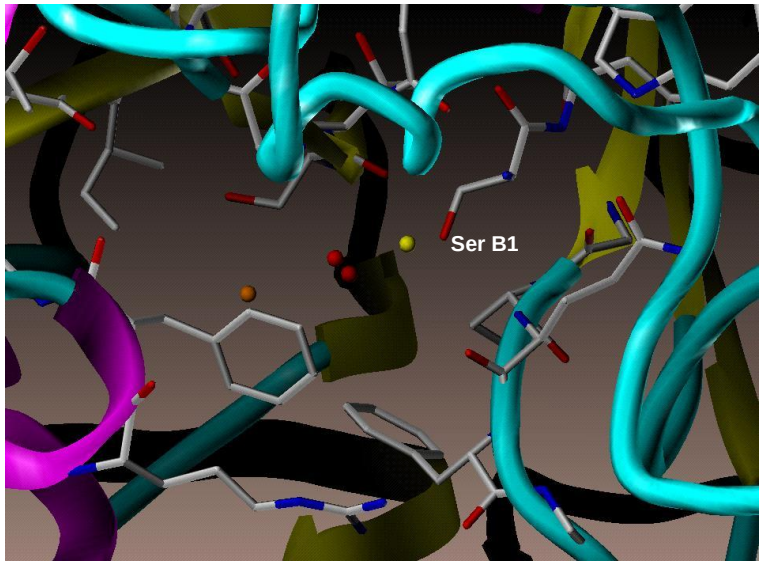
- Binding energy estimation (drugs strength; signal trasduction strength)
- Design and selection of new drugs
- Calculate interactions between molecules such as ligand-receptors; protein-protein, DNA-RNA; DNA-protein; etc.
- Selection and studies of substrates for enzymatic reactions

Mechanism of action of PGA



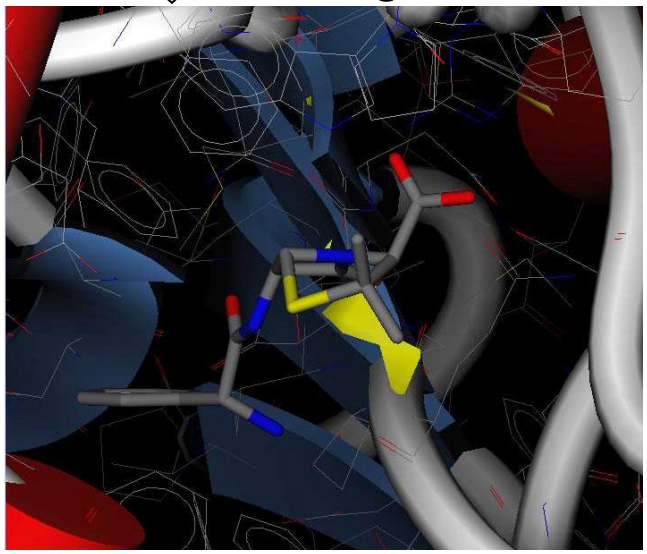
Tetrahedral intermediate

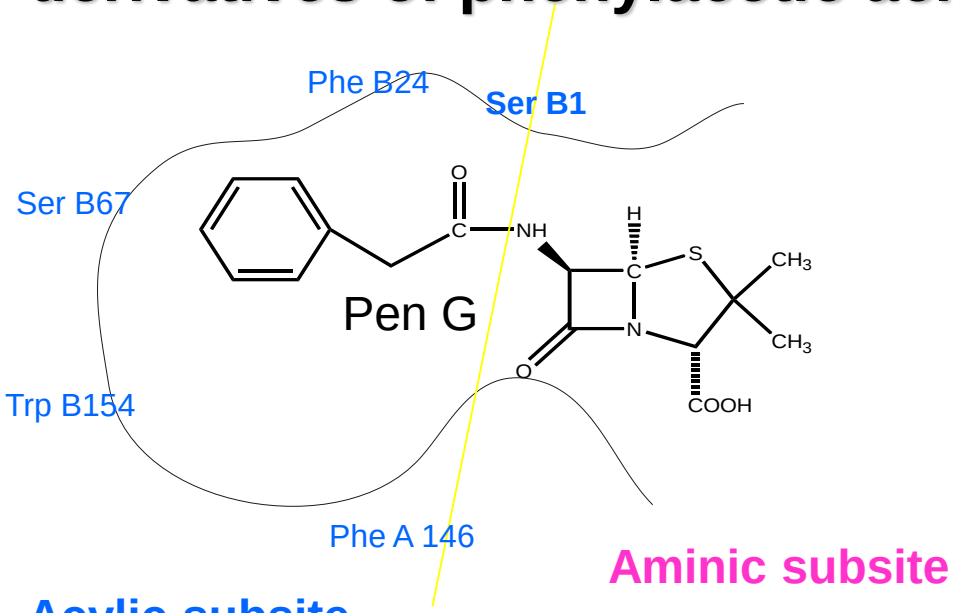
Penicillin G Amidase (PGA) recognizes substrates that are derivatives of phenylacetic acid



active site 3D structure with water molecules

Docking the substrate





Docking simulates enzyme-substrate recognition: affinity, Km

What is not accounted by docking?

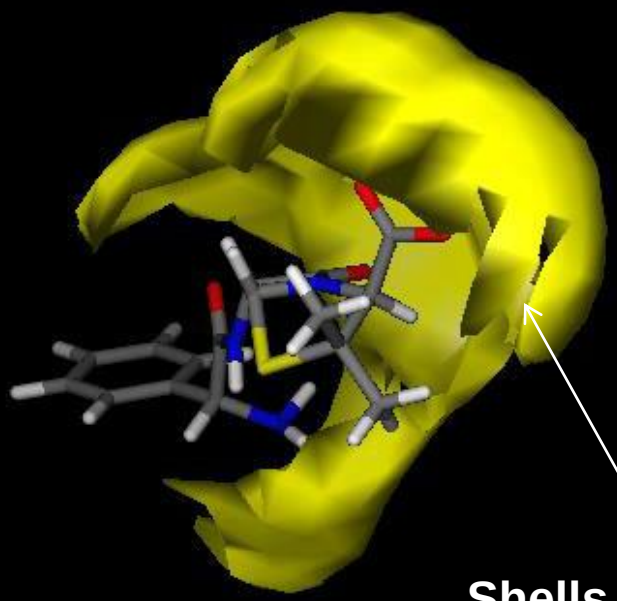
Physical chemical factors mostly related to solvation, desolvation, partition, diffusion

-***In the active site of the enzyme***
-***On the surface of the enzymes***

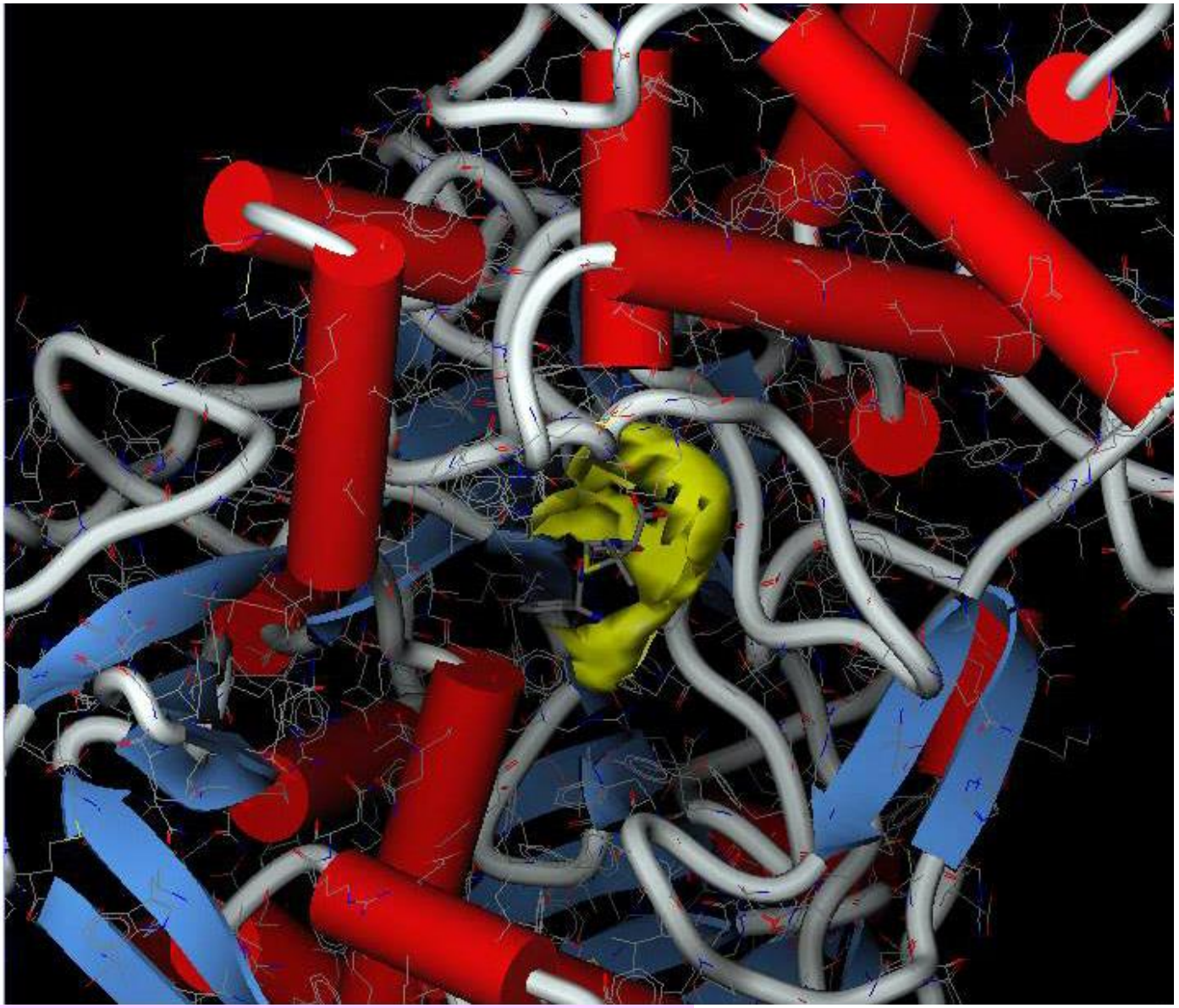
This is particularly relevant when enzymes work as heterogeneous catalysts, suspended in the medium

1.

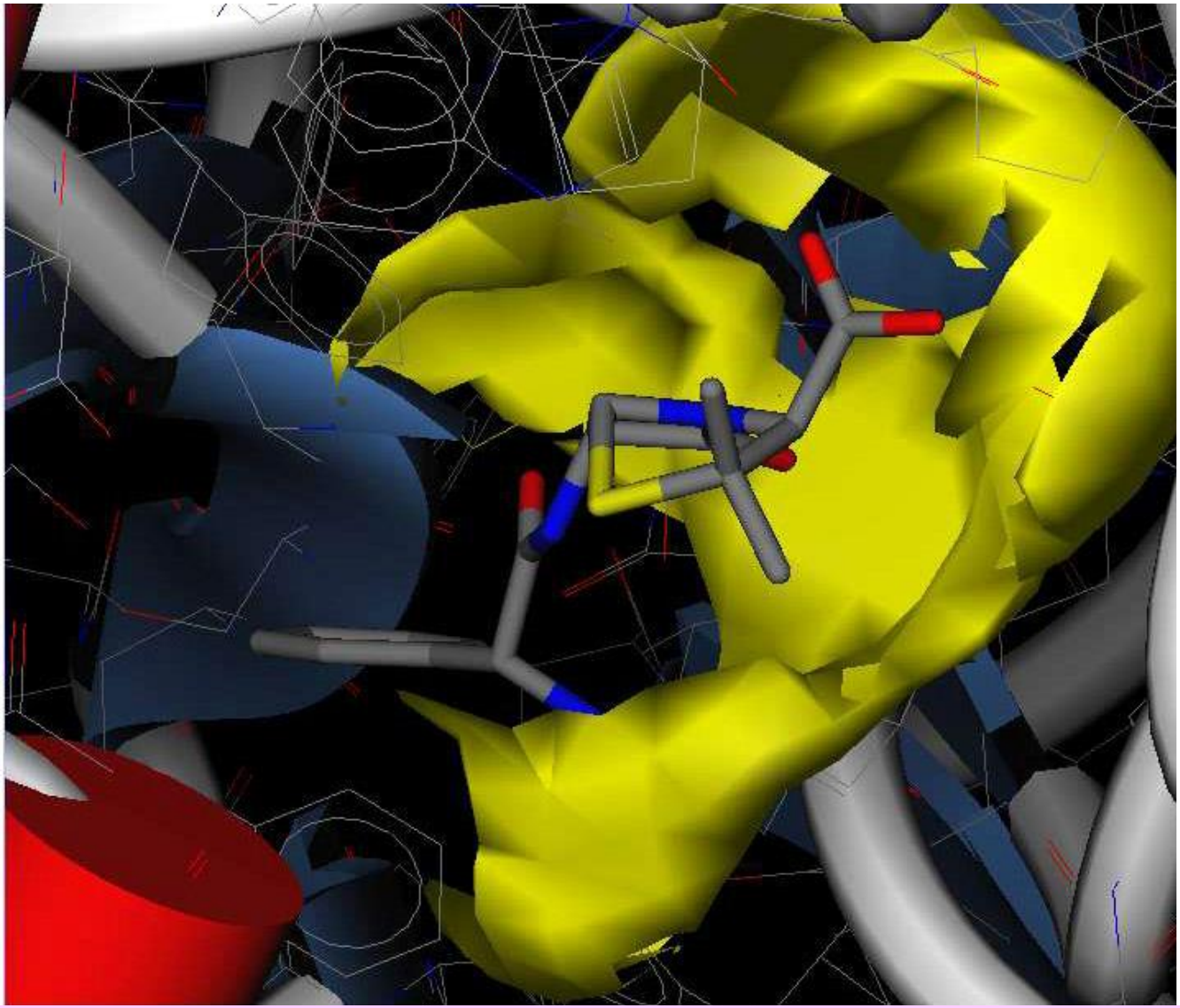
Penicillin before enzyme-substrate recognition



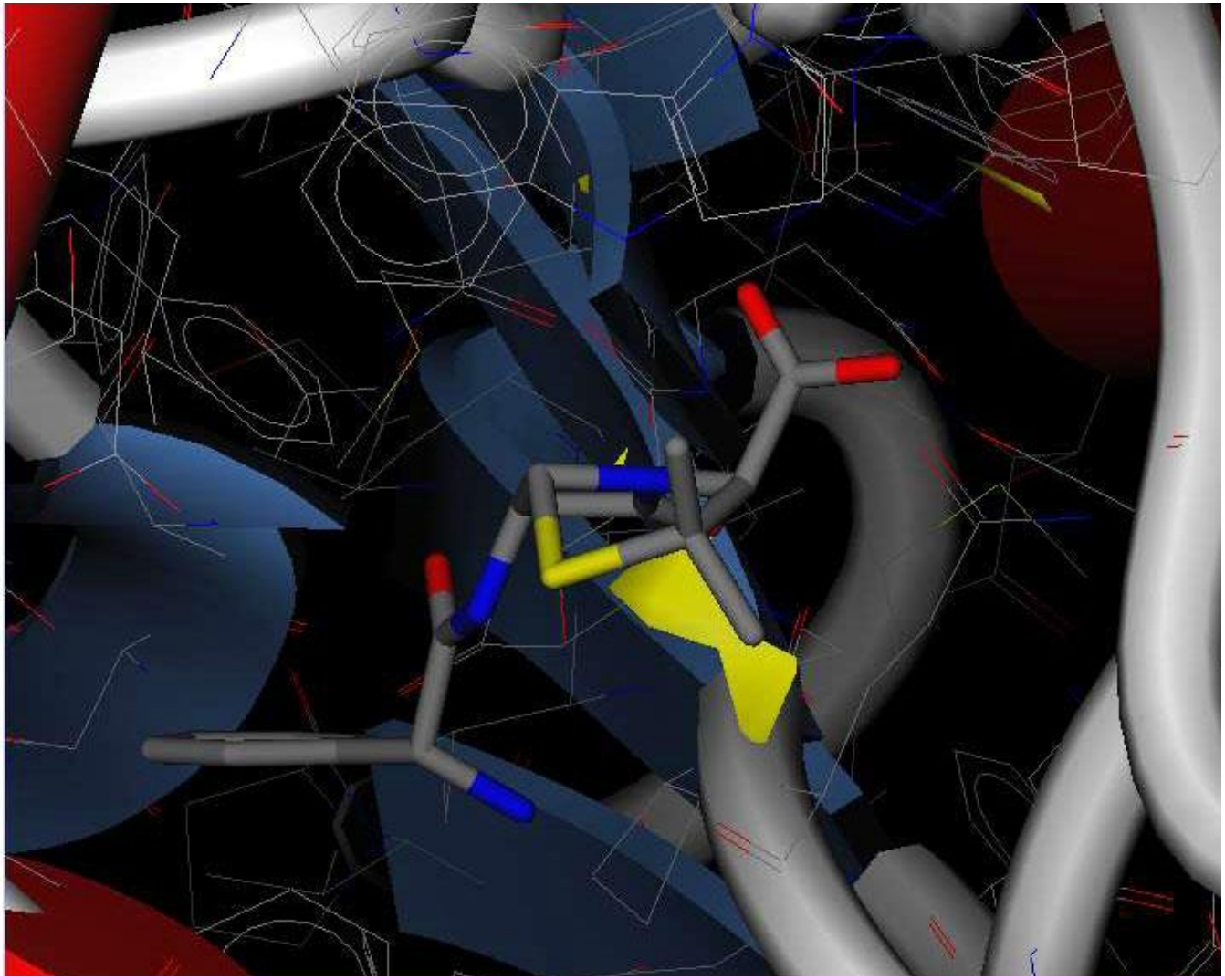
Shells of water solvating the substrate in the aqueous bulk medium



2. The substrate enters the active site

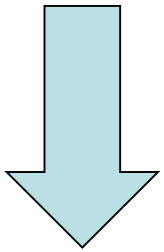


3. To enable enzyme substrate interactions/recognition, part of the water molecules must be displaced (desolvation)

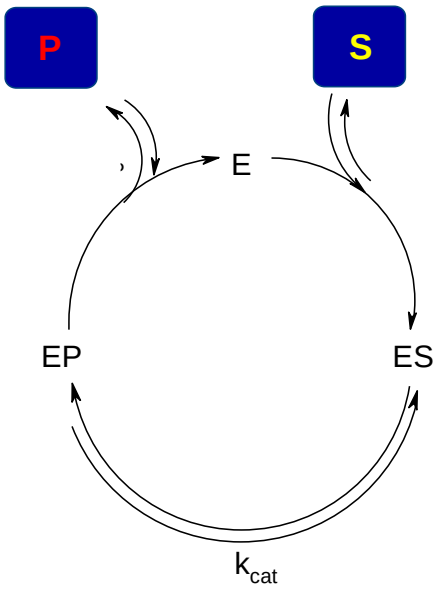


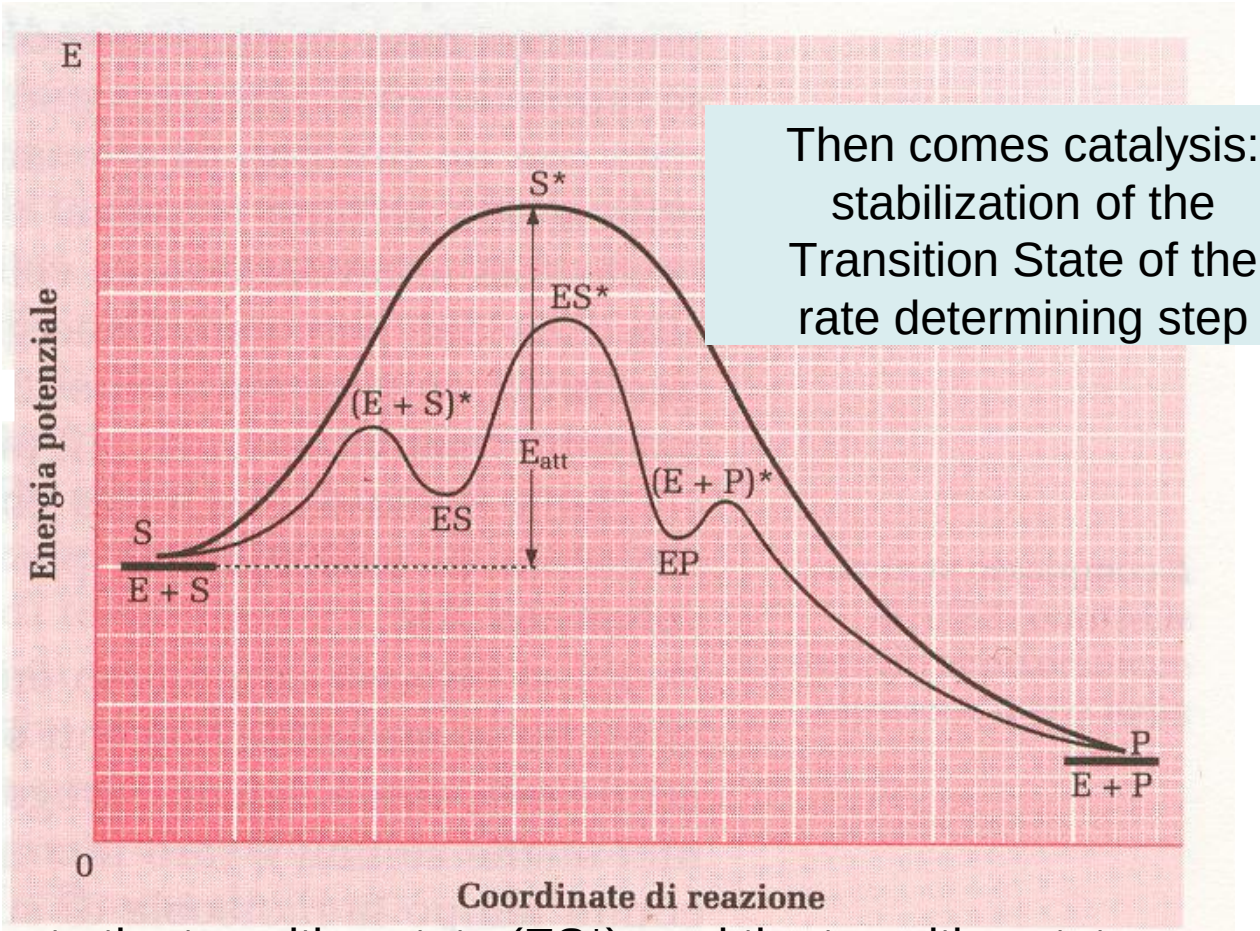
4. After displacement of water molecules, the substrate establish new electrostatic interactions with residues in the active site

Then comes catalysis: stabilization of the Transition State of the rate determining step



Kinetics !!!!



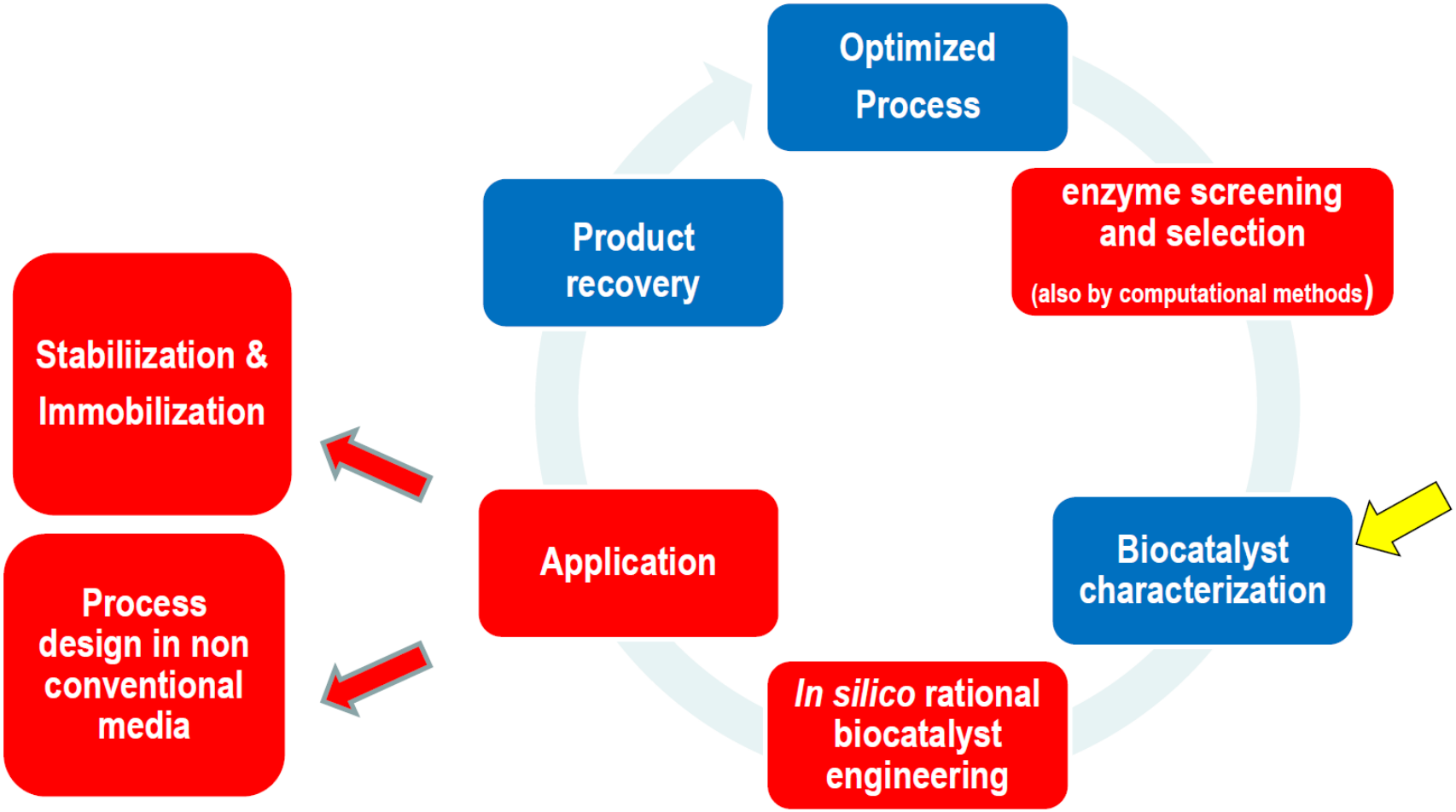


Then comes catalysis: stabilization of the Transition State of the rate determining step

ES complex must pass to the transition state (ES*); and the transition state complex must advance to an enzyme product complex (EP). The latter is finally competent to dissociate to product and free enzyme. The series of events can be shown as:

$$E + S \rightleftharpoons ES \longrightarrow ES^* \rightarrow EP \rightleftharpoons E + P$$

Enzymes are characterized by enzymologists in diluted buffers, therefore in conditions quite distant from the operational conditions in industry



Most often biocatalysts are used under non phisiological conditions, the so called:

non conventional media

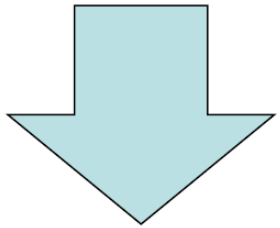
.....anything different from a dilute aqueous solution.

*Most often a **multi-phase system**, some examples:*

Neat substrates + Native enzyme + Traces of solvent

Neat substrates + Immobilized enzyme

The point of view of kinetic studies in fundamental enzymology: enzyme solubilized in buffer



Most often models developed by enzymologists cannot be used in applied biocatalysis

Mass transfer and diffusion limitations are most often rate determining

Parameters and equation refer to simplified ideal conditions

Equazione di Michaelis – Menten (1913)

Leonor Michaelis
(1875–1949)

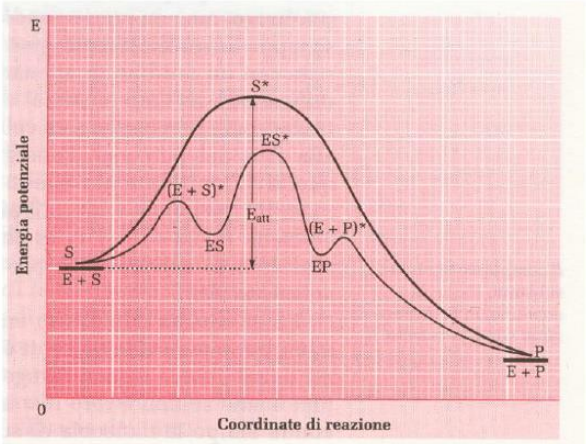
Maud Leonora Menten
(1879–1960)

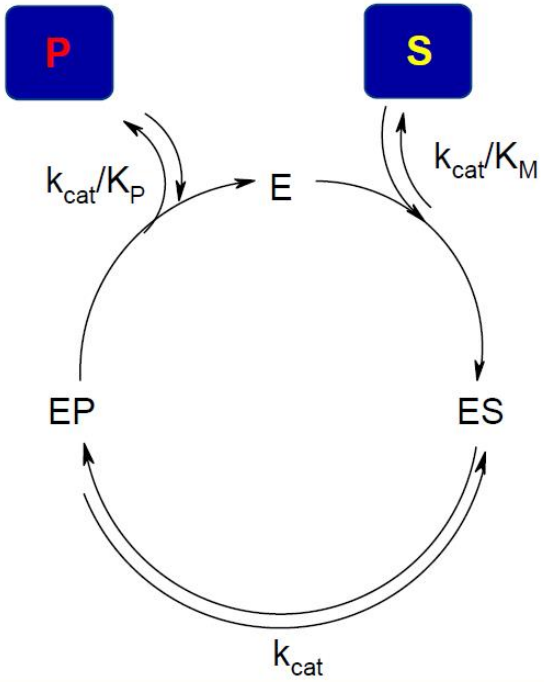
$$V = k_{cat}[E]_T \frac{[S]}{K_S + [S]}$$

oppure

$$v = V_{max} \frac{[S]}{K_M + [S]}$$

Parameters that describe enzyme efficiency in biocatalysis





$E + S \rightleftharpoons ES \longrightarrow ES^* \longrightarrow EP \rightleftharpoons E + P$

Specific activity: k_{cat}

k_{cat} (time^{-1})

First order rate constant for the conversion of ES, ES*, EP

All species involving E

What is really relevant in biocatalysis? Turnover number: biocatalyst productivity

In catalysis it refers to:

the number of moles of product formed per mole of catalyst over the reaction period.

dimensionless ratio:

(mol product) / (mol enzyme)

Specificity

k_{cat}/k_m specificity constant

Specific activity

affinity

K_{cat}/K_M must be calculated at low concentration of Substrate, when the enzyme is not saturated

Penicillin G amidase from different microorganisms display different specificity towards pen G

Enzima	Substrato	T(°C)	pH	k_{cat}/K_M ($M^{-1}s^{-1}$)
Penicillina amidasi				
E. coli	Penicillina G	25.0	7.8	4 800 000
A. faecalis	Penicillina G	25.0	7.8	10 000 000
K. citrofila	Penicillina G	25.0	7.8	3 000 000

Selectivity

k_{cat}/k_m specificity constant

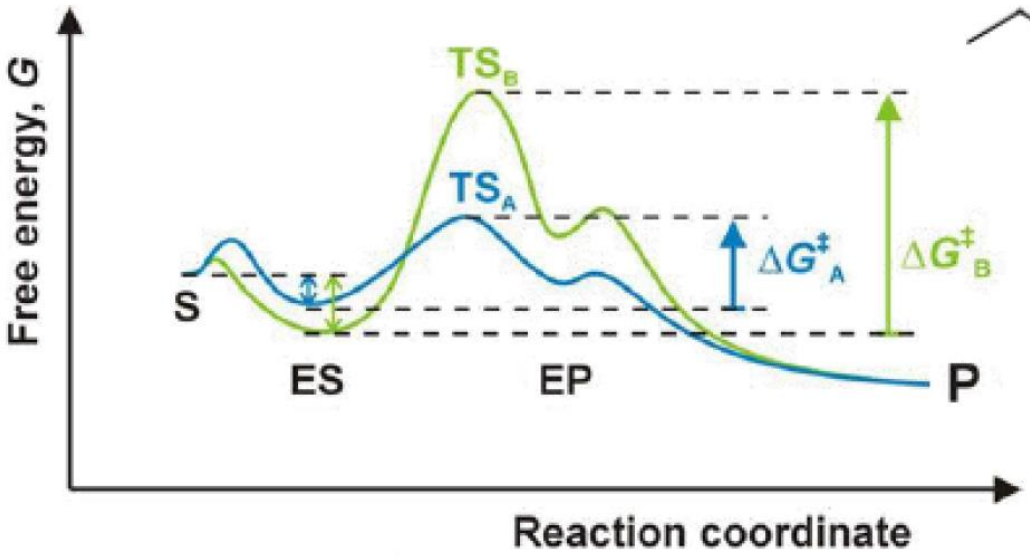
selectivity $\dfrac{(k_{cat}/k_m)_A}{(k_{cat}/k_m)_B}$

Ratio between the specificity constants

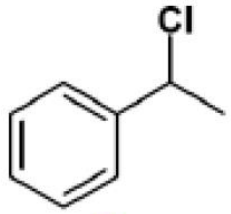
How the enzyme is able to discriminate between A and B

(chemo-regio-enantio-selectivity)

Enzyme selectivity: substrate A *vs* B



A



B

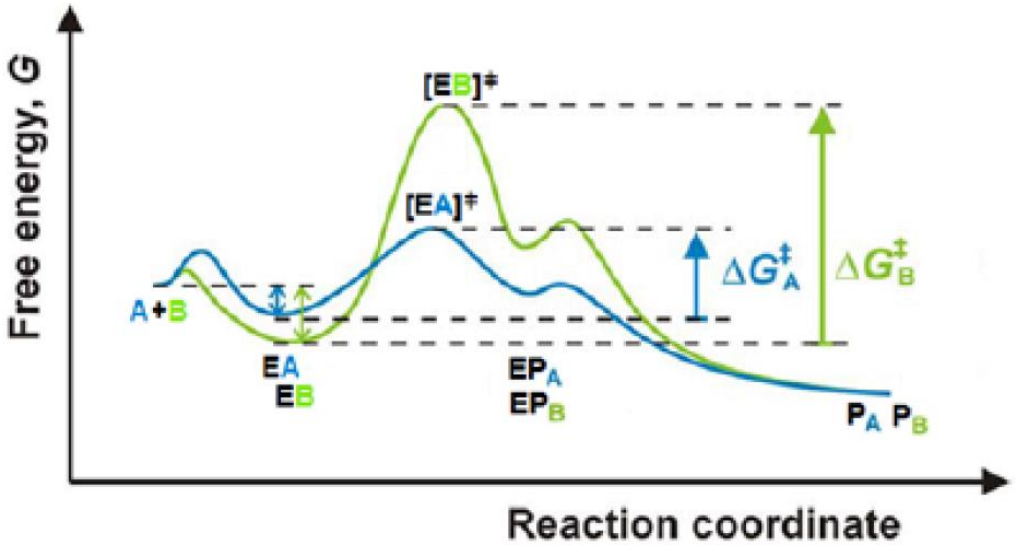
selettività: $\dfrac{(k_{cat}/k_m)_A}{(k_{cat}/k_m)_B}$

Selectivity depends on the ΔΔG activation of the RATE DETERMING STEP!

$\Delta\Delta G^{\#}$ di 5,71 kJmol^{-1} *v* x 10

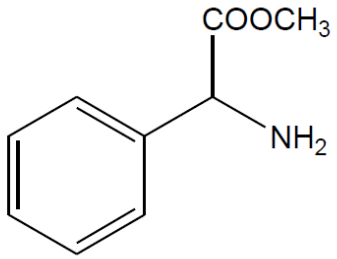
$\Delta\Delta G^{\#}$ di 34 kJmol^{-1} *v* x 10^6

Enzyme enantioselectivity: enantiomer S *vs* R

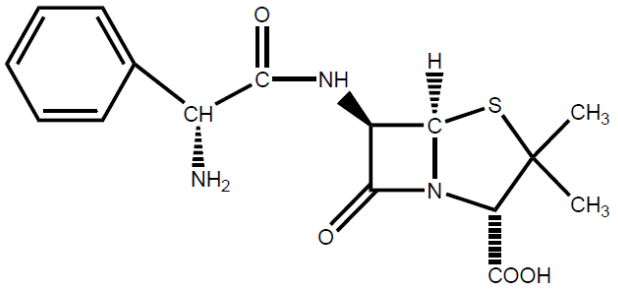


$\Delta\Delta G^{\neq}$ [kcal/mol]	v_A/v_B	e.e. [%]
0.118	1.2	10
0.651	3	50
1.74	19	90
2.17	39	95
3.14	199	99
4.50	1,999	99.9

Example:
enantioselectivity of different hydrolases against R or S phenylglycine (R) used in ampicillin synthesis)



Phenylglycine

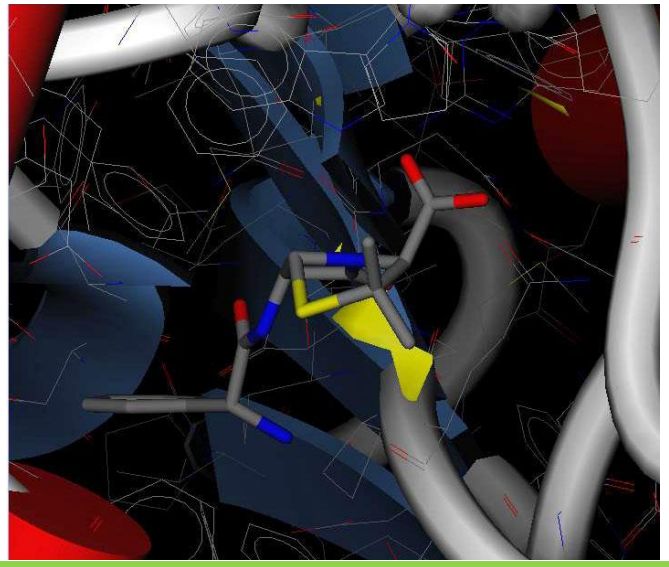


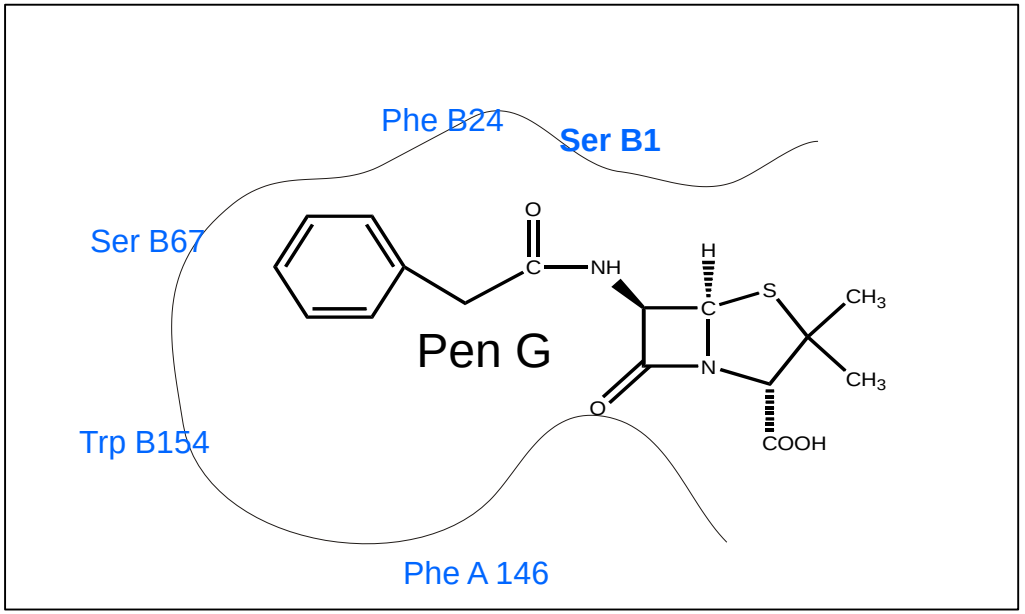
Only the R-enantiomer used for ampicillin semi-synthesis

Substrato	Enzima	k_{cat}/K_M ($M^{-1}s^{-1}$)	Stereoselettività $(k_{cat}/K_M)_S/(k_{cat}/K_M)_R$
(*S*)-fenilglicil↓OMe (*R*)-fenilglici↓OMe	Penicillina amidasi da *E. coli*	550 1100	0.5
(*S*)-fenilglicil↓OMe (*R*)-fenilglicil↓OMe	α-Chimotripsina (bovino)	9.2 0.57	14
(*S*)-fenilglicil↓OMe (*R*)-fenilglicil↓OMe	Proteinasi K da *Tritirachium album*	0.6 0.3	2

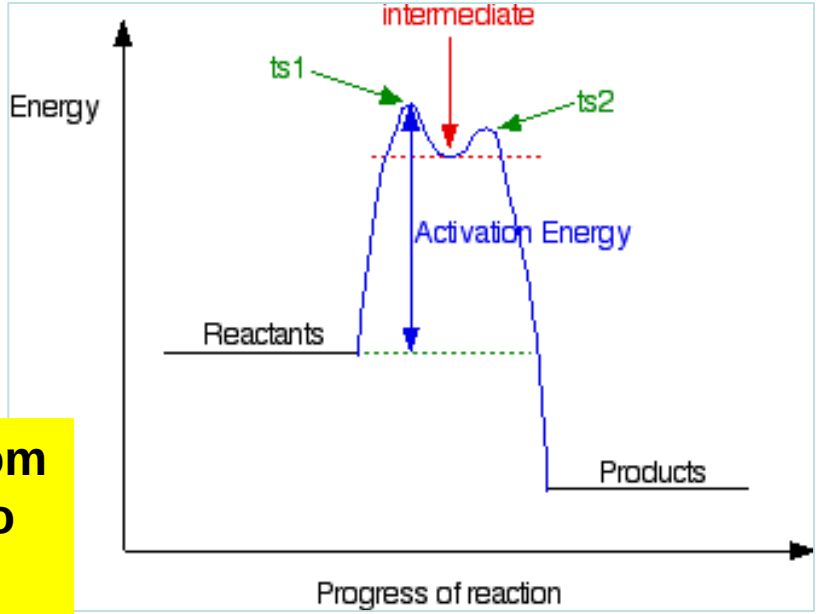
Different hydrolases are able to hydrolyze the methyl ester of (R) and (S) phenylglicine but with different enantioselectivity

Can we use computational methods to predict specific activity or selectivity?



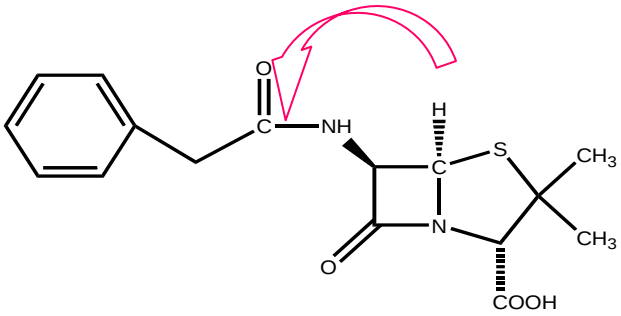


Docking of the ground state gives information only on enzyme-substrate recognition: affinity, Km



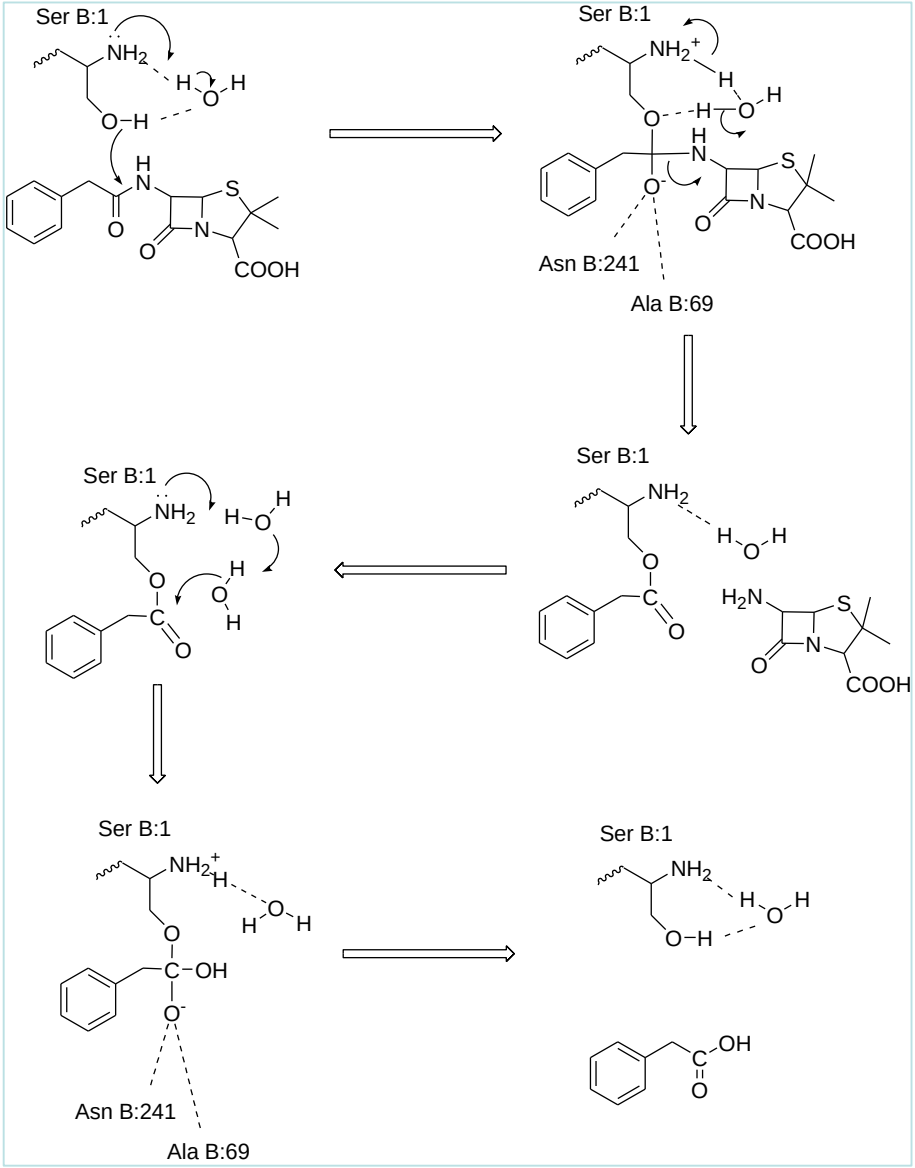
Information on Kcat derives only from the study of energies associated to the transition state of the rate determining step of the reaction

Example:
The enzymatic hydrolysis of penicillin G

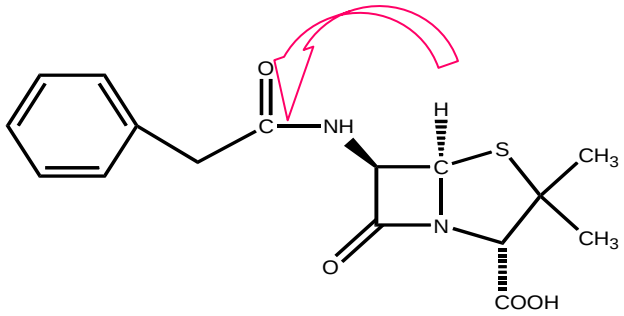


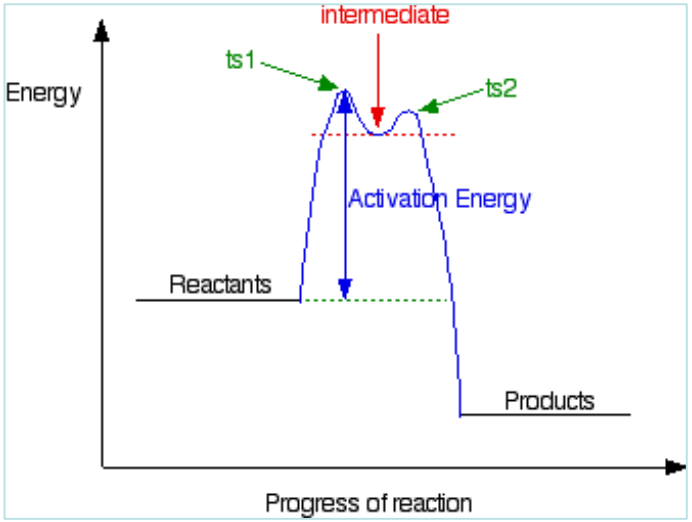
PGA is a serin hydrolase where the catalytic Ser corresponds to the N terminal residue

Generally, the formation of the acyl-enzyme is assumed as the rate determining state of the reaction



Example:
The enzymatic hydrolysis of penicillin G:



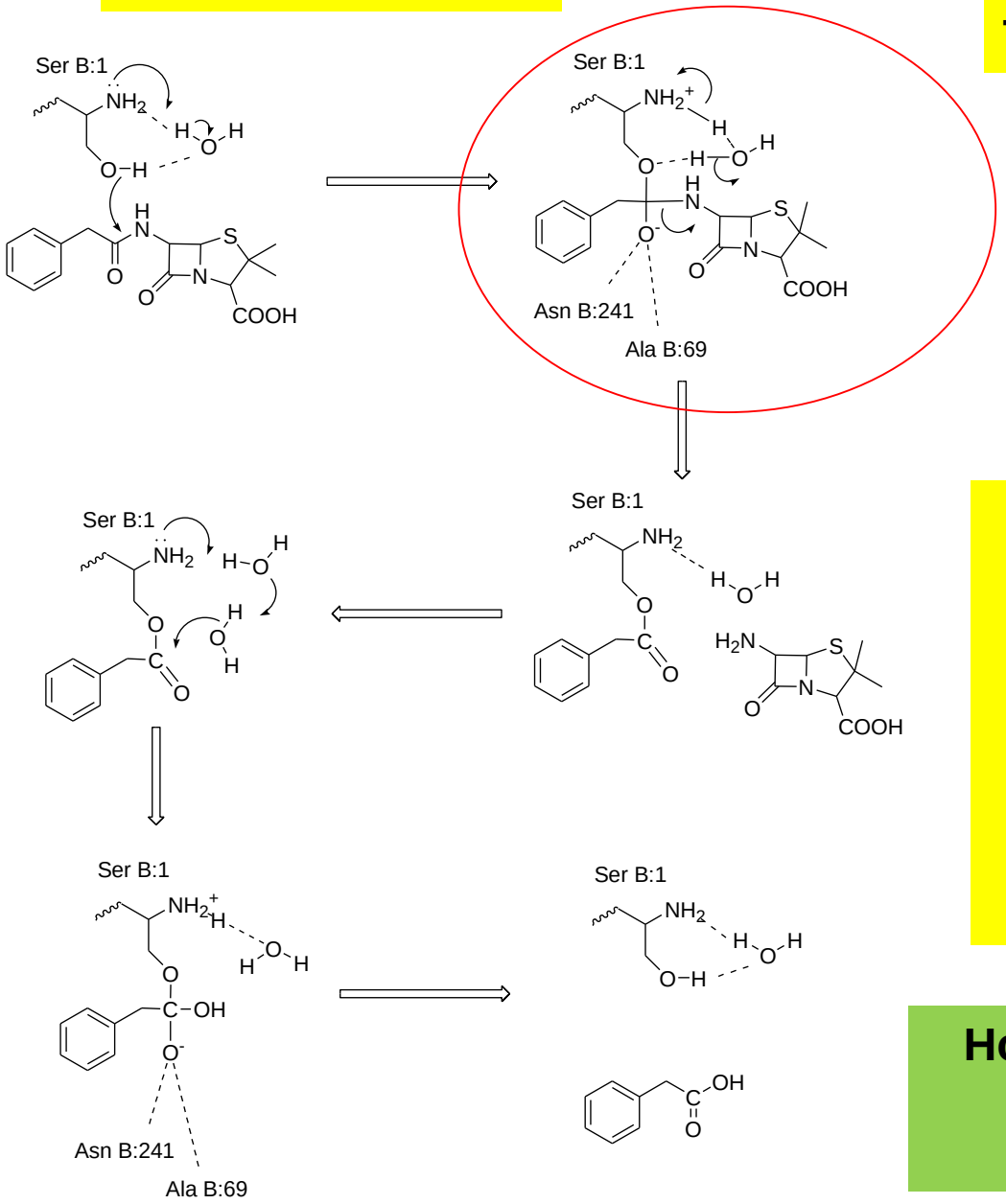


Transition state can be simulated only using Quanto Mechanics methods: they are able to simulate partially formed or broken bonds.

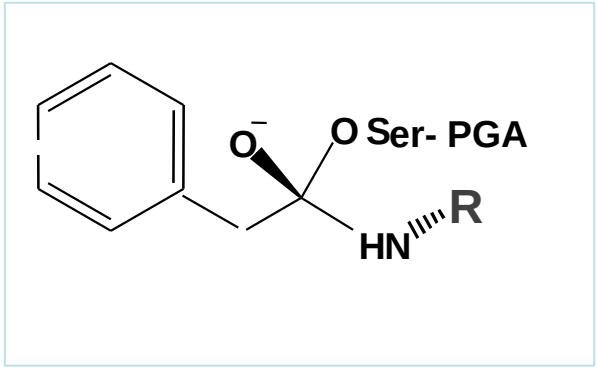
However the transition states of enzymatic reactions are too complex and big to be simulated by QM

Basso, A., et al., *Biochim. Biophys. Acta – Protein structure and Molecular Enzymology* 2002,1601, 85-92

Mechanism of PGA



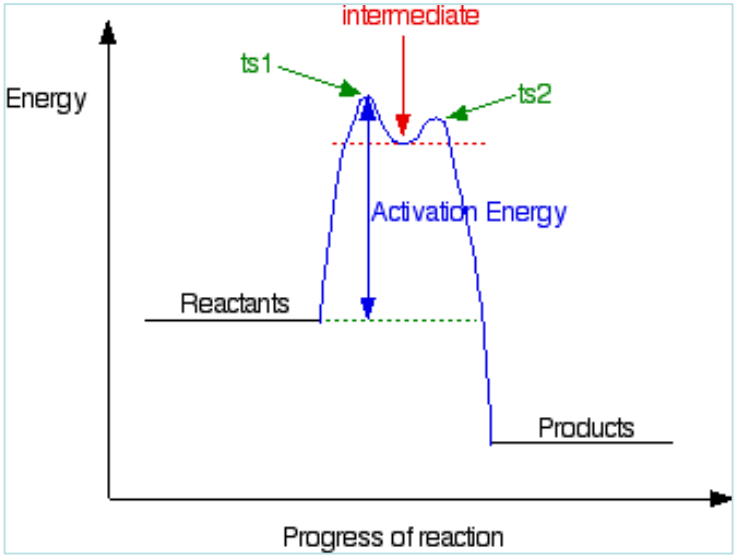
Tetrahedral intermediate



We can use a simplified approach by simulating a reaction intermediate that we suppose to be close to the transition state of the RATE DETERMINING STATE of the reaction

However: the rate determining step must be identified unequivocally !!

Construction of the tetrahedral intermediate in the active site of the enzymes



Covalent bond with Ser B:1

↓

Conformational analysis

↓

$$\Delta E = E_{inter.} - (E_{enzyme} + E_{substrates})$$

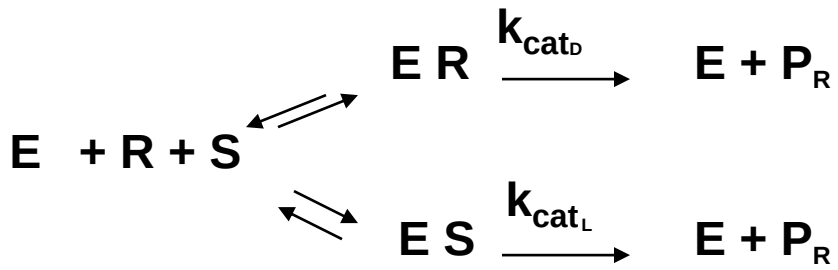
↓

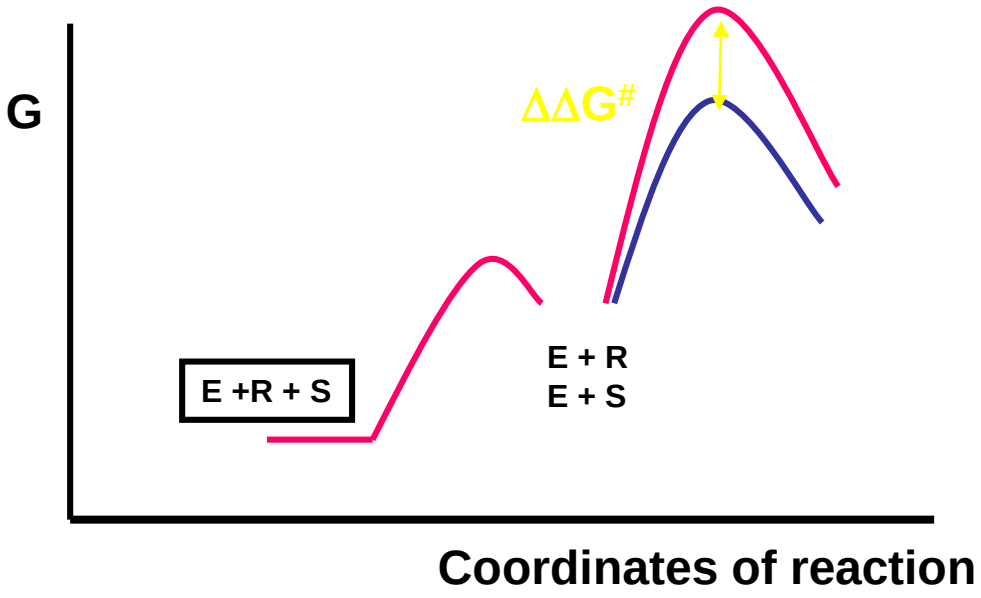
- Ony enthalpic contributions
- No entropy accounted
- No quantitative correlation to Kcat

Basso, A., Braiuca, P., Ebert, C., Gardossi, G., Linda, P., *Biochim. Biophys. Acta – Protein structure and Molecular Enzymology* 2002,*1601,* 85-92

The method can provide ONLY some semiqualitative hints on enzyme selectivity by comparing the enthalpy associated to two different tetrahedral intermediates

Enantioselectivity: $\dfrac{(k_{cat}/k_m)_S}{(k_{cat}/k_m)_R}$





$$\Delta\Delta G^{\#} = \Delta\Delta H^{\#} - T\,\Delta\Delta S^{\#}$$

Annexes

Equazione di Michaelis – Menten (1913)

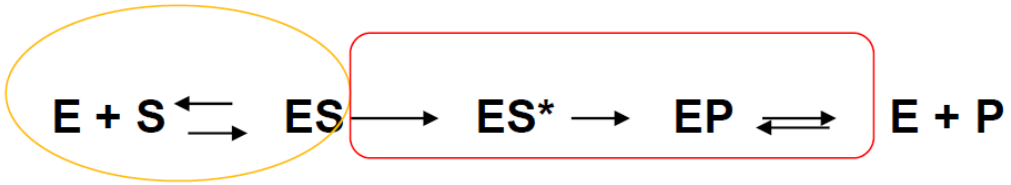
Leonor Michaelis
(1875–1949)

Maud Leonora Menten
(1879–1960)

$$V = k_{cat}[E]_T \frac{[S]}{K_S + [S]}$$

oppure

$$v = V_{max} \frac{[S]}{K_M + [S]}$$



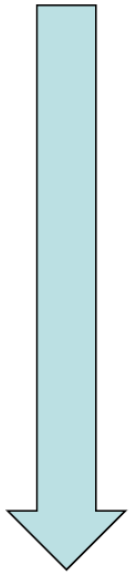
Cinetica semplificata

- Un enzima catalizza la reazione che dal substrato **S** porta al prodotto **P** attraverso il complesso enzima-substrato **ES** e il complesso enzima-prodotto **EP**:

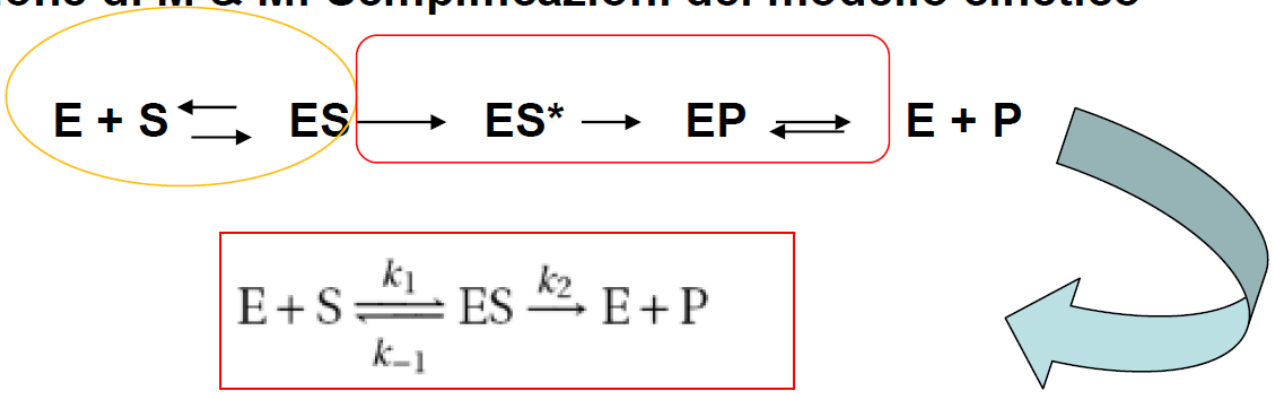
$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \rightleftharpoons \mathrm{EP} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{E} + \mathrm{P}$$

- Si consideri la dissociazione di EP molto rapida e irreversibile; lo stadio successivo:

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_2} \mathrm{E} + \mathrm{P}$$



L'equazione di M & M: Semplificazioni del modello cinetico



1. il primo stadio è rapido e reversibile, si forma il complesso enzima-substrato ES, la cui costante di dissociazione è K_S.

$$K_S = \frac{[E][S]}{[ES]}$$

2. Nel secondo stadio hanno luogo i processi chimici governati da una costante di velocità $k_2 = k_{cat}$

$$V = k_2[ES] = k_{cat}[ES] = \frac{d[P]}{dt} = -\frac{dS}{dt}$$

Come ricavare l'equazione di M & M

Scrivere il bilancio delle masse per ogni specie. Per l'enzima è:

$$[E]_T = [E]_{lib} + [ES]$$

$$[E]_{lib} = [E]_T - [ES]$$

$$K_S = \frac{[E][S]}{[ES]} = \frac{\{[E]_T - [ES]\}[S]}{[ES]}$$

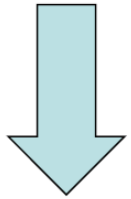
$$[ES]K_S = [E]_T[S] - [ES][S]$$

$$[ES]\{K_S + [S]\} = [E]_T[S]$$

$$[ES] = \frac{[E]_T[S]}{K_S + [S]}$$

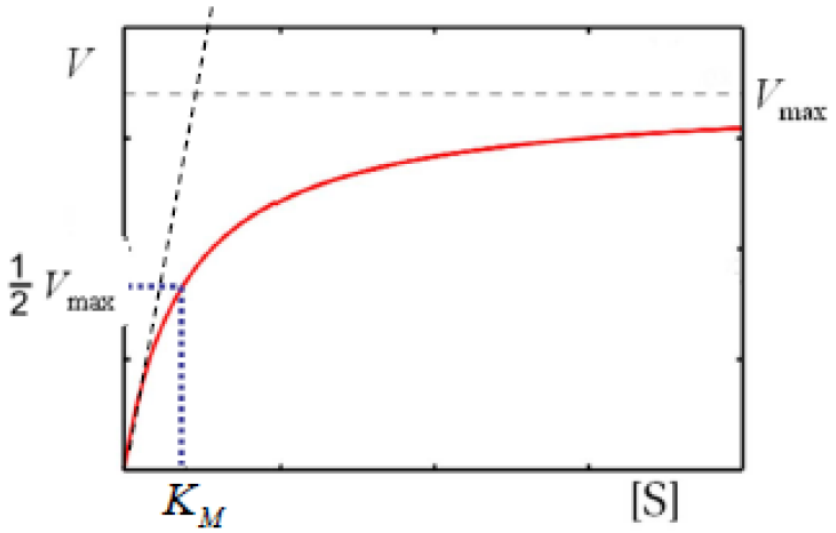
$$V = k_2[ES] = k_{cat}[E]_T \frac{[S]}{K_S + [S]}$$

L'enzima non deve essere saturato dal substrato



- Misurando V a diverse concentrazioni di S si trova la curva seguente:

$$V = k_{cat}[E]_T \frac{[S]}{K_S + [S]}$$



- Ad alte concentrazioni di S, K_S diventa trascurabile e V diventa:

$$V = V_{max} = k_{cat}[E]_T \quad \Rightarrow \quad V = V_{max} \frac{[S]}{K_S + [S]}$$

- Alla condizione nella quale $V = \frac{V_{max}}{2}$ $\Rightarrow$ $V = \frac{V_{max}}{2} = \frac{V_{max}[S]}{K_S + [S]}$

$$[S] = K_S = K_M$$

costante di Michaelis – concentrazione del substrato necessaria a raggiungere una velocità pari a $V = ½\ V_{max}$

$$V = V_{max} \frac{[S]}{K_M + [S]}$$

equazione di Michaelis-Menten

Significati di K_M

$$v = V_{max} \frac{[S]}{K_M + [S]}$$

velocità è indipendente da [S].

- [S] = K_M $V = V_{max}/2$, ossia
  - *K_M rappresenta la concentrazione di substrato che determina metà della velocità massima.*
  - K_M rappresenta pure *la concentrazione di substrato alla quale metà dei siti attivi sono occupati*, dando una misura della [S] richiesta affinché la catalisi avvenga in modo significativo
  - K_M è una misura dell'energia di binding ΔG_b