



Comparative ligand binding of Neuraminidase structures

Luciano Milanesi

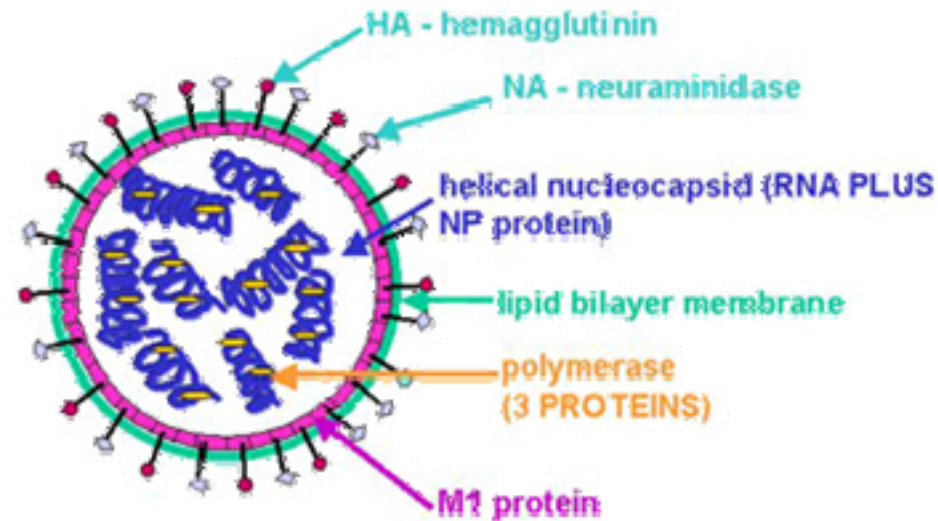
Institute for Biomedical Technology

National Research Council





Structure of the Influenza A virus



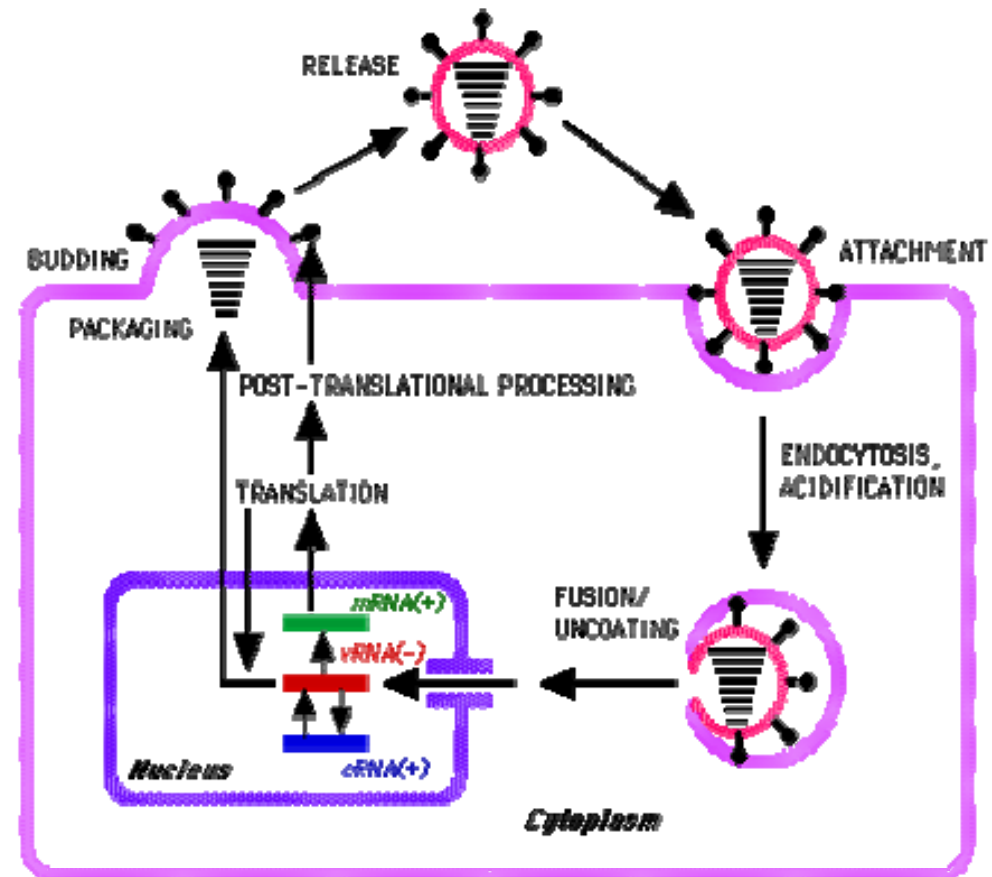
The outer shell of the influenza A virus is composed of a lipid bilayer with inserted proteins:

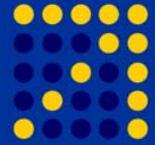
- the function of **Hemagglutinin (HA)** is to attach the virus to a host cell
- **Neuraminidase (NA)** is involved in facilitating the release of newly produced virus particles from the host cell
- M1 protein has multiple regulatory functions during the infectious cycle



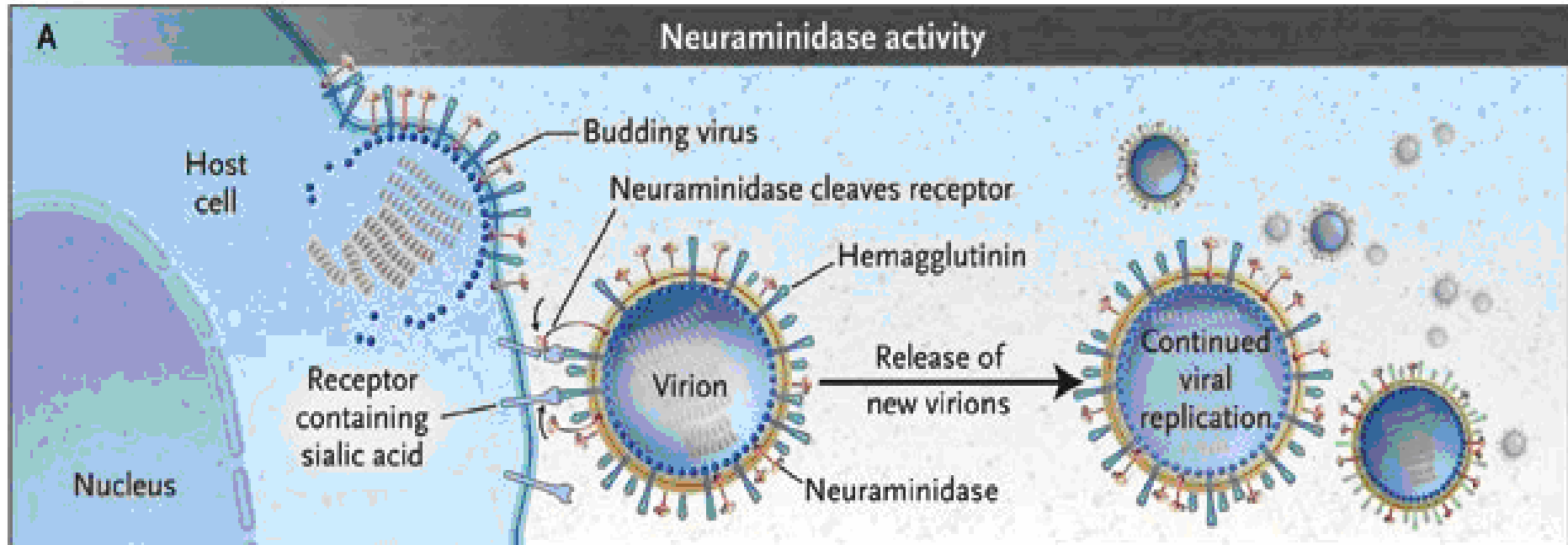
Replication cycle of influenza virus

- **Attachment:** the virus binds the receptors on the surface of the host cell utilizing the HA
- **Endocytosis:** after the binding, the particle is internalised into endosomes
- **Fusion and uncoating events:** are pH dependent, resulting in the release of the viral genome into the cytoplasm, the **vRNPs** are then imported into the nucleus for replication
- **Protein synthesis:** Viral messenger RNAs (**mRNAs**) are exported out of the nucleus into the cytoplasm for protein synthesis
- **Budding:** progeny viruses come together and bud from the plasma membrane





Neuraminidase activity



The hemagglutinin binds the sialic acid of the receptor localized on the surface of the host cell

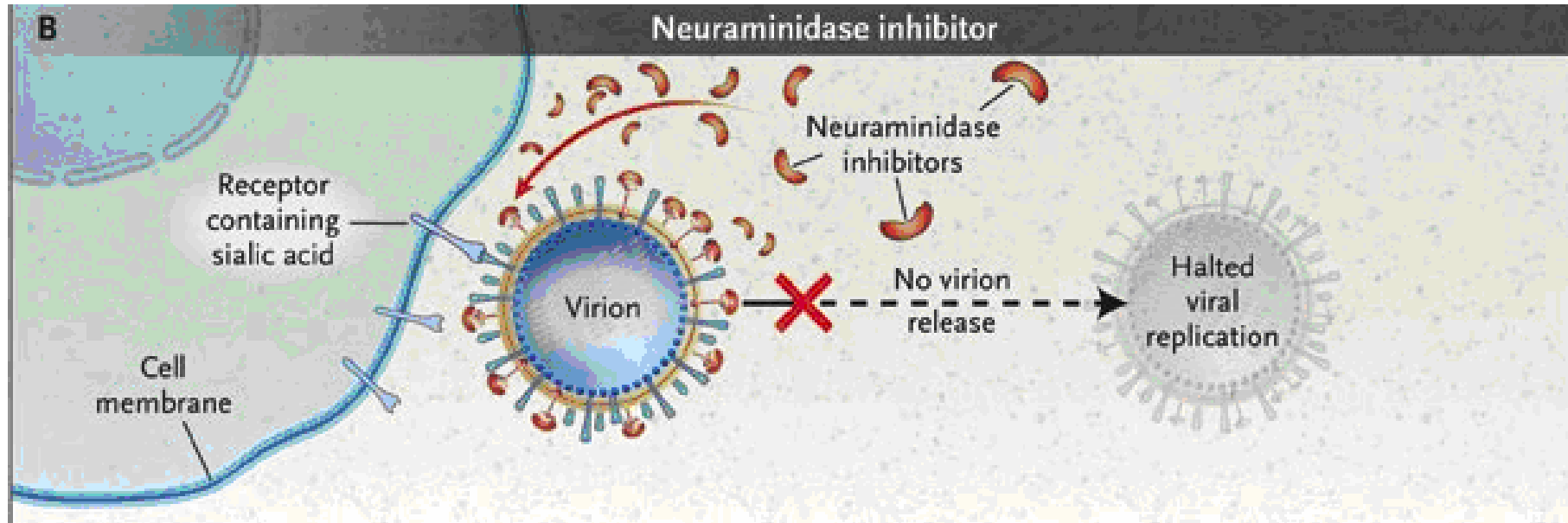
The interaction hemagglutinin-sialic acid receptor interferes with the release of the progeny influenza virus from infected host cells

The neuraminidase cleaves the sialic acid substrate of the host cell receptor

This cleavage releases the viruses, which can now infect new cells



Neuraminidase inhibitors



Neuraminidase inhibitors (NAIs) interfere with the release of the progeny influenza virus from the infected host cell

Potent Neuraminidase inhibitors fitting into the active site pocket prevent the neuraminidase activity and interfere with the release of progeny influenza virus from infected host cells

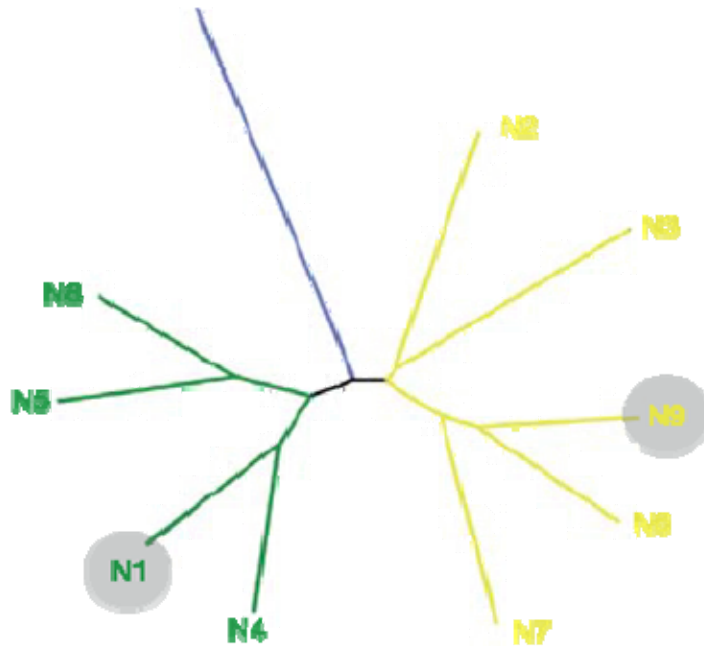
The neuraminidase viruses is considered a valid target for antiviral drugs



Neuraminidase subtypes

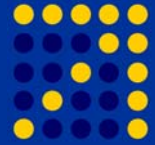
At present, two distinct groups of Neuraminidase are known:

- **group 1** includes subtypes N1, N4, N5 and N8
- **group 2** containig the subtypes N2, N3, N6, N7 and N9



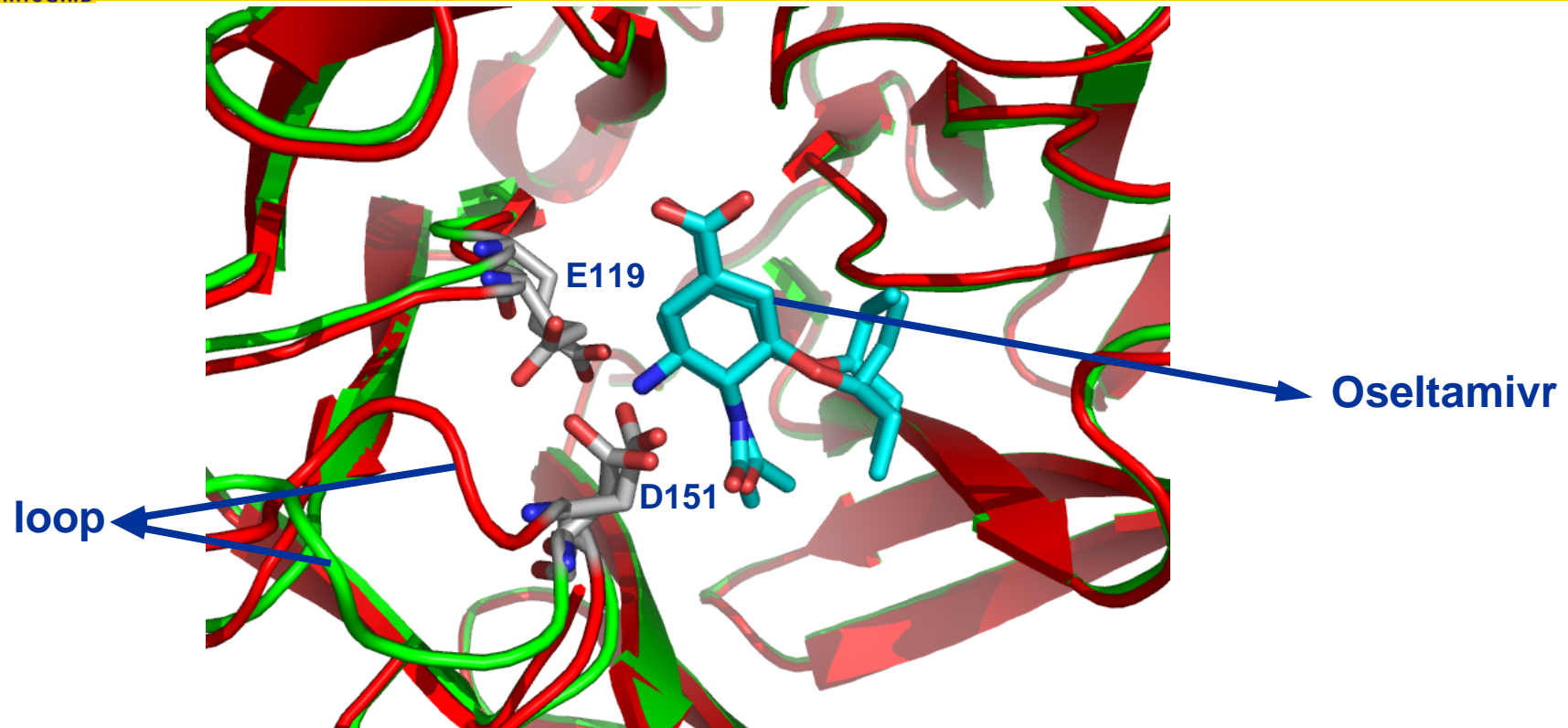
Pylogenetic tree of the neuraminidase of influenza A group 1 (green) and group 2 (yellow)

The design of neuraminidase inhibitors is based on the structure of group 2, assuming that the overall characteristics of the active site were conserved among the subtype.

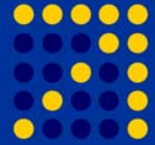


X-ray structures for Neuraminidase of group 1

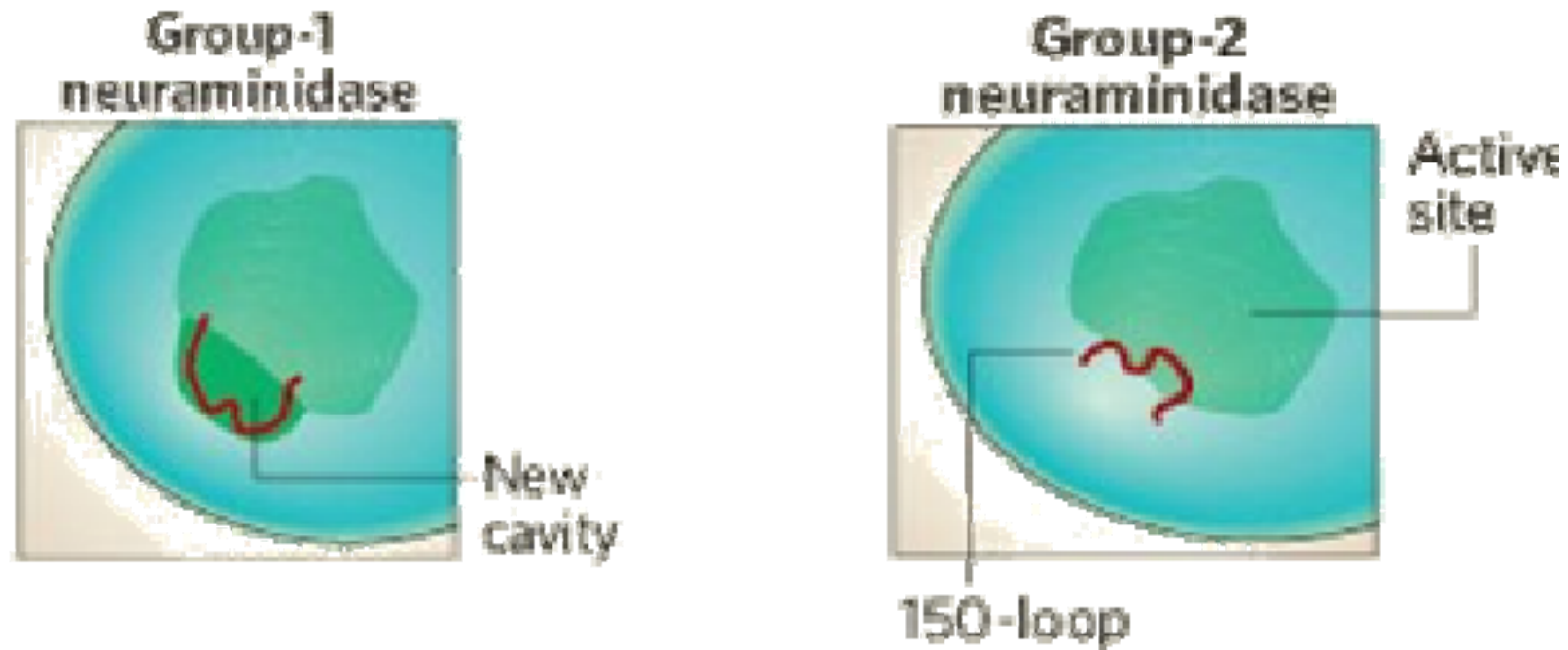
BioinfoGRID



X ray structures of group 1 neuraminidase showed that two different conformations of a **loop adjacent to the active site** can be detected, depending on ligand concentration. The loop is in the **open conformation** in the **apo-form or at low ligand concentration**, that forms a large cavity near the active site (**green structure**) **High ligand concentration** induces a conformational change of loop 150, that **closes** up the cavity on the ligand (**red structure**)

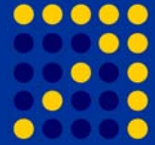


Structures of Neuraminidase subtypes



In **group 1** the loop is in the **open conformation** (apo-form and in low ligand concentration)

In **group 2** the loop is in the **closed conformation**



- We have screened, by docking simulation, a large set of compounds in the open and closed conformation of N1 neuraminidase

- The purpose of these studies is to compare the relative binding mode of the set within the two structures and to identify binders and functional groups of ligands that make contacts with the key residues in the open form



The dataset used for the screening contains:

- **49 known inhibitors**
- **1746 decoy molecules** designed in order to match the physical properties of specific ligands for the neuraminidase in the closed form (pdb code 1a4g)

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

DUD

A Directory of Useful Decoys

Welcome to DUD, a **directory of useful decoys for benchmarking virtual screening**. DUD is designed to help test docking algorithms by providing challenging decoys. It contains:

- A total of 2,950 active compounds against a total of 40 targets
- For each active, 36 "decoys" with similar physical properties (e.g. molecular weight, calculated LogP) but dissimilar topology.

DUD is provided by the [Shoichet Laboratory](#) in the [Department of Pharmaceutical Chemistry](#) at the [University of California, San Francisco \(UCSF\)](#). To cite DUD, please reference [Huang, Shoichet and Irwin, J. Med. Chem., 2006, 49\(23\), 6789-6801, doi 10.1021/jm060835c](#). There is a [DUD web page](#) where you can discuss DUD. We thank [NICMS](#) for financial support (GM71896). For correspondence about DUD, please write John Irwin ji@cgl.ucsf.edu.

[Go to Download Page](#)

DUD is drawn from [ZINC](#), a database of commercially available compounds for virtual screening, so compounds in DUD are purchasable, although some may become depleted over time. You may download DUD either in packages (some of which are large!) or you may browse the files and download them individually. Anticipating that problems will be found, and corrected, in DUD, we number our releases as follows:

Release	Date	Comments
1	Oct 1, 2006	Original release
2	Oct 22, 2006	Proofreading corrections accompanying the final manuscript.

Last updated by John Irwin ji@cgl.ucsf.edu
On October 22, 2006

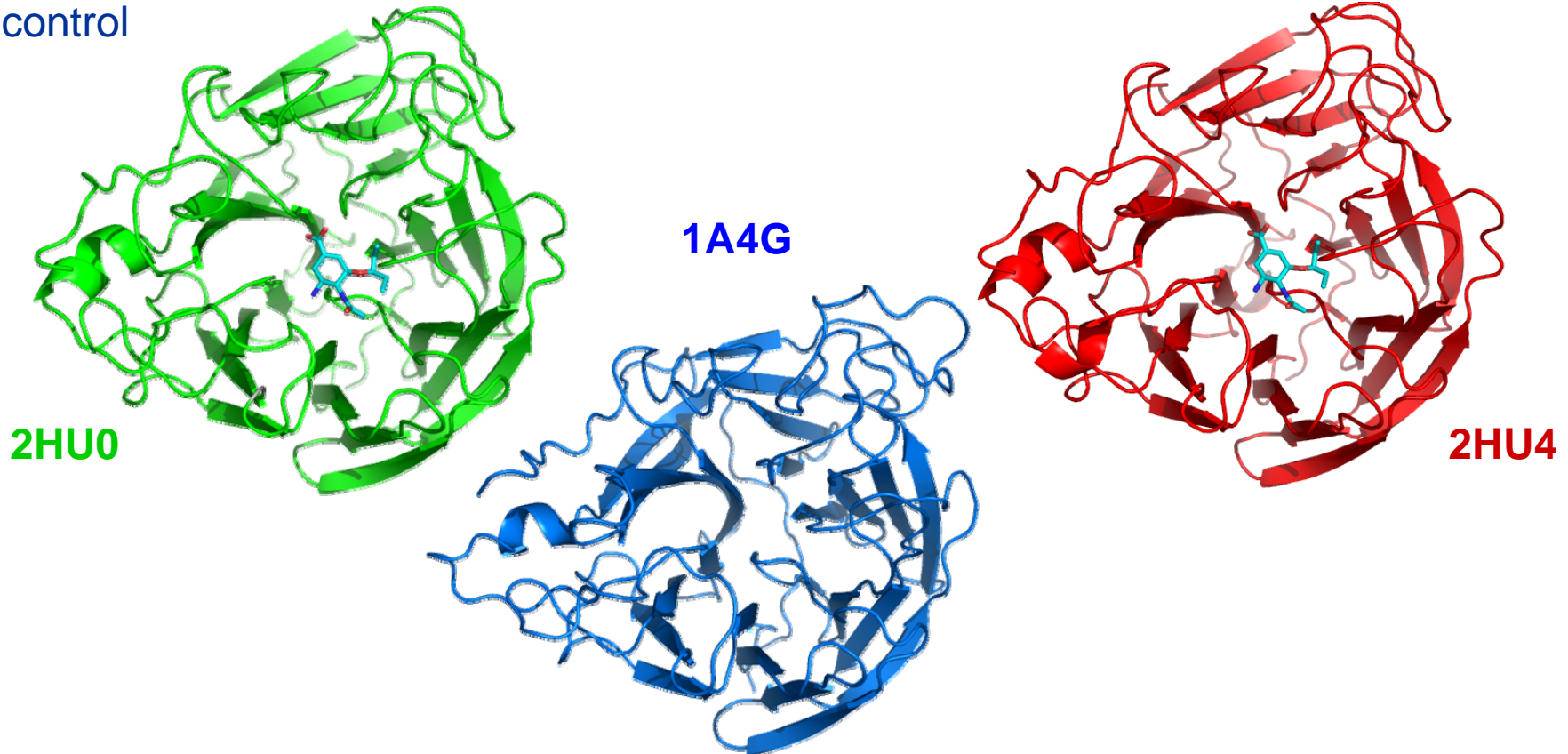
The dataset was downloaded from the site <http://blaster.docking.org.dud>

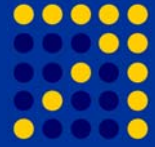


Target Structures for screening

The target structures for the screening were:

- Neuraminidase N1 subtype in open (pdb code **2HU0**) and closed conformation (pdb code **2HU4**)
- pdb code **1A4G** (Neuraminidase of influenza B in closed conformation) as control





The docking experiments were performed with Autodock 3.05

Lamarckian Genetic Algorithm:

Population size: 150

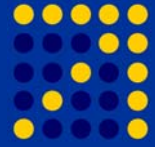
Mutation rate: 0.02

Crossover rate: 0.8

50 Run

Grid points: 60 x 60 x 40

grid points spacing: 0.375 Å



- Due to the high computational load needed to screen a ligand library against a set of mutated proteins, a high throughput approach can be very useful to reduce the simulation time
- The implementation of the protein domain analysis consists in creating an efficient system to coordinate the jobs submission, to check the computation status and to collect the results



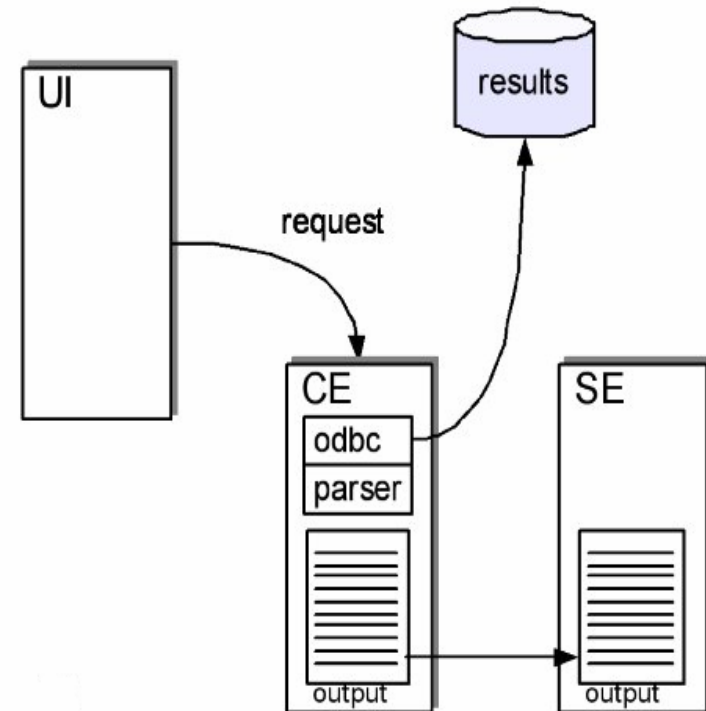
- The whole simulation challenge infrastructure relies on a relational database management system
- The input section of the database consists in a couple of tables that store structural information about both the ligands and the proteins involved in the simulation
- The output section organizes output data in order to collect information about the best poses of each simulation



- Another important section the of database is related to the job coordination on the grid platform
- Using a specific tool, called *vnas*, it is possibile to monitor the status of the job, and retrieve the results when a job is correctly finished or resubmit it again in case of failure
- In this context the relational database plays a crucial role in keeping track of the status of each grid job



- An important issue for docking screening in grid is the post process analysis of the results that can be **computationally very expensive**
- We face this problem when performing both the parsing of the output results and the storage of the data in the output database directly from the distributed computational resource
- For each simulation the results have been filtered and only the docking poses with an energy below a specific threshold and with a cluster population of more than one were stored in the database
- Using this approach for each docking simulation we stored a number of hits because the best poses can be different from the top ranked by the docking software





In order to make results rapidly accessible a **web interface has been developed** on the top of the output database

- It allows the user to browse docking results and query data both in relation to specific proteins or specific ligands
- For each different best hits of a simulation it is possible to view the docking energy, the cluster number and the mean energy
- Moreover for each poses it is possible to visualize the histogram of the whole simulation to verify the global distribution of the results
- For each result it is possible to download the coordinate of the docked ligand in order to perform further analysis

Protein 2hu0

1 to 20 of 66148 2hu0_id_bestsits_DESC,0

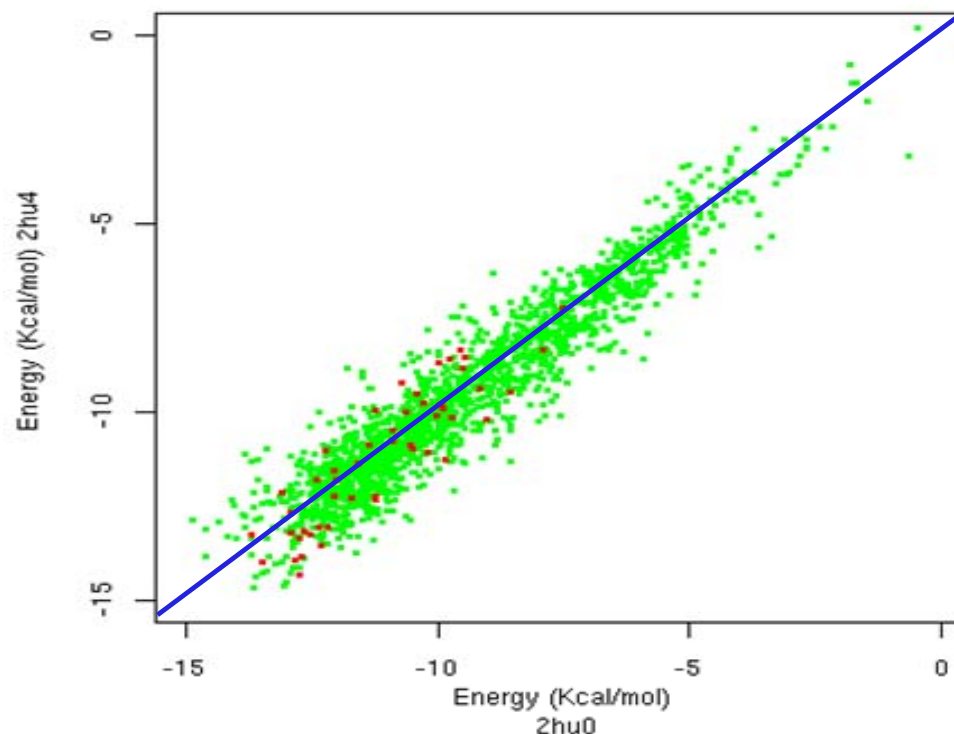
ID Besthits	ID Simulation	Ligando	Rank	Energy level	Run	Mean Energy	Num Cluster	File coordinate	File histogram	Info
564618	16063	4134498	2	-12.5	8	-12.16	2	Coordiante	Histogram	Info
564585	16042	4134497	31	-10.44	41	-10.43	2	Coordiante	Histogram	Info
564584	16042	4134497	22	-10.81	47	-10.52	2	Coordiante	Histogram	Info
564583	16040	4134495	31	-10.96	8	-10.83	2	Coordiante	Histogram	Info
564582	16037	4134492	11	-11.39	21	-11.39	2	Coordiante	Histogram	Info
564581	16037	4134492	9	-11.44	42	-11.26	2	Coordiante	Histogram	Info
564580	16037	4134492	8	-11.56	7	-11.49	2	Coordiante	Histogram	Info
564579	16037	4134492	6	-11.72	24	-11.57	3	Coordiante	Histogram	Info
564578	16037	4134492	5	-11.81	49	-11.72	2	Coordiante	Histogram	Info
564577	16037	4134492	3	-11.89	29	-11.73	4	Coordiante	Histogram	Info
564576	16037	4134492	2	-11.97	46	-11.73	3	Coordiante	Histogram	Info
555890	13709	4591827	50	-7.63	4	-7.63	1	Coordiante	Histogram	Info
555889	13709	4591827	49	-7.69	1	-7.69	1	Coordiante	Histogram	Info
555888	13709	4591827	48	-7.77	44	-7.77	1	Coordiante	Histogram	Info
555887	13709	4591827	47	-7.79	32	-7.79	1	Coordiante	Histogram	Info
555886	13709	4591827	46	-7.8	50	-7.8	1	Coordiante	Histogram	Info
555885	13709	4591827	45	-7.88	22	-7.88	1	Coordiante	Histogram	Info
555884	13709	4591827	44	-7.91	7	-7.91	1	Coordiante	Histogram	Info
555883	13709	4591827	43	-8.02	47	-8.02	1	Coordiante	Histogram	Info
555882	13709	4591827	42	-8.04	9	-8.04	1	Coordiante	Histogram	Info



Scatter plot of Docking Energy

From this analysis, we observed that the tested ligands bind the **open and closed conformation with comparable docking energies**

- The scatter plot reveals that a large number of ligands lie along the diagonal
- Few points are found above and below the diagonal corresponding to the compounds that bind preferentially to the open and closed conformation, respectively

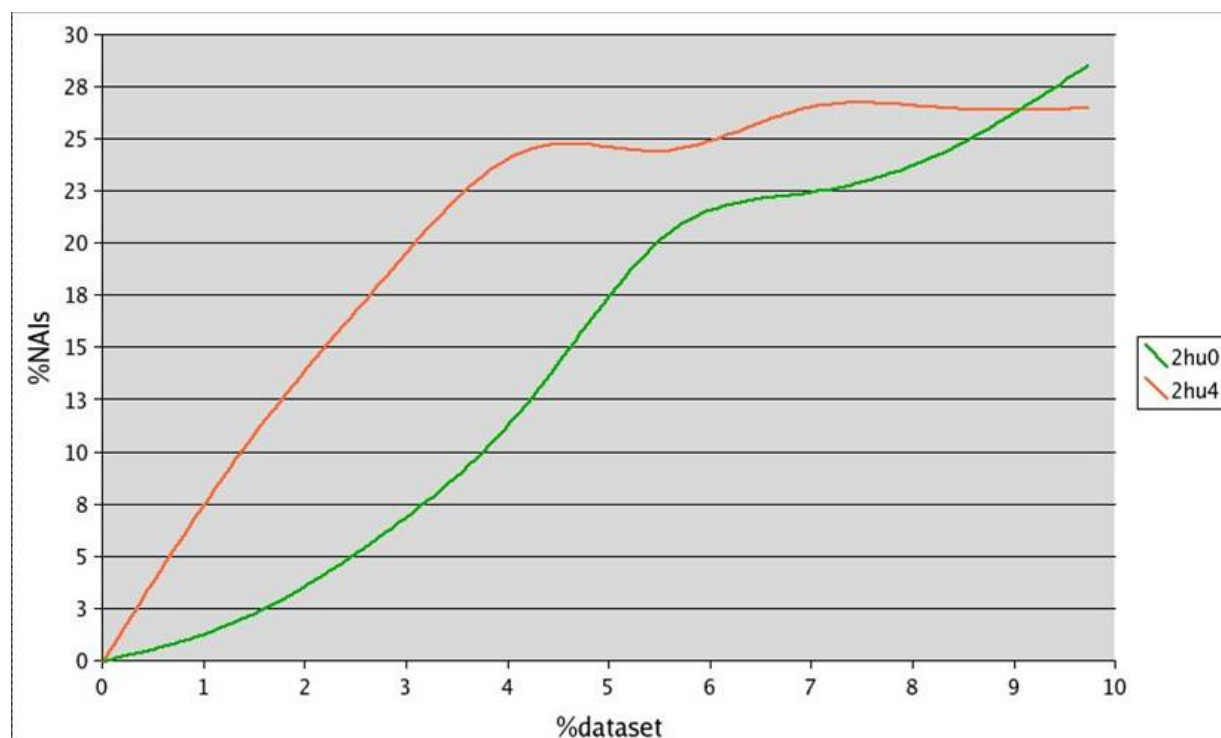


*Docking energy of the whole dataset with the open (**2hu0**) versus closed (**2hu4**) conformation of N1. NAIs are in red and decoys in green*



Enrichment analysis

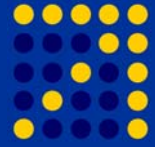
The ranking of ligands has been found to be different in the two structures



The top 5% of the docked ligands was enriched with

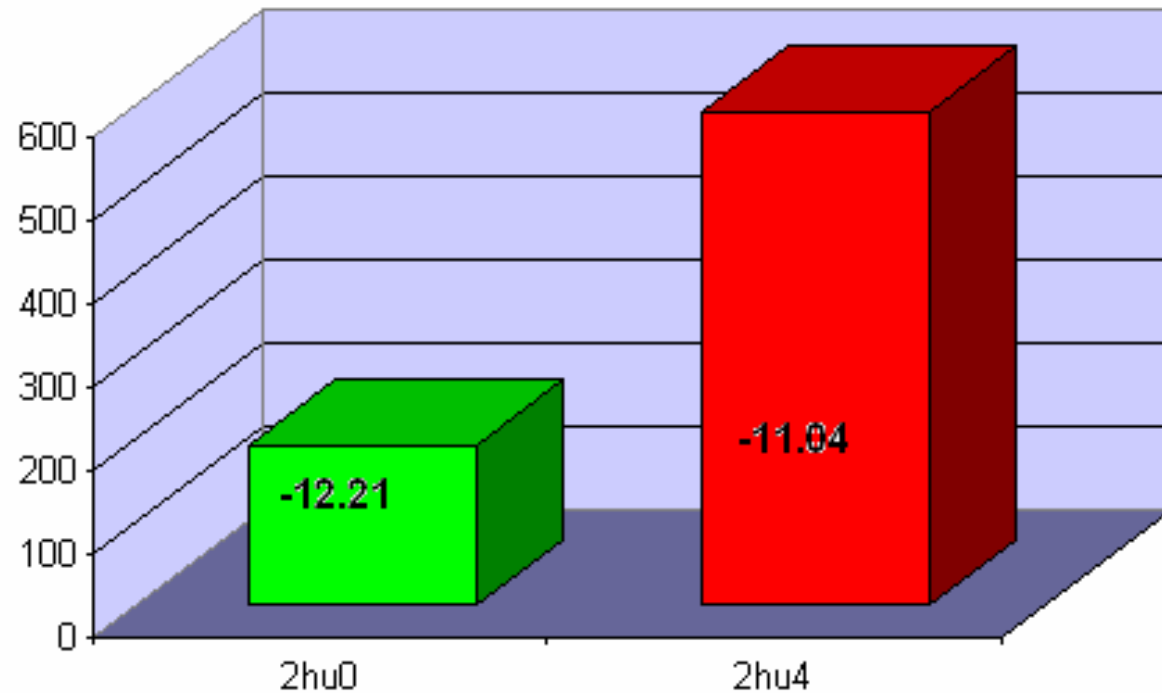
- **25% of NAIs in the closed (2hu4)**
- **18% of NAIs in the open form (2hu0)**

This suggested the possibility that the active site in the open and closed conformations have different binding properties



With the aim to explore the specific binding properties, a more detailed analysis of the contact interactions was performed on a selected set of the binders showing energy values lower than substrate, sialic acid

- 193 binders for the **open form** (2hu0)
- 594 binders for the **closed form** (2hu4)





The interaction with two high polar regions of the active site:

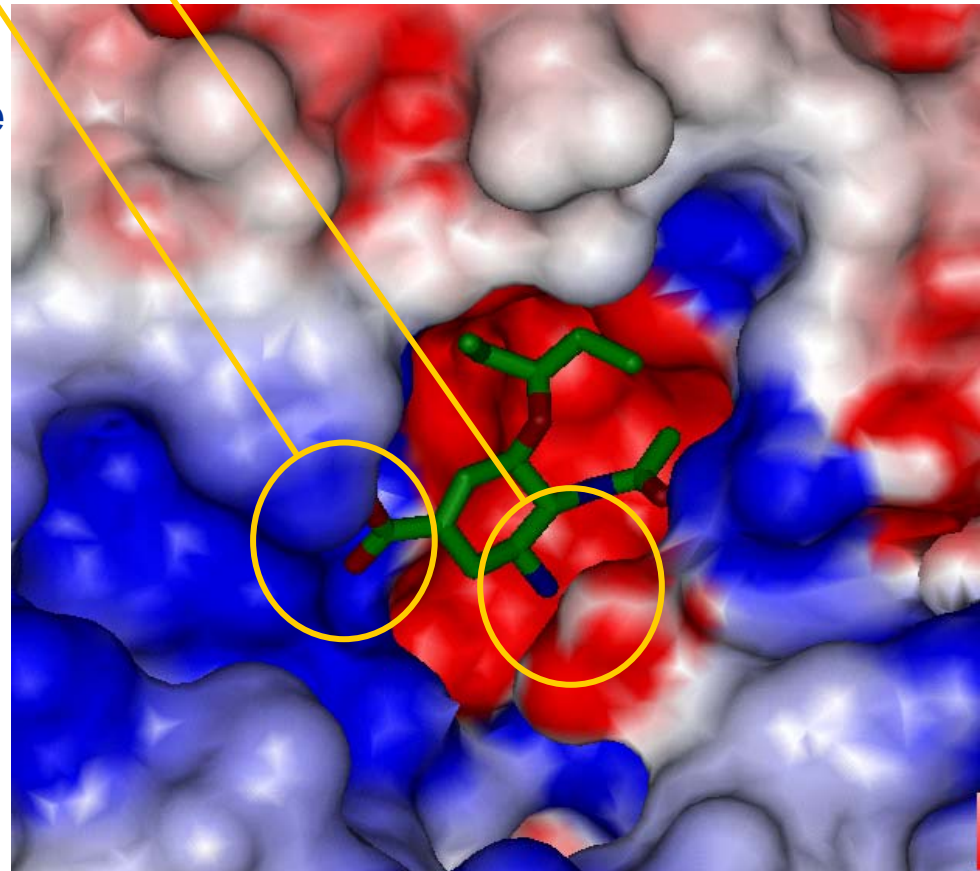
➤ **acid pocket: Glu119, Asp151, Glu227**

➤ **basic pocket: Arg118, Arg292, Arg371**

are indispensable to the binding affinity of a POTENT neuraminidase inhibitor

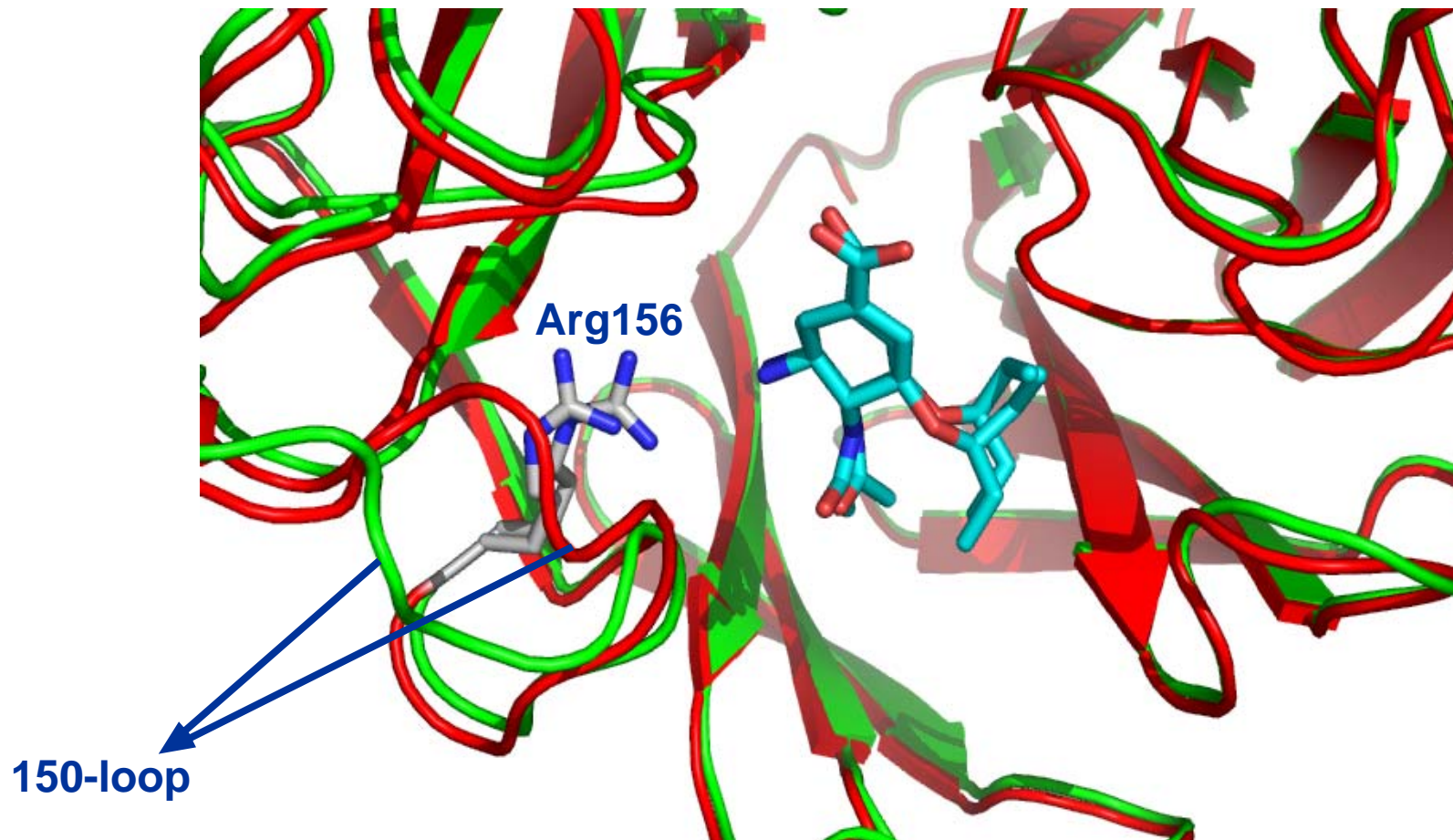
Connolly surface electrostatic potential distribution of neuraminidase.

*Acid (negative potential) in red
Basic (positive potential) in blue*





The interaction with **150-loop** and **Arg156**, at the base of the 150 cavity, are markers of cavity filling binders in the open conformation





An automatic procedure was set up to calculate the interactions for the selected docking complexes using Ligplot and to withdraw the subgroup of ligands establishing specific interactions with:

- **acid pocket: Glu119, Asp151, Glu227**
- **basic pocket: Arg118, Arg292, Arg371**
- **150-loop: resid 147-152**
- **Arg156**

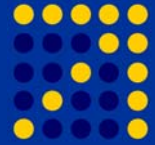


Conclusion - Work in progress

The analysis performed identified:

- **56 binders**, all of them belonging to the group of **decoys**
- **None of the known inhibitors** established interactions **with the loop**
- **4 NAIs and sialic acid** establish **Hbond with Arg156**

Structural analysis is in progress to identify the functional groups of the ligands that form key interactions with the subsites (**acid and basic pockets, 150-loop and Arg156**) of the binding site of neuraminidase in the open conformation



This project has been supported by:

- European "Specific Support Action BioinfoGRID" and "EGEE" projects
- Italian FIRB-MIUR project "ITALBIONET"

I would like to thank the people that have worked on this project:

- Pasqualina D'Ursi
- Federica Chiappori
- Erika Salvi
- Ivan Merelli
- Ermanna Rovida
- **WISDOM, EGEE, BioinfoGRID Partners**