BT6320

Protein Interaction: Computational Techniques
QSAR II

Shreeharsha G Bhat | BE21B037
Department of Biotechnology
Indian Institute of Technology Madras

TABLE OF CONTENTS

Pharmacophore modeling of FDA approved drugs for HIV protease inhibitors

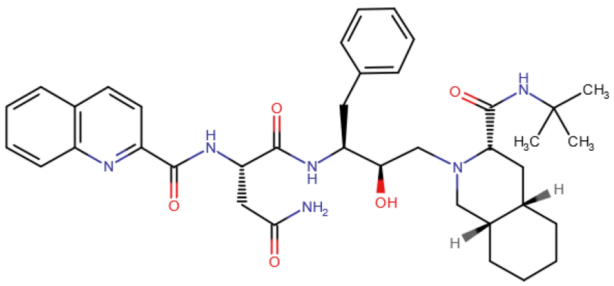
Introduction:

In the previous study we performed a QSAR study on FDA approved drugs for retrovirus Human immunodeficiency virus. We observed a high correlation of drug flexibility, polar properties with the drug's activity. This study will focus on performing pharmacophore modeling on a few of these drugs to identify other ligands that have the same binding features as the current drugs in the market. This analysis will be further exemplified with docking done on all the current drugs and drugs that have similar pharmacophore binding features with the protease inhibitor. The drugs include Saquinavir, Ritonavir, Indinavir, Nelfinavir , Amprenavir, Lopinavir , Atazanavir, Tipranavir, Darunavir and Telinavir, which is a Phase 2 drug. The interactions of these drugs with protease inhibitors include the Asp 25 and Asp 25' residues on the inhibitor. Most of the interactions being hydrophobic or hydrogen bond, donor-acceptor interactions. The drugs all act as hydrogen bond acceptors for the receptors. Nearby residues such as Gly27, Ile50, Val32, Ile47 also perform hydrophobic interactions to facilitate a binding.

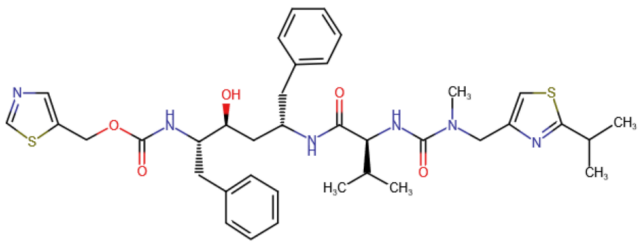
Literature Review:

Structure and Functioning of the 10 drugs:

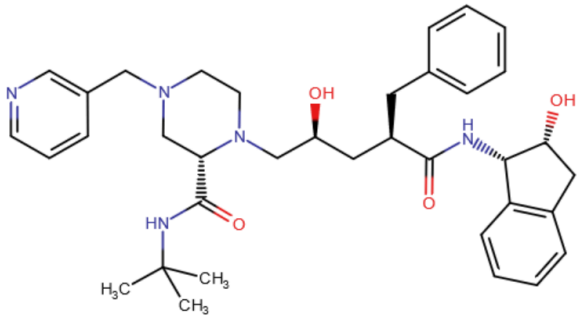
To understand the interactions necessary for the binding of these drugs to the protease inhibitor we study their structure. The following table provides the PDB IDs of all the 10 drugs in their bound state. The important interactions apart from the Asp25 and Asp 25' interaction with the inhibitor are listed for each of these drugs. (Asp25 and Asp25' are hydrogen bond interactions)



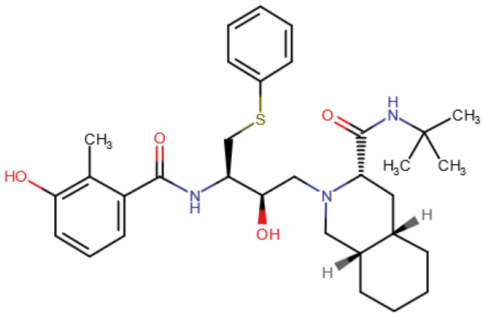
Saquinavir



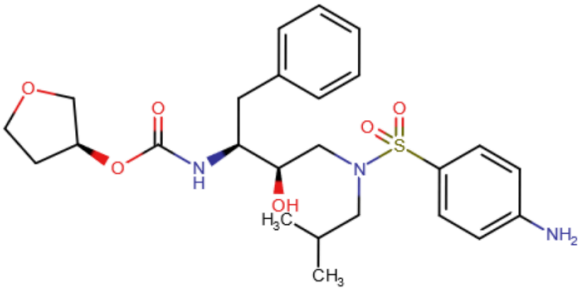
Ritonavir



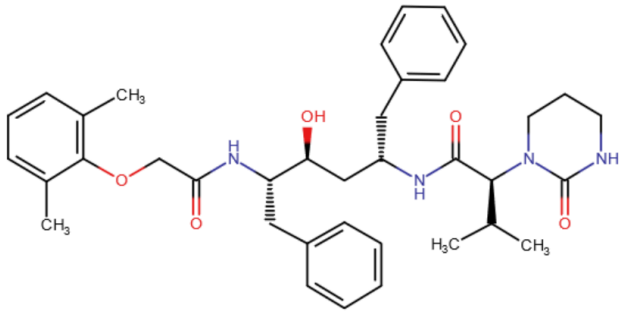
Indinavir



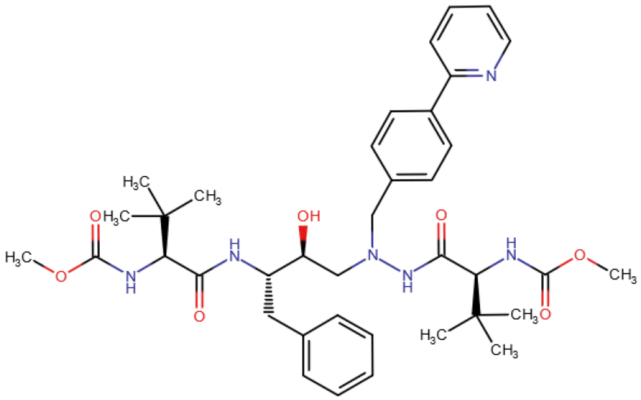
Nelfinavir



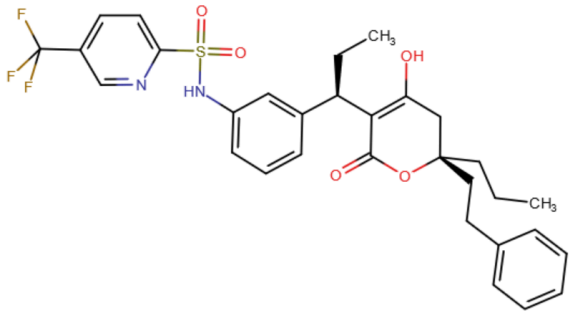
Amprenavir



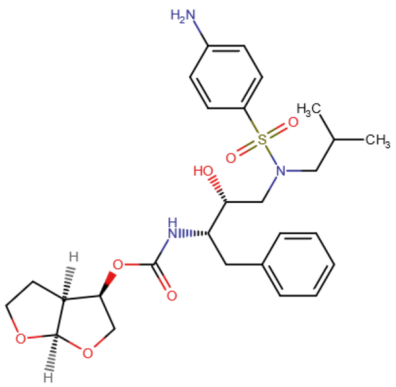
Lopinavir



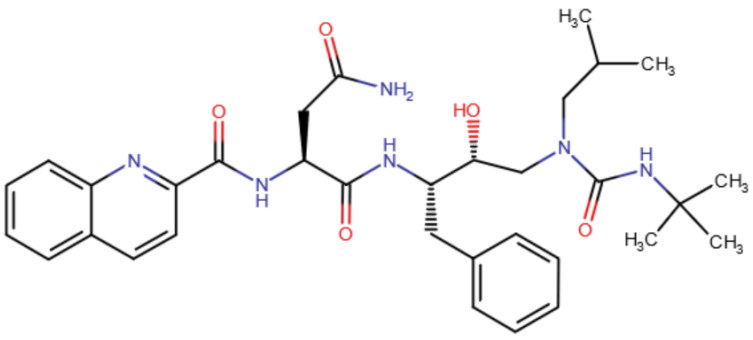
Atazanavir



Tipranavir



Darunavir



Telinavir

The interactions are

Sl.no	Drug	PDB ID	Interacting residue	Interaction Type
1	Saquinavir	1HXB	Ile50, Val82	Hydrophobic
2	Ritonavir	1HXW	Val32, Ile47, Ile50	Hydrophobic
3	Indinavir	1HSG	Gly27,Ile50,Val32,Ile47	Hydrophobic
4	Nelfinavir	2PYM	Gly27, Ile47, Ile50	Hydrophobic
5	Amprenavir	1HPV	Ile50, Pro81, Val82	Hydrophobic
6	Lopinavir	1MUI	Ile50, Val82	Hydrophobic
7	Atazanavir	2AQU	Val32, Ile47, Ile50	Hydrophobic
8	Tipranavir	1TW7	Extensive contact	Mainly van der waals
9	Darunavir	2IEN	Gly27, Ile50	Van der Waals
10	Telinavir	–	–	–

The majority of the interactions of the drugs involve Hydrophobic and hydrogen bond interactions.

Methodology:

Pharmacophore modeling:

Pharmacophore modeling is an essential technique in drug discovery and molecular drug design which helps in identifying and characterizing essential chemical features required for molecular recognition by the target. It defines the spatial arrangement of the functional groups that are necessary for optimal binding. The drug's essential pharmacophore is a template for finding other drugs with similar drug design.

Biovia Discovery studio Visualizer:

Biovia is a commercial-grade graphics visualization tool for analyzing and viewing protein interactions. The molecular overlay tool allows us to overlap multiple drugs to see the similarity in their structure. Though the 10 drugs have similar functions the analysis showed that they have poor overlapping similarity. Their structural differences are too varied.

Zinc Pharmer:

ZINCPharmer is free pharmacophore search software for screening the purchasable subset of the ZINC database. ZINCPharmer can import LigandScout and MOE pharmacophore definitions, as well as identify pharmacophore features directly from structure. As seen from the literature review we shall keep a higher importance on hydrophobic and essential hydrogen bond donors while performing ZINCPharmer searches. We enable only the hydrogen bond acceptors known from literature review responsible for interaction with the Asp 25 and Asp 25'. And since hydrophobic interactions were largely responsible for the stability of other drugs we also enable those.

AutoDock (Vina):

AutoDock Vina. AutoDock Vina is an open-source program for doing molecular docking of multiple ligands simultaneously. The grid box used was centered at the receptor and encompassed the entire protease inhibitor. The PDB ID for the receptor is 1TW7. The protease inhibitor was prepped using Autodock, removing all hetatms and converting it to a PDBQT file. The ligands themselves were also prepared similarly, making their amide bonds rotatable to allow for more flexibility. Rigid docking was performed for all the 10 drugs and the resultant hits from Zinc Pharmer.

Grid Box parameters:
X center : 3.167
Y center : 18.949
Z center : -13.246

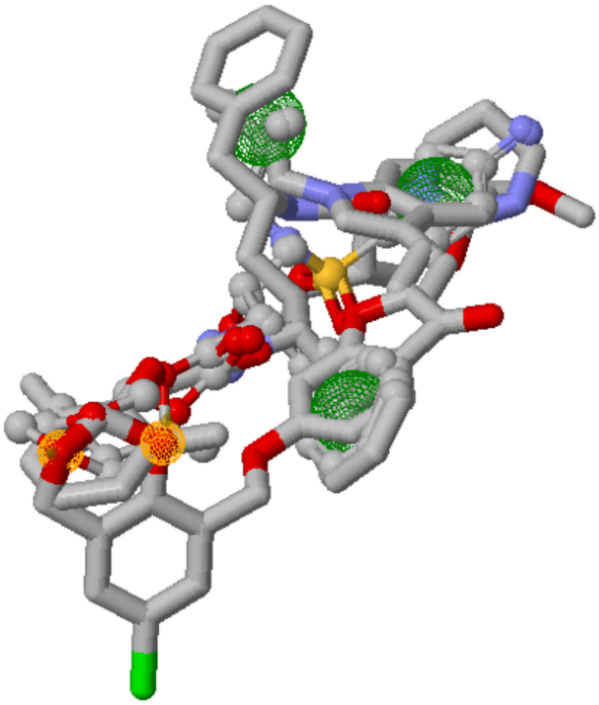
The box dimensions : 63 Å × 47 Å × 40 Å
Exhaustiveness : 10

Results:

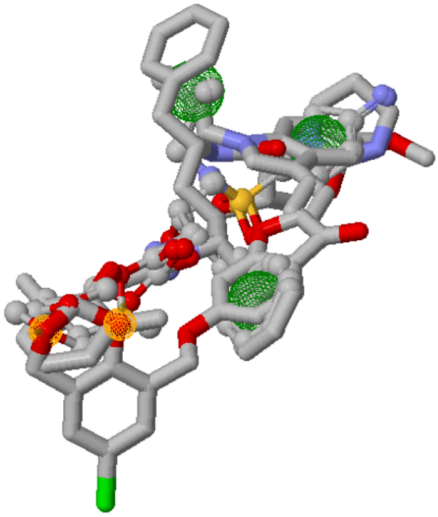
While the 10 drugs have similar functions, the analysis reveals poor overlap within all the drugs. To perform pharmacophore modeling three key drugs **Darunavir, Amprenavir and Lopinavir** were chosen since they have high structural overlap. With more than 70% similarity between the three we pick the overlapped molecule as our template for the pharmacophore model. The overlapping matrix is shown below

Drug	Darunavir	Amprenavir	Lopinavir
Darunavir	1	0.930587	0.737763
Amprenavir	0.930587	1	0.731096
Lopinavir	0.737763	0.731096	1

Given below is the overlapped model



The interaction enabled were



Pharmacophore | Filters | Viewer | Submit Query

>	HydrogenAcceptor	-1.82	-1.96	6.74	0.50	☐
>	HydrogenAcceptor	-4.04	-2.05	10.13	0.50	☑
>	HydrogenAcceptor	-6.01	-0.90	10.61	0.50	☑
>	HydrogenAcceptor	-1.13	-4.46	0.64	0.50	☐
>	HydrogenAcceptor	-1.94	-3.43	-1.49	0.50	☐
>	Hydrophobic	-0.82	-4.76	6.64	1.00	☑
>	Hydrophobic	1.86	-3.47	-1.48	1.00	☑
>	Hydrophobic	-0.48	0.48	-0.83	1.00	☑

The query search from ZINCPharmer revealed 14 unique hits for the model and we tested the docking for all the 14 regardless of their ADME properties.

Name	RMSD	Mass	RBnds
ZINC41507961	0.471	518	8
ZINC41507752	0.472	502	7
ZINC41507953	0.472	488	6
ZINC11573260	0.509	412	11
ZINC14887233	0.538	428	9
ZINC74257768	0.546	324	10
ZINC39590391	0.644	400	5
ZINC39590367	0.644	386	4
ZINC36719240	0.661	429	12
ZINC07508964	0.673	416	7
ZINC25858990	0.705	416	5
ZINC12757891	0.717	487	4
ZINC39590373	0.737	402	4
ZINC08453112	0.747	649	15

Docking results:

The simulation output log files for the initial 10 drugs are shown below

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.0	0.000	0.000
2	-6.0	7.047	11.747
3	-5.6	3.495	6.760
4	-5.6	1.982	3.782
5	-5.3	2.315	4.200
6	-5.3	3.297	7.665
7	-5.3	4.125	6.527
8	-5.3	2.904	4.777
9	-5.2	2.174	6.716

Amprenavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.3	0.000	0.000
2	-6.1	3.312	7.491
3	-6.0	2.578	7.012
4	-6.0	2.529	6.653
5	-6.0	3.030	5.535
6	-6.0	3.862	8.956
7	-5.9	3.796	9.708
8	-5.8	3.365	7.845
9	-5.8	3.843	9.290

Atazanavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.7	0.000	0.000
2	-6.5	3.864	6.092
3	-6.5	3.228	5.246
4	-6.4	3.988	6.375
5	-6.4	8.257	11.013
6	-6.4	3.444	5.779
7	-6.3	2.959	5.735
8	-6.3	10.016	12.620
9	-6.2	8.654	11.472

Darunavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.6	0.000	0.000
2	-7.5	4.069	8.016
3	-7.5	2.753	9.948
4	-7.3	3.248	5.891
5	-7.2	1.921	4.166
6	-7.2	3.153	9.180
7	-7.2	2.722	9.420
8	-7.2	5.105	7.681
9	-7.2	2.212	6.418

Indinavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.9	0.000	0.000
2	-6.8	1.225	2.503
3	-6.7	2.175	10.664
4	-6.7	2.629	10.327
5	-6.6	3.597	5.914
6	-6.6	2.493	8.623
7	-6.5	3.534	5.618
8	-6.5	2.570	10.557
9	-6.5	2.150	9.504

Lopinavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.3	0.000	0.000
2	-7.2	3.138	9.667
3	-7.2	3.427	10.008
4	-7.2	3.706	10.085
5	-7.0	3.050	9.847
6	-6.9	4.522	9.096
7	-6.9	1.568	2.150
8	-6.8	1.971	2.680
9	-6.8	3.509	7.245

Nelfinavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.9	0.000	0.000
2	-6.8	1.773	2.178
3	-6.7	2.776	9.438
4	-6.7	2.585	6.358
5	-6.6	2.958	8.037
6	-6.6	3.287	10.244
7	-6.6	5.541	9.564
8	-6.6	3.124	10.334
9	-6.6	4.072	7.450

Ritonavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-8.7	0.000	0.000
2	-8.6	2.385	11.393
3	-8.4	2.641	3.897
4	-8.3	2.799	6.381
5	-8.3	2.882	4.158
6	-8.3	2.759	5.564
7	-8.3	2.886	10.388
8	-8.2	2.759	7.533
9	-8.2	2.677	10.528

Saquinavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.4	0.000	0.000
2	-7.3	3.064	11.026
3	-7.2	1.417	2.100
4	-6.9	2.826	8.506
5	-6.9	2.871	7.784
6	-6.8	3.074	10.979
7	-6.8	2.829	6.280
8	-6.8	3.555	9.529
9	-6.7	3.707	9.548

Telinavir

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.0	0.000	0.000
2	-7.0	4.505	7.351
3	-7.0	7.344	11.194
4	-7.0	6.710	10.601
5	-7.0	7.390	12.372
6	-6.9	3.525	6.145
7	-6.9	3.555	8.155
8	-6.9	7.542	11.367
9	-6.9	4.217	9.054

Tipranavir

The docking results of the 14 drug hits are shown below:

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-5.0	0.000	0.000
2	-4.8	22.357	23.525
3	-4.7	21.516	23.558
4	-4.7	17.458	19.179
5	-4.6	14.475	17.003
6	-4.6	22.680	23.768
7	-4.5	4.532	5.983
8	-4.5	15.131	17.871
9	-4.5	15.619	18.330

ZINC07508964

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.8	0.000	0.000
2	-6.5	1.305	4.124
3	-6.5	4.042	9.853
4	-6.5	2.933	6.337
5	-6.3	4.464	9.451
6	-6.3	4.576	9.880
7	-6.3	4.413	8.231
8	-6.3	2.819	6.149
9	-6.3	2.804	5.375

ZINC08453112

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-5.3	0.000	0.000
2	-5.2	13.011	14.760
3	-5.2	13.125	14.900
4	-5.1	2.679	5.814
5	-5.0	13.236	15.405
6	-4.9	8.529	11.330
7	-4.9	19.629	22.595
8	-4.9	3.674	6.923
9	-4.9	9.609	12.077

ZINC11573260

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-8.8	0.000	0.000
2	-8.6	4.250	11.281
3	-8.6	5.029	6.678
4	-8.6	9.672	13.968
5	-8.6	4.050	10.929
6	-8.5	2.122	10.266
7	-8.5	3.965	11.122
8	-8.3	1.477	2.163
9	-8.2	3.663	10.931

ZINC12757891

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.2	0.000	0.000
2	-7.1	5.235	11.605
3	-7.0	1.975	2.685
4	-7.0	1.915	2.548
5	-6.9	3.050	10.325
6	-6.8	5.802	7.576
7	-6.8	1.673	2.773
8	-6.7	6.073	9.640
9	-6.7	7.677	11.500

ZINC14887233

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.9	0.000	0.000
2	-7.8	2.921	10.164
3	-7.8	2.967	10.581
4	-7.7	3.362	10.530
5	-7.6	2.671	10.504
6	-7.6	6.491	9.022
7	-7.6	6.984	8.992
8	-7.5	8.636	12.116
9	-7.5	3.116	4.375

ZINC25858990

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.7	0.000	0.000
2	-6.6	9.048	13.098
3	-6.5	10.444	13.633
4	-6.3	2.107	3.580
5	-6.3	12.064	15.514
6	-6.2	4.511	8.578
7	-6.2	8.761	12.998
8	-6.2	2.856	4.785
9	-6.1	5.409	8.860

ZINC36719240

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.4	0.000	0.000
2	-6.3	3.101	9.466
3	-6.3	9.708	15.444
4	-6.3	4.406	5.451
5	-6.3	1.439	1.812
6	-6.2	2.281	2.866
7	-6.2	3.131	9.929
8	-6.2	3.173	4.275
9	-6.2	3.378	9.795

ZINC39590367

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.9	0.000	0.000
2	-6.9	3.981	10.371
3	-6.7	11.177	13.642
4	-6.6	4.279	5.899
5	-6.5	2.652	8.725
6	-6.4	1.194	1.979
7	-6.3	2.881	8.374
8	-6.3	5.474	8.356
9	-6.2	2.881	9.001

ZINC39590373

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.6	0.000	0.000
2	-6.6	6.608	12.788
3	-6.5	4.534	10.891
4	-6.5	8.565	12.294
5	-6.3	5.886	7.048
6	-6.3	5.117	9.905
7	-6.2	12.696	16.981
8	-6.1	5.788	6.943
9	-6.1	8.581	10.875

ZINC39590391

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-8.7	0.000	0.000
2	-8.5	4.470	10.955
3	-8.4	4.390	10.766
4	-8.1	3.307	5.008
5	-8.0	1.532	2.426
6	-7.9	4.189	10.380
7	-7.9	4.036	9.934
8	-7.9	3.840	5.533
9	-7.9	7.824	9.657

ZINC41507752

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.9	0.000	0.000
2	-7.8	5.736	7.118
3	-7.8	0.775	2.029
4	-7.8	5.961	9.716
5	-7.8	6.467	9.072
6	-7.8	6.252	9.349
7	-7.8	6.087	7.875
8	-7.8	3.663	8.822
9	-7.7	5.654	7.462

ZINC41507953

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-9.0	0.000	0.000
2	-8.7	4.225	10.587
3	-8.7	4.590	12.309
4	-8.6	6.535	11.951
5	-8.5	5.813	10.073
6	-8.4	5.143	9.308
7	-8.3	1.971	2.442
8	-8.2	5.960	9.035
9	-8.2	5.571	9.934

ZINC41507961

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-5.0	0.000	0.000
2	-4.8	22.357	23.525
3	-4.7	21.516	23.558
4	-4.7	17.458	19.179
5	-4.6	14.475	17.003
6	-4.6	22.680	23.768
7	-4.5	4.532	5.983
8	-4.5	15.131	17.871
9	-4.5	15.619	18.330

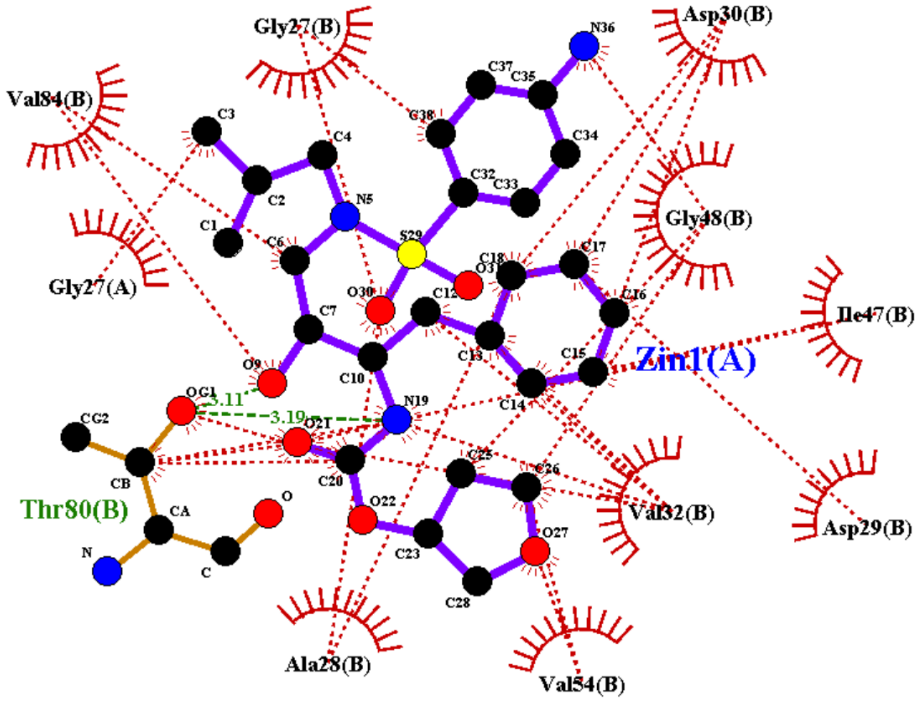
ZINC74257768

We first check the lowest binding affinity of the three drugs chosen for the pharmacophore modeling, they are Darunavir, Amprenavir and Lopinavir having the respective lowest binding affinity as –6.7, –6.0 and –6.9 kcal/mol.

The zinc drugs that have the lowest binding affinity are ZINC41507961, ZINC41507752, ZINC12757891 having binding affinity of –9.0, –8.7 and –8.8 kcal/mol respectively.

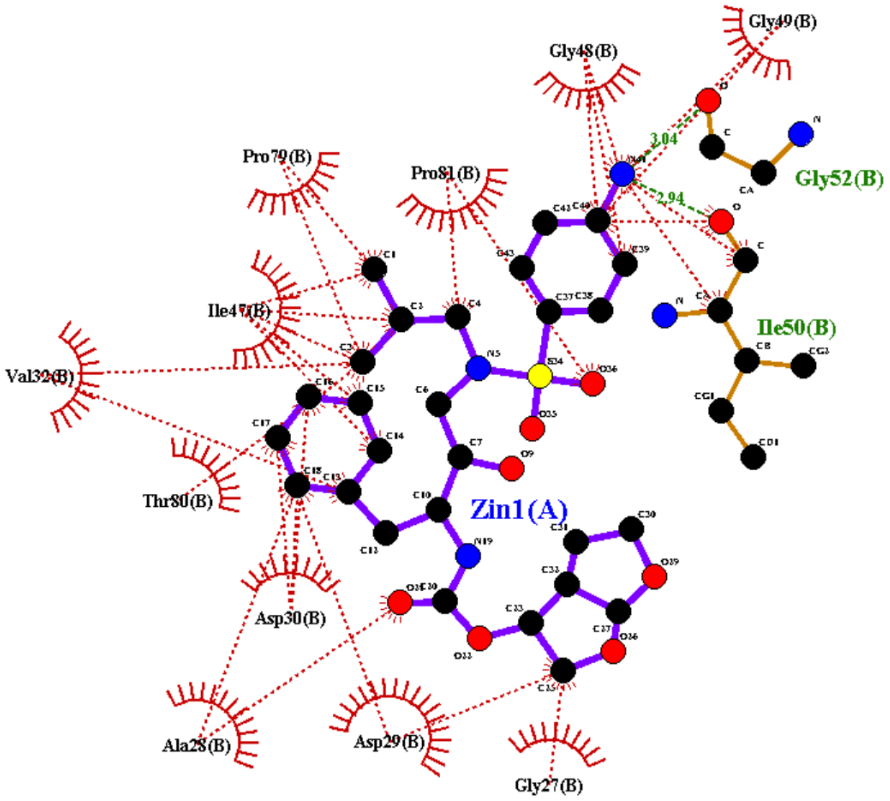
LigPlot Visualization:

The interaction was plotted on LigPlot to visualize the binding. Shown below are the ligplot results and the interactions post docking. We concern ourselves with just the hydrogen bonded interactions, but the rest of the interactions are available in the appendix



Amprenavir

```
PDB code: Amprenavir
==============


Hydrogen bonds
--------------

      <----- A T O M   1 ----->       <----- A T O M   2 ----->

      Atom Atom Res  Res              Atom Atom Res  Res
       no  name name  no  Chain        no  name name  no  Chain  Distance
   1    17  N19  ZIN    1    A    ---    42  OG1 THR   80    B     3.186
   2     8  O9   ZIN    1    A    ---    42  OG1 THR   80    B     3.114


Non-bonded contacts
-------------------

      <----- A T O M   1 ----->       <----- A T O M   2 ----->

      Atom Atom Res  Res              Atom Atom Res  Res
       no  name name  no  Chain        no  name name  no  Chain  Distance
   1    78  O    GLY   27    B    ---    35  C38 ZIN    1    A     3.836
   2    72  CA   GLY   48    B    ---    33  N36 ZIN    1    A     3.446
   3    70  CB   ALA   28    B    ---    27  O30 ZIN    1    A     3.436
   4    67  CA   ALA   28    B    ---    27  O30 ZIN    1    A     3.577
   5    77  C    GLY   27    B    ---    27  O30 ZIN    1    A     3.633
   6    85  CG2  VAL   54    B    ---    24  O27 ZIN    1    A     3.550
   7    85  CG2  VAL   54    B    ---    23  C26 ZIN    1    A     3.667
   8    72  CA   GLY   48    B    ---    23  C26 ZIN    1    A     3.494
   9    72  CA   GLY   48    B    ---    22  C25 ZIN    1    A     3.844
  10    40  CB   THR   80    B    ---    19  O21 ZIN    1    A     3.487
  11    65  CD1  ILE   47    B    ---    19  O21 ZIN    1    A     3.342
  12    48  CG1  VAL   32    B    ---    19  O21 ZIN    1    A     3.289
  13    47  CB   VAL   32    B    ---    19  O21 ZIN    1    A     3.819
```



Darunavir

PDB code: Darunavir

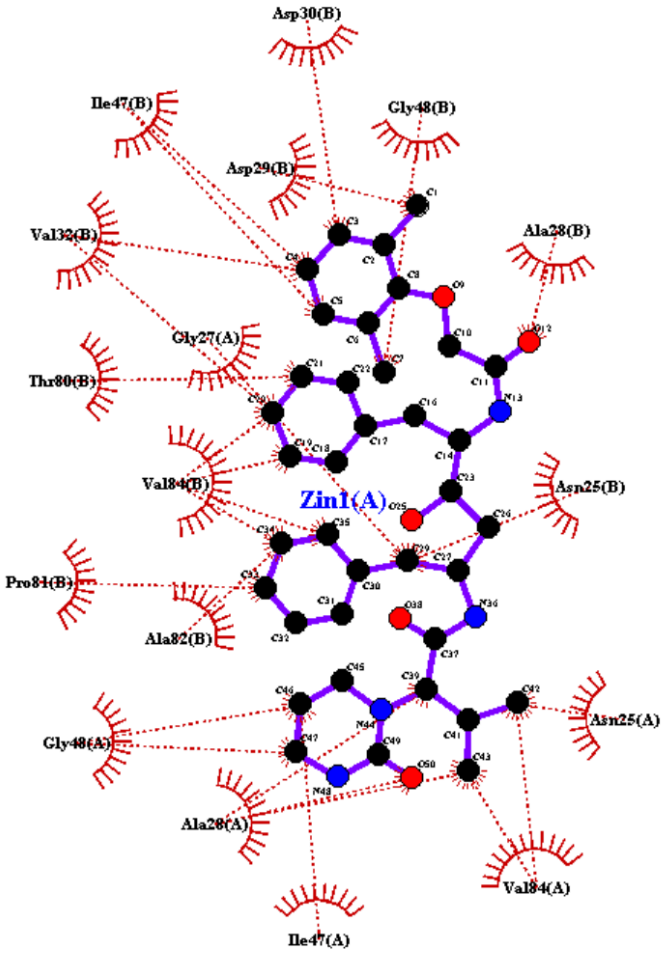
==========

Hydrogen bonds

	<----- A T O M 1 ----->					<----- A T O M 2 ----->					
	Atom no	Atom name	Res name	Res no	Chain		Atom no	Atom name	Res name	Res no	Chain	Distance
1	36	N41	ZIN	1	A	---	70	O	GLY	52	B	3.036
2	36	N41	ZIN	1	A	---	42	O	ILE	50	B	2.939

Non-bonded contacts

	<----- A T O M 1 ----->					<----- A T O M 2 ----->					
	Atom no	Atom name	Res name	Res no	Chain		Atom no	Atom name	Res name	Res no	Chain	Distance
1	41	C	ILE	50	B	---	36	N41	ZIN	1	A	3.632
2	40	CA	ILE	50	B	---	36	N41	ZIN	1	A	3.832
3	78	C	GLY	49	B	---	36	N41	ZIN	1	A	3.681
4	77	CA	GLY	49	B	---	36	N41	ZIN	1	A	3.864
5	65	C	GLY	48	B	---	36	N41	ZIN	1	A	3.750
6	64	CA	GLY	48	B	---	36	N41	ZIN	1	A	3.694
7	42	O	ILE	50	B	---	35	C40	ZIN	1	A	3.871
8	76	N	GLY	49	B	---	35	C40	ZIN	1	A	3.843
9	64	CA	GLY	48	B	---	35	C40	ZIN	1	A	3.645
10	64	CA	GLY	48	B	---	34	C39	ZIN	1	A	3.779
11	86	CD	PRO	81	B	---	31	O36	ZIN	1	A	3.801
12	85	CG	PRO	81	B	---	31	O36	ZIN	1	A	3.642
13	87	N	ASP	29	B	---	22	C25	ZIN	1	A	3.690
14	119	O	GLY	27	B	---	22	C25	ZIN	1	A	3.476
15	75	CB	ALA	28	B	---	19	O21	ZIN	1	A	3.453

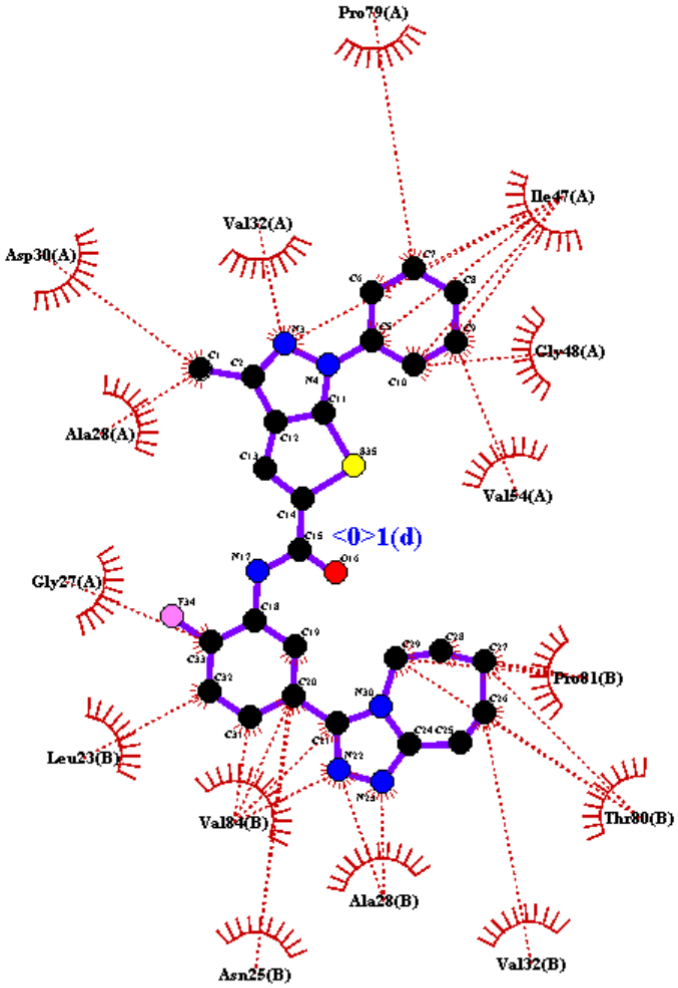


Lopinavir

PDB code: Lopinavir

Non-bonded contacts

	<----- A T O M 1 ----->					<----- A T O M 2 ----->					
	Atom no	Atom name	Res name	Res no	Chain		Atom no	Atom name	Res name	Res no	Chain	Distance
1	63	CA	ALA	28	A	---	46	O50	ZIN	1	A	3.374
2	67	N	GLY	48	A	---	43	C47	ZIN	1	A	3.525
3	67	N	GLY	48	A	---	42	C46	ZIN	1	A	3.751
4	133	CB	ILE	47	A	---	42	C46	ZIN	1	A	3.872
5	76	CG1	VAL	84	A	---	39	C43	ZIN	1	A	3.849
6	66	CB	ALA	28	A	---	39	C43	ZIN	1	A	3.689
7	77	CG2	VAL	84	A	---	38	C42	ZIN	1	A	3.642
8	123	ND2	ASN	25	A	---	38	C42	ZIN	1	A	3.632
9	66	CB	ALA	28	A	---	36	C39	ZIN	1	A	3.809
10	53	CG2	VAL	84	B	---	32	C35	ZIN	1	A	3.679
11	53	CG2	VAL	84	B	---	31	C34	ZIN	1	A	3.579
12	155	O	ALA	82	B	---	31	C34	ZIN	1	A	3.880

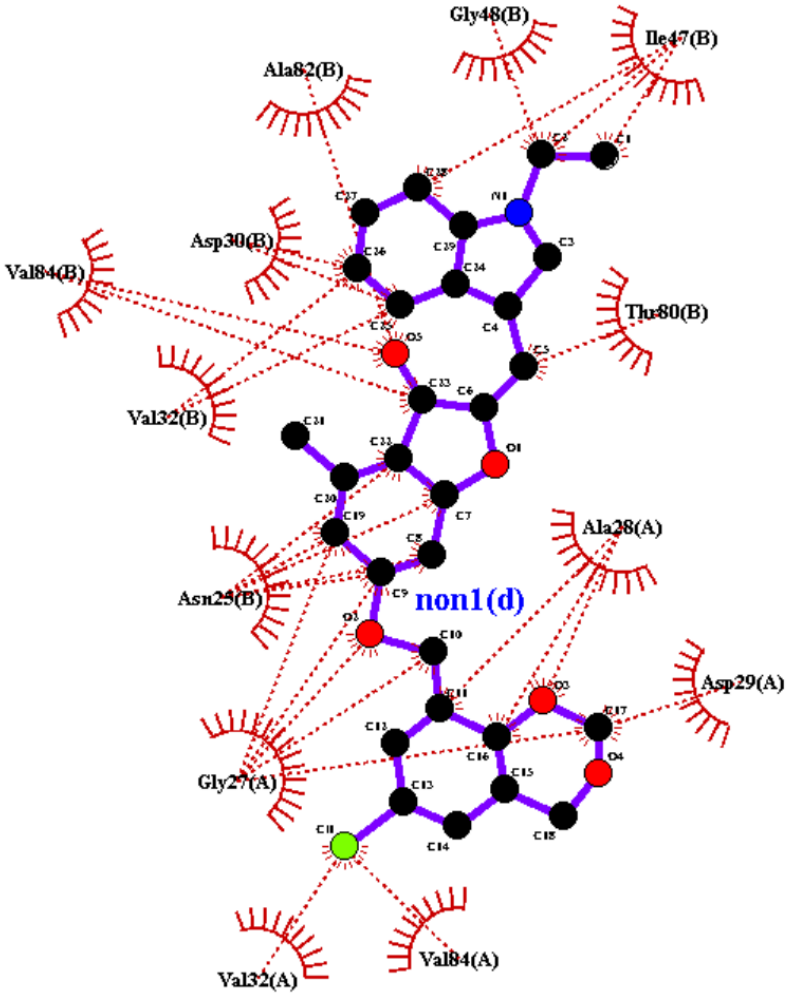


ZINC12757891

PDB code: ZINC12757891
==========

Non-bonded contacts

	<----- A T O M 1 ----->					<----- A T O M 2 ----->				
	Atom no	Atom name	Res name	Res no	Chain		Atom no	Atom name	Res name	Res no	Chain	Distance
1	93	O	GLY	27	A	---	33	C33	<0>	1	d	3.885
2	127	CD2	LEU	23	B	---	32	C32	<0>	1	d	3.794
3	50	CG2	VAL	84	B	---	31	C31	<0>	1	d	3.135
4	64	CD	PRO	81	B	---	29	C29	<0>	1	d	3.853
5	57	OG1	THR	80	B	---	29	C29	<0>	1	d	3.726
6	64	CD	PRO	81	B	---	28	C28	<0>	1	d	3.829
7	64	CD	PRO	81	B	---	27	C27	<0>	1	d	3.792
8	57	OG1	THR	80	B	---	27	C27	<0>	1	d	3.468
9	57	OG1	THR	80	B	---	26	C26	<0>	1	d	3.532
10	134	CG2	VAL	32	B	---	26	C26	<0>	1	d	3.866

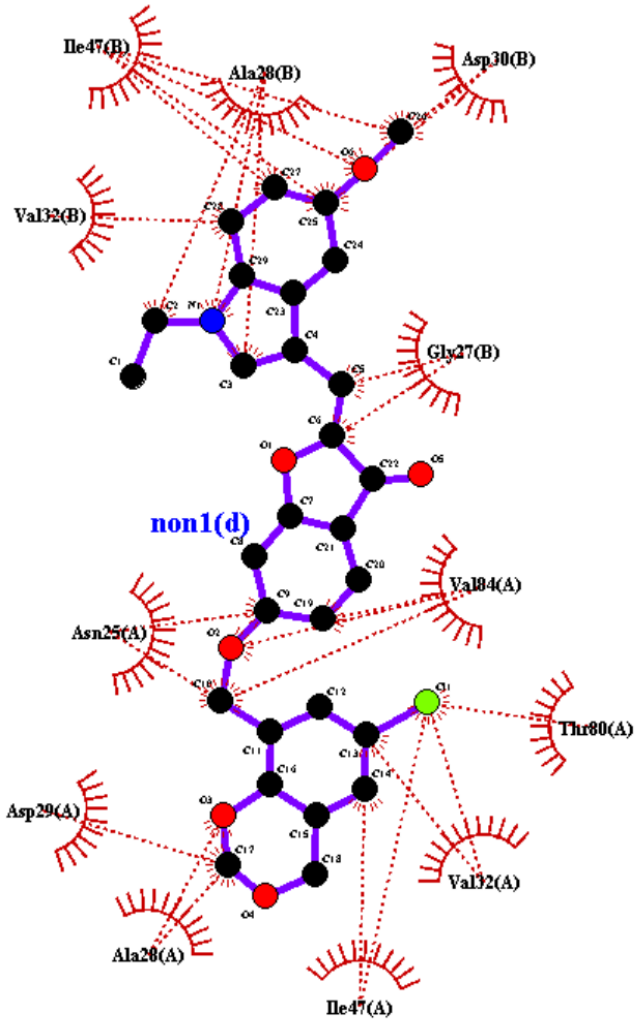


ZINC41507752

PDB code: ZINC41507752

Non-bonded contacts

	<----- A T O M 1 ----->					<----- A T O M 2 ----->					
	Atom no	Atom name	Res name	Res no	Chain		Atom no	Atom name	Res name	Res no	Chain	Distance
1	68	CG2	VAL	84	A	---	36	C11	non	1	d	3.588
2	66	CB	VAL	84	A	---	36	C11	non	1	d	3.806
3	109	CG2	VAL	32	A	---	36	C11	non	1	d	3.637
4	56	CD1	ILE	47	B	---	34	C28	non	1	d	3.645
5	83	CG2	VAL	32	B	---	32	C26	non	1	d	3.672
6	72	O	ASP	30	B	---	32	C26	non	1	d	3.420
7	83	CG2	VAL	32	B	---	31	C25	non	1	d	3.607
8	72	O	ASP	30	B	---	31	C25	non	1	d	3.843
9	94	CG2	VAL	84	B	---	29	C23	non	1	d	3.710
10	47	ND2	ASN	25	B	---	28	C22	non	1	d	3.785
11	47	ND2	ASN	25	B	---	25	C19	non	1	d	3.871
12	40	O	GLY	27	A	---	25	C19	non	1	d	3.000



ZINC41507961

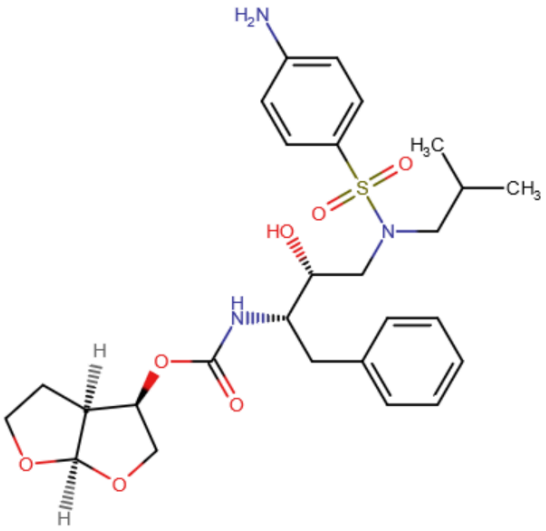
PDB code: ZINC41507961

Non-bonded contacts

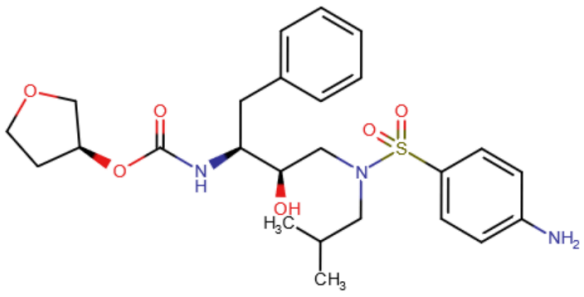
	<----- A T O M 1 ----->					<----- A T O M 2 ----->				
	Atom no	Atom name	Res name	Res no	Chain	Atom no	Atom name	Res name	Res no	Chain	Distance
1	110	CB	THR	80	A	37	Cl1	non	1	d	3.640
2	85	CD1	ILE	47	A	37	Cl1	non	1	d	3.636
3	57	CG2	VAL	32	A	37	Cl1	non	1	d	3.627
4	56	CG1	VAL	32	A	37	Cl1	non	1	d	3.652
5	55	CB	VAL	32	A	37	Cl1	non	1	d	3.728
6	119	CG2	VAL	32	B	35	C28	non	1	d	3.736
7	45	CD1	ILE	47	B	34	C27	non	1	d	3.755
8	44	CG2	ILE	47	B	34	C27	non	1	d	3.768
9	42	CB	ILE	47	B	34	C27	non	1	d	3.638
10	44	CG2	ILE	47	B	33	C26	non	1	d	3.837
11	42	CB	ILE	47	B	33	C26	non	1	d	3.892

Discussion:

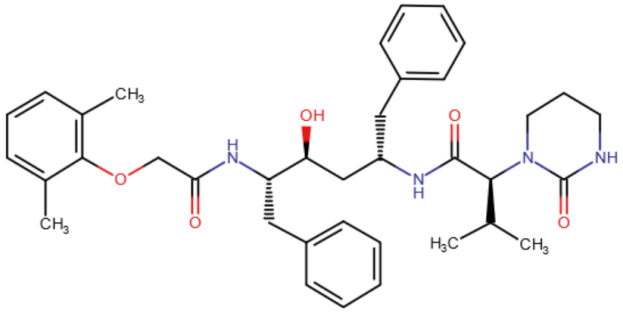
Due to the lack of current drugs in the market that are available for HIV–protease inhibitor as the target, we see that the 10 drugs are structurally different from each other. Hence it was prudent to use just 3 drugs for overlapping during pharmacophore modeling.
The three drugs are structurally the most similar among the 10



Darunavir



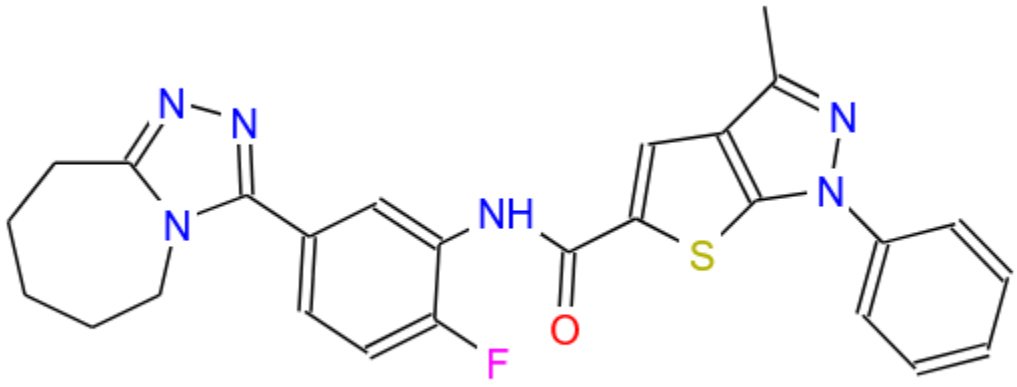
Amprenavir



Lopinavir

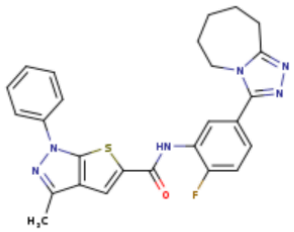
Due to lack of hits being restricted to 14 and not 20. We check the ADME properties of the 3 drugs that showed the least binding affinity among all the hits. The structure,the smiles representations and the ADME properties of the 3 drugs are shown below

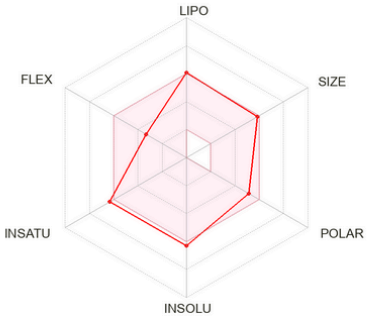
ZINC12757891



SMILES: Cc1nn(-c2ccccc2)c2sc(C(=O)Nc3cc(-c4nnc5n4CCCCC5)ccc3F)cc12

Molecule 1





SMILES O=C(c1sc2c(c1)c(nn2c1ccccc1)C)Nc1cc(ccc1F)c1nnc2n1CCCCC2

Physicochemical Properties	
Formula	C26H23FN6OS
Molecular weight	486.56 g/mol
Num. heavy atoms	35
Num. arom. heavy atoms	25
Fraction Csp3	0.23
Num. rotatable bonds	5
Num. H-bond acceptors	5
Num. H-bond donors	1
Molar Refractivity	135.64
TPSA	105.87 Å²

Lipophilicity	
Log $P_{o/w}$ (iLOGP)	3.95
Log $P_{o/w}$ (XLOGP3)	5.15
Log $P_{o/w}$ (WLOGP)	6.00
Log $P_{o/w}$ (MLOGP)	4.27
Log $P_{o/w}$ (SILICOS-IT)	5.13
Consensus Log $P_{o/w}$	4.90

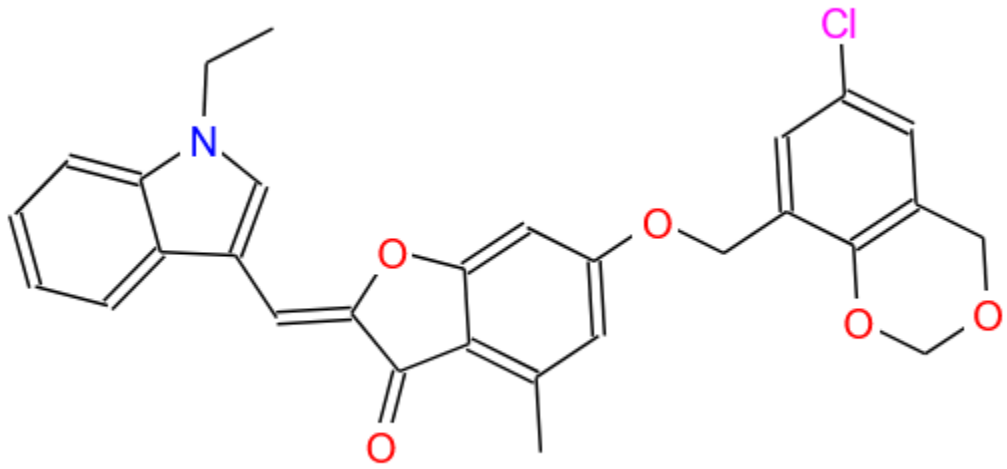
Water Solubility	
Log *S* (ESOL)	-6.30
Solubility	2.44e-04 mg/ml ; 5.01e-07 mol/l
Class	Poorly soluble
Log *S* (Ali)	-7.12
Solubility	3.70e-05 mg/ml ; 7.61e-08 mol/l
Class	Poorly soluble
Log *S* (SILICOS-IT)	-8.44
Solubility	1.77e-06 mg/ml ; 3.65e-09 mol/l
Class	Poorly soluble

Pharmacokinetics	
GI absorption	Low
BBB permeant	No
P-gp substrate	Yes
CYP1A2 inhibitor	No
CYP2C19 inhibitor	Yes
CYP2C9 inhibitor	Yes
CYP2D6 inhibitor	No
CYP3A4 inhibitor	Yes
Log K_p (skin permeation)	-5.61 cm/s

Druglikeness	
Lipinski	Yes; 1 violation: MLOGP>4.15
Ghose	No; 3 violations: MW>480, WLOGP>5.6, MR>130
Veber	Yes
Egan	No; 1 violation: WLOGP>5.88
Muegge	No; 1 violation: XLOGP3>5
Bioavailability Score	0.55

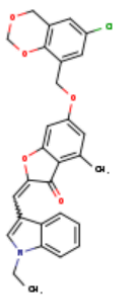
Medicinal Chemistry	
PAINS	0 alert
Brenk	0 alert
Leadlikeness	No; 2 violations: MW>350, XLOGP3>3.5
Synthetic accessibility	3.93

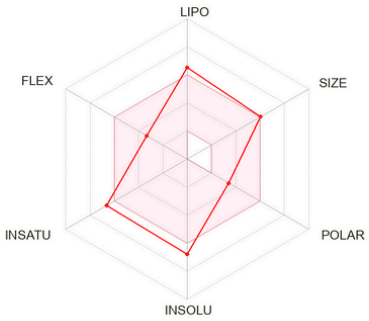
ZINC41507752



SMILES: CCn1cc(/C=C2Oc3cc(OCc4cc(Cl)cc5c4OCOC5)cc(C)c3C2=O)c2ccccc21

Molecule 1





SMILES CCn1cc(c2c1cccc2)C=C1Oc2c(C1=O)c(C)cc(c2)OCc1cc(Cl)cc2c1OCOC2

Physicochemical Properties	
Formula	C29H24ClNO5
Molecular weight	501.96 g/mol
Num. heavy atoms	36
Num. arom. heavy atoms	21
Fraction Csp3	0.21
Num. rotatable bonds	5
Num. H-bond acceptors	5
Num. H-bond donors	0
Molar Refractivity	138.77
TPSA	58.92 Å²

Lipophilicity	
Log $P_{o/w}$ (iLOGP)	4.83
Log $P_{o/w}$ (XLOGP3)	5.93
Log $P_{o/w}$ (WLOGP)	6.27
Log $P_{o/w}$ (MLOGP)	3.39
Log $P_{o/w}$ (SILICOS-IT)	6.87
Consensus Log $P_{o/w}$	5.46

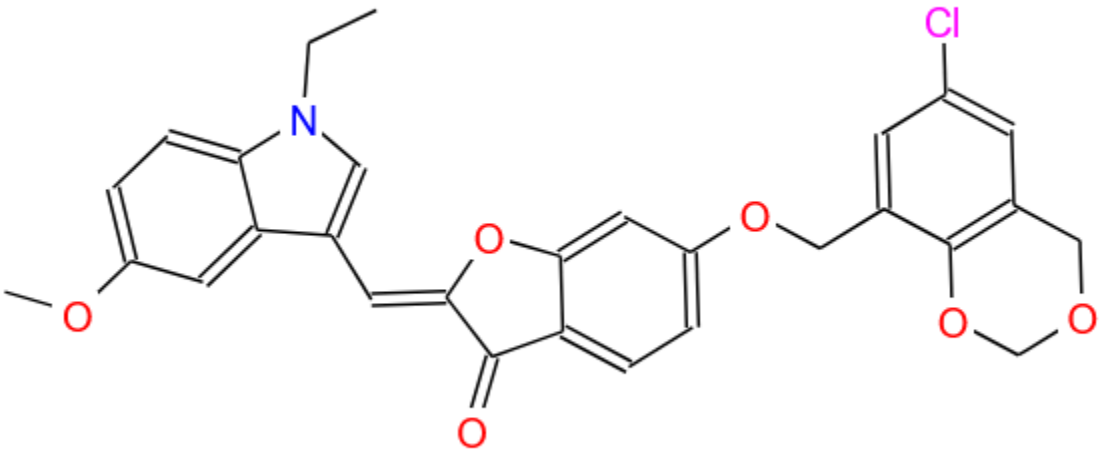
Water Solubility	
Log *S* (ESOL)	-6.79
Solubility	8.15e-05 mg/ml ; 1.62e-07 mol/l
Class	Poorly soluble
Log *S* (Ali)	-6.94
Solubility	5.74e-05 mg/ml ; 1.14e-07 mol/l
Class	Poorly soluble
Log *S* (SILICOS-IT)	-9.49
Solubility	1.61e-07 mg/ml ; 3.21e-10 mol/l
Class	Poorly soluble

Pharmacokinetics	
GI absorption	High
BBB permeant	No
P-gp substrate	Yes
CYP1A2 inhibitor	No
CYP2C19 inhibitor	Yes
CYP2C9 inhibitor	Yes
CYP2D6 inhibitor	No
CYP3A4 inhibitor	No
Log K_p (skin permeation)	-5.15 cm/s

Druglikeness	
Lipinski	Yes; 1 violation: MW>500
Ghose	No; 3 violations: MW>480, WLOGP>5.6, MR>130
Veber	Yes
Egan	No; 1 violation: WLOGP>5.88
Muegge	No; 1 violation: XLOGP3>5
Bioavailability Score	0.55

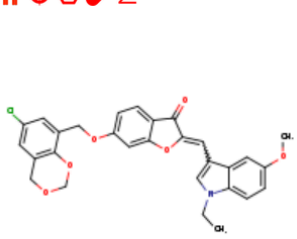
Medicinal Chemistry	
PAINS	0 alert
Brenk	1 alert: michael_acceptor_1
Leadlikeness	No; 2 violations: MW>350, XLOGP3>3.5
Synthetic accessibility	4.08

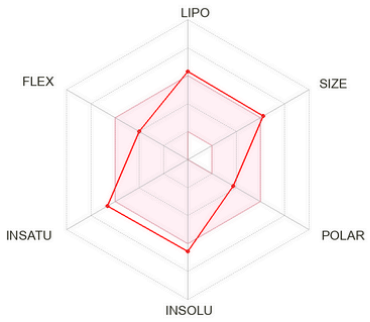
ZINC41507961



SMILES: CCn1cc(/C=C2\Oc3cc(OCc4cc(Cl)cc5c4OCOC5)ccc3C2=O)c2cc(OC)ccc21

Molecule 1





SMILES COc1ccc2c(c1)c(/C=C/1\Oc3c(C1=O)ccc(c3)OCc1cc(Cl)cc3c1OCOC3)cn2CC

Physicochemical Properties	
Formula	C29H24ClNO6
Molecular weight	517.96 g/mol
Num. heavy atoms	37
Num. arom. heavy atoms	21
Fraction Csp3	0.21
Num. rotatable bonds	6
Num. H-bond acceptors	6
Num. H-bond donors	0
Molar Refractivity	140.29
TPSA	68.15 Å²

Lipophilicity	
Log $P_{o/w}$ (iLOGP)	4.92
Log $P_{o/w}$ (XLOGP3)	5.53
Log $P_{o/w}$ (WLOGP)	5.97
Log $P_{o/w}$ (MLOGP)	2.86
Log $P_{o/w}$ (SILICOS-IT)	6.41
Consensus Log $P_{o/w}$	5.14

Water Solubility	
Log *S* (ESOL)	-6.56
Solubility	1.43e-04 mg/ml ; 2.76e-07 mol/l
Class	Poorly soluble
Log *S* (Ali)	-6.72
Solubility	9.85e-05 mg/ml ; 1.90e-07 mol/l
Class	Poorly soluble
Log *S* (SILICOS-IT)	-9.22
Solubility	3.15e-07 mg/ml ; 6.08e-10 mol/l
Class	Poorly soluble

Pharmacokinetics	
GI absorption	High
BBB permeant	No
P-gp substrate	Yes
CYP1A2 inhibitor	No
CYP2C19 inhibitor	Yes
CYP2C9 inhibitor	Yes
CYP2D6 inhibitor	No
CYP3A4 inhibitor	No
Log K_p (skin permeation)	-5.53 cm/s

Druglikeness	
Lipinski	Yes; 1 violation: MW>500
Ghose	No; 3 violations: MW>480, WLOGP>5.6, MR>130
Veber	Yes
Egan	No; 1 violation: WLOGP>5.88
Muegge	No; 1 violation: XLOGP3>5
Bioavailability Score	0.55

Medicinal Chemistry	
PAINS	0 alert
Brenk	1 alert: michael_acceptor_1
Leadlikeness	No; 2 violations: MW>350, XLOGP3>3.5
Synthetic accessibility	4.12

Conclusion:

The pharmacophore modeling for the HIV protease inhibitor drugs was performed and three new drugs that show similar pharmacophores were identified. Docking was performed on these drugs with the protease inhibitor to identify the binding of these drugs. Each of the drugs did not form the necessary hydrogen bond required (Asp25 and Asp25') to inhibit the target molecule. Further study regarding the drug structure revealed that ZINC41507961 and ZINC41507752 have similar structure and pharmacophores as well. For future study, running the docking for a larger number of steps would be more effective. The drugs themselves seem to be structurally varied which makes it difficult to predict other drugs that could be effective.

Appendix:

Files and Links

The FDA approved drugs (ZINC, mol2 files) – here.
ZINCPharmer hits (ZINC, mol2 files) – here.
The protease inhibitor (PDB, PDBQT files) – here.
All drugs (PDBQT files) – here.
AutoDock Vina output (log files) – here.
LigPlot Visualization and Interactions (PDB, txt files) – here.

Acknowledgements:

I would like to thank the course faculty Prof.Mukesh Doble and his teaching assistants, for their invaluable inputs throughout the course.

****