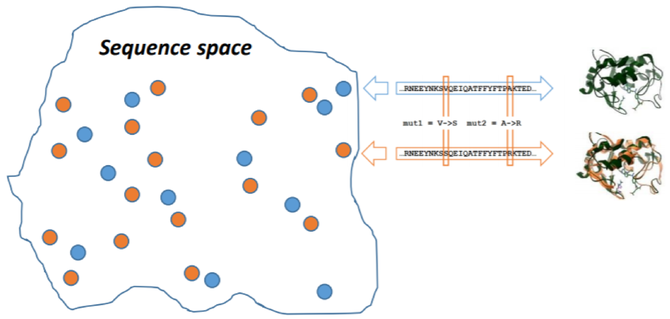
1

Evolution escape room: Make the virus great again

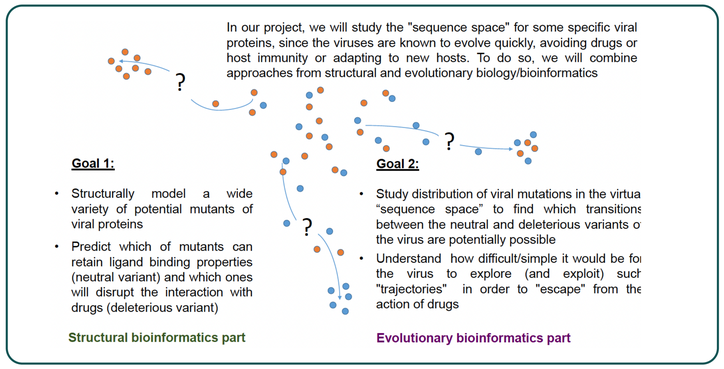
Peter Vlasov
Yakov Bogantcev
Sonya Beliaeva
Dmitriy Korkin
Ivanna Ostapchuk
Polina Avdiunina
Angelika Dodonova
Marlen Toktomamatov

Alexey Kim
Anna Metreveli
Timur Terekhov
Olga Kuzmich
Nikita Radaev
Elizabeth Lesher
Yegor Chudovskiy
Lina Mameshova

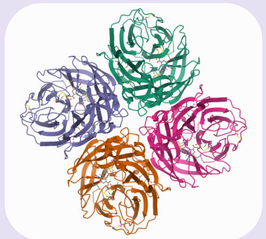
Among effective and widely used antiviral drugs, there are some that work by direct inhibition of a functional site of the target protein by the drug molecule. If this interaction is strong, it leads to undesirable—for the viral life cycle—structural changes. Because of numerous random mutations, which is typical for viruses, target proteins can acquire variants that generally retain the protein's functionality, but may also locally change its structure.



Each point in this space corresponds to a protein sequence, and the transitions between the points correspond to evolutionary events — mutations — that often look like substitutions, insertions, or deletions of amino acid residues.
Of course, many of these points turn out to be physiologically "inappropriate" - due to the inability to fold into a stable structure or perform the function by the corresponding proteins.
And the related questions — the ratio of "good" vs "bad" variants (for the virus' success), how difficult it is for the evolutionary process to "find" a new suitable state, and how this space is generally arranged — remain open to science.



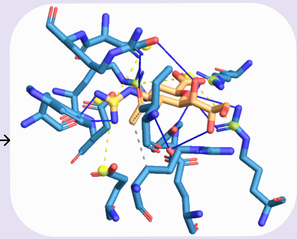
RCSB PDB PROTEIN DATA BANK



neuraminidase complex with Oceltamivir - 2hu4

PDB is a database for structures of proteins and protein complexes. We have used it to search for the annotated structures of proteins of our interest to use them as templates for modelling of mutants.

Protein-Ligand Interaction Profiler

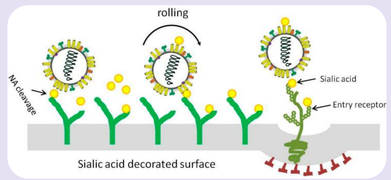


neuraminidase complex with Zanamivir - 4b7q

We browsed PLIP to find annotated complexes of proteins of our interest with small molecules in order to identify ligand-binding site for further docking and generation of mutants.

Materials and methods

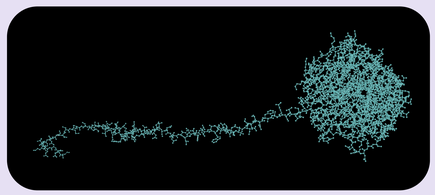
Target protein



The NA could degrade sialic acids (SA) near the attachment site, resulting in the SA density gradient that promotes virus moving until successful infection occur. Anti-viral drugs block the active site of neuraminidase and leaves uncleaved sialic acid residues on the surfaces of host cells and influenza viral envelopes.

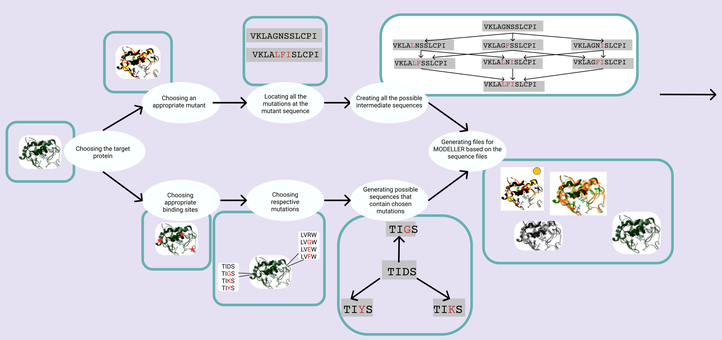
Antivirals block the neurominidase's active site, preserving sialic acid on the host cells, and preventing viral entry and infection.

UCSF Chimera/PyMol



Are irreplaceable tools for compairing native and mutant proteins.

Pipiline for high throughput generation of mutants

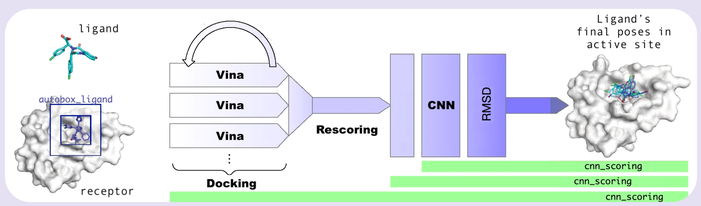



```
>1 1 117V
MNPNQKIITIGSVCMTIGMANLILQIGNIISIWISHSIQLGNQNQIETCNQSVITYENNTWVNQTYVNISNTNFAAGQSV
VSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFVIREPFISCSPLECRTFFLTQGALLNDKHSNGTIKDRSPYRTLMSC
PIGEVPSPYNSRFESVAWSASACHDGINWLTIGISGPDNGAVAVLKYNGIITDTIKSWRNNILRTQESECACVNGSCFTV
MTDGPSNGQASYKIFRIEKGKIVKSVEMNAPNYHYEECSCYPDSSEITCVCRDNWHGSNRPWVSFNQNLEYQIGYICS
GIFGDNPRPNDKTGSCGPVSSNGANGVKGFSFKYGNGVWIGRTKSISSRNGFEMIWDPNGWTGTDNNFSIKQDIVGI
NEWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKENTIWTSGSSISFCGVNSDTVGWSWPDGAELPFTIDK

>25 3 117V 223R 427T
MNPNQKIITIGSVCMTIGMANLILQIGNIISIWISHSIQLGNQNQIETCNQSVITYENNTWVNQTYVNISNTNFAAGQSV
VSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFVIREPFISCSPLECRTFFLTQGALLNDKHSNGTIKDRSPYRTLMSC
PIGEVPSPYNSRFESVAWSASACHDGINWLTIGISGPDNGAVAVLKYNGIITDTIKSWRNNLLRTQESECACVNGSCFTV
MTDGPSNGQASYKIFRIEKGKIVKSVEMNAPNYHYEECSCYPDSSEITCVCRDNWHGSNRPWVSFNQNLEYQIGYICS
GIFGDNPRPNDKTGSCGPVSSNGANGVKGFSFKYGNGVWIGRTKSISSRNGFEMIWDPNGWTGTDNNFSIKQDIVGI
NEWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKENTIWTSGSSISFCGVNSDTVGWSWPDGAELPFTIDK
```

MODELLER software was used for the mutant generation and for transition of *.fasta* files to the *.pdb* format.

Pipiline for high throughput molecular docking

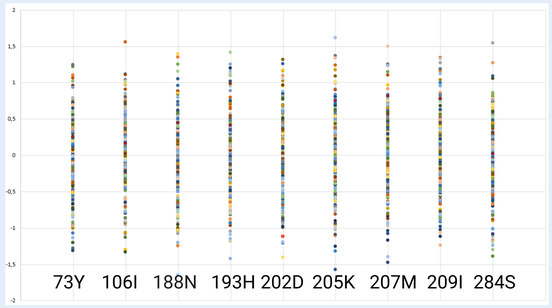


ZIMIN FOUNDATION

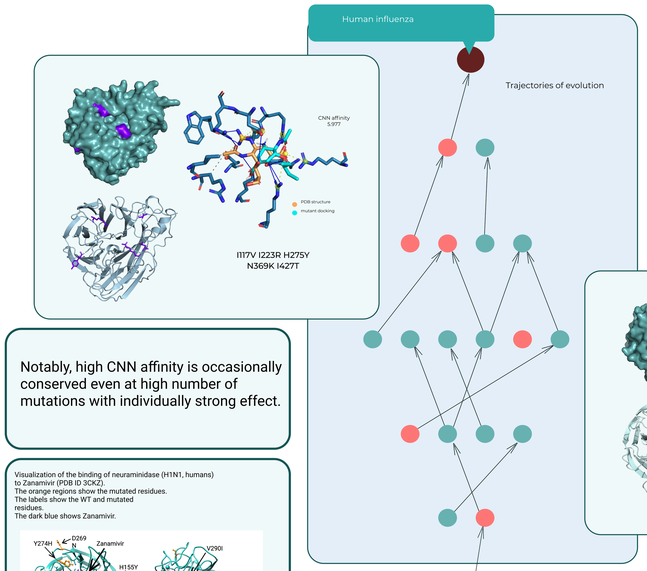
SMTB

Petrovax

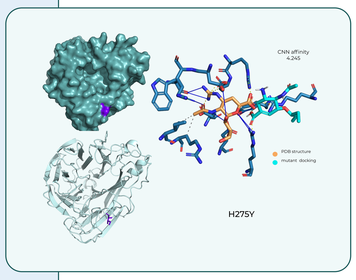
2 Data collection, analysis, and discussion



Notably, any of the mutants can have a wildly different effect on the binding energy, both negative and positive.



Notably, high CNN affinity is occasionally conserved even at high number of mutations with individually strong effect.

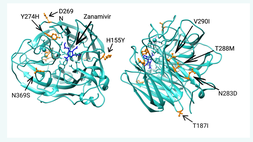


Dataset 1: avian influenza to human influenza jump. H7N1 avian influenza's NA gene (GEN.Bank JQ973644.1) was BLASTed against PDB. Human influenza NA (3CKZ) was chosen as it had ligand in the complex. All the differing aminoacids that weren't close to the active site were set to human influenza sequence. Mutation trajectrories were then made from human to modified avian influenza.

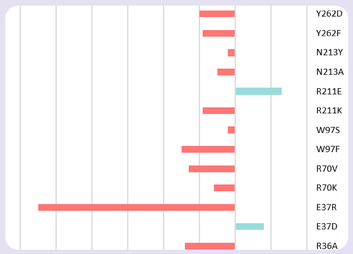
https://www.frontiersin.org/articles/10.3389/fmicb.2016.01066/full

Special project "Angry birds" is dedicated to a jump of an avian influenza to human population. We studied a possibility of accumulation of mutations which would make neurominidase more similar to its human variant. Notably, while accumulating mutations to infect humans the virus can pick up a mutation making it resistant to NA inhibitors.

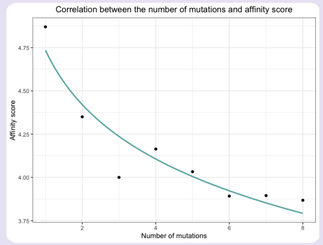
Visualization of the binding of neuraminidase (H1N1, humans) to Zanamivir (PDB ID 3CKZ).
The orange regions show the mutated residues.
The labels show the WT and mutated residues.
The dark blue shows Zanamivir.



E37D	0,08
E37R	-0,55
R70K	-0,06
R70V	-0,13
W97F	-0,15
W97S	-0,02
R211K	-0,09
R211E	0,13
N213A	-0,05
N213Y	-0,02
Y262F	-0,09
Y262D	-0,1

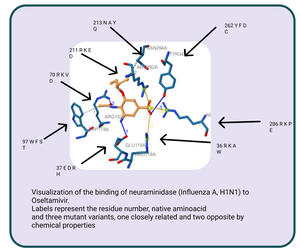


Notably, the best binding energy for ligand wasn't the one with the most mutations. This mutant had substitutions: 36K, 37R, 97S, 213Y, 262F, 286K and CNN affinity -3.28



The binding energy decreases rapidly with each successive mutation.

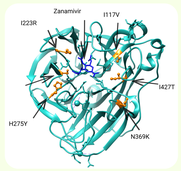
Dataset 2: surroundings of a protein. Human influenza neuraminidase (PDB ID 2HU4) was taken as the wt. Aminoacids (aa) labelled as in active site were mutated to either a chemically similar aa or a chemically dissimilar aa. All versions were generated using a script.



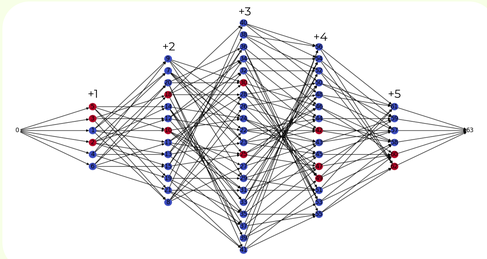
Visualization of the binding of neuraminidase (Influenza A, H1N1) to Oseltamivir.
Labels represent the residue number, native aminoacid and three mutant variants, one closely related and two opposite by chemical properties

Dataset 3: H1N1 resistance occurance. Reports of clinically resistant mutants were compiled and a range of mutations were combined to make a mutant containing 6 substitutions relative to a reported wt (PDB ID: 2HU4) (1, 2). The combination of all possible mutants were then generated.

1. https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(11)70318-8/fulltext
2. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0210119

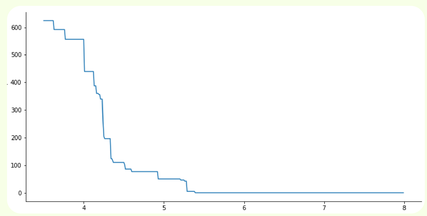


Visualization of the binding of neuraminidase (H1N1) to Zanamivir drug molecule. The orange regions show the mutated amino acids.
Mutations positions are shown on the image.
Labels represent the wild type aminoacid, the residue number and the mutant aminoacid respectively.



All the possible trajectories between the WT protein and the multiple-mutant protein carrying several single resistant mutations. Red points mark the mutants with high binding affinity, while blue points mark those with low binding affinity. Edges represent transitions from one state to another in one aminoacid residue mutation.

Number of "good" trajectories



Barrier threshold fo CNN affinity

Amount of "good" (consisted of resistent mutants) trajectories as function from threshold of mutant division to resistant and non-resistant.
Notably, in 4 и 5 CNN affinity range we can observe radical failure of probability to go from initial point to final.

ZIMIN FOUNDATION

SMTB

Petrovax