Biomodeling *Biotech*

by Stefano Moro

Molecular Modeling Section (MMS)

Department of Pharmaceutical and Pharmacological Sciences

University of Padova

Post-translational modifications

The amazing role of post-translational modifications:

1. Phosphorylation;
2. N-acetylation;
3. N- ed O-glycosylation;
4. N-methylation;
5. Nitrosilation;
6. Acylation (palmitoylation, farnesylation, miristoylation);
7. Ubiquitintion e SUMOylation;

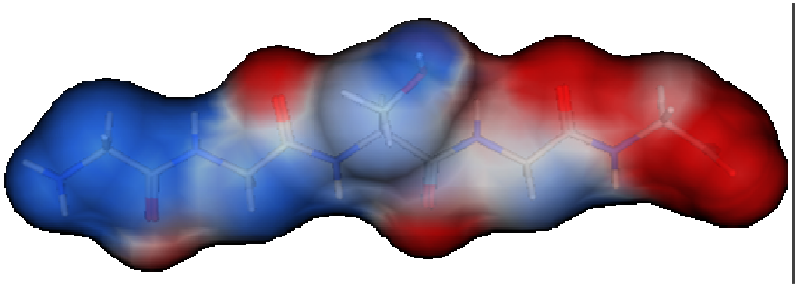
Phosphorylation… *in summary*

Modified AA: Ser, Thr, Tyr
∆PM: 80 Da
∆local charge: neutral → negative
Reversibility: ☺

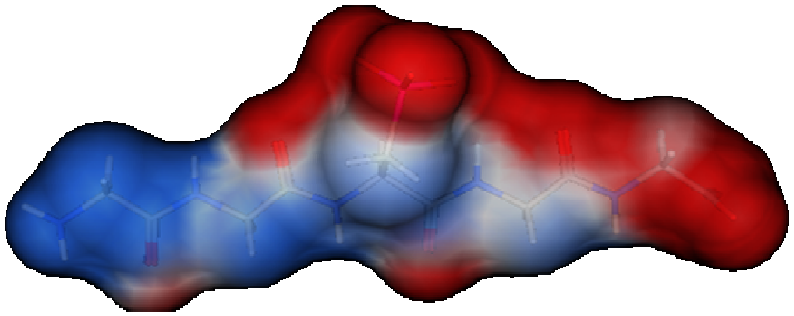
>30% of all proteins are phosphorylated and very often poli-phosphorylated

***Function*:**
Ser, Thr kinase association and activation
Tyr dimerization and receptor activation

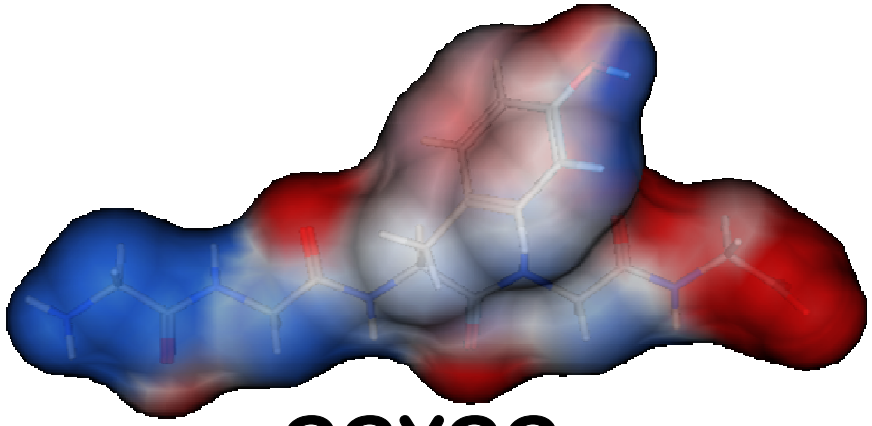
Here is the modification of local surface properties after phosphorylation…



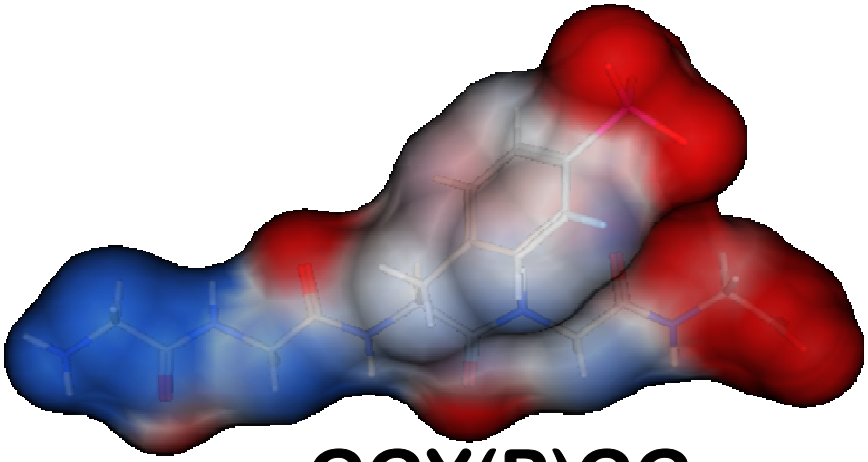
GGSGG



GGS(P)GG



GGYGG



GGY(P)GG

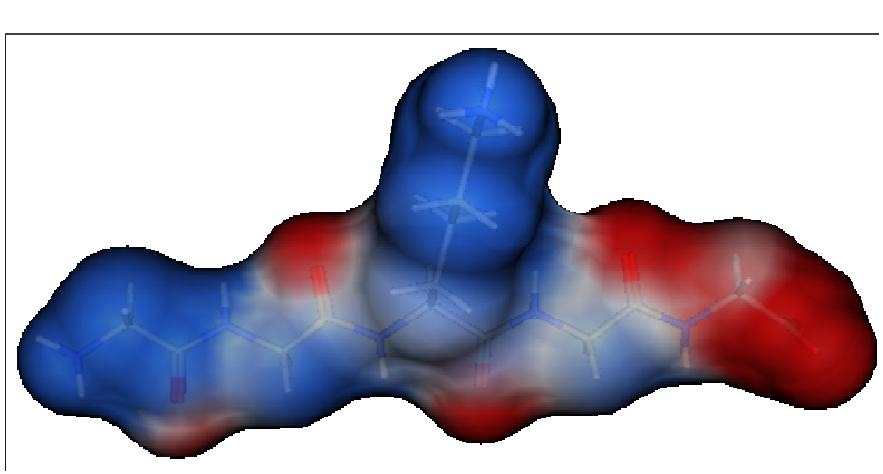
Acetylation… *in summary*

Modified AA: N-ter residue, Lys
ΔPM: 42 Da
Δlocal charge: positive → neutral
Reversibility: ☺

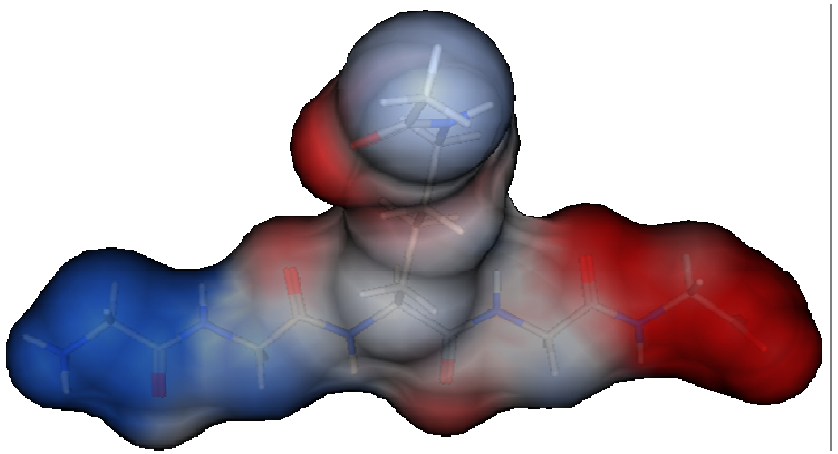
Function:

Acetylation is an important modification of proteins in cell biology; and proteomics studies have identified thousands of acetylated mammalian proteins. Acetylation occurs as a co-translational and post-translational modification of proteins, for example, histones, p53, and tubulins. Among these proteins, chromatin proteins and metabolic enzymes are highly represented, indicating that acetylation has a considerable impact on gene expression and metabolism. In bacteria, 90% of proteins involved in central metabolism of Salmonella enteric are acetylated.

Here is the modification of local surface properties after Lys-acetylation…



GGKGG



GGK(Ac)GG

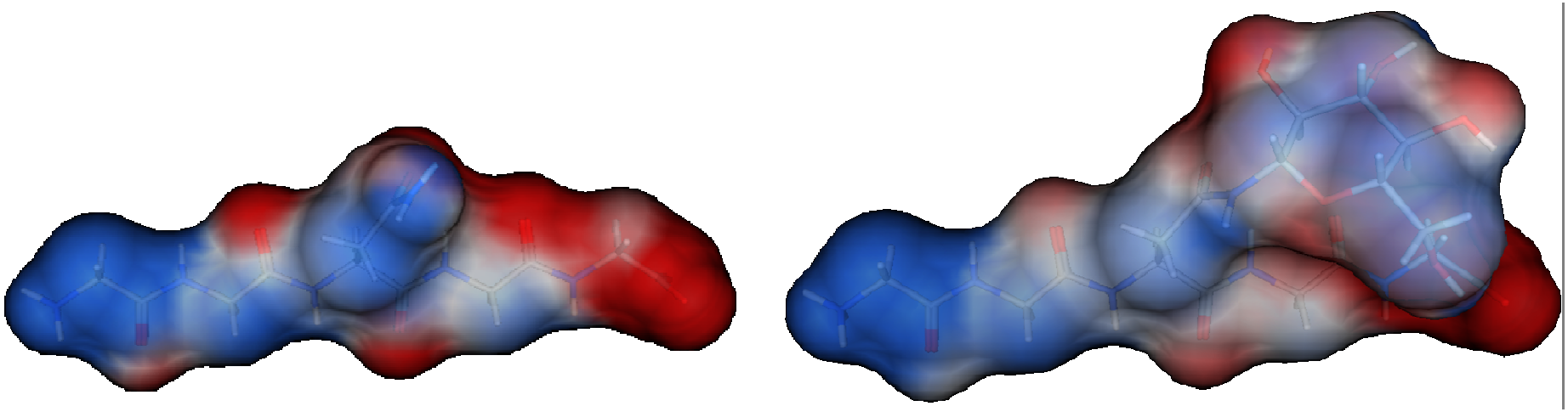
Glycosylation… *in summary*

Modified AA: N-glico, Asp
ΔPM: >800 Da
Δlocal charge: neutral → variable
Reversibility: ☺

Modified AA: O-glico, Ser e Thr
ΔPM: 203 - >800 Da
Δlocal charge: neutral → variable
Reversibility: ☺

Around 50% of human proteins are glycosylated, mainly formed in Golgi and RE.
***Functions*:**
stability and water solubility, recognition signalling,

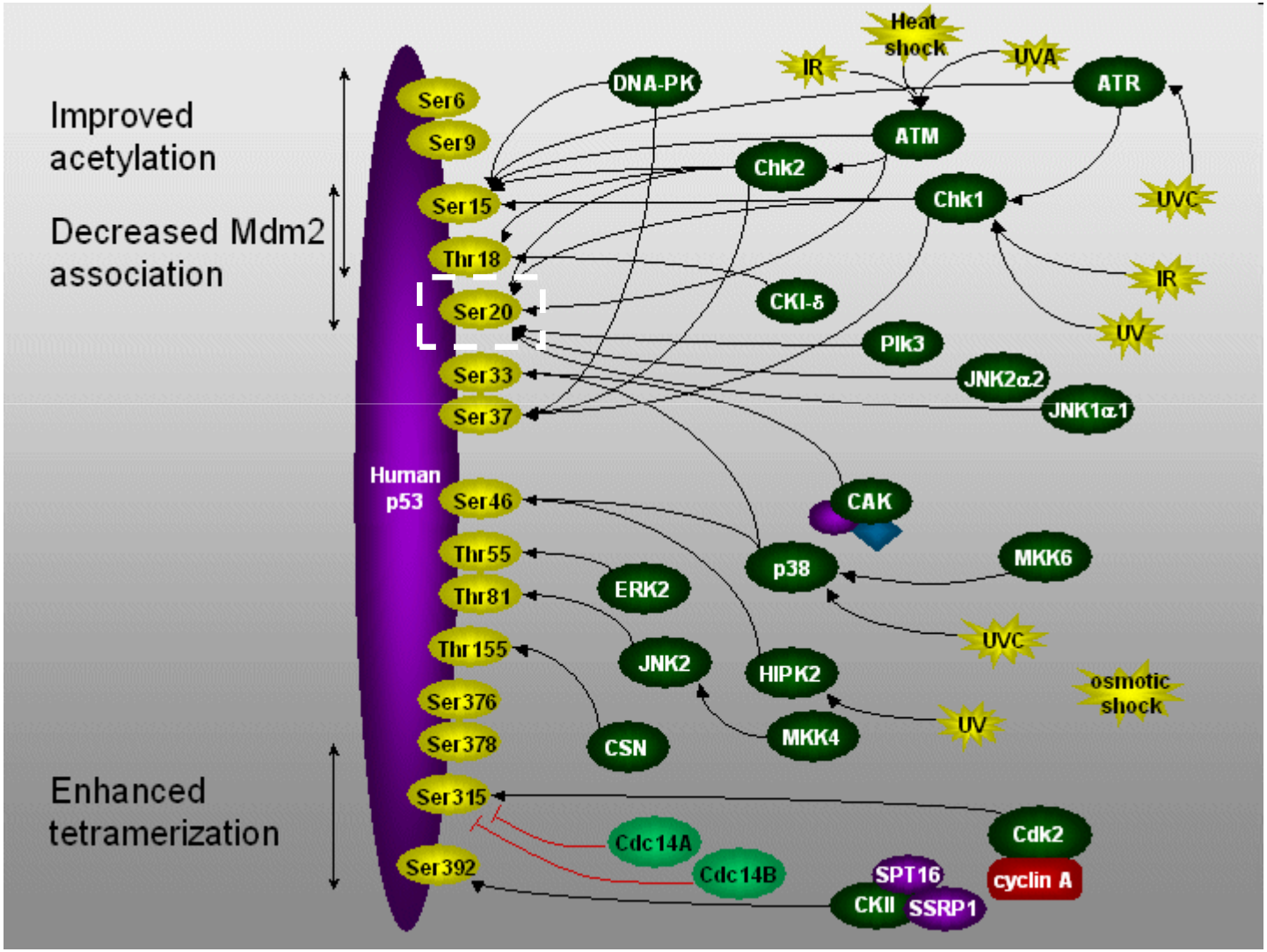
Here is the modification of local surface properties after Gln-glucosylation…



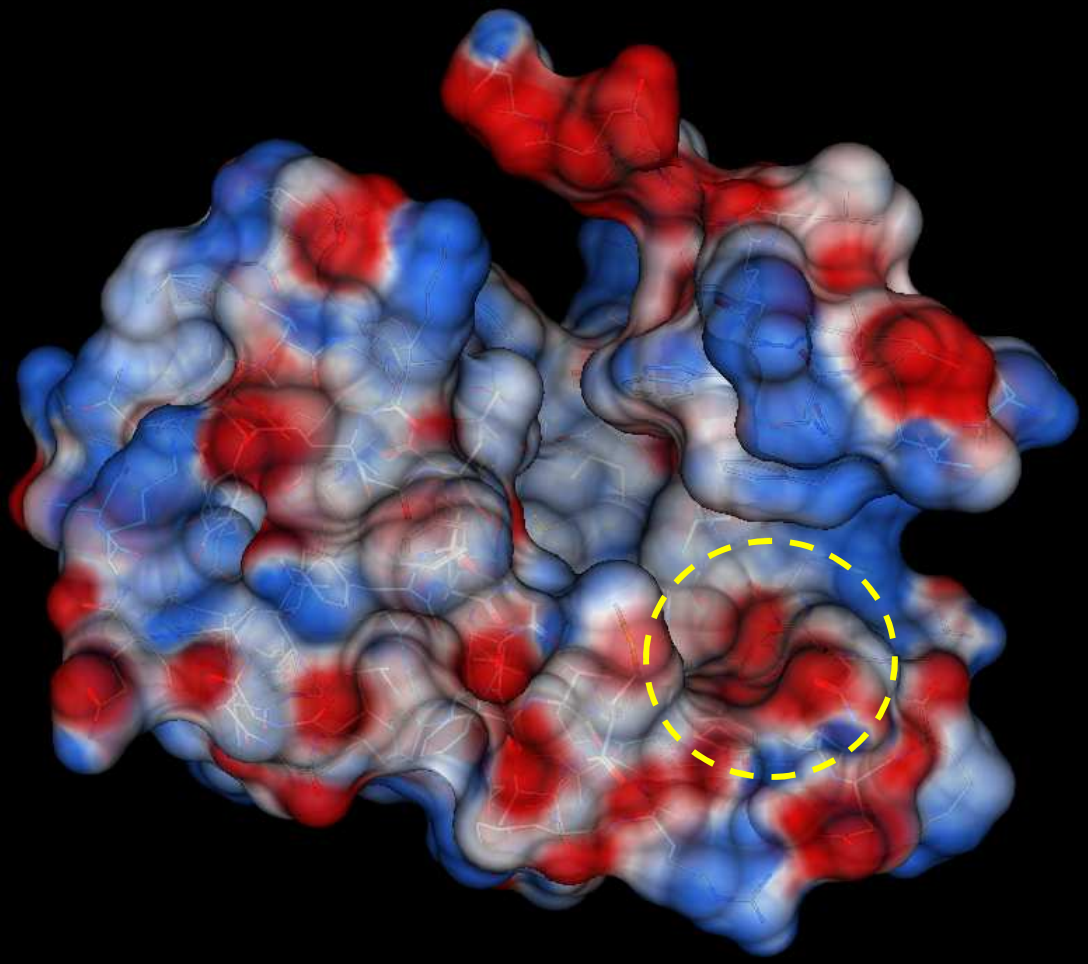
GGNGG

GGN(glu)GG

Back to our p53…



Do you remember… this is the non-phosphorylated p53-MDM2 recognition:



Back at the beginning of the story: let's do Ser20 phosphorylation!

Ser20

Thr18

Chk1/2

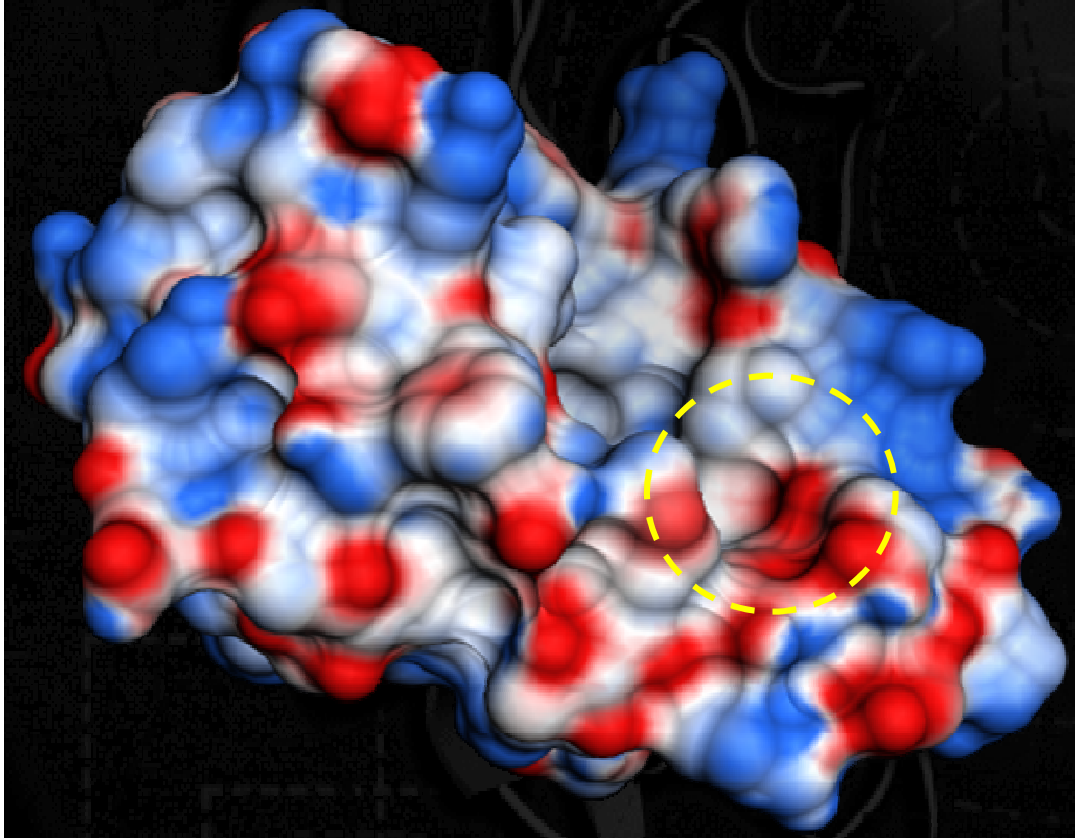
Ser(P)20

Now it is understandable the role of the Ser20-phosphorylation!!!

Now it is very clear!

Steric and Electrostatic repulsion

Now, can you predict the effect of these mutations of p53?



Ser20Asp
Ser20Ala
Ser20Lys

BRAVI!!!

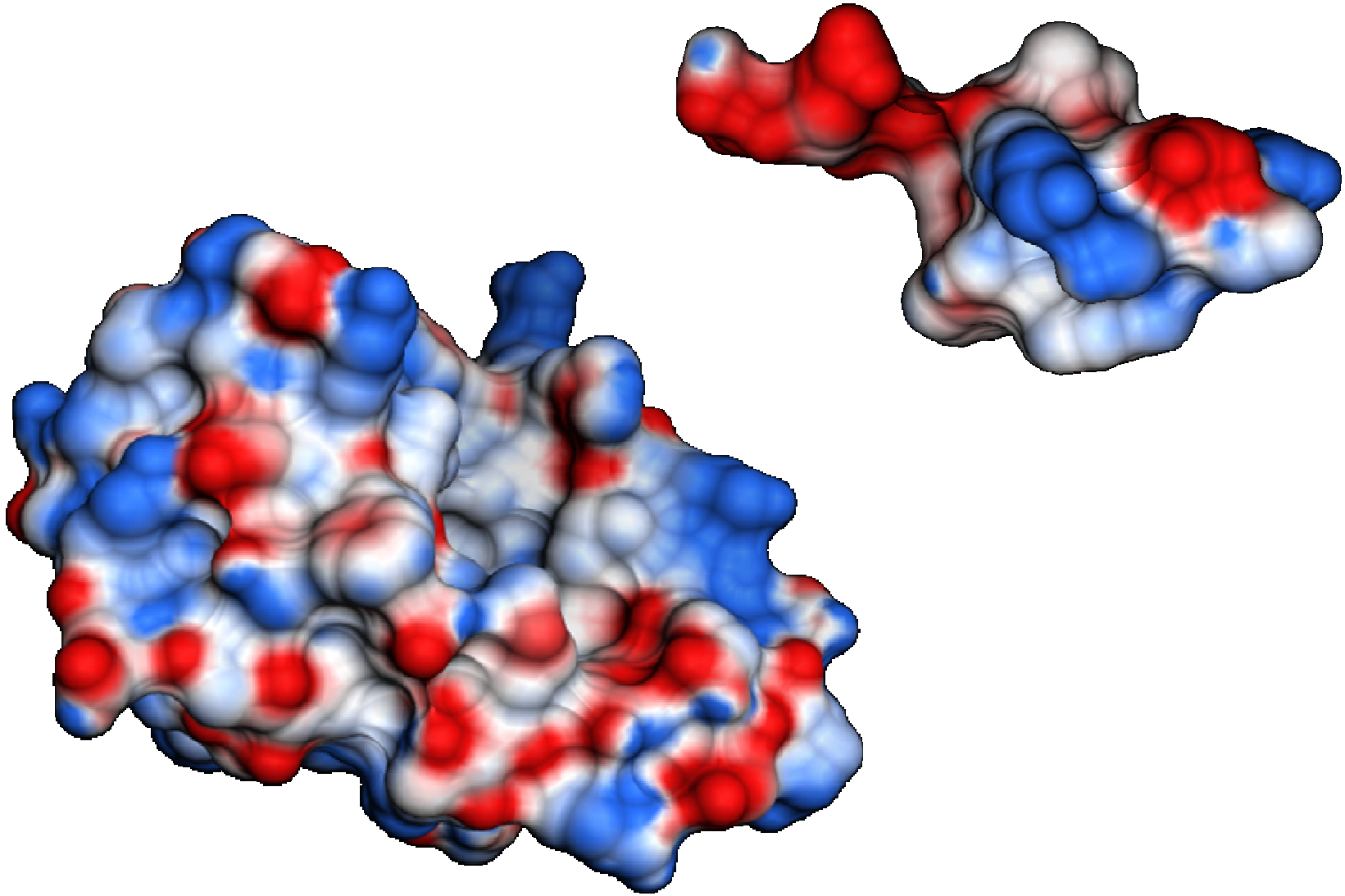
wt

Ser20Asp

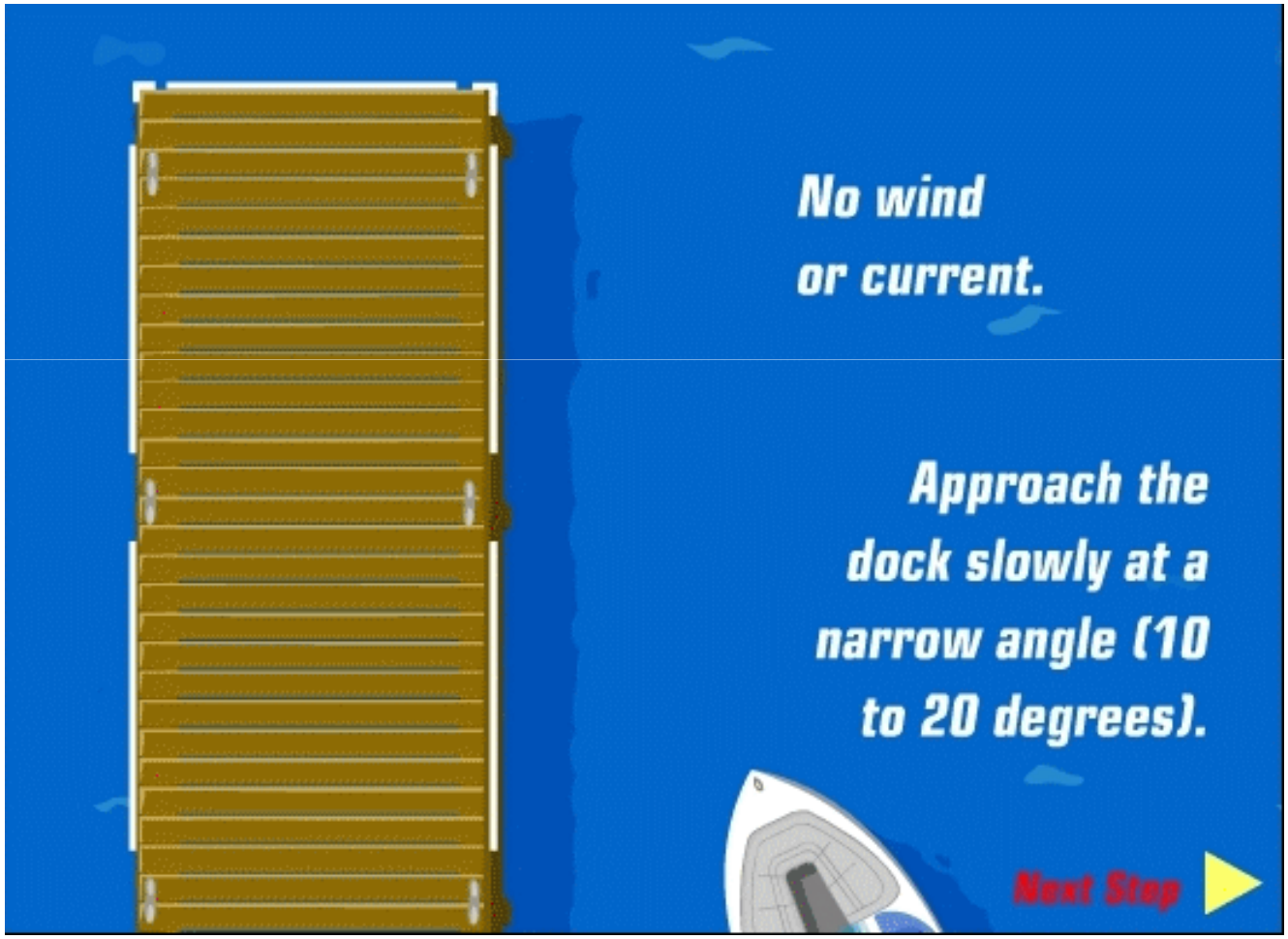
Ser20Ala

Ser20Lys

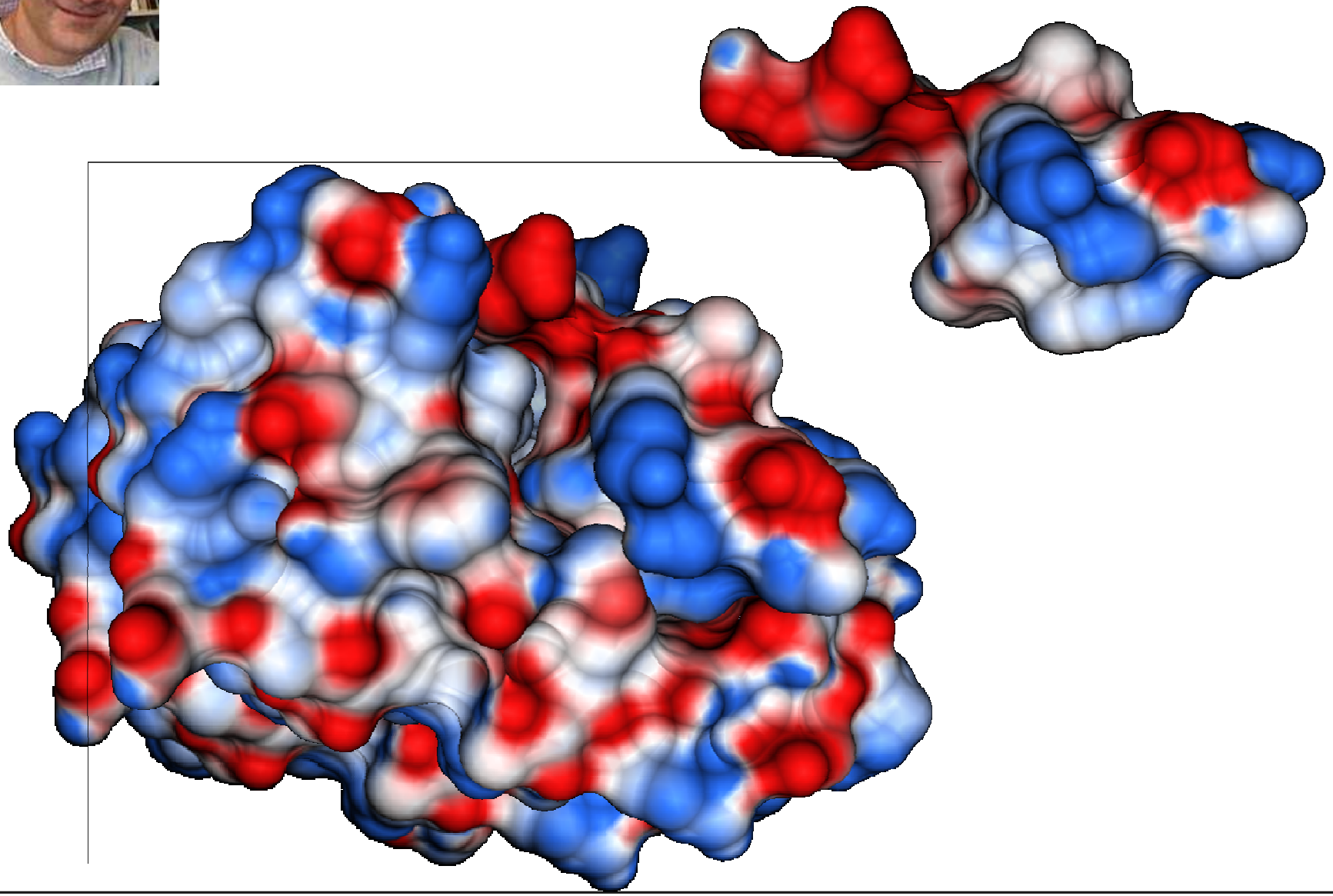
Protein-Protein Docking



There is no better description of the concept of *docking* of the following:



that translated at the molecular level :



How we can virtualize the protein-protein docking process :

Generation of plausible protein-protein complexes on the basis of electrostatic potential and topological complementarities (*docking*);

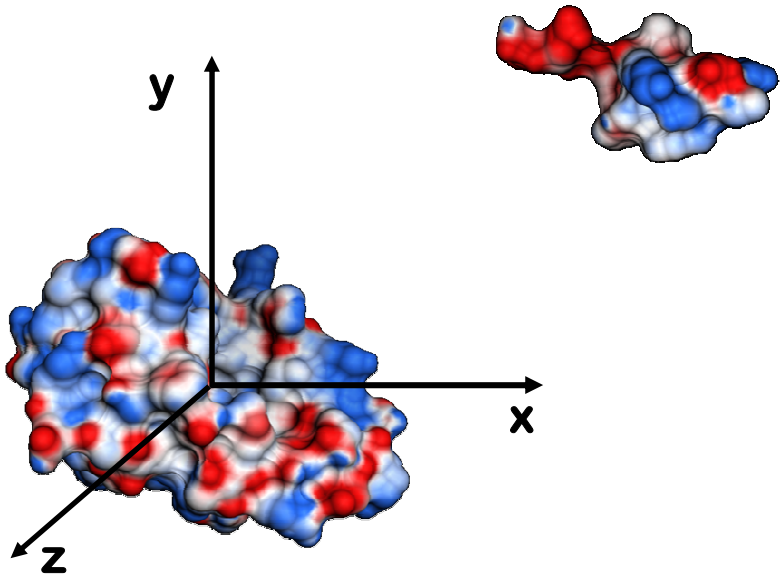
Stability assessment of individual complexes with the intent to infer on the corresponding value of ΔG_{bind} (*scoring*).

Some general considerations:

If the formation of protein-protein complexes does NOT imply a conformational modification of individual monomers virtualization process is called *rigid docking*;

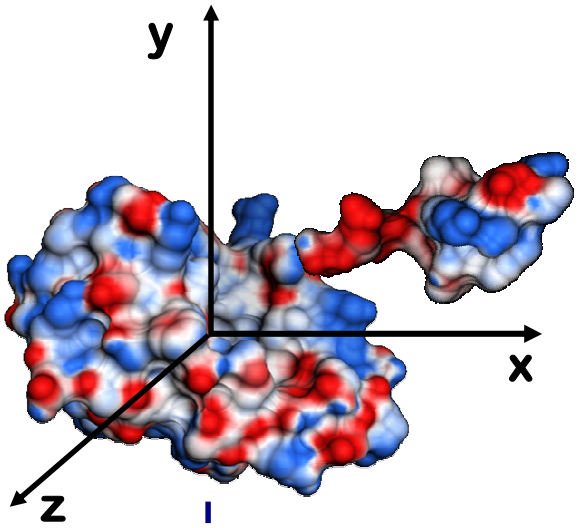
Otherwise, the procedure of virtualization (for high degree of conformational flexibility) is defined as *flexible docking*.

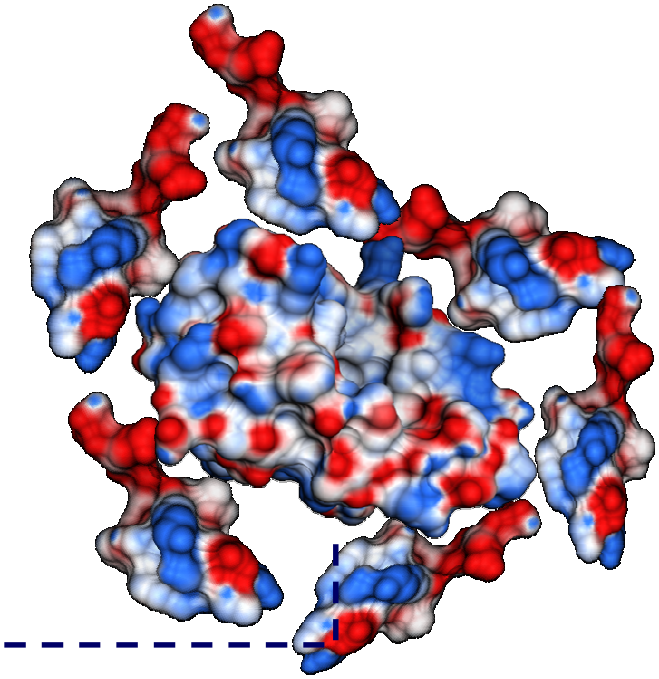
Let's try to virtualize a *rigid docking!*



Considered a reference system attached to the center of mass of one of the two objects, the degrees of freedom allowed are simply *3 translational motions* and *3 rotational motions*.

Here is how we can virtualize a docking process:

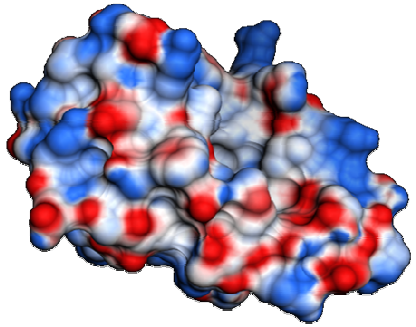




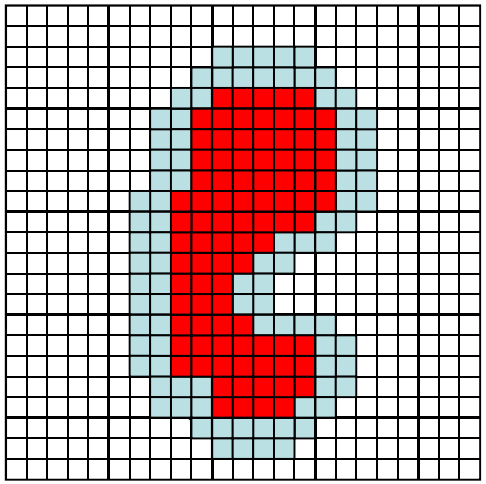
COMPLEMENTARITY

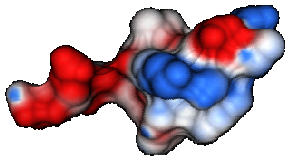
***f*(time)** ***f*(space)**

A simple method to analyze shape complementarity:

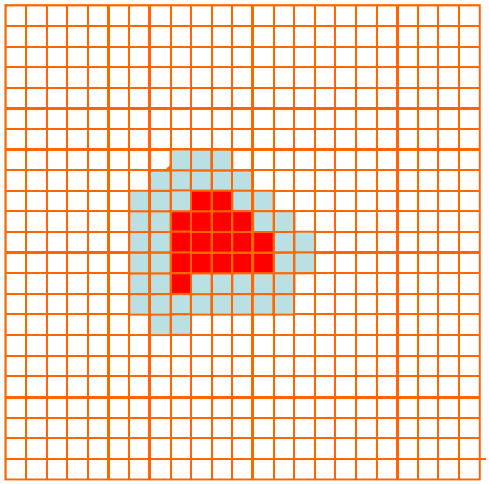


R





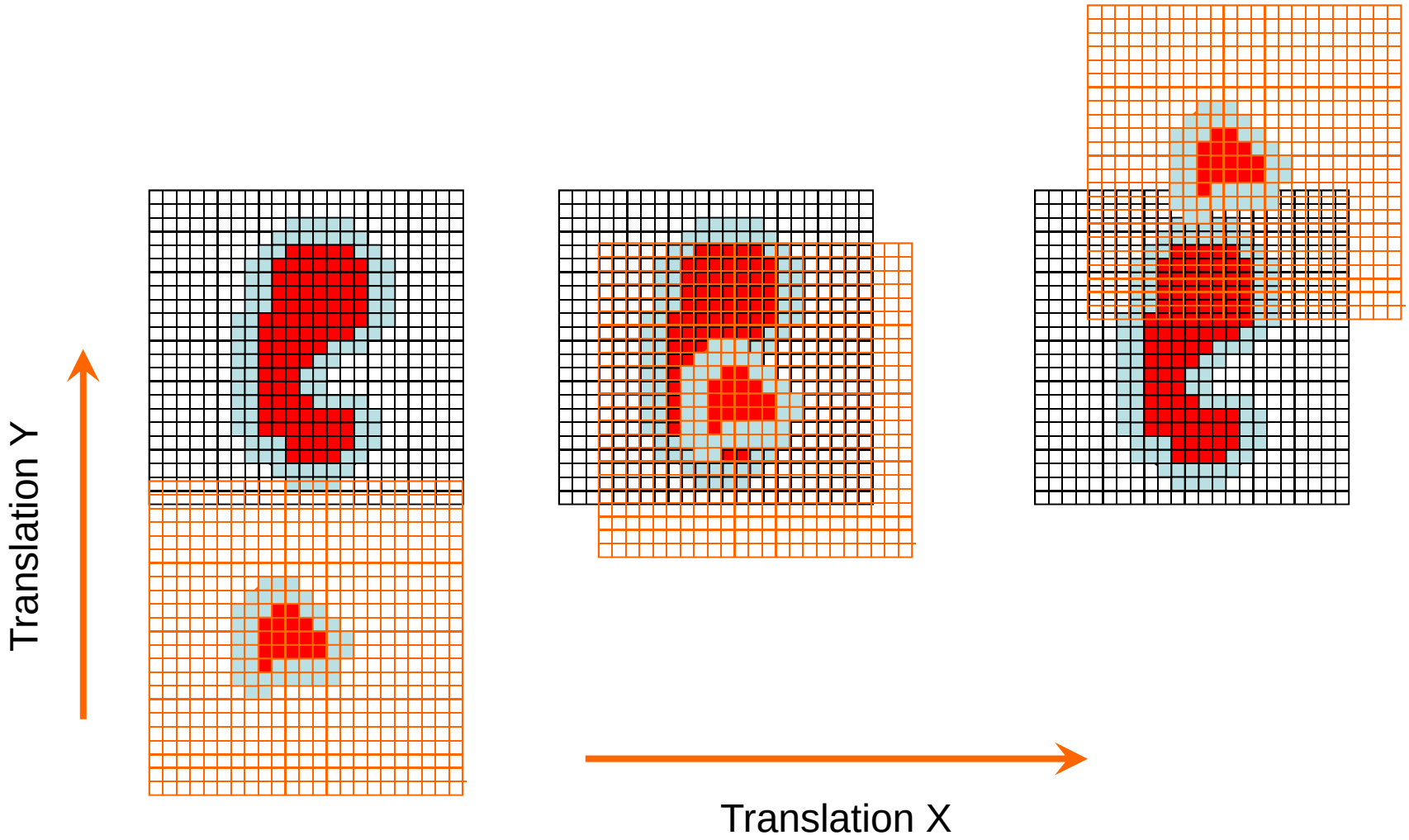
L



At each cell we can attribute:

- Ext: a(i,j) = 0
- Surf: a(i,j) = +1
- Int: a(i,j) = -15

- Ext: b(i,j) = 0
- Surf: b(i,j) = +1
- Int: b(i,j) = -15



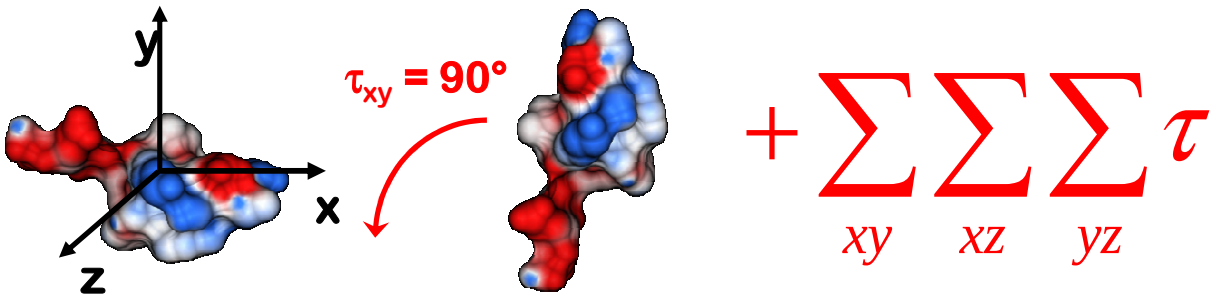
$$Score = \sum_i \sum_j a(i,j)b'(i,j)$$

***for each Y-translation step calculate all X-traslations**

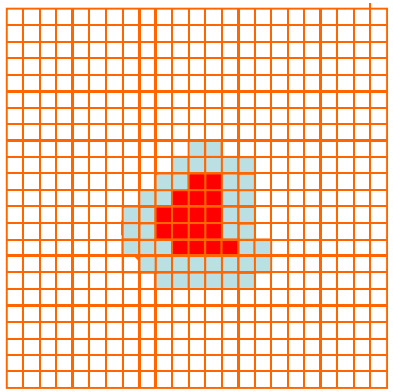
This is not a trivial process from a numerical point of view:

In fact, we have to compute all possible positions of L respect R:

- ***for example, we have to consider all possible rotations of L;***
- ***and for each rotation of L, we have to take into account all translation of L on R;***
- ***after that we can calculate the Score(i,j) = R (i,j)*L (i,j).***

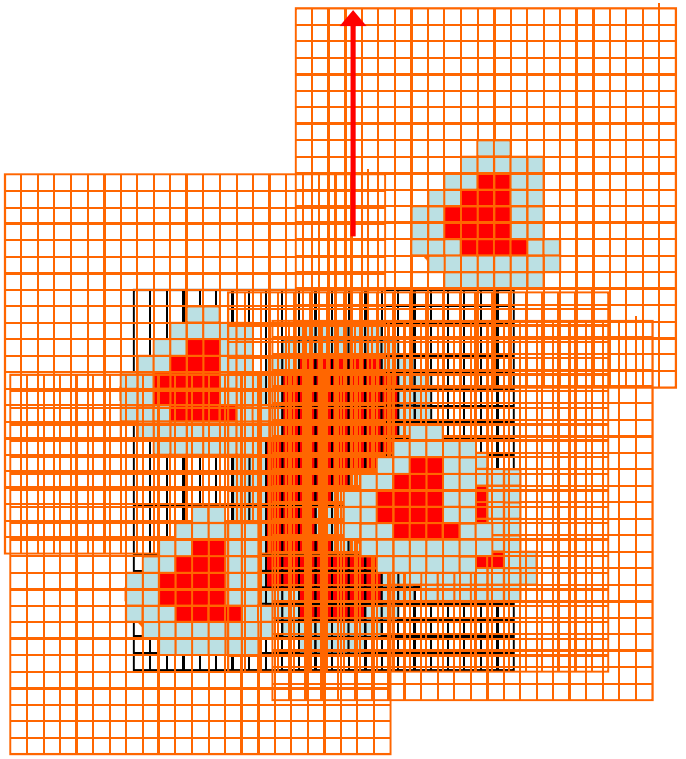


$$+\sum_{xy}\sum_{xz}\sum_{yz}\tau$$

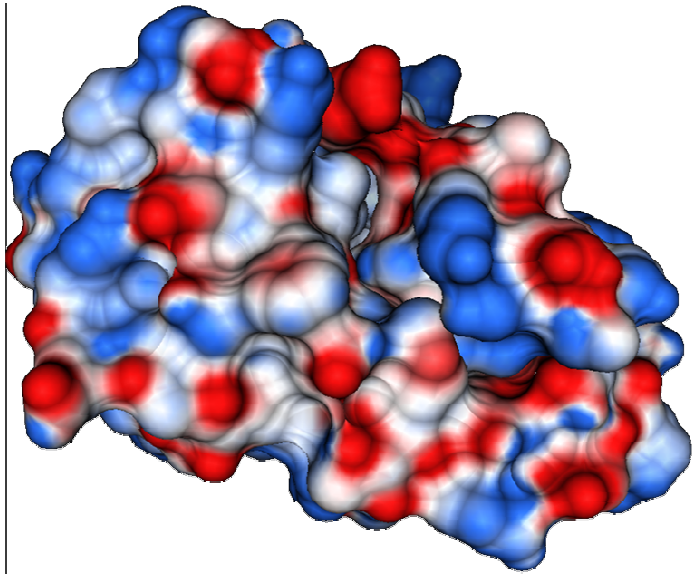


$$S(R,T)=\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}a(i,j,k)b'\left(i+T_x,j+T_y,k+T_k\right)$$

Traslations
x,y,z



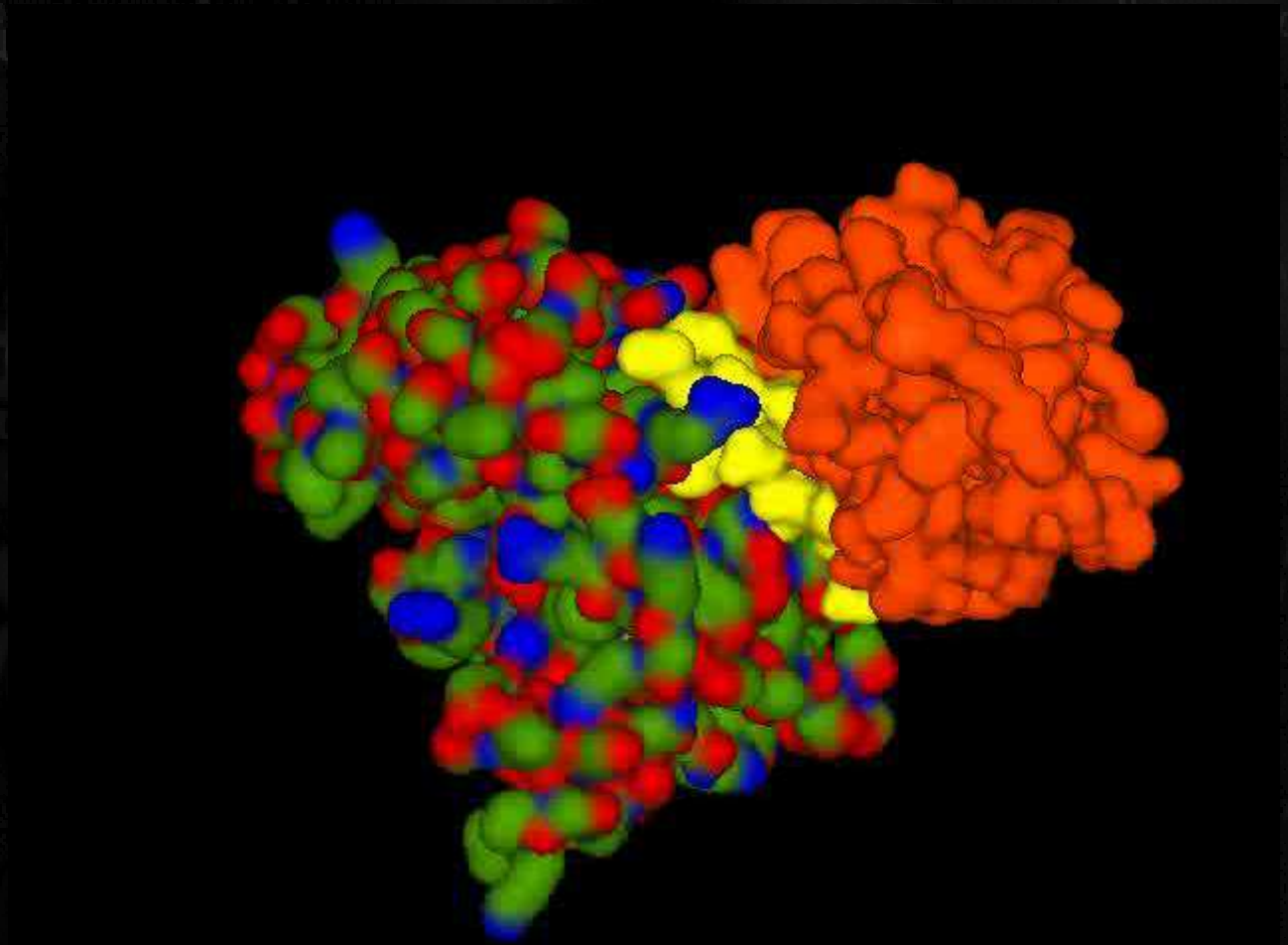
Electrostatic scoring is now easy to calculate:



For each previously selected complex based on topological complementarity, we can apply the Coulomb equation of inter-molecular interactions:

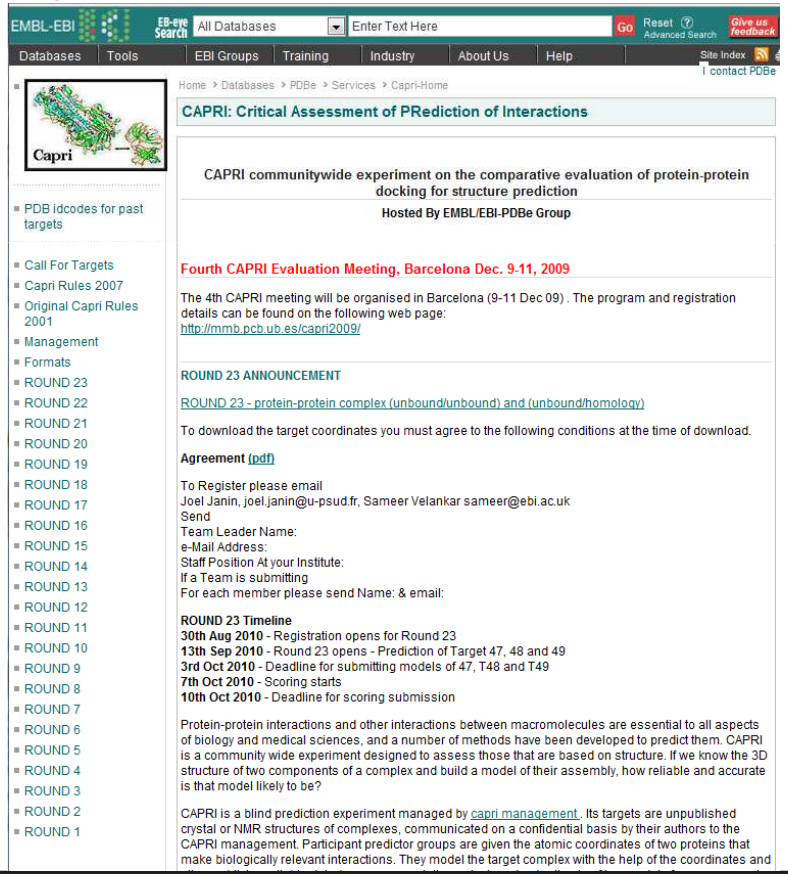
$$E_{el}(r) = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

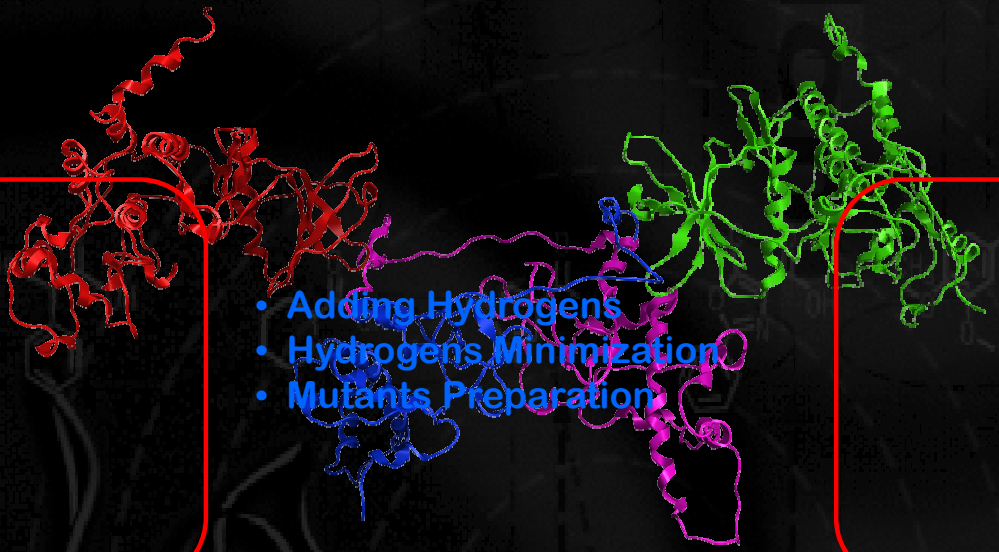
An example:



CAPRI…

CAPRI: Critical Assessment of PRediction of Interactions (http://www.ebi.ac.uk/msd-srv/capri/)





- Adding Hydrogens
- Hydrogens Minimization
- Mutants Preparation

Crystallographic Protein-Protein Complexes

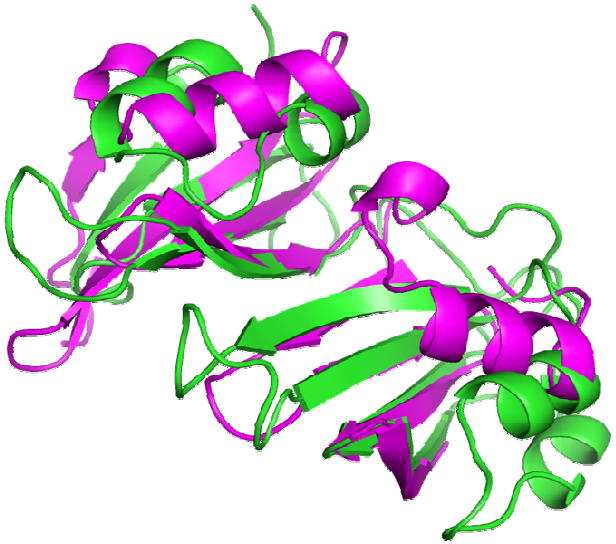
Subunits Splitting

Monomer preparation

How we can compute the accuracy (performance) of a protein-protein docking protocol?

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n} (x_{1,i} - x_{2,i})}{n}}$$

RMSD = 4.5 Å



RMSD = 0.5 Å

RMSD = root mean square deviation

have a nice docking…

- ClusPro http://cluspro.bu.edu/
- Ftdock http://www.sbg.bio.ic.ac.uk/docking/
- Zdock http://zlab.bu.edu/zdock/
- Gramm http://vakser.bioinformatics.ku.edu/main/resources_gramm.php
- Haddock http://www.nmr.chem.uu.nl/haddock/
- Hex http://hex.loria.fr/
- Rosetta http://graylab.jhu.edu/docking/rosetta/