



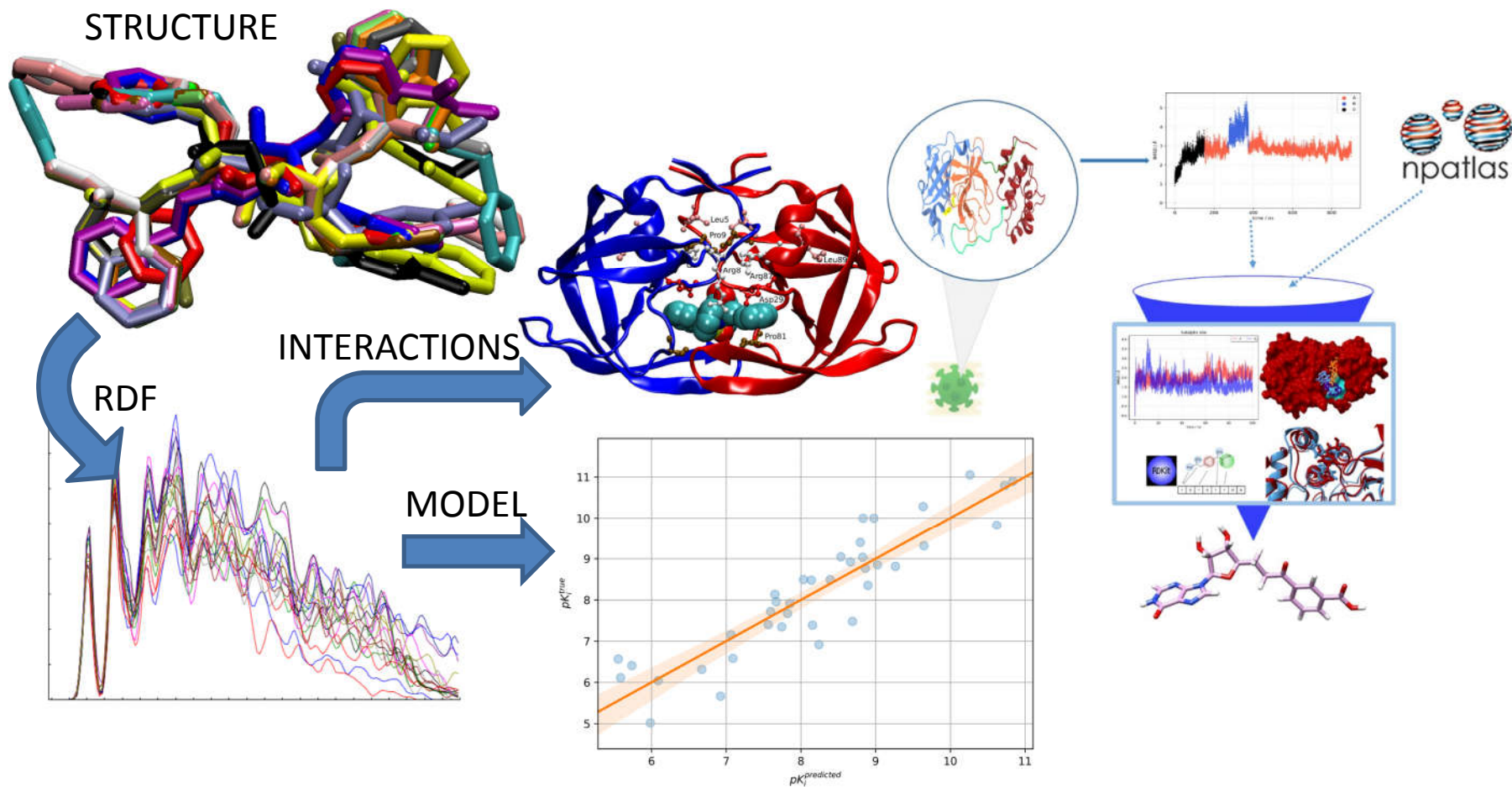
Южно-Уральский
государственный
университет
Национальный
исследовательский
университет

Computational chemistry artillery against viruses

Jurica Novak

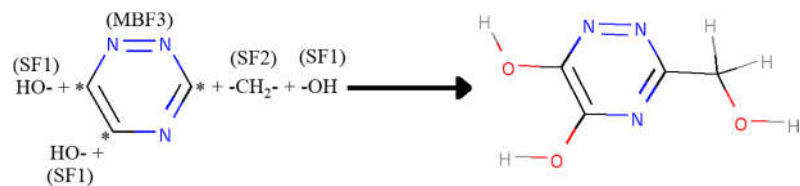
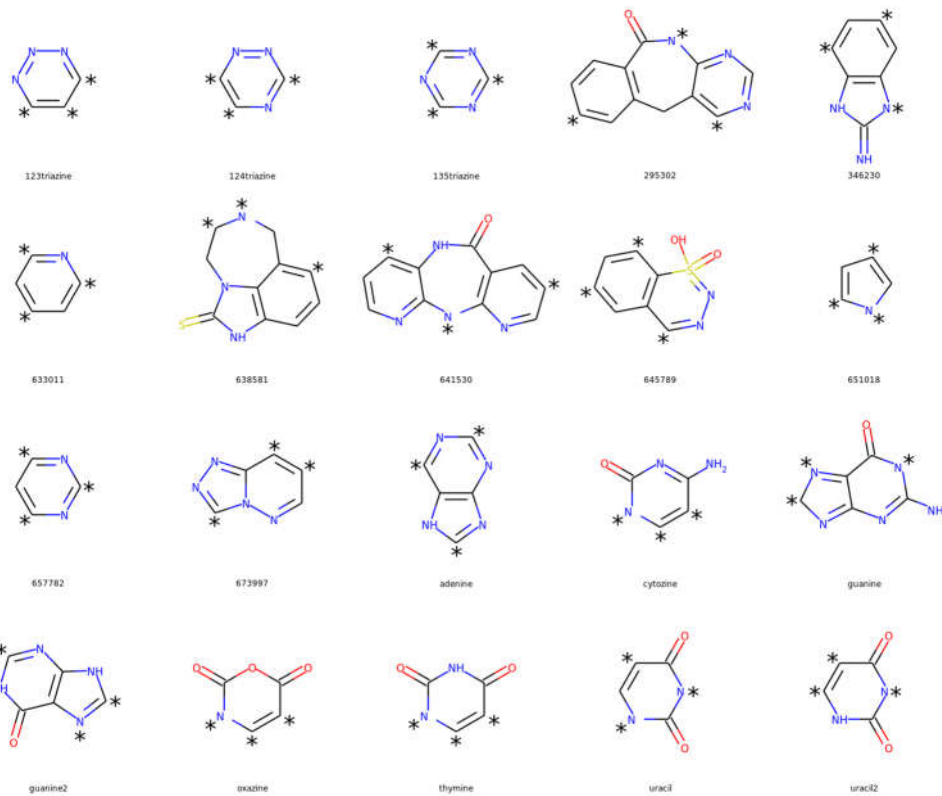
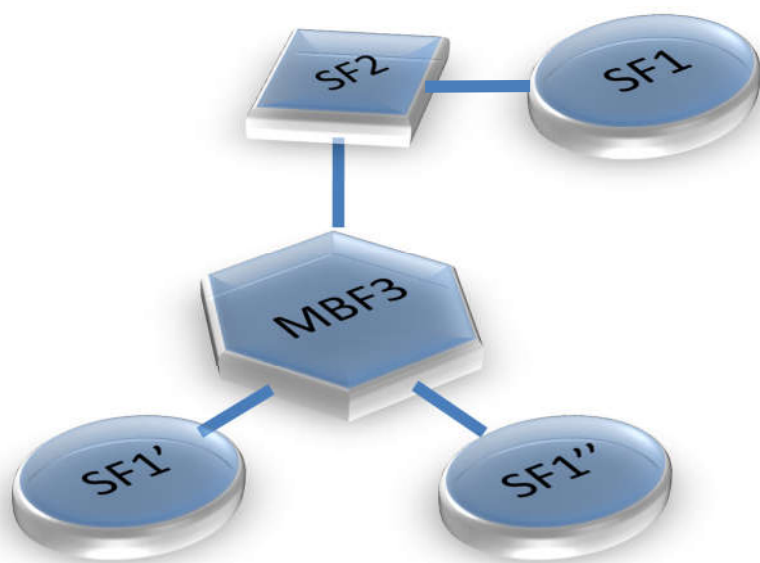
September 2022., Zagreb

Introduction



- [1] JN *et al.* The Performance of Radial Distribution Function Based Descriptors in the Chemoinformatic Studies of HIV-1 Protease, FMC 12 (2020) 299
 [2] JN *et al.* Novel radial distribution function approach in the study of point mutations: the HIV-1 protease case study, FMC 12 (2020) 1025
 [3] JN *et al.* Can bacterial and fungal natural products stop SARS-CoV-2 virus? A virtual screening and molecular dynamics study of a natural product database, FMC 13 (2021) 363-378
 [4] JN *et al.* Proposition of a new allosteric binding site for potential SARS-CoV-2 3CL protease inhibitors by utilizing molecular dynamics simulations and ensemble docking, JBSD

HIV

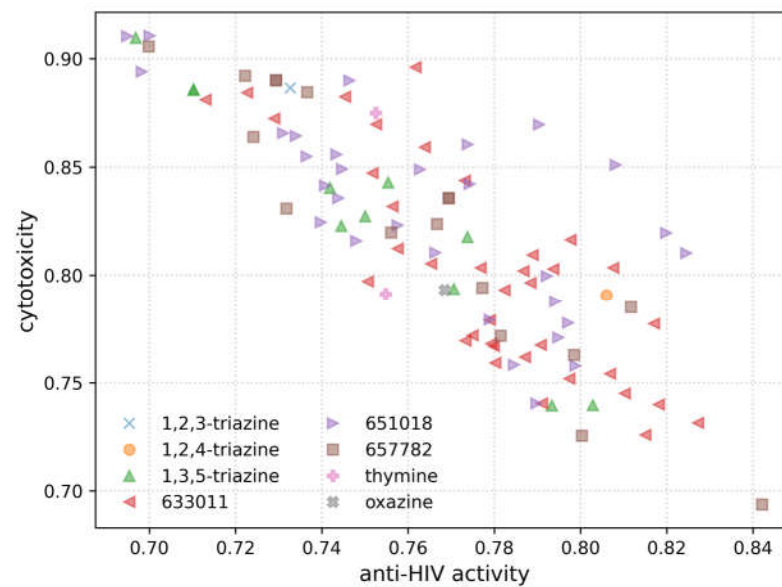
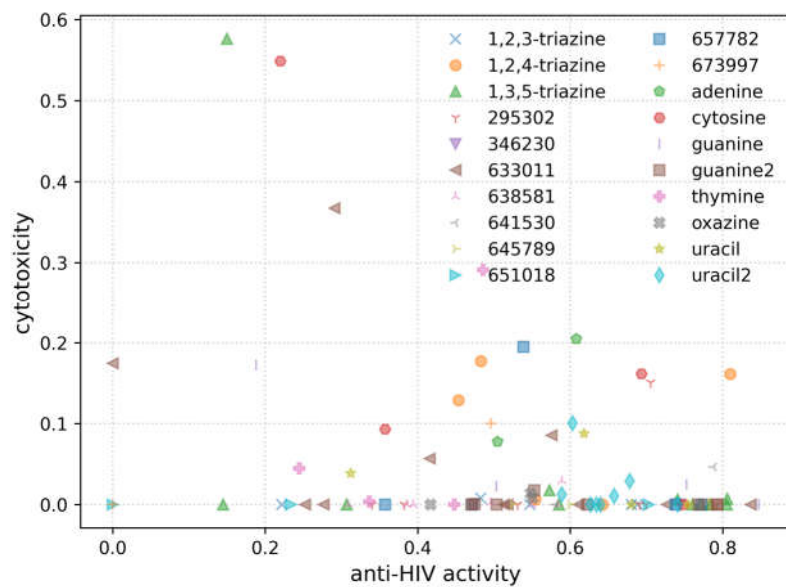


HIV

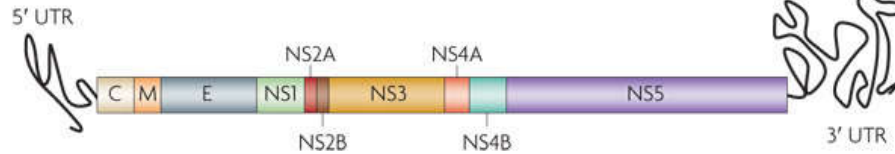
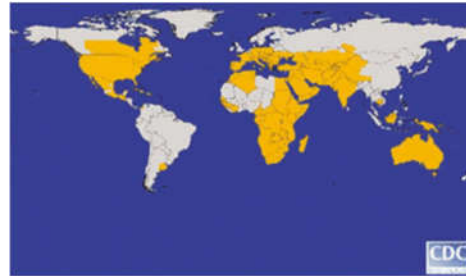
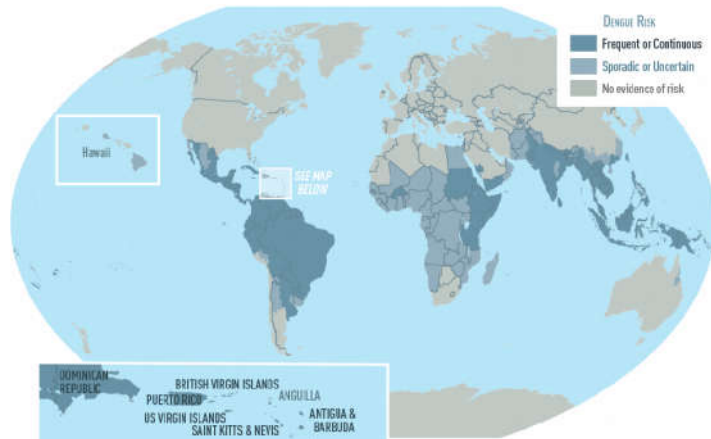
	MBF3	SF2	SF1	SF1'	SF1''
Parent1	6	25	10	19	73
Parent2	2	39	46	24	6
Child1	2	25	10	19	73
Child2	6	39	46	24	6
Mutant1	2	25	10	19	79
Mutant2	6	39	46	60	6

HIV

$$g = \left(\prod_{i=1}^{n=18} p_{anti-HIV,i} \prod_{j=1}^{m=12} p_{tox,j} \right)^{\frac{1}{n+m}}$$



Flavivirus



60 70 80 90

KFDV QLTAEWCGEVEWNPDLVNEGGEVNLKVRQDAMGNLHLTEVEKEER
 ZIKA GLELKKLGEVSWEEAEISGSSARYDVALSEQGEFKLLSEEKVPW
 WNV DLELERAADVKWEDQAEISGSSPILSITISEDGSMISKNEEEQ
 JEV DMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEEDGPPM
 YFV DMWIERTADITWESDAEITGSSERVDRLDDGNFQLMNDPGAPW
 DENGUE DMWLERAADISWEMDAAITGSSRRLDVKLDDGGFHLIDDPGVPW

Consensus DLLEERADWE DAEI GSS RLDV LDED GFL EGGW
 D++LERA+D++WE+DAEITGSSPRLDV+LDEDG+F+L++EEG+PW

12 22 32 42

KFDV DGAVRIYSPGLLWGHRIIGVGVYGAQVLTMMWHVTRGAALVVDVA
 ZIKA DGVYRVMTRRLGSGTQVGVGMQEGVFHTMWHVTRGAALRSSEG
 WNV TGVYRIMTRGLLGSYQAGAGVMVEGFHTLWHTTRGAALMSGEG
 JEV TGVYRIMARGILGTYQAGVVMYENVFHTLWHTTRGAALMSGEG
 YFV DGIYGFQSTFLGASGRVGVVAQGGVFHTMWHVTRGAFVLRNGK
 DENGUE DGAVRIKQKQILGYSIGAGVYKEGTFHTMWHVTRGAVLRHKKG

Consensus DGVYRI RGLLWGS+Q+GVGVMEGVFHTMWHVTRGAALMSGEG
 DGVYRIM+RGLLWGS+Q+GVGVMEGVFHTMWHVTRGAALMSGEG

52 62 72 82

KFDV ISGPIYWADVREDVVCGGAWSLESRWGE-TVGVHAFPPGRPQET
 ZIKA RLDPIYWDVQDLVSCGPWKLDAAWDGLSEVLLAVPPGERARN
 WNV RLDPIYWGSVKEBRLCYGGPWKLQHKWNGHDEVMIIVVEPGKNVK
 JEV KLTPIYWGSVKEBRVYGGPWRFRKRWNGTDDVVIVVEPGKAAN
 YFV KLI PSWASVKEDLVAYGGSWKLEGRWDGEEVOLI AAVPGKNVVN
 DENGUE RIEPSWADVKKDLISVGGGWKLEGEWKEGEEVVLALVPGKNPRA

Consensus PWA VKEDLV YGGWLE WGEVQVAEPGKNIN
 RLDPIYW++VKEDLV+YGGPWKLEGEW+V+GE+EVQVIAVEPGKN++N

102 112 122 132

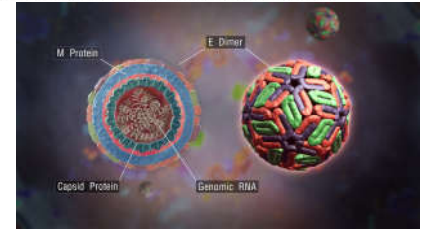
KFDV HCCQPGELILENGRKLGAVPIDL SKGTSGPSII NAQGEVVGLYGN
 ZIKA IOTLPGIFKTKDGDIGAVALDYPAGTSGSPI LDKCGRVIGLYGN
 WNV VQTKPGVFKTPEGEIGAVTLDYPGTSGSPIVDKNGDVIIGLYGN
 JEV IOTKPGVFRT-PFGEVAVSLDYPGTSGSPI LDSNGDIIGLYGN
 YFV VQTKPSLQKVRNGGEIGAVALDYPGTSGSPIVNRNGEVIIGLYGN
 DENGUE VQTKPGLFKT-NAGTIGAVSLDFSPGTSGSPIIDKKKVVVGLYGN

Consensus VQTKPG+FKT+NGGEIGAV+LDYP+GTSGPSI+DKNG+VIGLYGN

142

KFDV GLKT-NEAYVSSIAE
 ZIKA GVIKNGSYVSAITE
 WNV GVIMPNGSYVSAIVG
 JEV GVELDGGSYVSAIVG
 YFV GILVGDNSFVSAISG
 DENGUE GVVTRSGAYVSAIAE

Consensus GVVVTGNGSYVSAI+Q



Flavivirus

KFDV NS3	100	48.98	46.26	44.22	50.68	53.06
ZIKA NS3	48.98	100	69.33	62.67	56.67	57.33
WNV NS3	46.26	69.33	100	76	57.33	56.67
JEV NS3	44.22	62.67	76	100	54.67	52
YFV NS3	50.68	56.67	57.33	54.67	100	60.93
DENV NS3	53.06	57.33	56.67	52	60.93	100
	KFDV NS3	ZIKA NS3	WNV NS3	JEV NS3	YFV NS3	DENV NS3

```

KFDV  Q L T A E W C G E V E W N P D L V N E G G E V N L K V R Q D A M G N L H L T E V E K E E R
ZIKA  G L E L K K L G E V S W E E A E I S G S S A R Y D V A L S E Q G E F K L L S E E K V P W
WNV   D L E L E R A A D V K W E D Q A E I S G S S P I L S I T I S E D G S M S I K N E E E E Q T
JEV   D M Y I E R A G D I T W E K D A E V T G N S P R L D V A L D E S G D F S L V E E D G P P M
YFV   D M W I E R T A D I T W E S D A E I T G S S E R V D R L D D D G N F Q L M N D P G A P W
DENGUE D M W L E R A A D I S W E M D A A I T G S S R R L D V K L D D D G F H L I D D P G V P W

Consensus D L E R A A D I W E D A E I G S S R L D V L D E D G F L E G G P W
D + + L E R A + D + + W E + D A E I T G S S P R L D V + L D E D G + F + L + + E E G + P W
    
```

```

KFDV  D G A Y R I Y S P G L L W G H R D I G V G Y G A K G V L H T M W H V T R G A A L V V D E A
ZIKA  D G V Y R V M T R R L L - G S T D V G V G M Q E G V F H T M W H V T K G A A L R S G E G
WNV   T G V Y R I M T R G L L - G S Y D A G A G V M E G V F H T L W H T T K G A A L M S G E G
JEV   T G V Y R I M A R G I L - G T Y D A G V G V M Y E N V F H T L W H T T R G A A I M S G E G
YFV   D G I Y G I F Q S T F L - G A S R G V G V A O G G V F H T M W H V T R G A F L V R N G K
DENGUE D G A Y R I K Q K G I L - G Y S I G A G Y K E G T F H T M W H V T R G A V L M H K G K

Consensus D G Y R I R G L L W G S + Q + G V G V M Q E G V F H T M W H V T R G A A L M S G E G
D G V Y R I M + R G L L W G S + Q + G V G V M Q E G V F H T M W H V T R G A A L M S G E G
    
```

```

KFDV  I S G P Y W A D V R E D V V C Y G G A W S L E S R W R G E - T V G V H A F P P G R P Q E T
ZIKA  R L D P Y W G D V K Q D L V S Y C G P W K L D A A W D G L S E V Q L L A V P P G E R A R N
WNV   R L D P Y W G S V K E B R L C Y G G P W K L Q H K W N G H D E V M I V V E P G K N V K N
JEV   K L T P Y W G S V K E B R I A Y G G P W R F D R K W N G T D D V V I V V E P G K A A V N
YFV   K L I P S W A S V K E D L V A Y G G S W K L E G R W D G E E E V L I A A V P G K N V V N
DENGUE R I E P S W A D V K K D L I S Y G G G W K L E G E W K E G E E V V L A L E P G K N P R A

Consensus R L D P Y W + + V K E D L V + Y G G P W K L E G + W + G E + E V Q V I A V E P G K N + + N
R L D P Y W + + V K E D L V + Y G G P W K L E G + W + G E + E V Q V I A V E P G K N + + N
    
```

```

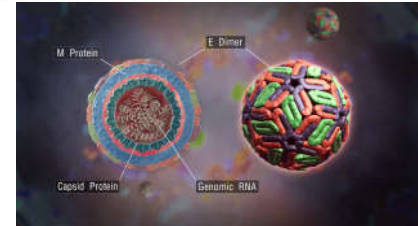
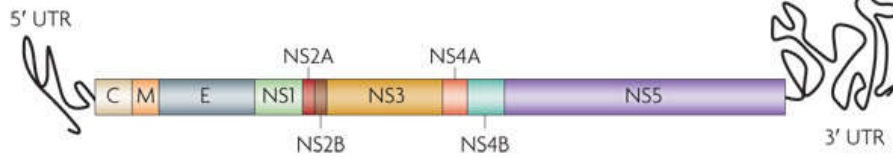
KFDV  H C C P G G E L I L E N G R K L G A V P I D L S K G T S G S P I I N A O G E V V G L Y G N
ZIKA  I O T L P G I F K T - K D G D I G A V A L D Y P A G T S G S P I L D K C G R V I G L Y G N
WNV   V O T K P G V F K T - P E G E I G A V T L D Y P T S T S G S P I V D K N G D V I G L Y G N
JEV   I O T K P G V F R T - P F G E V G A V S L D Y P R G T S G S P I L D S N G D I I G L Y G N
YFV   V O T K P S L F K V R N G G E I G A V A L D Y P S G T S G S P I V N R N G E V I G L Y G N
DENGUE V O T K P G L F K T - N A G T I G A V S L D F S P G T S G S P I I D K K K V V G L Y G N

Consensus V O T K P G + F K T + N G G E I G A V + L D Y P + G T S G S P I + D K N G + V I G L Y G N
V O T K P G + F K T + N G G E I G A V + L D Y P + G T S G S P I + D K N G + V I G L Y G N
    
```

```

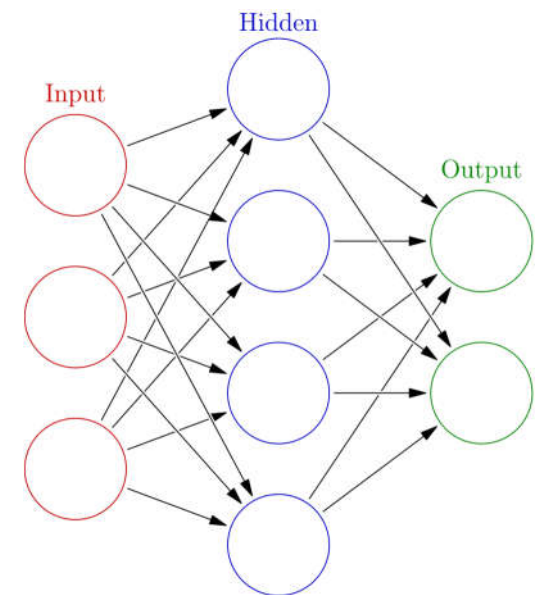
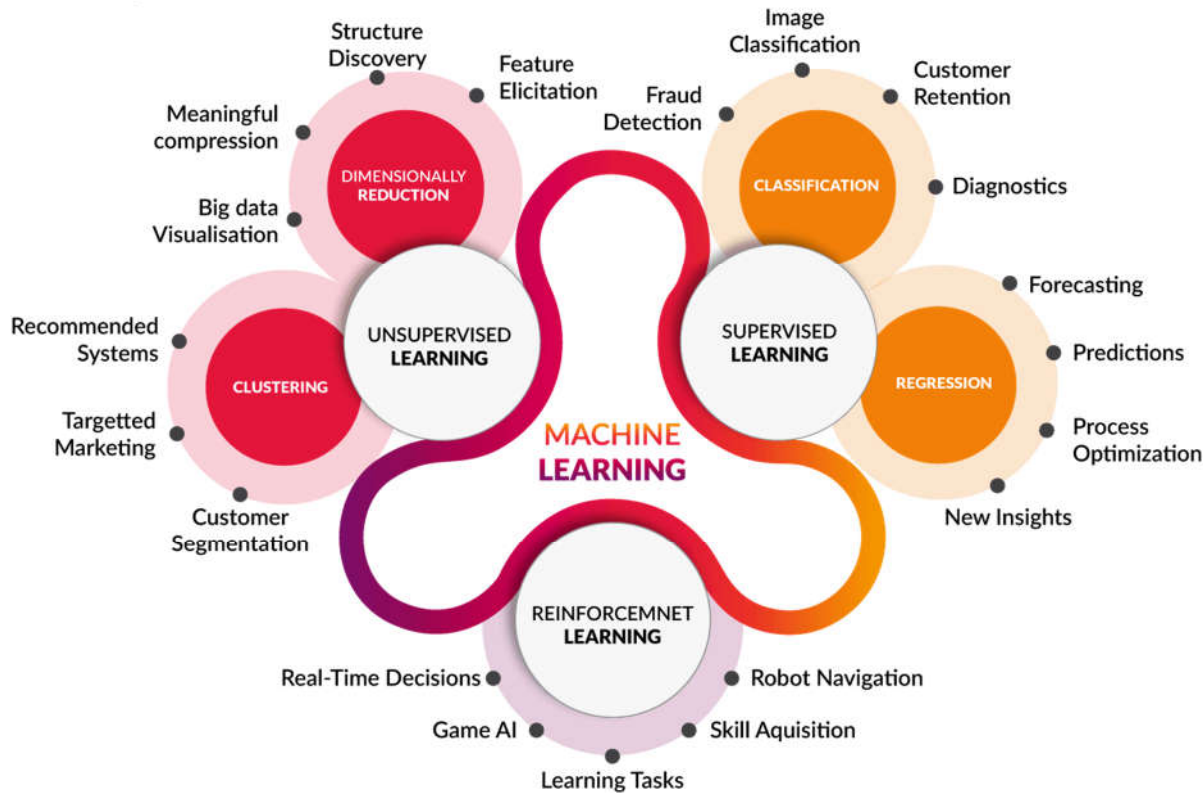
KFDV  G L K T - N E A Y V S S I A G
ZIKA  G V V I K N G S Y V S A I T E
WNV   G V I M P N G S Y I S A I V G
JEV   G V E L G D G S Y V S A I V G
YFV   G I L V G D N S F V S A I S G
DENGUE G V V T R S G A Y V S A I A G

Consensus G V V T R N G S Y V S A I + Q
G V V T R N G S Y V S A I + Q
    
```



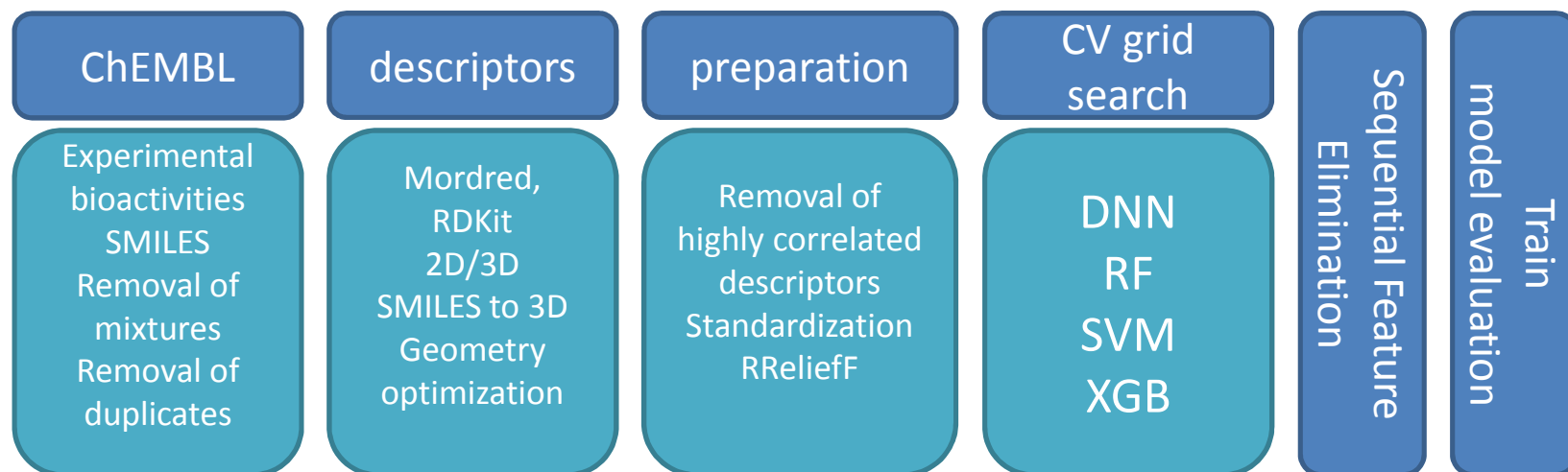
Machine learning

- a technique which uses computers to discover patterns or information about data

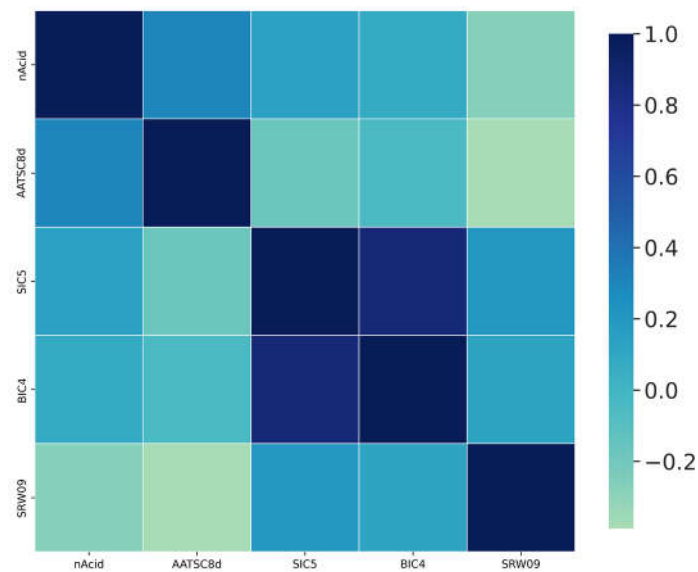
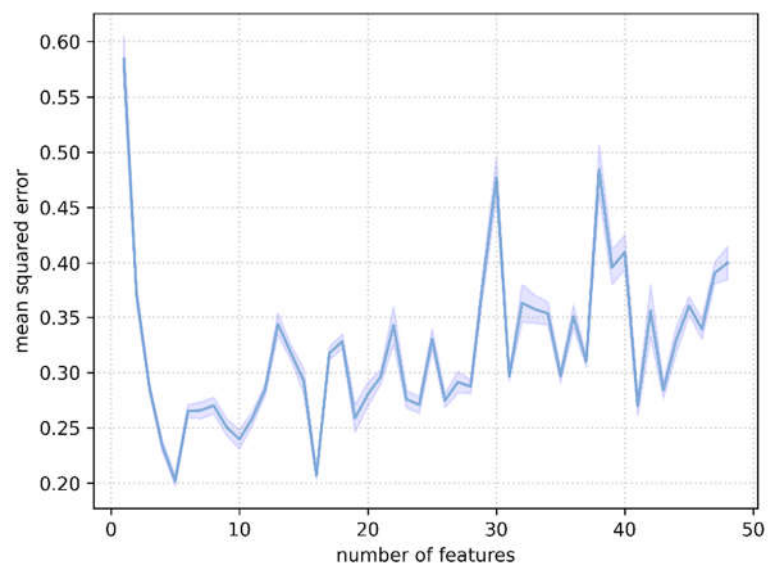
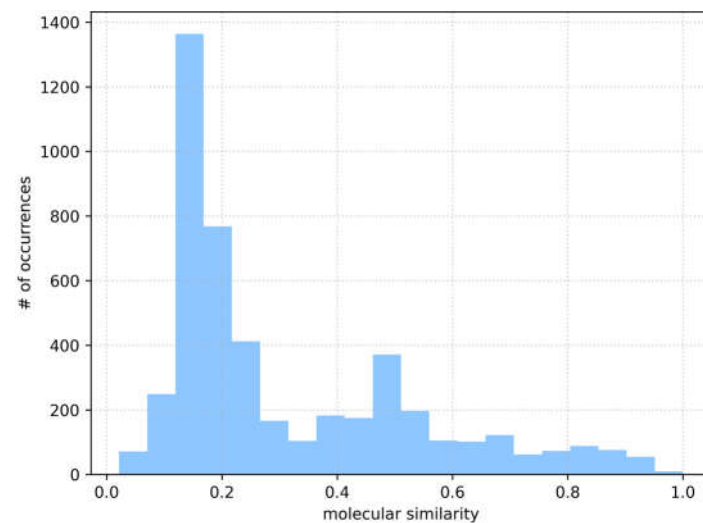
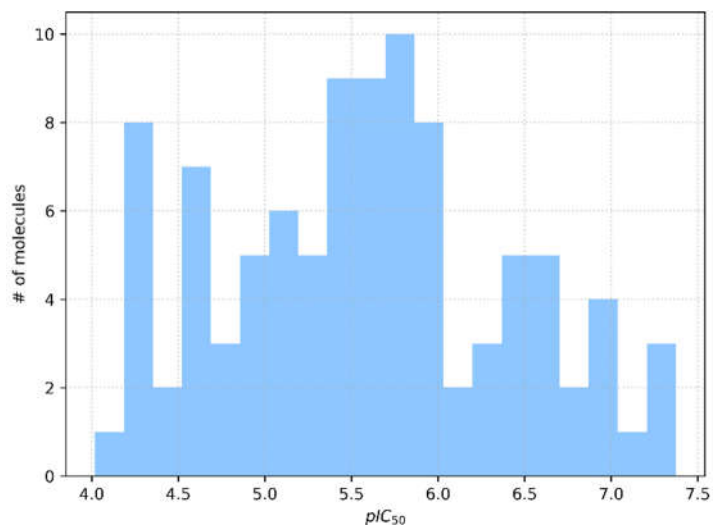


<https://auginvite.com/introduction-to-machine-learning/>

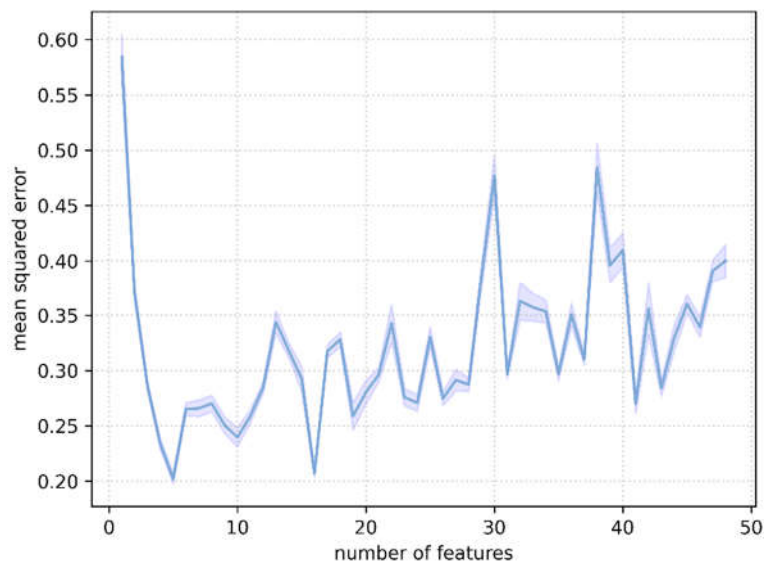
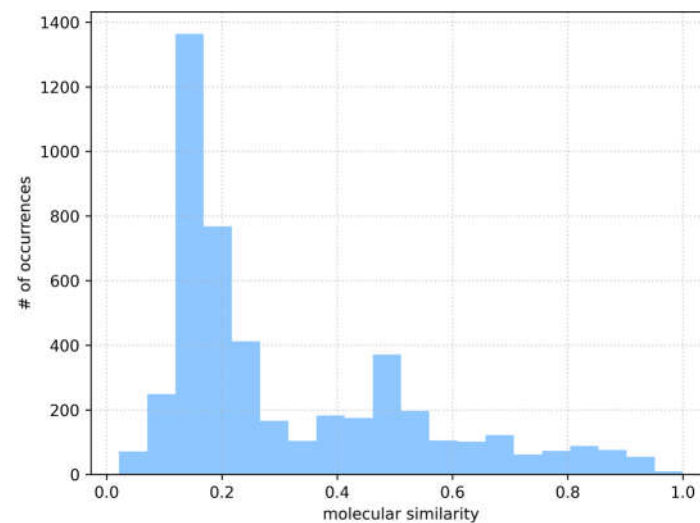
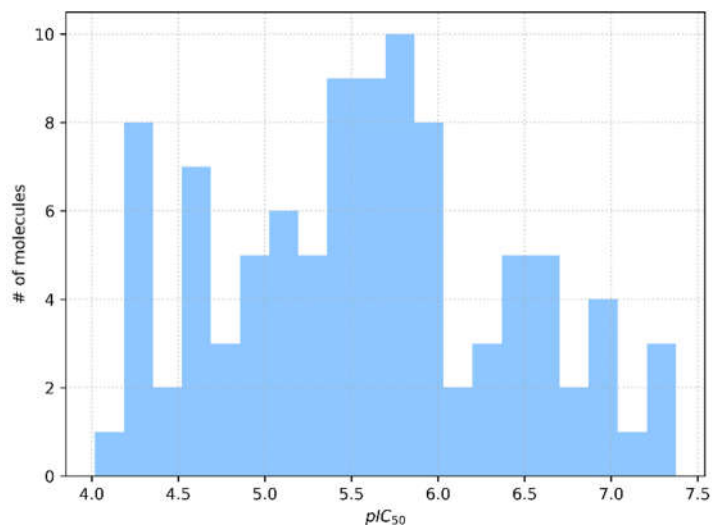
ML scheme



DNN



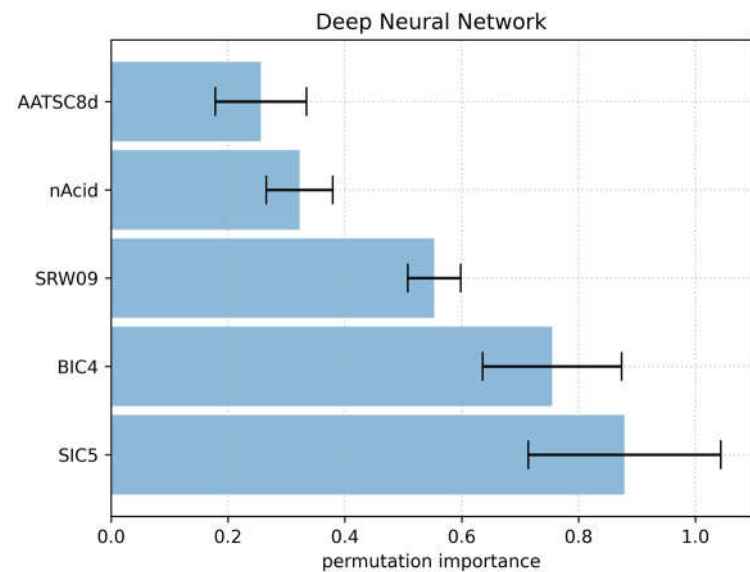
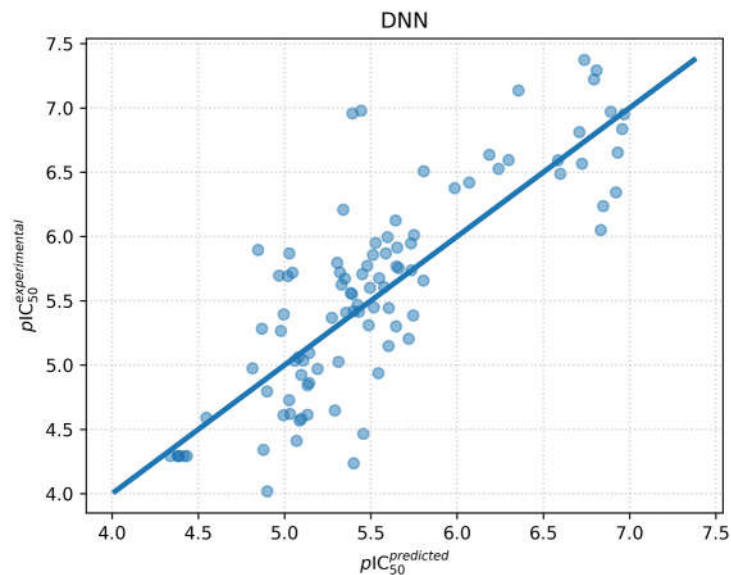
DNN



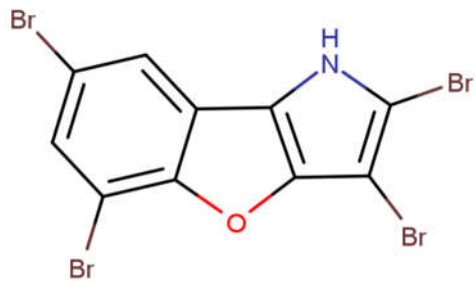
dropout_rate: [0.0, 0.1, 0.2, 0.3]
n_layers: [8, 16, 32, **64**, 128]
n_neurons: [32, 64, 128, **256**]
L2 weight_decay: [0.0, 0.1, 0.001, **1e-05**]
Early stopping
Trainable params: 4080897

WNV NS2B/NS3

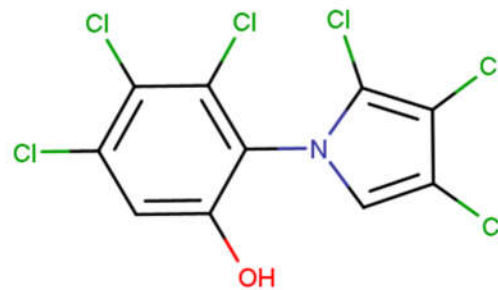
model	N	train			test		
		MAE	MSE	RMSE	MAE	MSE	RMSE
DNN	5	0.250	0.116	0.341	0.253	0.112	0.335
RF	5	0.159	0.049	0.221	0.413	0.285	0.534
SVM	12	0.188	0.082	0.286	0.404	0.300	0.548
XGB	8	0.134	0.032	0.184	0.352	0.232	0.482



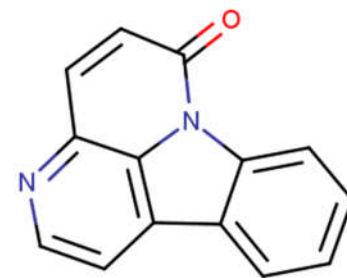
WNV NS2B/NS3



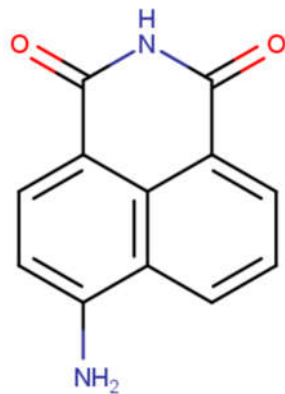
$pIC_{50} = 12.5$



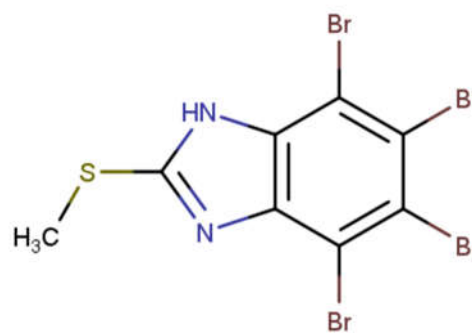
$pIC_{50} = 11.3$



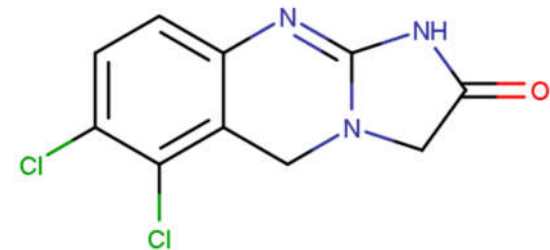
$pIC_{50} = 10.7$



$pIC_{50} = 10.0$

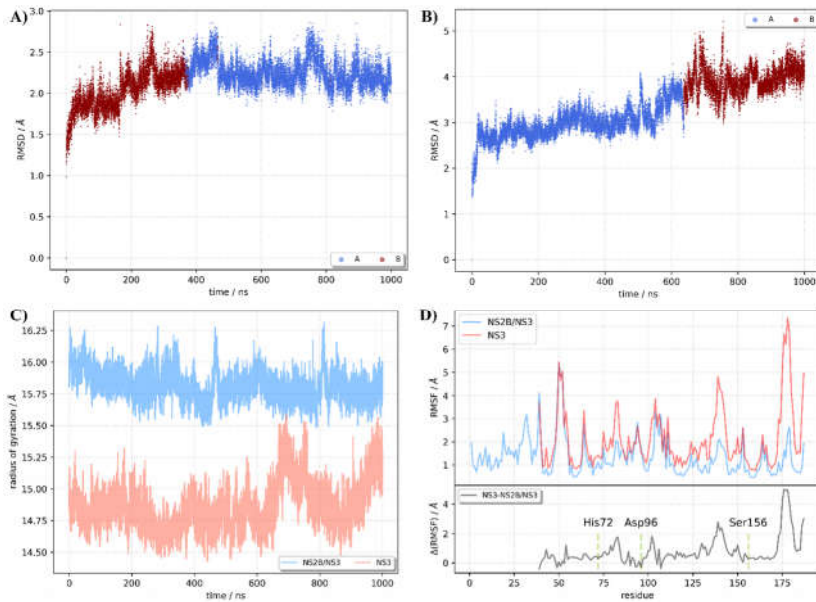
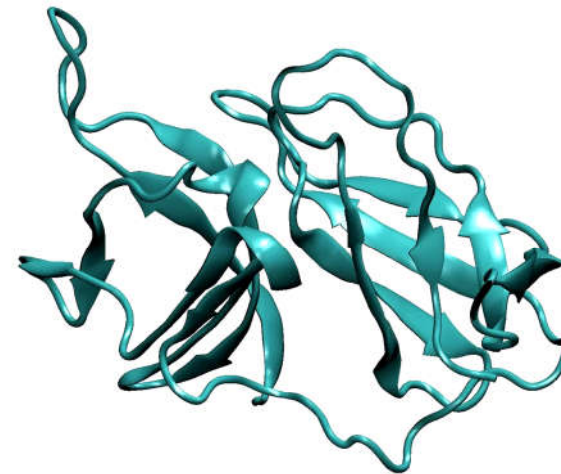
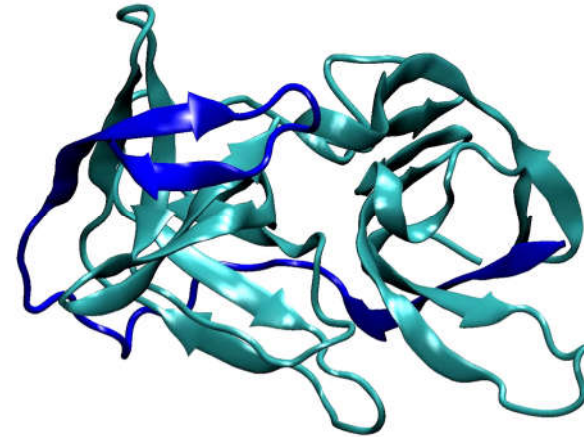
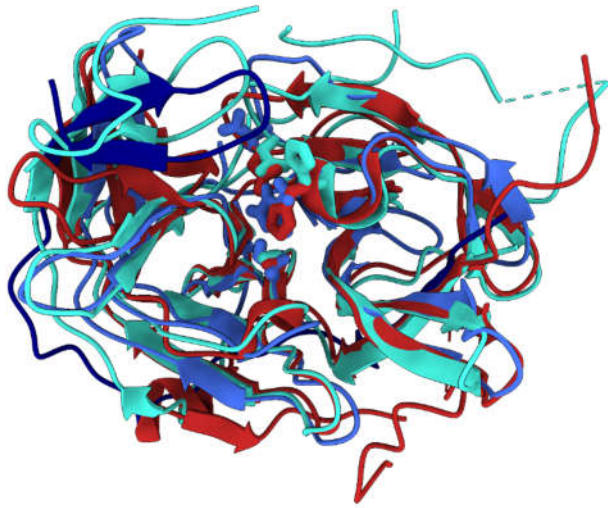


$pIC_{50} = 10.0$



$pIC_{50} = 9.9$

KFDV NS2B/NS3



[
si
[
K₁...

s
e

Thanks

- Shivananda Kandagalla
- Hrvoje Rimac
- Vladimir Potemkin
- Vladislav Naumovich
- Prateek Pathak
- Maria Grishina
- Gordan Janeš
- Ivan Mitrović

