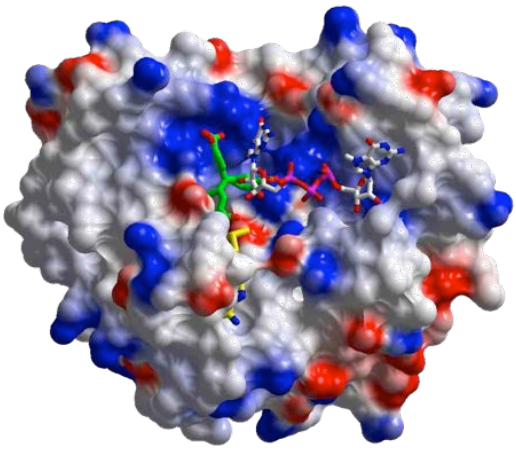
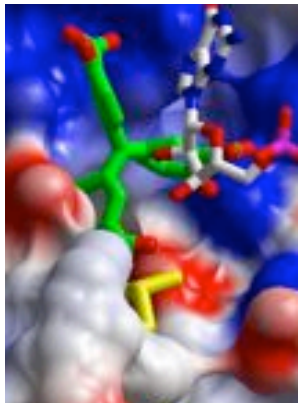
3. Proteine virali e ricerca razionale di farmaci antivirali

Mario Milani
mario.milani@mi.infm.it
http://digilander.libero.it/mario.milani/teaching.html





Proteine non strutturali (NS) e replicazione dei flavivirus

• Traduzione: Ribosomi umani, costruzione della poliproteina virale

• Attivazione delle proteine virali: Proteasi virale NS2B/NS3 (NS3 N-ter) + proteasi umane

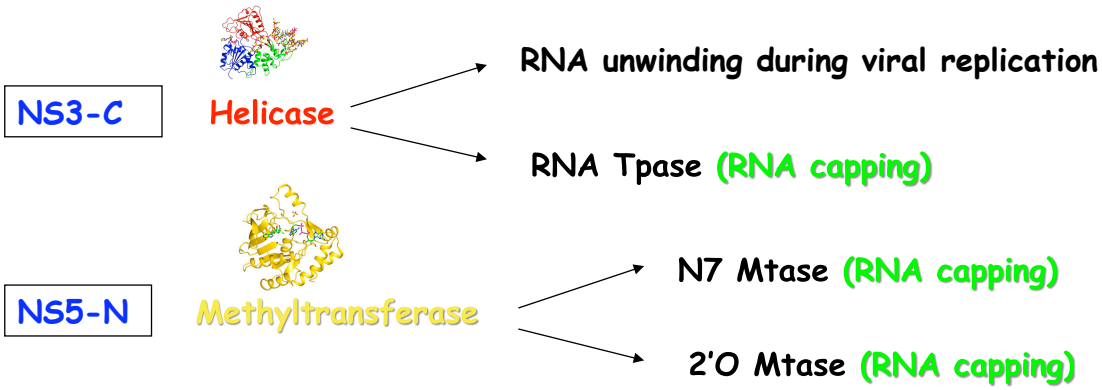
• Replicazione del genoma (trascrizione):. RNA polimerasi (NS5 C-ter) + elicasi (NS3 C-ter)

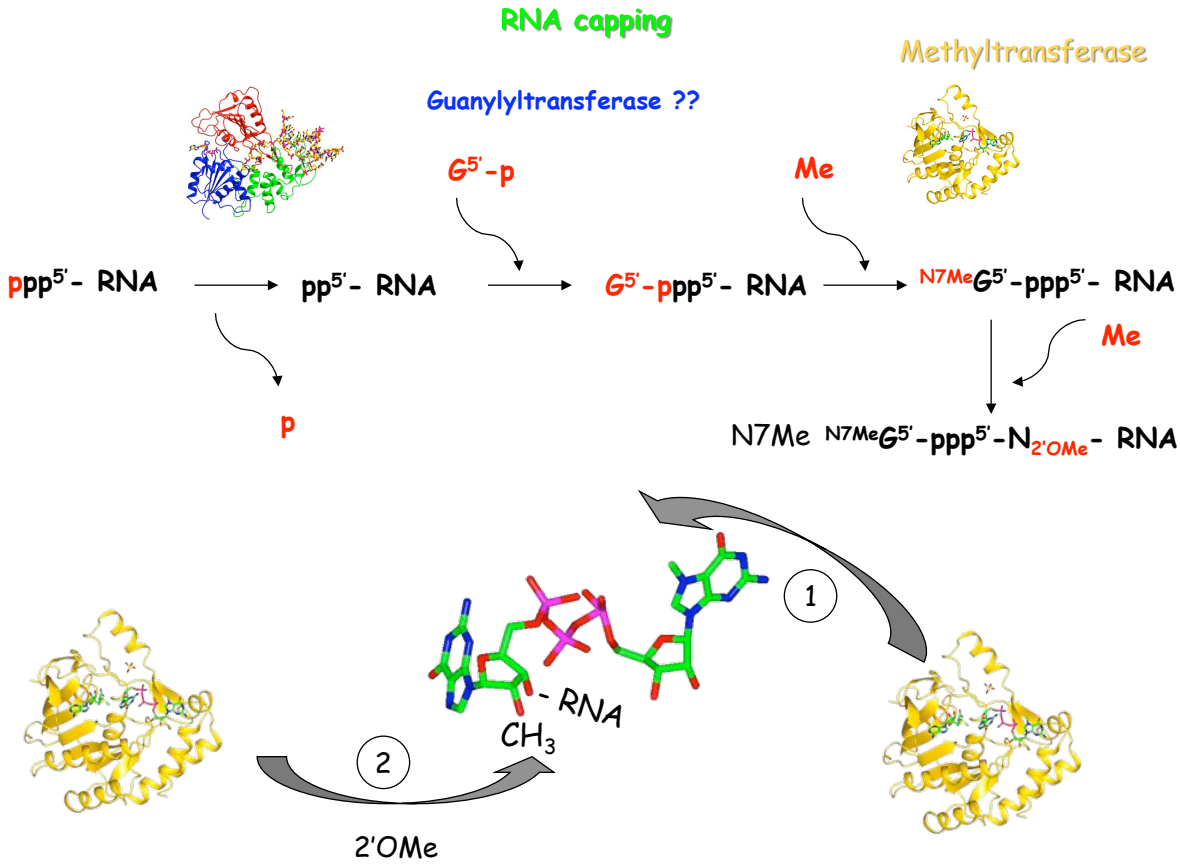
• Capping dell'RNA: NTPasi (NS3 C-ter) + guanilil trasferasi (?) + Mtasi (NS5 N-ter)

• Costruzione e assemblaggio di nuove particelle virali: (?)

NS3: 70 kDa, N-ter (1-180) Ser proteasi, C-ter (181-630) elicasi e RNA trifosfatasi

NS5: 104 kDa, N-ter Mtasi (1-265), C-ter (270-930) RNA polimerasi



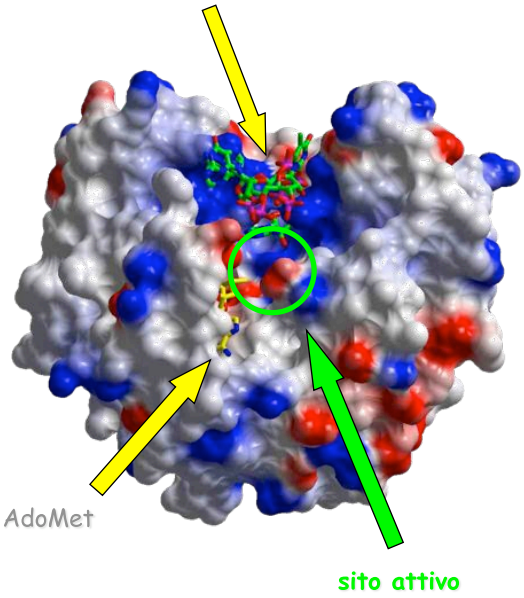


Ray *et al.*, 2006

come funziona la MTase ?

1. N7 metilazione

bassa affinita'/primo sito di legame dell'RNA



Modello

AdoMet

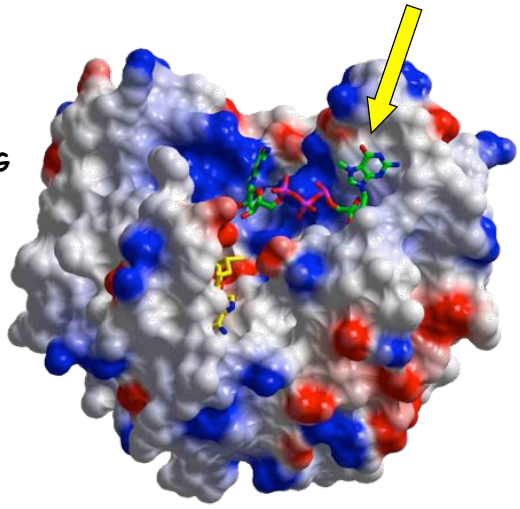
RNA-^{N7Me}GpppG

l'RNA scorre

AdoHcy

2. 2'O methylation

alta affinita'/secondo sito di legame dell'RNA



Struttura 3D

Mastrangelo *et al.*, 2007
Bollati *et al.*, 2009

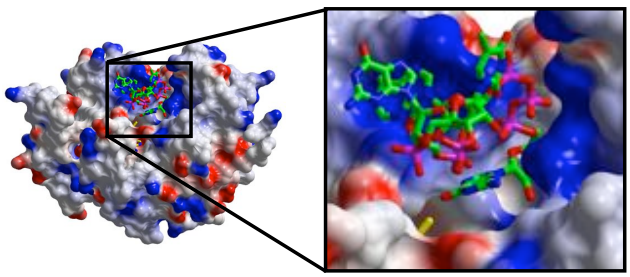
MTase molecular dynamics simulation

Autodock4 with short capped RNA

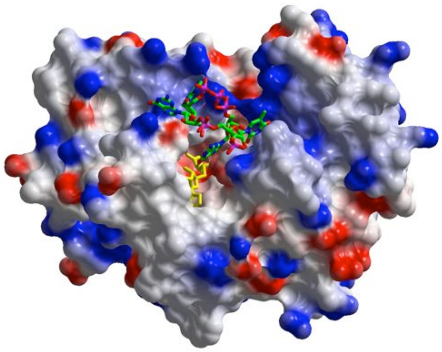
GpppAGTp

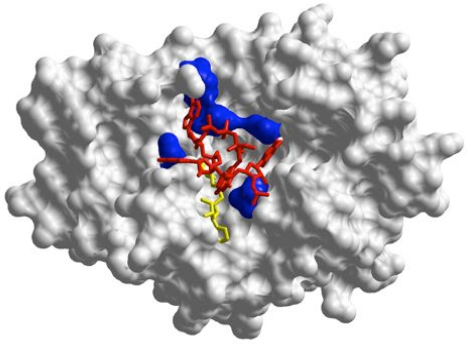
-11.86 kcal/mol

Ki = 2.01 nM



After 9 ns of MD simulation (gromacs)





5 residues showed to be crucial for N7 MTase activity:
Arg37, Arg57, Arg84, Glu149, and Arg213
Dong *et al.*, 2008

Protein structure-function analysis

↓

Hypothesis on activity

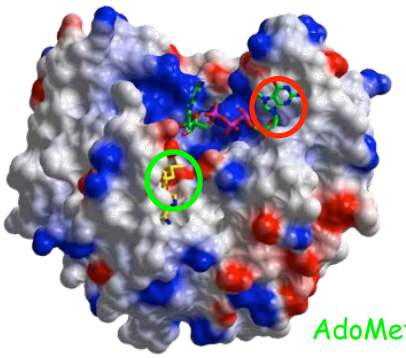
↓

New inhibition site

↓

in silico docking search

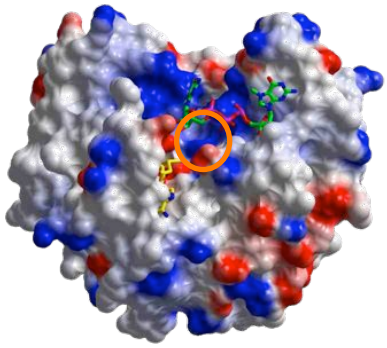
Mtase 'classical' inhibition sites:



High affinity cap RNA binding site

AdoMet binding site

Mtase new inhibition site:



Active site

Programmi utili per il docking

Preparazione dei ligandi

MarvinBeans http://www.chemaxon.com/marvin/download.html

babel http://cds.dl.ac.uk/cds/interface_and_utilities/babel.html

banche dati http://www.qsarworld.com/free-databases.php

Prodrg http://davapc1.bioch.dundee.ac.uk/programs/prodrg/

Preparazione della proteina e docking

MGLTools http://mgltools.scripps.edu/downloads

autodock4 http://autodock.scripps.edu/downloads

Preparazione della proteina: amminoacidi fissi

Proteina: formato .pdb + eventuali ligandi (i.e. FAD, AdoMet)

↓ pmv (parte di MGLTools): aggiunta degli idrogeni e delle cariche

formato .pdbqt, +H e cariche

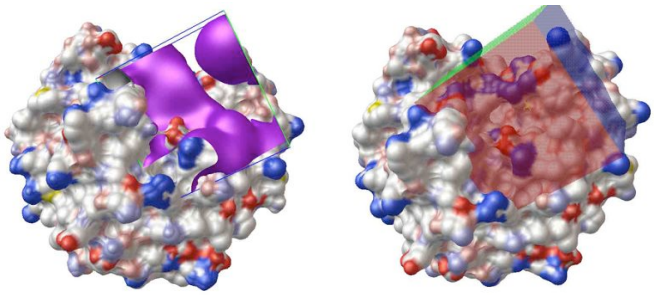
si costruisce una griglia intorno alla zona che interessa

gpf = grid parameter file

```
npts 54 54 54                           # num.grid points in xyz
gridfld bir2.maps.fld                   # grid_data_file
spacing 0.375                           # spacing(A)
receptor_types A C H HD N OA SA         # receptor atom types
ligand_types A C HD N NA OA SA          # ligand atom types
receptor bir2_rigid.pdbqt               # macromolecule
gridcenter 32.779 33.786 53.122         # xyz-coordinates or auto
```

autogrid4

autogrid4 -p ./bir2.gpf -l ./bir2.glg &



isosuperficie energia di e

preparazione dei ligandi

Ligandi: formato .sdf in 2D

```
1.6250   -1.8792    0.0000 N   0  3  3  0  0  0  0  0  0  0  0  0
9.4417   -3.7417    0.0000 N   0  3  0  0  0  0  0  0  0  0  0  0
4.3917   -4.8875    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
1.8292   -3.1167    0.0000 C   0  0  3  0  0  0  0  0  0  0  0  0
2.4667   -2.5500    0.0000 C   0  0  3  0  0  0  0  0  0  0  0  0
5.1042   -4.4667    0.0000 C   0  0  3  0  0  0  0  0  0  0  0  0
…
```

(Pymol)

↓ molconvert

/Applications/ChemAxon/MarvinBeans/bin/molconvert -vv -3 sdf LO3300.sdf -o Lo3D.sdf

.sdf in 3D

↓ babel (convert.csh)

```
#!/bin/csh -f
set n = 1
echo $n
while ( $n < 1268 )
babel  -isdf Lo3D.sdf -f "$n" -l "$n"  -opdb pdb/di"$n".pdb
@ n = $n + 1
end
```

.pdb

↓ prepare_ligand4.py (prepare_ligand.csh)

```
#!/bin/csh -f
foreach f (`ls ../pdb/*.pdb`)
echo $f
./prepare_ligand4.py -A bonds_hydrogens -l $f -o "$f"qt
end
```

.pdbqt, ligando + H e carica, e torsioni

Docking

autodock4 -p bir2.dpf -l bir2.dlg &

dpf = docking parameter file

```
…
ligand_types A C HD N NA OA           # atoms types in ligand
fld bir2.maps.fld                     # grid_data_file
map bir2.A.map                        # atom-specific affinity map
…
move smac10.pdbqt                     # small molecule
about -21.5599 9.9909 48.6289         # small molecule center
…
ga_run 1                              # do this many hybrid GA-LS runs
```

semiempirical free energy force field

variazione di energia nel ligando

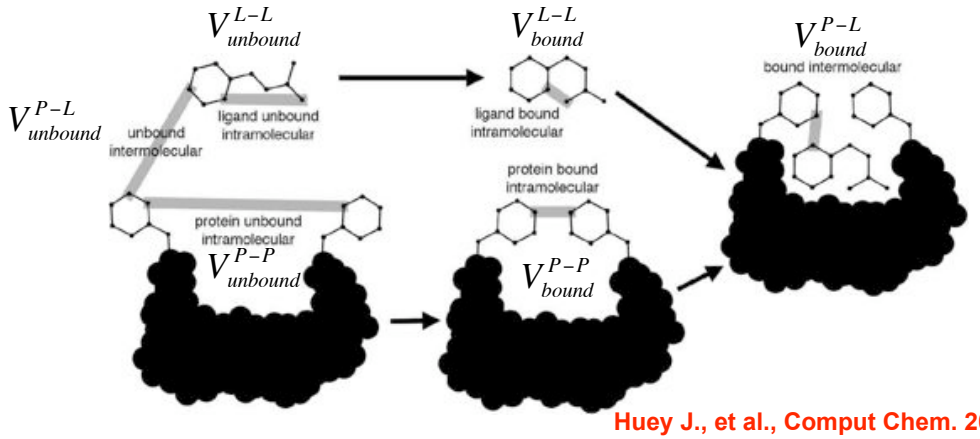
variazione di energia nella proteina

variazione di energia nell'interazione

$$\Delta G = \left(V^{L-L}_{bound} - V^{L-L}_{unbound}\right) + \left(V^{P-P}_{bound} - V^{P-P}_{unbound}\right) + \left(V^{P-L}_{bound} - V^{P-L}_{unbound} + \Delta S_{conf}\right)$$

variazione di energia libera del sistema proteina-ligando

entropia conformazionale perduta nel legame



Huey J., et al., Comput Chem. 2007, 28, 1145-52.

$$\Delta G = \left(V_{bound}^{L-L} - V_{unbound}^{L-L}\right) + \left(V_{bound}^{P-P} - V_{unbound}^{P-P}\right) + \left(V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S_{conf}\right)$$

se la proteina non si muove

proteina e ligando non interagiscono nello stato non legato

$$\Delta S_{conf} = W_{conf} N_{tors}$$

cost. empiriche che pesano i vari termini

Amber force field

angolo del legame idrogeno

$$V = W_{vdw} \sum_{i,j} \left(\frac{A_{i,j}}{r_{i,j}^{12}} - \frac{B_{i,j}}{r_{i,j}^{6}} \right) + W_{h-bound} \sum_{i,j} E(\vartheta) \left(\frac{C_{i,j}}{r_{i,j}^{12}} - \frac{D_{i,j}}{r_{i,j}^{10}} \right) +$$

$$W_{elec} \sum_{i,j} \left(\frac{q_i q_j}{\varepsilon(r_{i,j}) r_{i,j}} \right) + W_{sol} \sum_{i,j} \left(S_i V_j + S_j V_i \right) \exp\left(-\frac{r_{i,j}^2}{2\sigma^2} \right)$$

int. regolata dalla dist. i,j

Volume degli atomi intorno a i

cost. dielettrica dip. dalla dist.

parametro di solvatazione dell'atomo j

termine di desolvatazione:

stima della quantita' di solvatazione

$$W_{sol} \sum_{i,j} \left(S_i V_j + S_j V_i \right) \exp\left(-\frac{r_{i,j}^2}{2\sigma^2} \right)$$

energia richiesta a trasferire un atomo da uno stato completamente idratato a uno completamente protetto

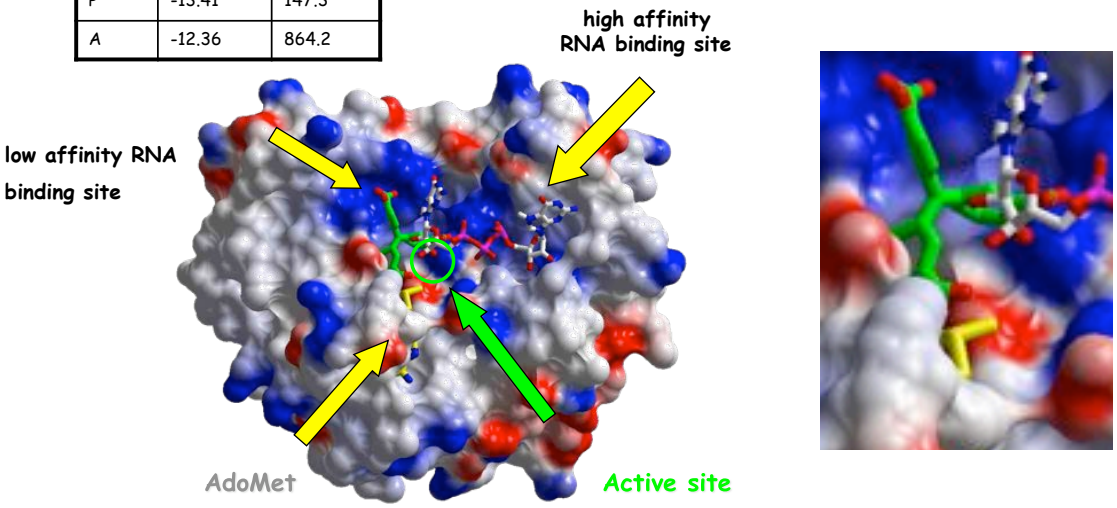
parametri fittati dal confronto con dati sperimentali

$$S_i = \left(A_i + B_i |q_i| \right)$$

MTase inhibition

MTase docking

Name	Energy (kcal/mol)	Ki (pM)
N	-16.16	1.4
P	-13.41	147.3
A	-12.36	864.2



ATA is a non-specific enzyme inhibitor (Bina-Stein & Tritton 1975)

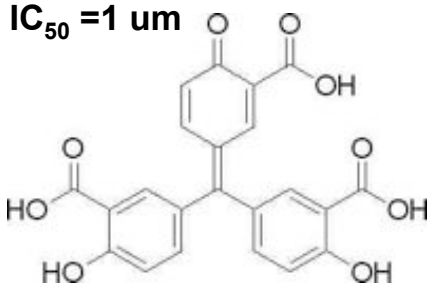
A recent patent claims that ATA is an inhibitor against SARS coronavirus (He *et al.*, 2004; Yap *et al.*, 2005)

Milani *et al.*, 2009

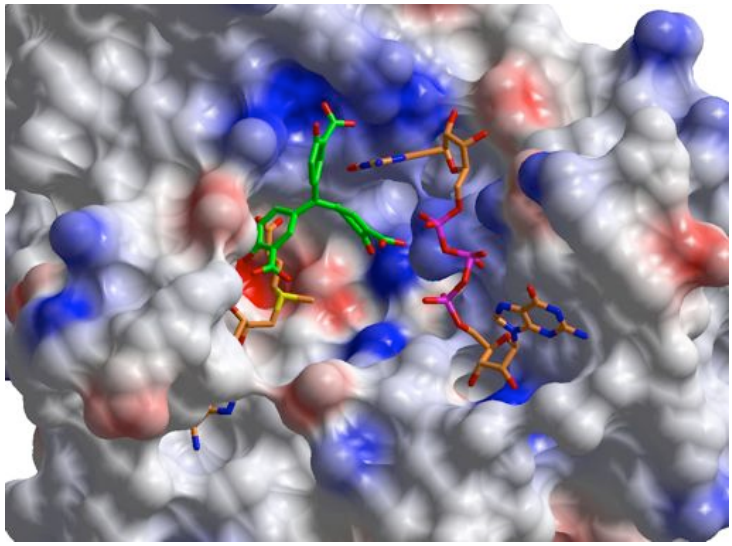
Inibitore Mtasi

Use of brain or high-throughput ?

da 1600 mol.
IC_{50} =1 um



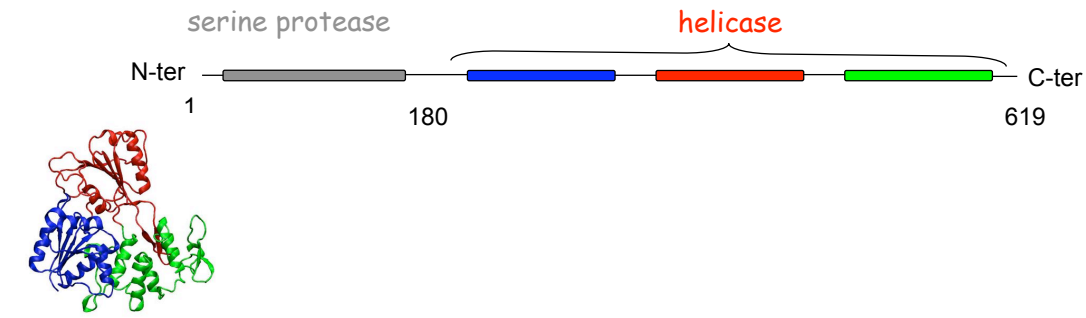
Aurintricarboxylic acid

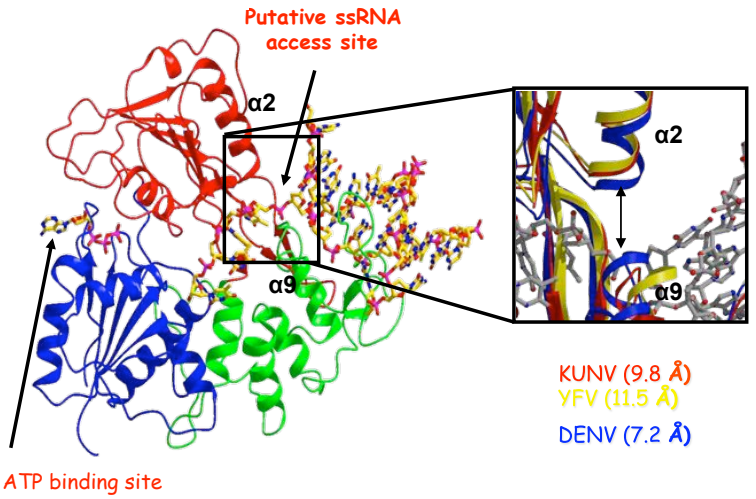


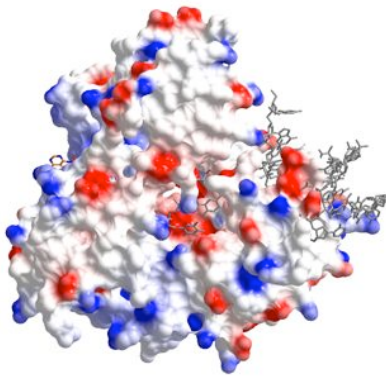
Luzhkov et al., 2007

Partendo da 2.1 M di molecole
IC_{50} =60 um

NS3 elicasi



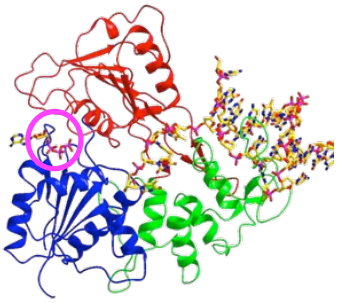




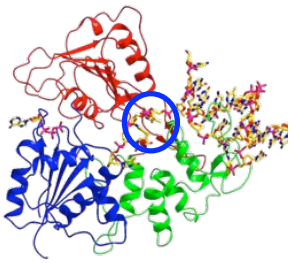
Mastrangelo et al., 2007

Hel 'classical' inhibition:

ATP binding site



Putative ssRNA access



Protein structure-function analysis

↓

Hypothesis on activity

↓

New inhibition site

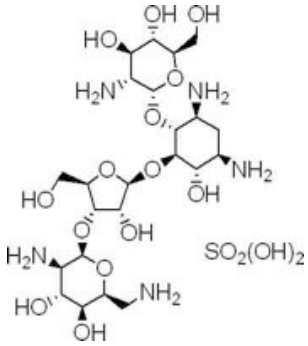
↓

***in silico* docking search**

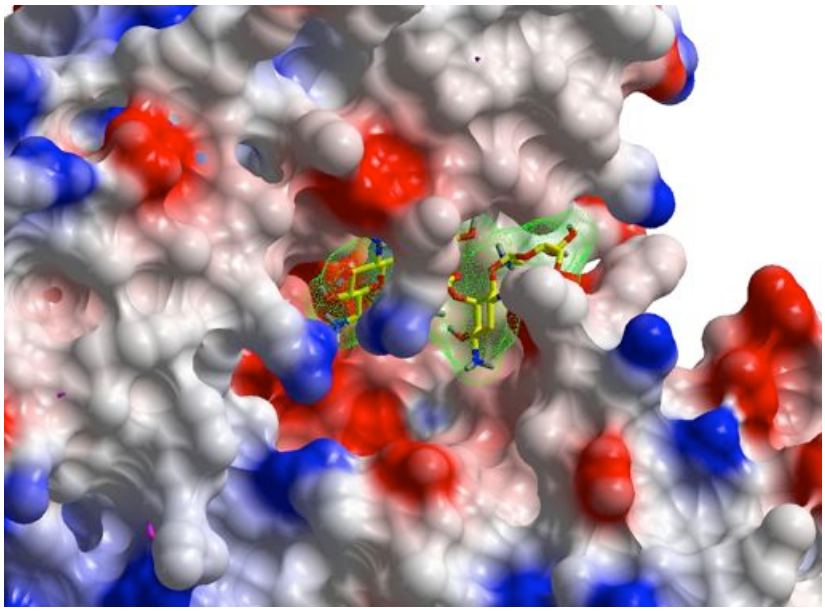
↓

New inhibitors + validation of the proposed protein function mechanism

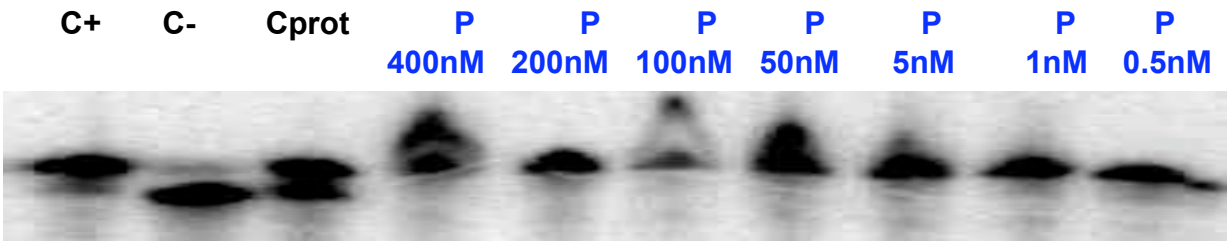
inibitori helicase

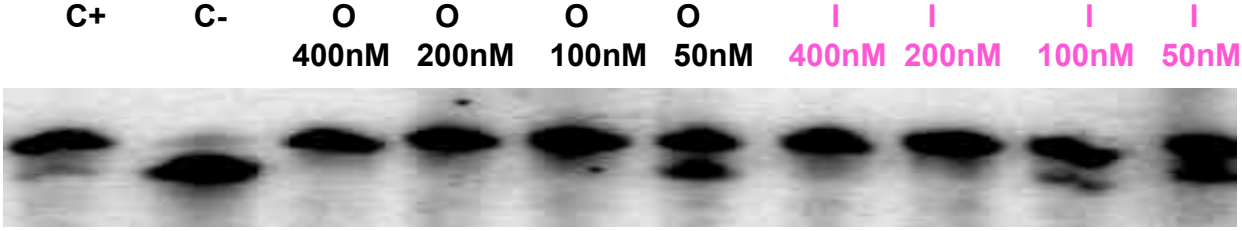


Paromomycin sulfate



Kokobera virus helicase inhibition assay
Proteina 200 nM





SOMMARIO

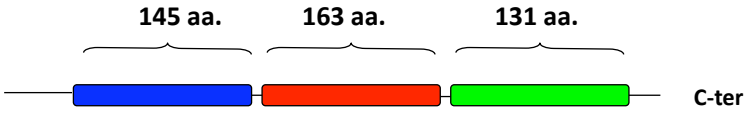
1. avere aspettative ragionevoli: dati sperimentali complementari sono essenziali !

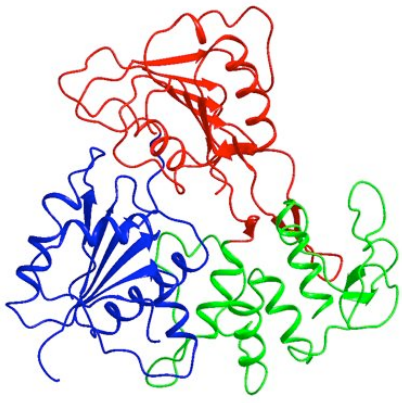
2. avere aspettative ragionevoli: il docking fornisce dei *modelli*, non e' la realtà !

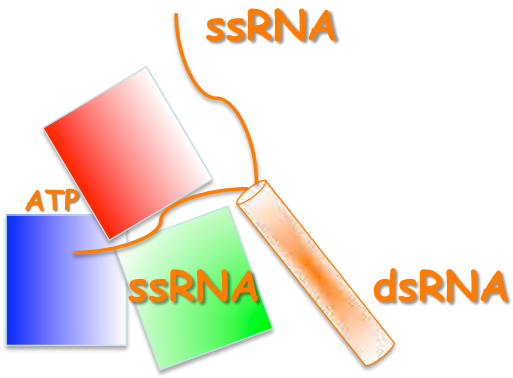
i *modelli*, dipendono dalle assunzioni teoriche che si fanno per descrivere l'attivita' della proteina

3. usare piu' la testa dell' *high-throughput* !

Flavivirus Helicase





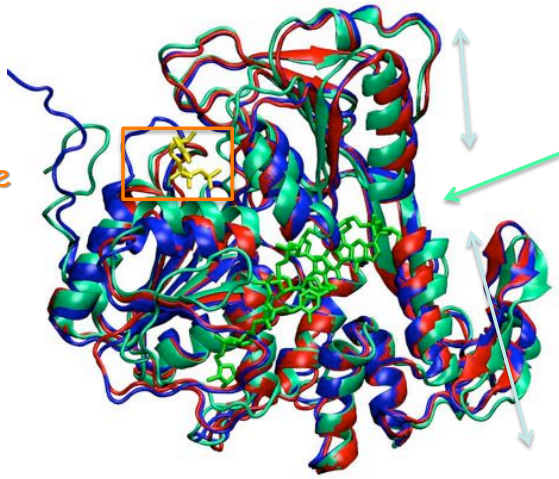


Crystal structures did not answer to the main questions:

Helicase is a motor protein: how does it work?

How ATP hydrolysis is related to protein motion?

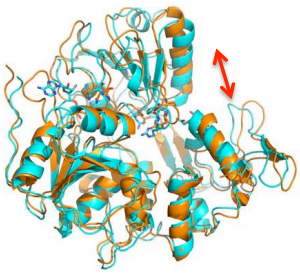
apo Hel
ATP Hel
RNA Hel

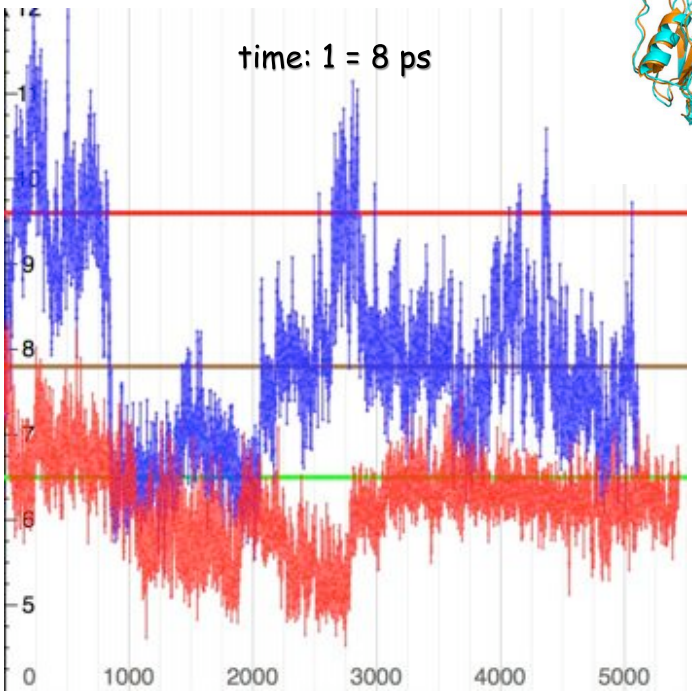


α helices distance: ssRNA access site

apo Hel 40.9 ns: open 20%, intermediate 54%, closed 24%

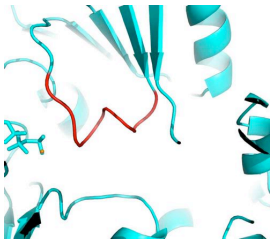
ATP Hel 43.5 ns: closed 100%

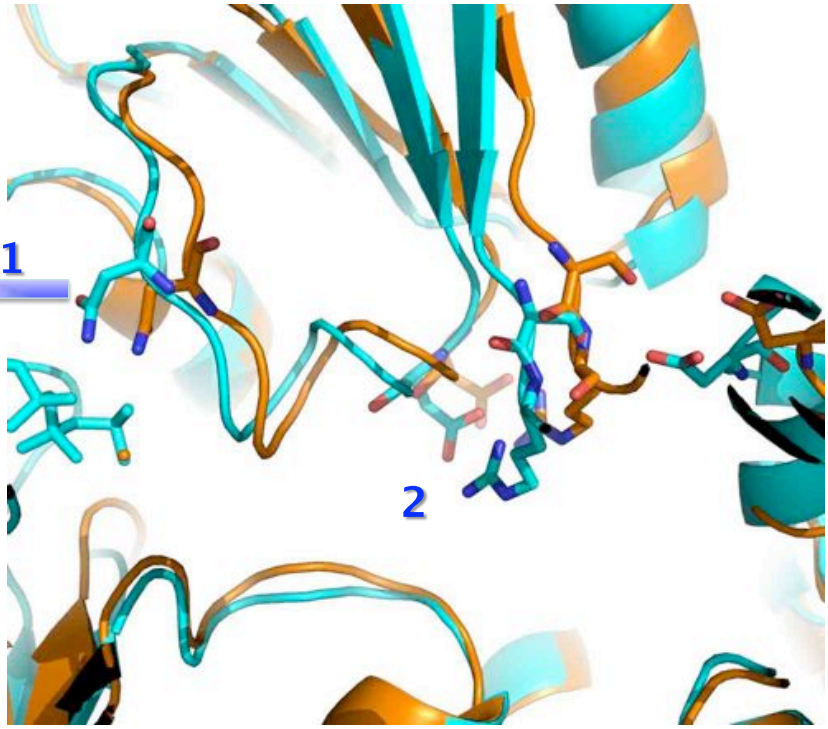




why does ATP induce a closed state ?

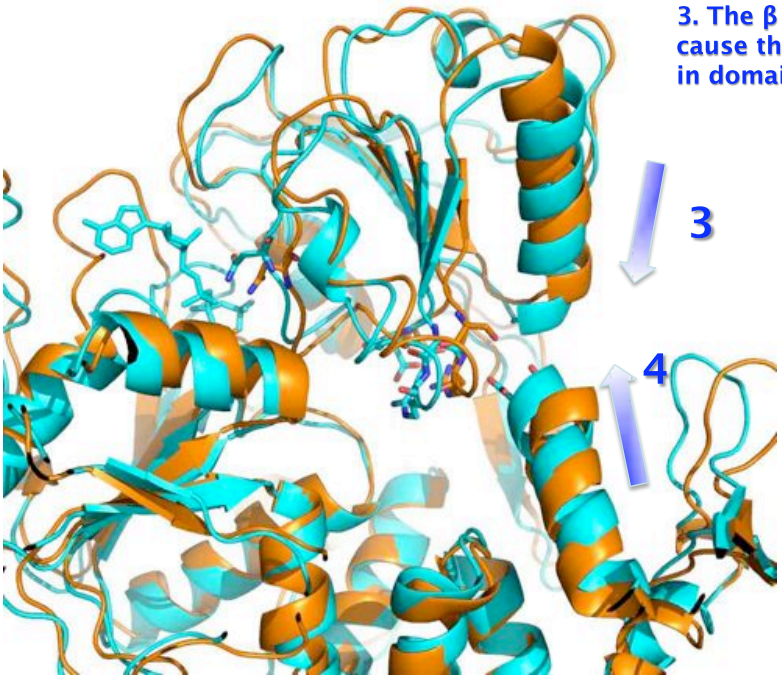
1. Asn416 interacts with ATP shifting conserved motive V





2. The shift is transmitted by salt bridge between Asp409 and Arg387 and amplified by the β sheet of domain 2

why does ATP induce a closed state ?



3. The β sheet translation cause the closure of helix in domain 2

4. Interaction between Ser386 and Asp603 induces the closure of helix in domain 3

how ATP hydrolysis is related to RNA unwinding?

1. ATP binding shifts an open-closed equilibrium toward a closed state blocking protein progression along ssRNA

2. ATP is hydrolyzed to ADP (lower affinity) by the protein and thus released

3. open-closed equilibrium of apo protein and Brownian motion causes protein progression onto ssRNA and RNA unwinding