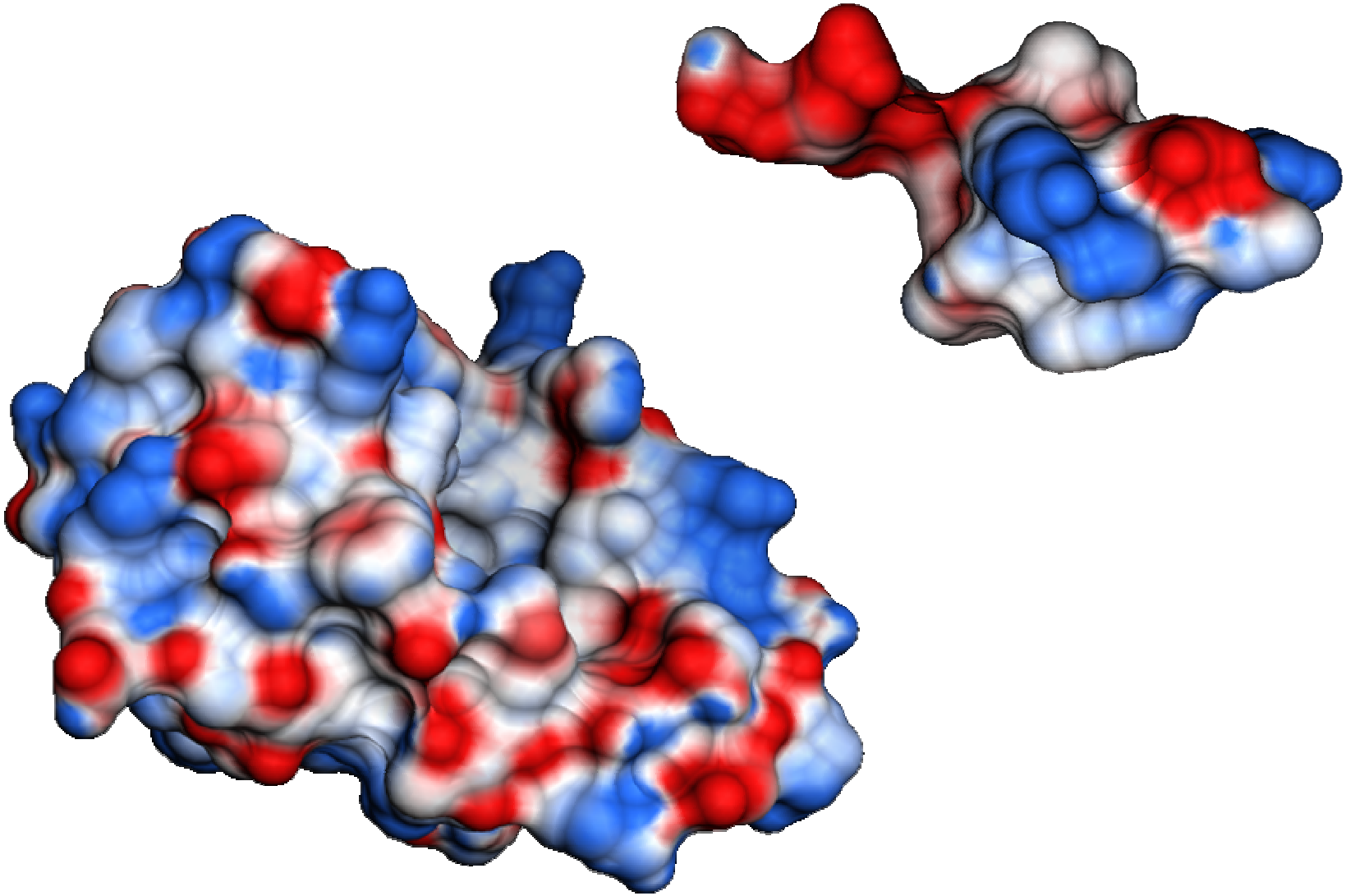
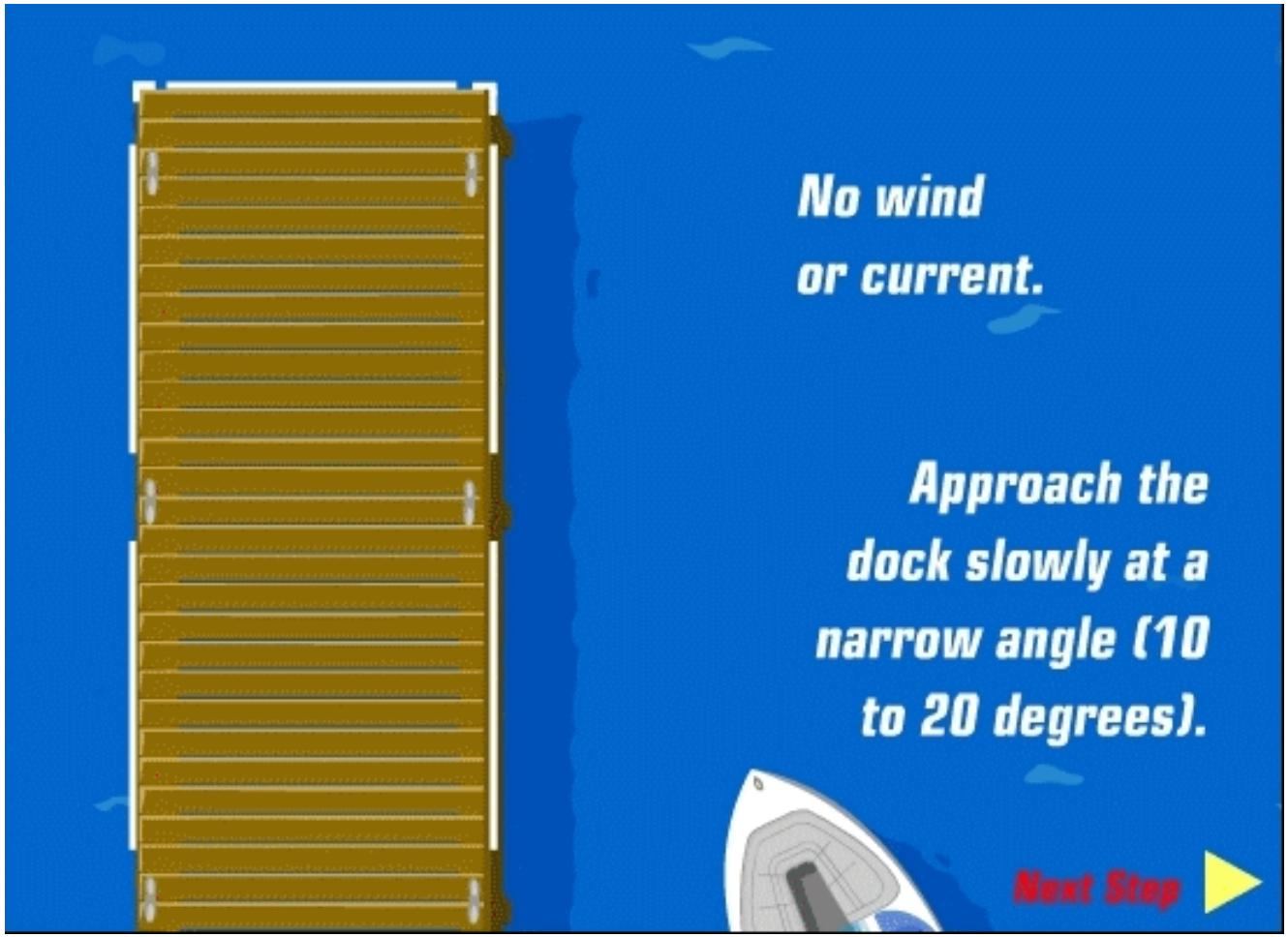
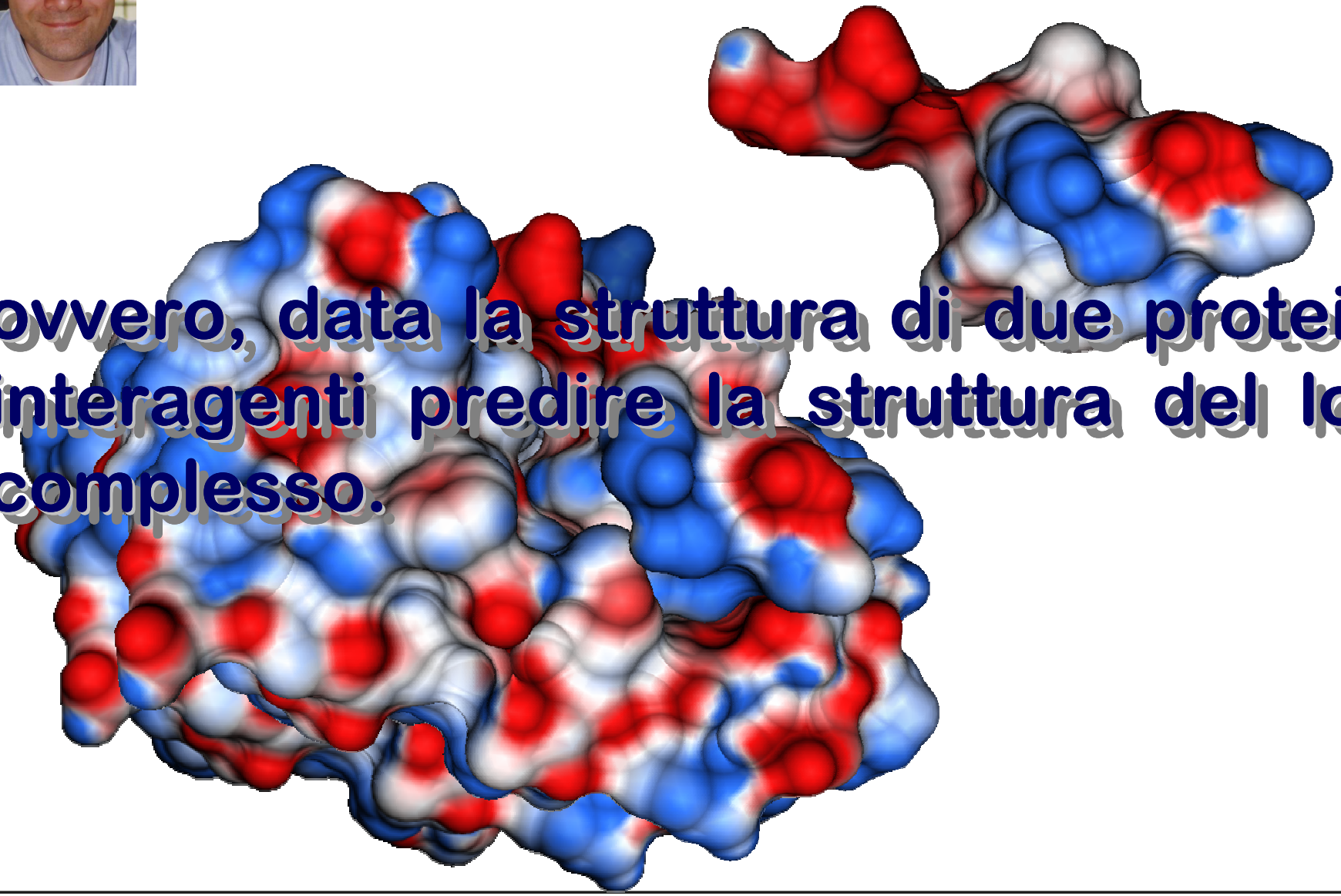
Protein-Protein Docking



Non c'è descrizione migliore del concetto di *docking* della seguente:



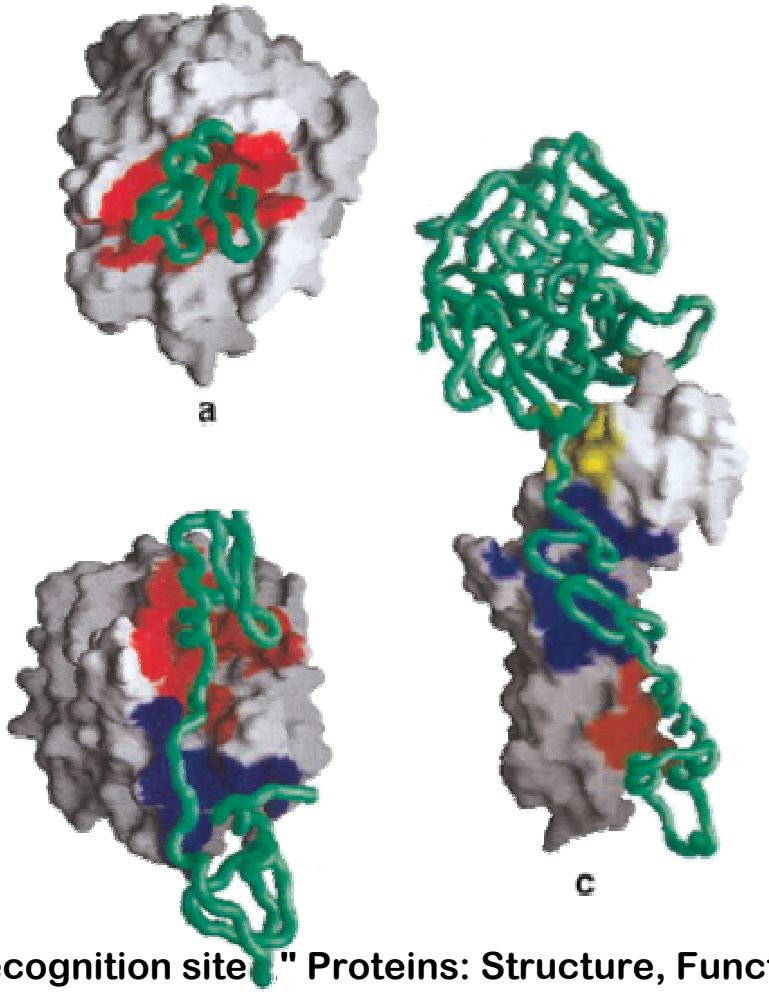
che tradotto a livello molecolare:



ovvero, data la struttura di due proteine interagenti predire la struttura del loro complesso.

Possiamo fare qualche considerazione sulle proprietà delle superfici all'interfaccia di riconoscimento:

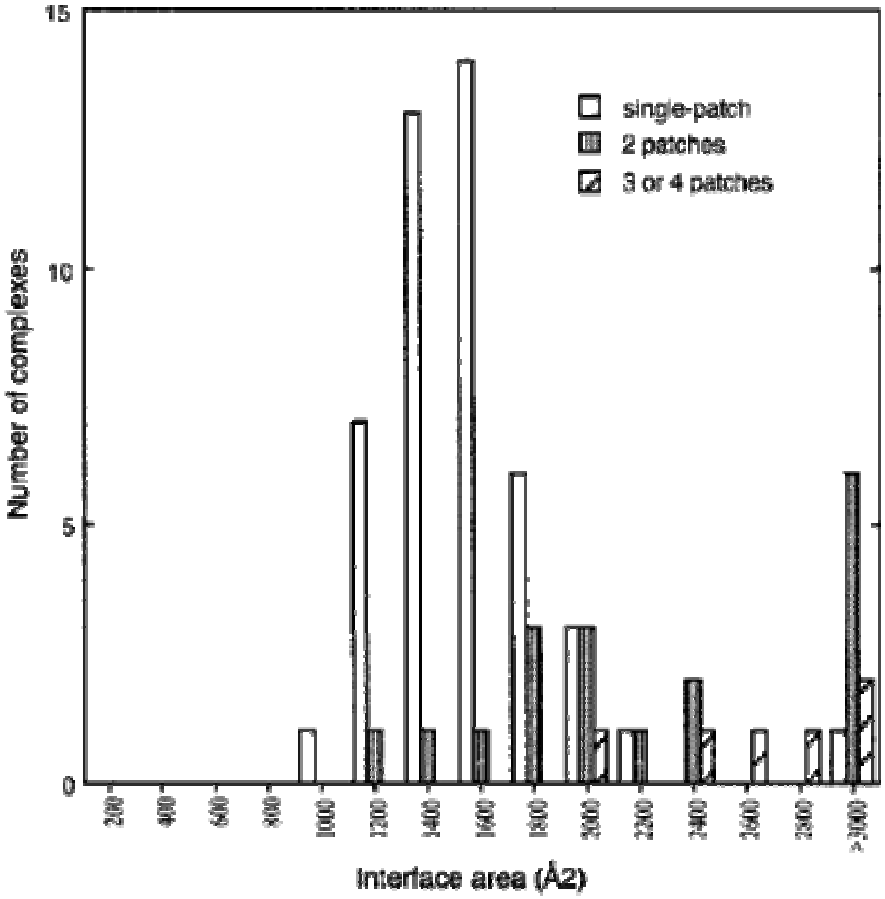
In primis, non tutte le interazioni proteina-proteina sono mediate da una sola interfaccia (one patch, *a*), ma possono essere due (two patches, *b*) oppure più di due (more patches, *c*)



P. Chakrabarti, J. Janin, "Dissecting protein-protein recognition sites," Proteins: Structure, Function, and Genetics, 47, 3, 2002, pp. 334-343.

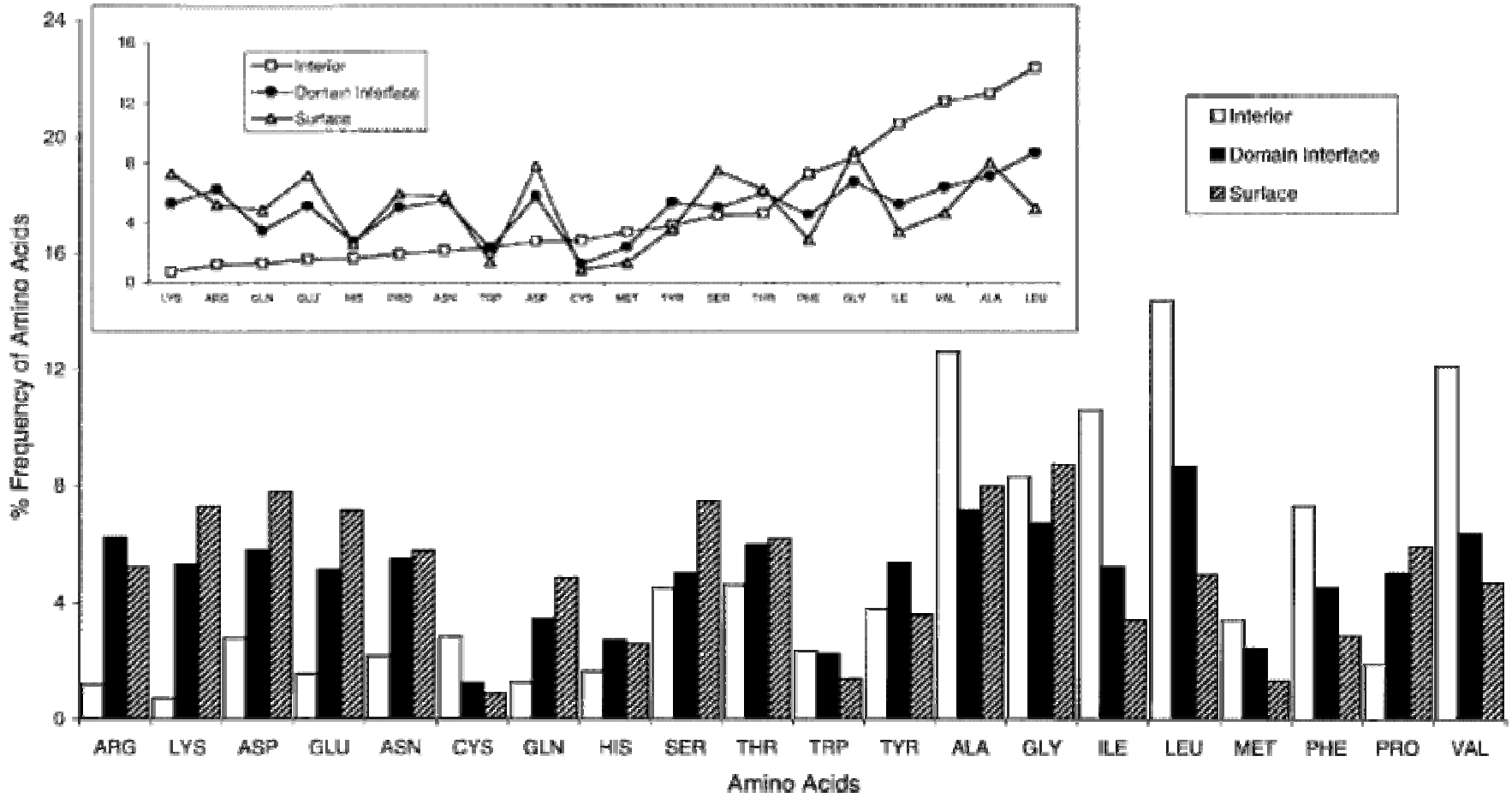
Possiamo fare qualche considerazione sulle proprietà delle superfici all'interfaccia di riconoscimento:

Una interazione proteina-proteina *one-patch* richiede una copertura di superficie pari a 1320 ± 520 Å²



P. Chakrabarti, J. Janin, "Dissecting protein-protein recognition sites," Proteins: Structure, Function, and Genetics, 47, 3, 2002, pp. 334-343.

Possiamo fare qualche considerazione sulle proprietà delle superfici all'interfaccia di riconoscimento:



S. Jones, A. Marin, J.M. Thornton, "Protein domain interfaces: characterization and comparison with oligomeric protein interfaces," Protein Engineering, 13, 2, 2000, pp. 77-82.

Come possiamo virtualizzare una procedura di docking:

Generazione di plausibili complessi proteina-proteina sulla base della complementarietà topologica e del potenziale elettrostatico (*docking*);

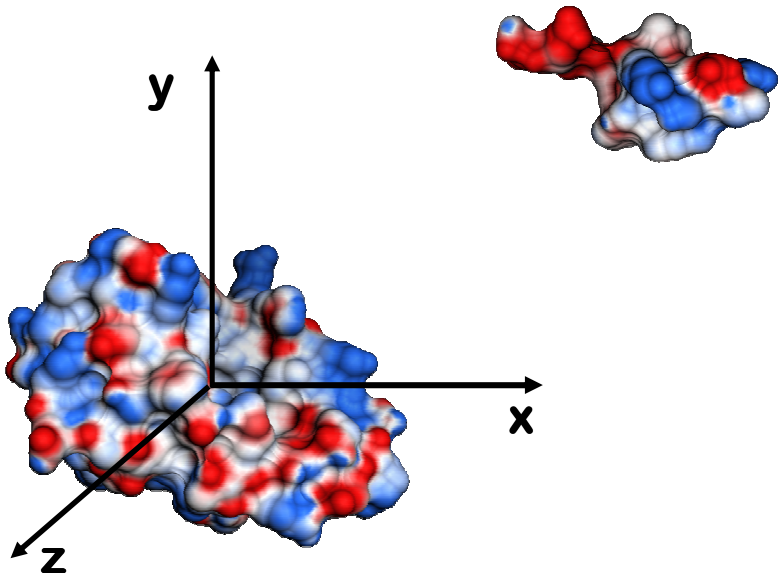
Valutazione della stabilità dei singoli complessi con l'intento di inferire sul corrispettivo valore di ΔG_{bind} (*scoring*).

Alcune considerazioni generali:

Se la formazione dei complessi proteina-proteina NON implica una modificazione conformazionale dei singoli monomeri la procedura di virtualizzazione viene definita *docking rigido*;

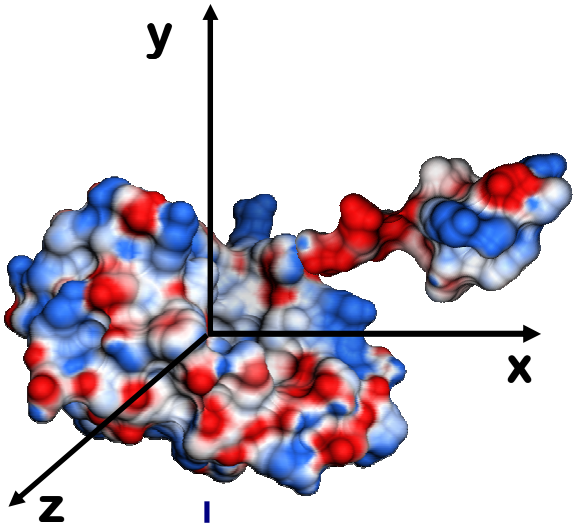
Nel caso contrario, la procedura di virtualizzazione (ad elevato grado di complessità intrinseco) viene definita *docking flessibile*.

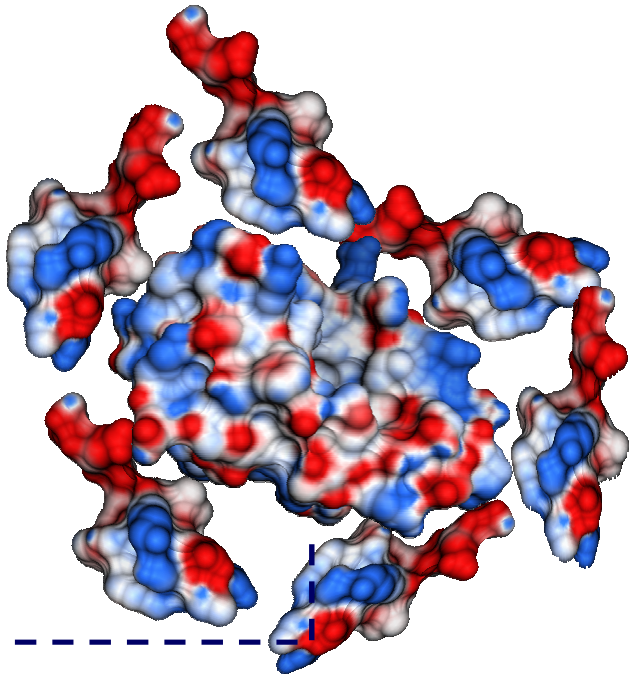
Proviamo a formalizzare la virtualizzazione di un *docking rigido*:



Considerato un sistema di riferimento solidale con il centro di massa di uno dei due oggetti, i gradi di libertà consentiti al secondo sono semplicemente 3 *moti traslazionali* e 3 *moti rotazionali*.

Proviamo a formalizzare la virtualizzazione di un *docking rigido*:



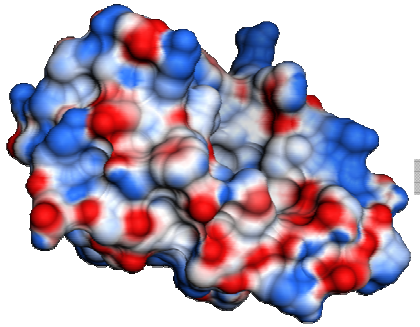


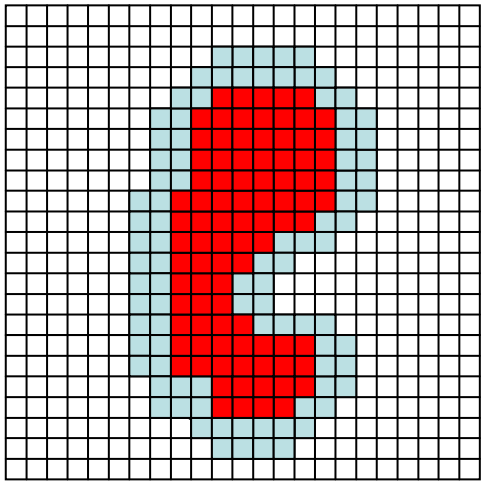
COMPLEMENTARIETA'

***f*(tempo)**

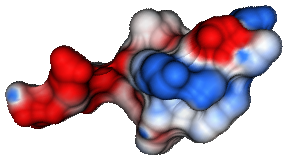
***f*(spazio)**

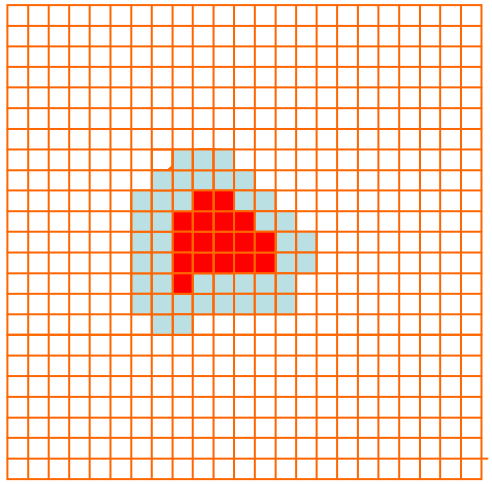
Ecco un modo per scalare l'efficacia del riconoscimento tra i due oggetti:





R



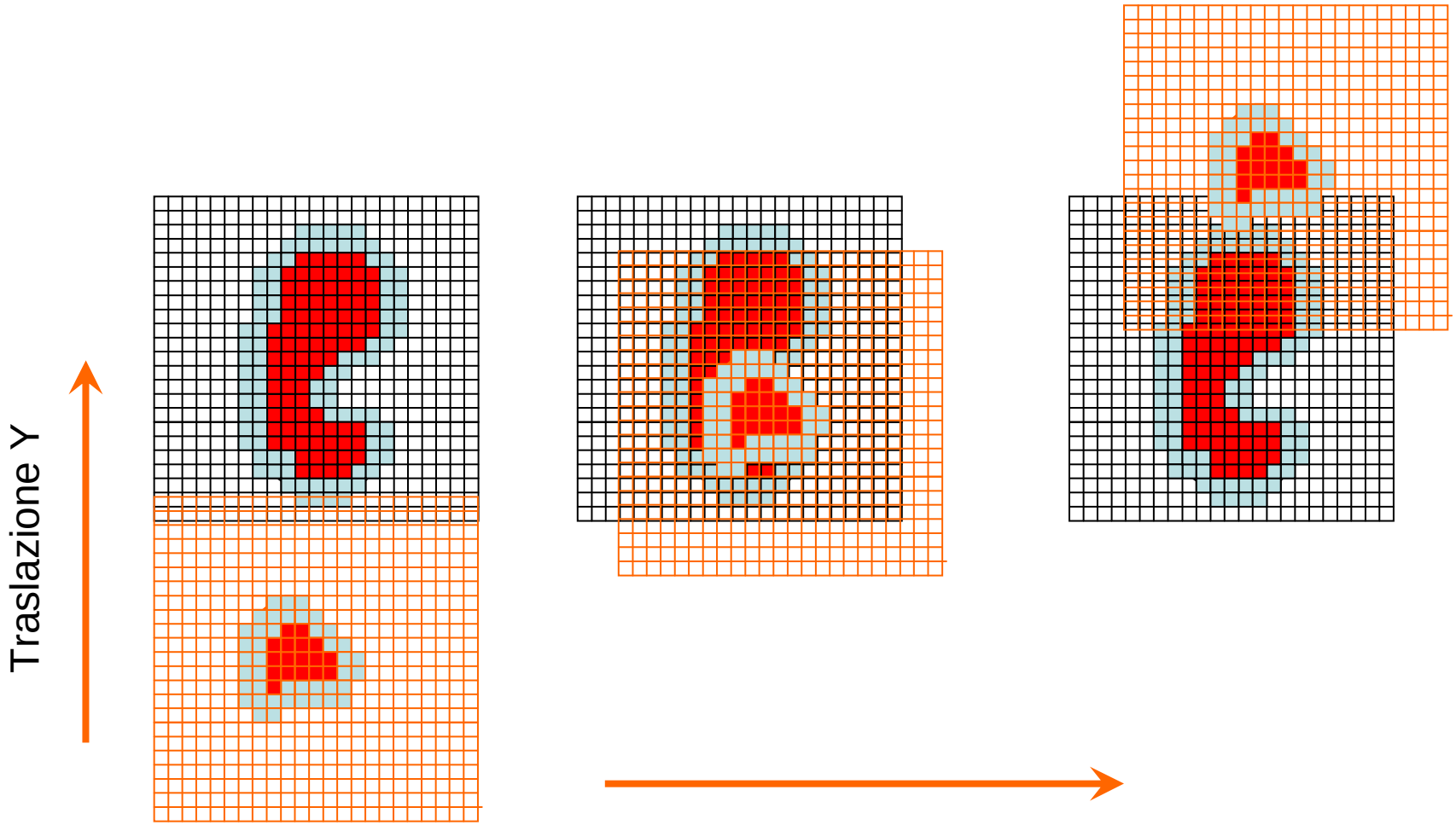


L

Assegnamo un valore ad ogni cella:

- 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



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

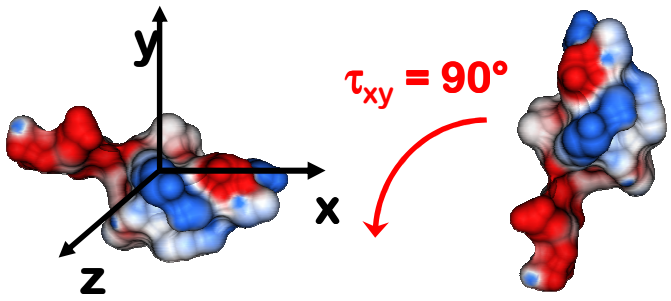
***ci ricordiamo che *b'* fa riferimento alla griglia dopo la traslazione**

Questo da un punto di vista *numerico* è davvero molto complesso però:

Bisogna infatti computare tutte le possibili posizioni di L rispetto ad R:

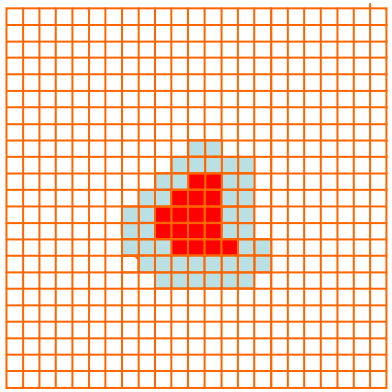
- ***per esempio tutte le rotazioni di L;***
- ***e per ogni rotazione di L computare tutte le transizioni della griglia di L sulla griglia di R;***

- ***da cui il computo della funzione di Score(i,j) = R (i,j)*L (i,j).***



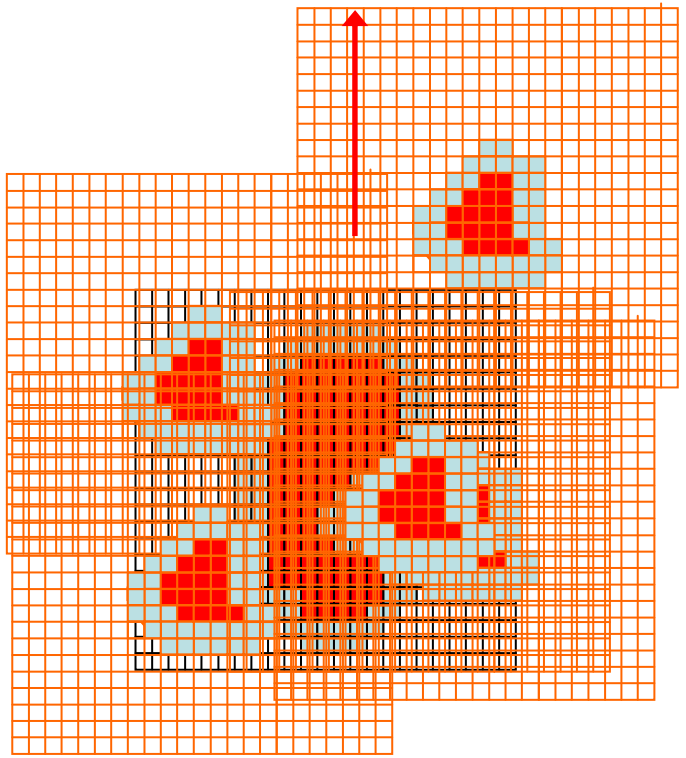
$\tau_{xy} = 90°$

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



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

Traslazioni
x,y,z

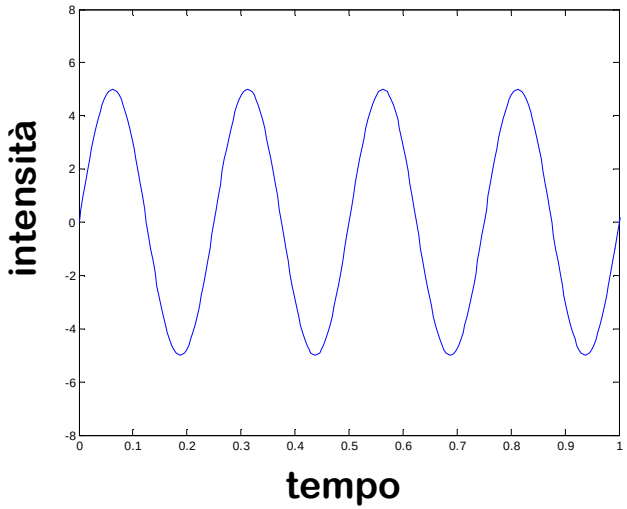


Meno male che Fourier c'è!

Funzioni, segnali e loro trasforzioni…

Un piccolo bignami di... Fourier!

Funzione seno

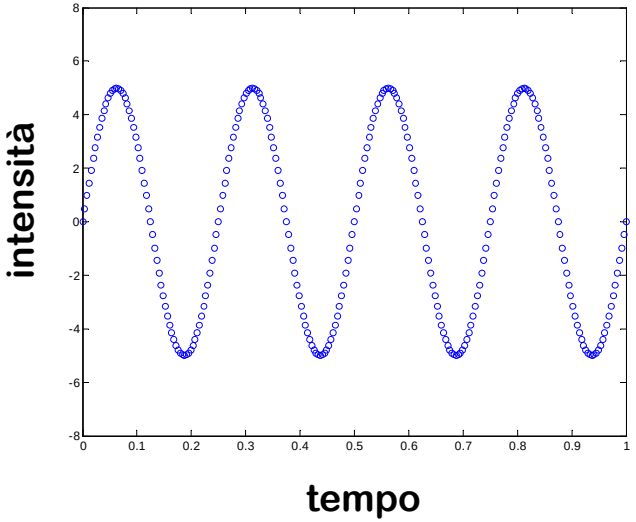


5*sin (2π4t)

Amplitude = 5

Frequency = 4 Hz

Segnale a funzione seno



5*sin(2π4t)

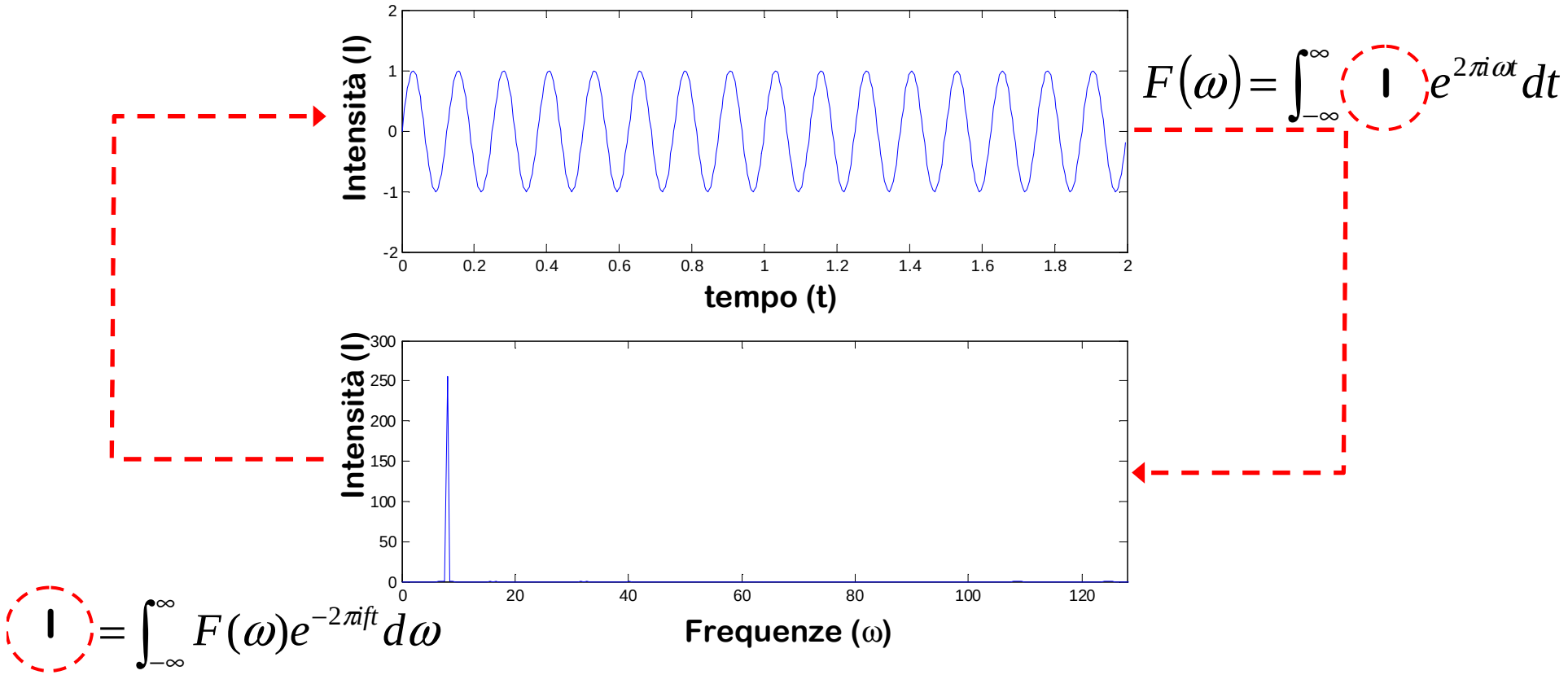
Amplitude = 5

Frequency = 4 Hz

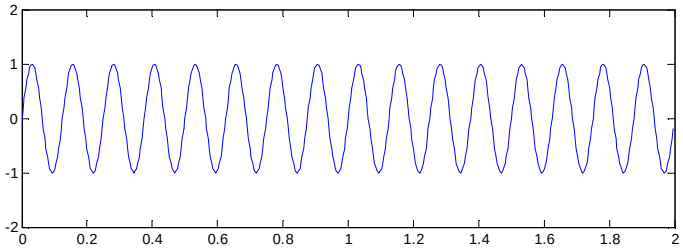
Sampling rate = 256 samples/second

Sampling duration = 1 second

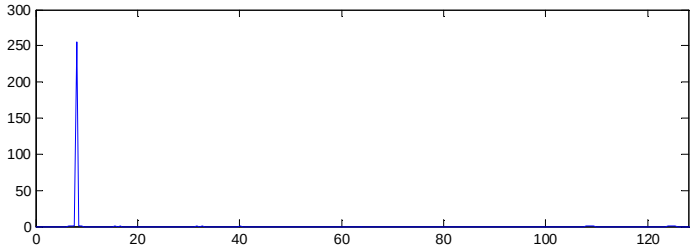
Possiamo definire trasformazione di Fourier un operatore matematico che trasforma una funzione (od un segnale) in un'altra funzione (od un altro segnale) in maniera reciprocamente reversibile.

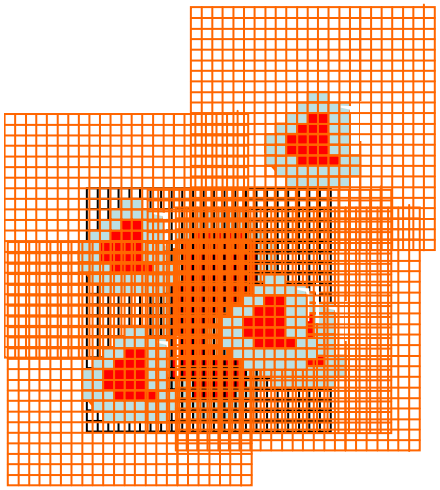


La semplificazione che deriva dall'applicazione della FT spero sia molto evidente… dato che siamo interessati proprio agli intorni dei massimi di intensità!

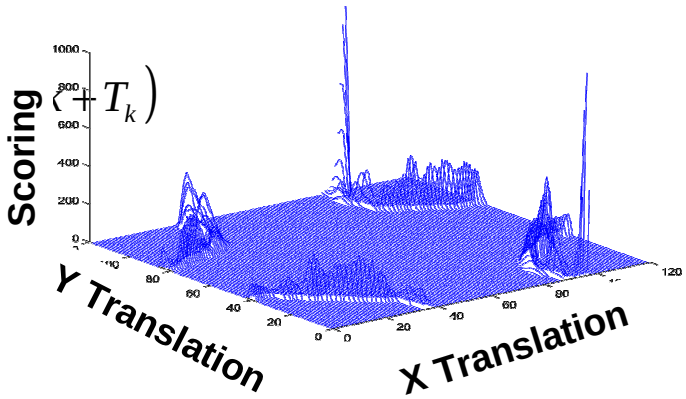


$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{2\pi i \omega t} dt$$

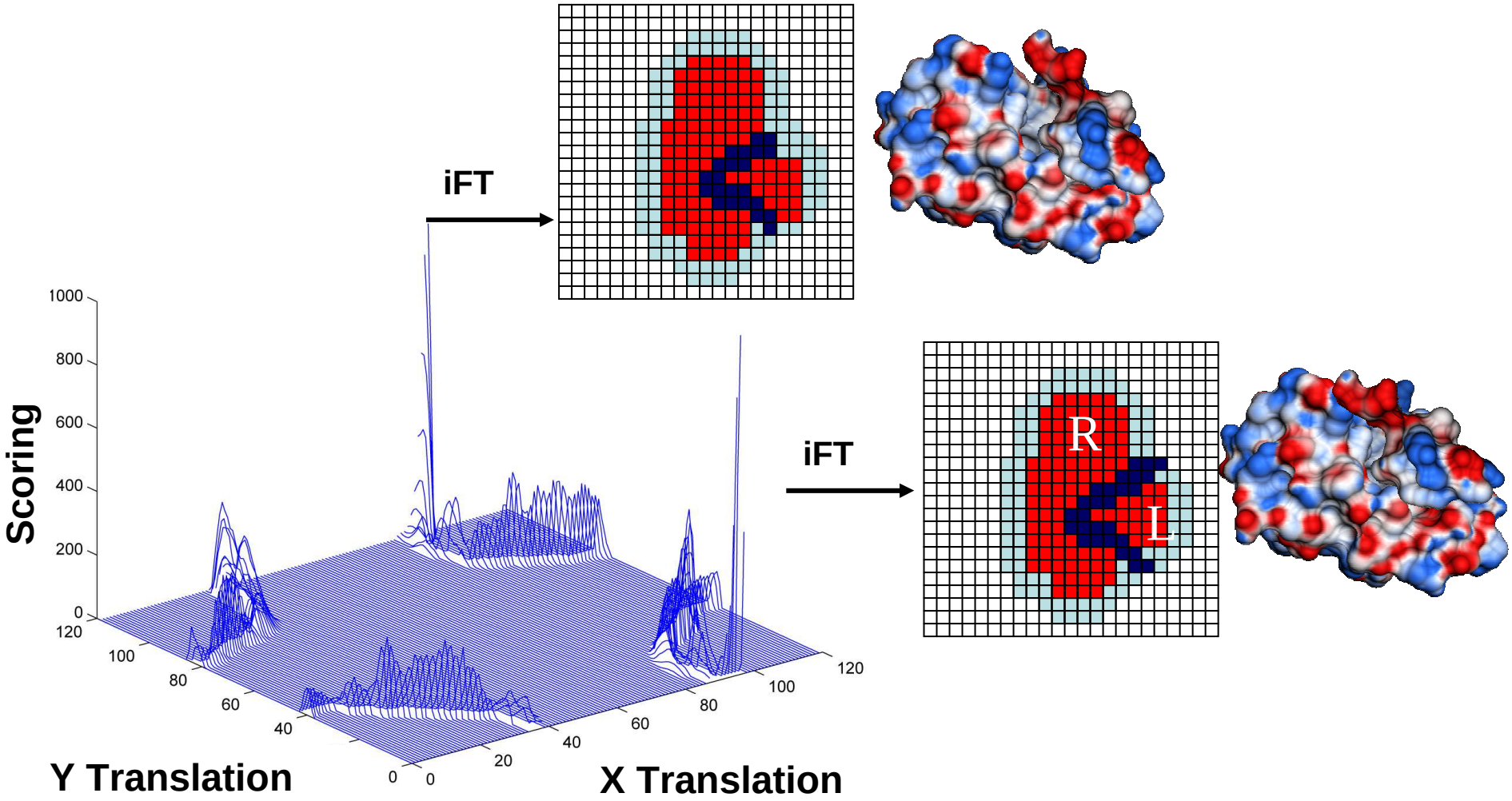




$$F\left(S(T_x, T_y)\right) = \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} a(i,j,k) b'\left(i + T_x, j +$$

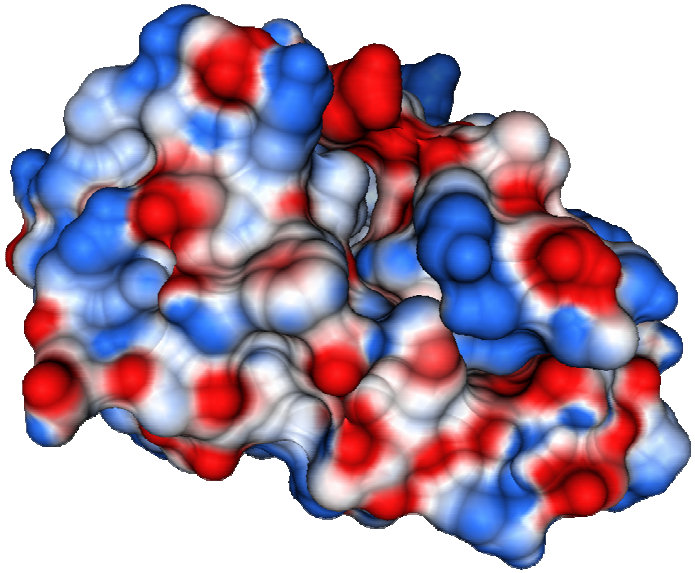


Incremento di velocità di un fattore pari 10^7



Surface Interior Binding Site

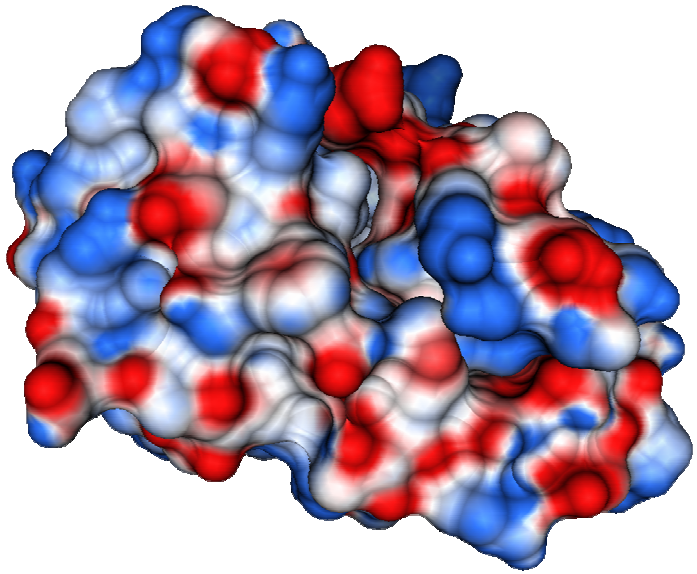
e dopo la complementarietà topologica quella elettrostatica vien da sé!

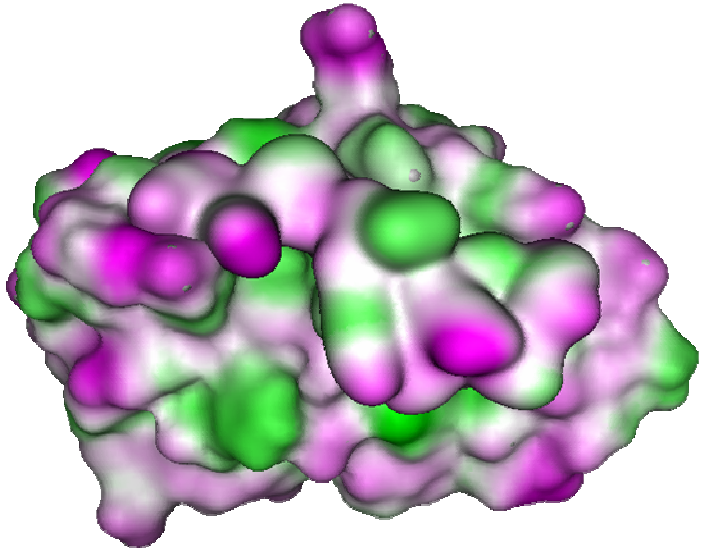


Per ogni complesso precedentemente selezionato in base alla complementarietà topologica, possiamo applicare l'equazione di Coulomb inter-molecolare:

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

Ma la sola componente elettrostatica ci aiuta a decodificare la stabilità dei vari complessi proteina-proteina?

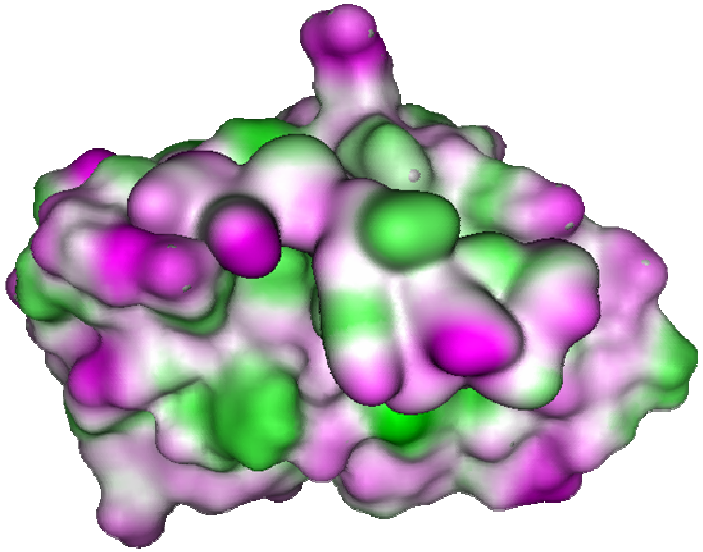




$$E_{el}(r) = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

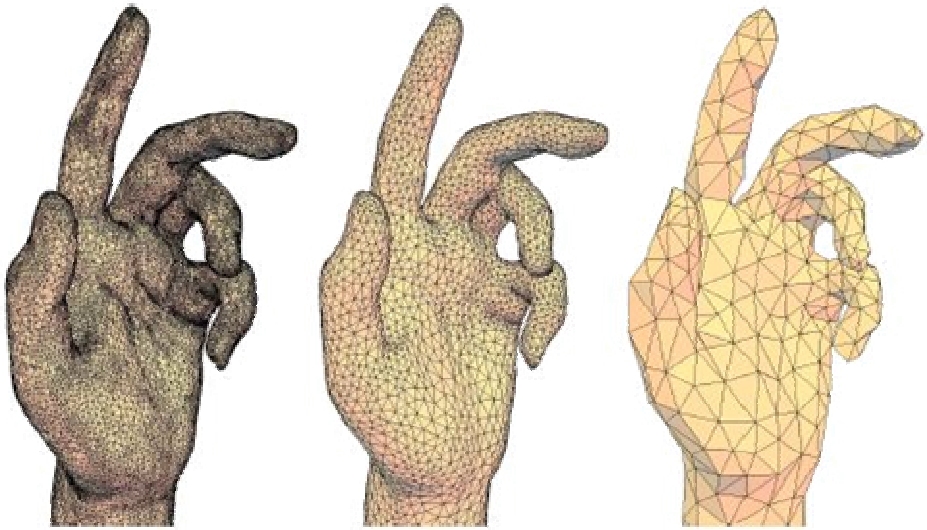
?

Area Accessibile al Solvente (ASA) ed Energia Libera di Solvatazione (ΔG_{sol}):

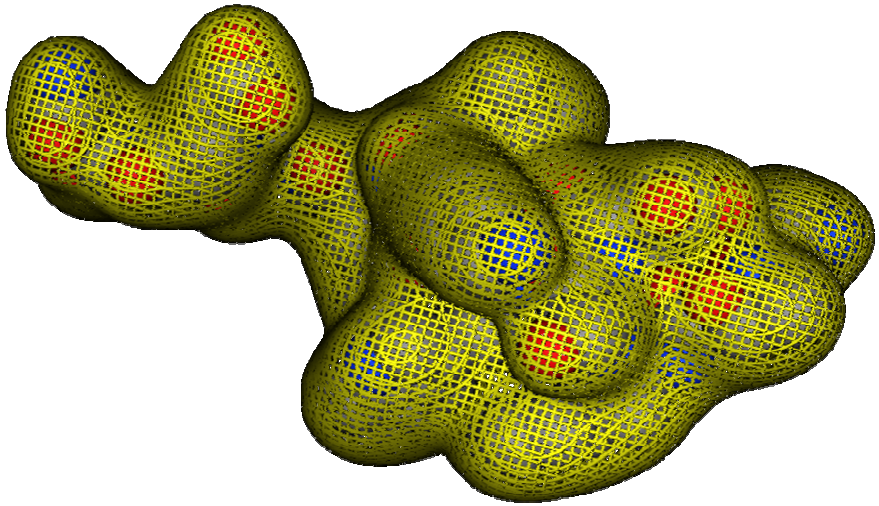


$$\Delta G_{sol} = f\,(ASA_{com}) = \Sigma_i\, \sigma_i\, ASA_i$$

Avere una stima della misura di una superficie non poi così difficile!

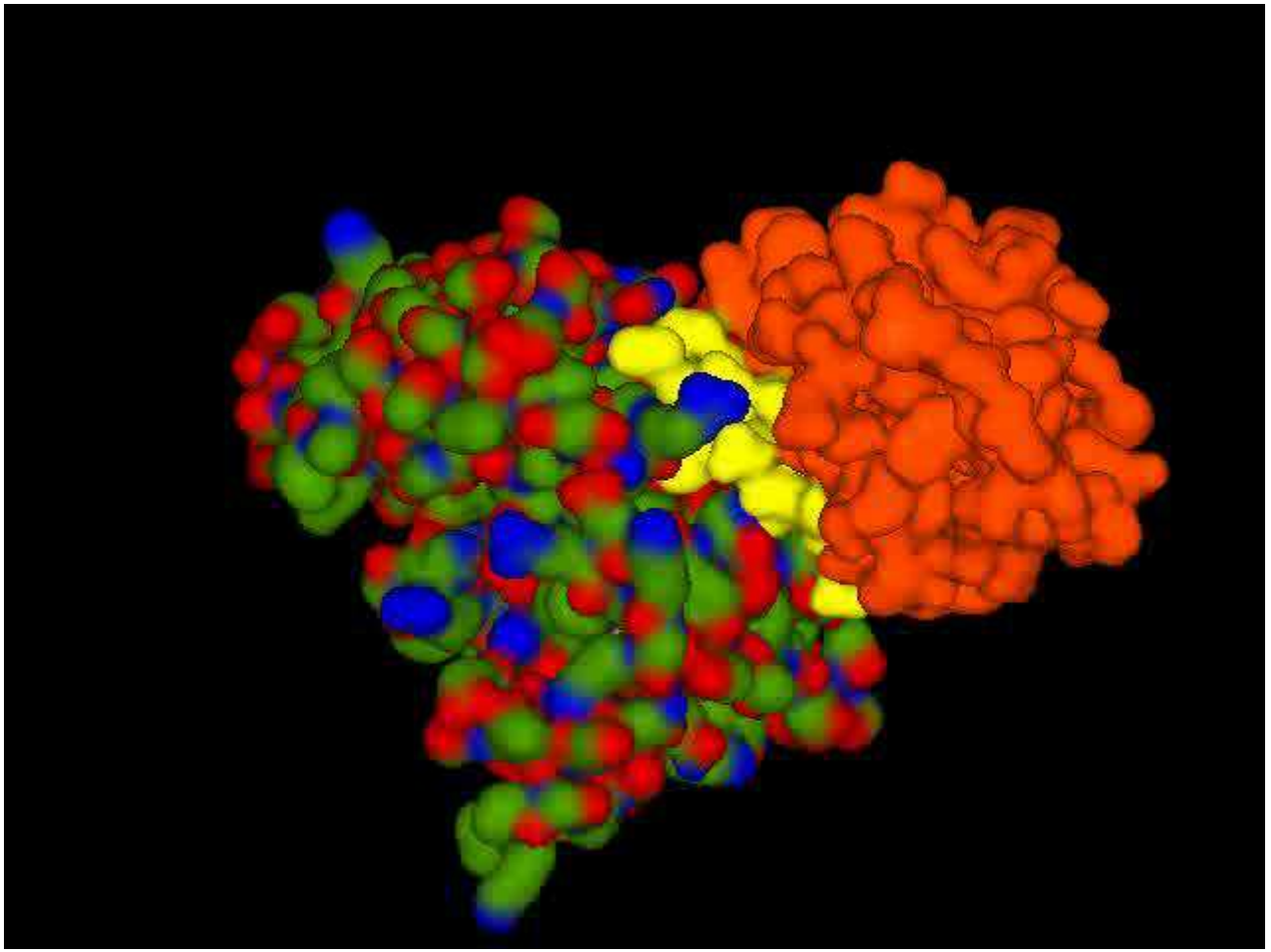


(a) 25,000 vertices. (b) 5,000 vertices. (c) 500 vertices.



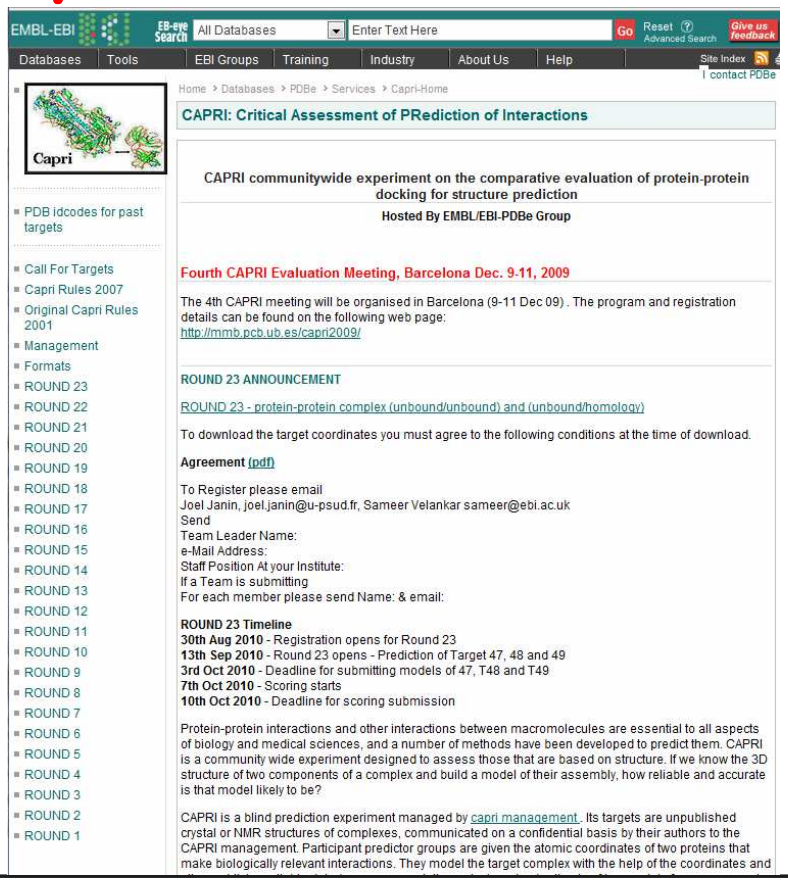
Superficie Coperta: $ASA_{RL} - ASA_R - ASA_L$

un esempio concreto....



CAPRI...

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



EMBL-EBI | EB-eye Search | All Databases | Enter Text Here | Go | Reset | Advanced Search | Give us feedback

Databases | Tools | EBI Groups | Training | Industry | About Us | Help | Site Index | I contact PDBe

Home > Databases > PDBe > Services > Capri-Home

Capri

- PDB idcodes for past targets
- Call For Targets
- Capri Rules 2007
- Original Capri Rules 2001
- Management
- Formats
- ROUND 23
- ROUND 22
- ROUND 21
- ROUND 20
- ROUND 19
- ROUND 18
- ROUND 17
- ROUND 16
- ROUND 15
- ROUND 14
- ROUND 13
- ROUND 12
- ROUND 11
- ROUND 10
- ROUND 9
- ROUND 8
- ROUND 7
- ROUND 6
- ROUND 5
- ROUND 4
- ROUND 3
- ROUND 2
- ROUND 1

CAPRI: Critical Assessment of PRediction of Interactions

CAPRI communitywide experiment on the comparative evaluation of protein-protein docking for structure prediction

Hosted By EMBL/EBI-PDBe Group

Fourth CAPRI Evaluation Meeting, Barcelona Dec. 9-11, 2009

The 4th CAPRI meeting will be organised in Barcelona (9-11 Dec 09) . The program and registration details can be found on the following web page:
http://mmb.pcb.ub.es/capri2009/

ROUND 23 ANNOUNCEMENT

ROUND 23 - protein-protein complex (unbound/unbound) and (unbound/homology)

To download the target coordinates you must agree to the following conditions at the time of download.

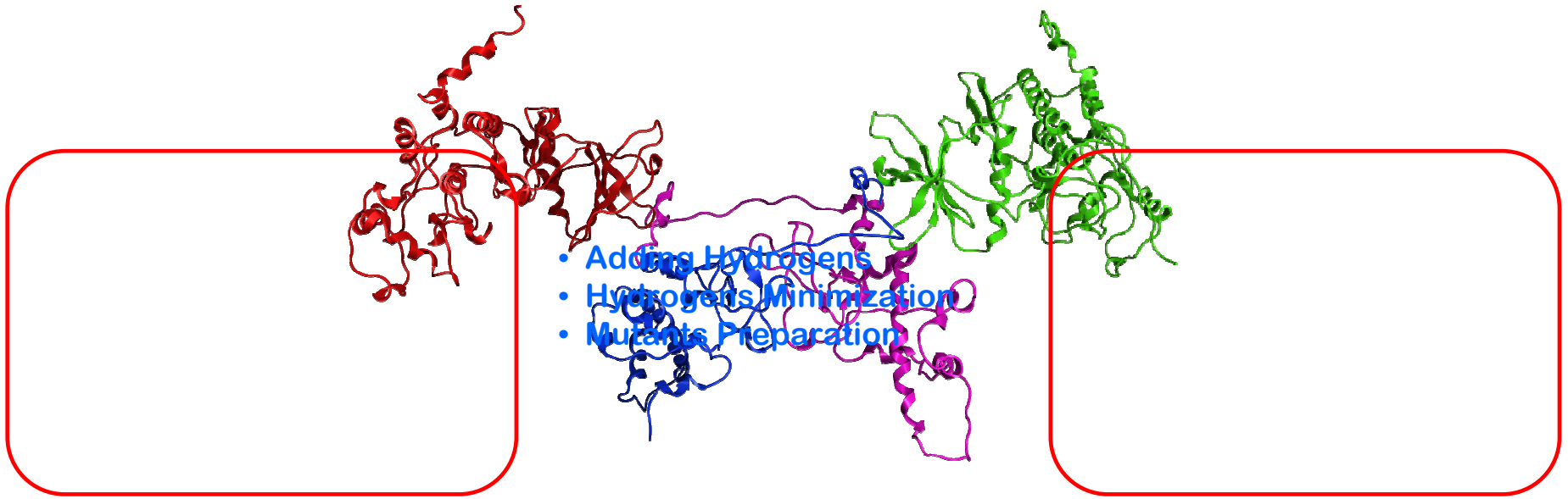
Agreement (pdf)

To Register please email
Joel Janin, joel.janin@u-psud.fr, Sameer Velankar sameer@ebi.ac.uk
Send
Team Leader Name:
e-Mail Address:
Staff Position At your Institute:
If a Team is submitting
For each member please send Name: & email:

ROUND 23 Timeline
30th Aug 2010 - Registration opens for Round 23
13th Sep 2010 - Round 23 opens - Prediction of Target 47, 48 and 49
3rd Oct 2010 - Deadline for submitting models of 47, T48 and T49
7th Oct 2010 - Scoring starts
10th Oct 2010 - Deadline for scoring submission

Protein-protein interactions and other interactions between macromolecules are essential to all aspects of biology and medical sciences, and a number of methods have been developed to predict them. CAPRI is a community wide experiment designed to assess those that are based on structure. If we know the 3D structure of two components of a complex and build a model of their assembly, how reliable and accurate is that model likely to be?

CAPRI is a blind prediction experiment managed by capri management . Its targets are unpublished crystal or NMR structures of complexes, communicated on a confidential basis by their authors to the CAPRI management. Participant predictor groups are given the atomic coordinates of two proteins that make biologically relevant interactions. They model the target complex with the help of the coordinates and

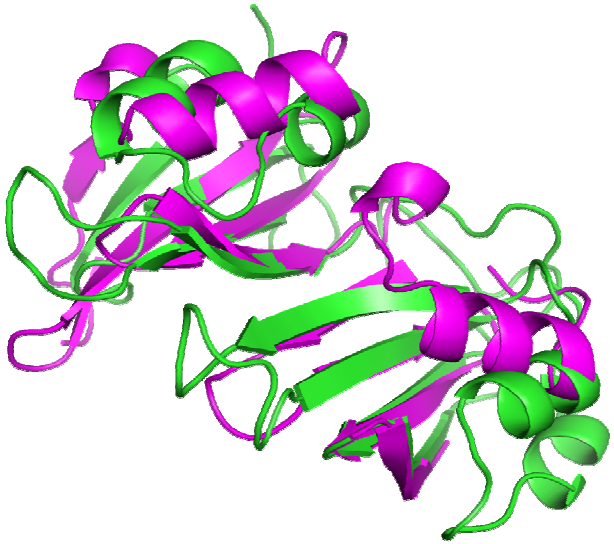


Crystallographic Protein-Protein Complexes → **Subunits Splitting** → **Monomer preparation**

e come giudichiamo chi vince?

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

RMSD = 4.5 Å



RMSD = 0.5 Å

RMSD = root mean square deviation

e buon 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/