POXVIRUS AND SMALLPOX COMPUTATIONAL ANALYSIS OF NATURAL COMPOUNDS AND SYNTHETIC MOLECULES ON VIRAL PROTEIN TARGETS

ALESSANDRO CAREGLIO[1]

[1]Department of Drug Science and Technology of the University of Turin (doctor in herbal medicine and student in chemistry and pharmaceutical technologies)

Correspondence: Alessandro Careglio Frazione Santa Maria 92/a La Morra 12064 (CN) Italy

alessandro.careglio@edu.unito.it

alessandrocareglio@gmail.com

Keywords: poxvirus, smallpox, computational chemistry, natural compounds

ABSTRACT

In this study, natural compounds and synthetic drugs, in particular antivirals, are computationally compared to some known and new protein targets of Poxviruses with Cresset Flare.

Several antiviral drugs have demonstrated anti-poxvirus activity *in vitro*, but only cidofovir, a nucleoside analogue approved for the therapy of cytomegalovirus retinitis in AIDS, has shown some usefulness, although its high nephrotoxicity seems to greatly limit its therapeutic use. [1]

The following compositional comparisons are performed:

- ligand-based screening using cidofovir as a reference on a library of about 1000 compounds of natural compounds and common drugs, and a library of 700 molecules with antiviral activity.
- screening structure based being present a cocristallized in cidofovir protein data bank with PDB ID 5KM8 of a library of natural compounds.
- alignment of S. pneumoniae topoisomerases in which there is a drug cocristallized with smallpoisomerase topoisomerase that has no ligands from Protein Data Bank.
- covalent screening on serine and tyrosine residues present at the mouth of the hydrophobic pocket with drugs and natural compounds containing nitrile groups

The L1 protein of the poxvirus, a molecule that is conserved throughout the poxvirus family and is almost identical in the vaccine smallpox virus and the human smallpox virus, L1 is a myristoylated envelope protein that is a powerful target for antibody neutralization and an important component of current experimental vaccines[4] . The L1 structure reveals a hydrophobic cavity located adjacent to its terminal nitrogen. The cavity would be able to shield the myristed part, which is essential for the assembly of virion[4].

INTRODUCTION

Poxviruses (*family Poxviridae),* among which the causative agent of smallpox for centuries one of the "scourges" of humanity, are the largest viruses capable of infecting vertebrate and invertebrate animals. Interest in these viruses, greatly decreased after the eradication of smallpox and the discontinuation of the vaccine in the early 80s

Poxviruses are linear bicatenary DNA viruses, oval or brick in shape and I with complex symmetry, whose large dimensions (170-200 x 300-450 nm), at the limits of the resolution power of the optical microscope.

The genes present in the central part of the genome (about 90 in chordopoxviruses) are rather preserved and encode almost all the proteins essential for the completion of the multiplicative cycle of which the enzymes required for the transcription and replication of DNA, such as RNA and DNA polymerase, kinase, phosphatase, etc. while the terminal genes, variable between the different Poxviruses, are involved in the definition of the host spectrum and, in the case of Chordopoxviruses, in the modulation of the inflammatory and immune responses of the host and in the pathogenesis ("virulence genes"). This allows viral multiplication in the cytoplasmic site and mechanisms for its control similar to those of eukaryotic cells.

Chordopoxviruses are epitheliotropic and therefore have a common tendency to produce *skin lesions.* Infections can range from localized forms, with benign skin lesions, to highly lethal generalized forms (it is believed that smallpox was the human infectious disease that caused the most deaths.

Smallpox was an ancient dreaded acute epidemic disease, the outcome of which was death or healing accompanied by lasting immunity.

Three Orthopoxviruses also assume epidemiological relevance in humans: the vaccine virus (VACV), the bovine smallpox virus (cowpoxvirus, CPXV) and monkey virus (monkeypoxvirus, MPXV)

Since many of the Poxvirus infections resolve spontaneously, no therapeutic intervention is usually required. In the case of localised infections (e.g. molluscum contagiosum) cryotherapy or removal by squeezing ("curettage") or surgical nodules, as well as topical application of antiviral agents (e.g. cantharidine and cidofovir) or immunomodulating agents (e.g. imiquimod) are possible.

In the case of severe systemic infections or in immuno-depressed individuals or with atopic dermatitis, systemic therapeutic approaches may be necessary. Several

antiviral drugs have demonstrated anti-poxvirus activity *in vitro*, but only cidofovir [(S)-l-(3-hydroxy-2-phosphonylmethoxypropyl) cytosine], a nucleoside analogue approved for the therapy of cytomegalovirus retinitis in AIDS, has shown some usefulness, although its high nephrotoxicity seems to greatly limit its therapeutic use. [1]

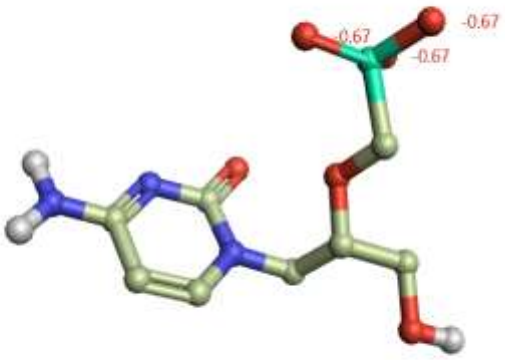


FIGURE 2 Cidofovir: nitrogen atoms in blue carbon green oxygen in bright green phosphorus red.

MATERIALS AND METHODS AND RESULTS

Computational comparisons are performed with Flare Cresset: ligand and structure based screening, and protein alignment

Ligand-based

1. The molecules with the greatest analogy referring to the electrostatic fields generated by the cidofovir molecule compared to the molecules present in a database (selleckchem.com)[2] are sought, comparing 3000 compounds both natural and very common drugs. (https://www.selleckchem.com/screening/natural-product-library.html)

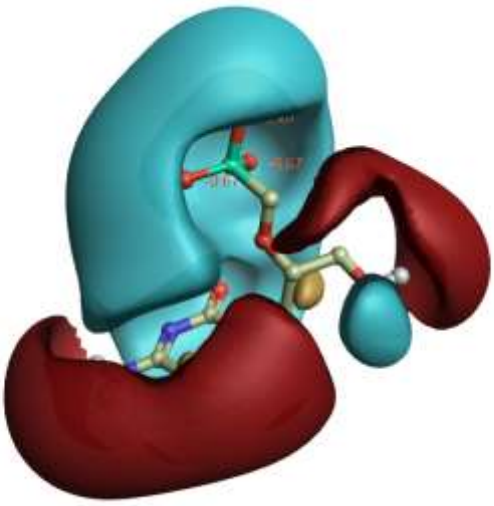
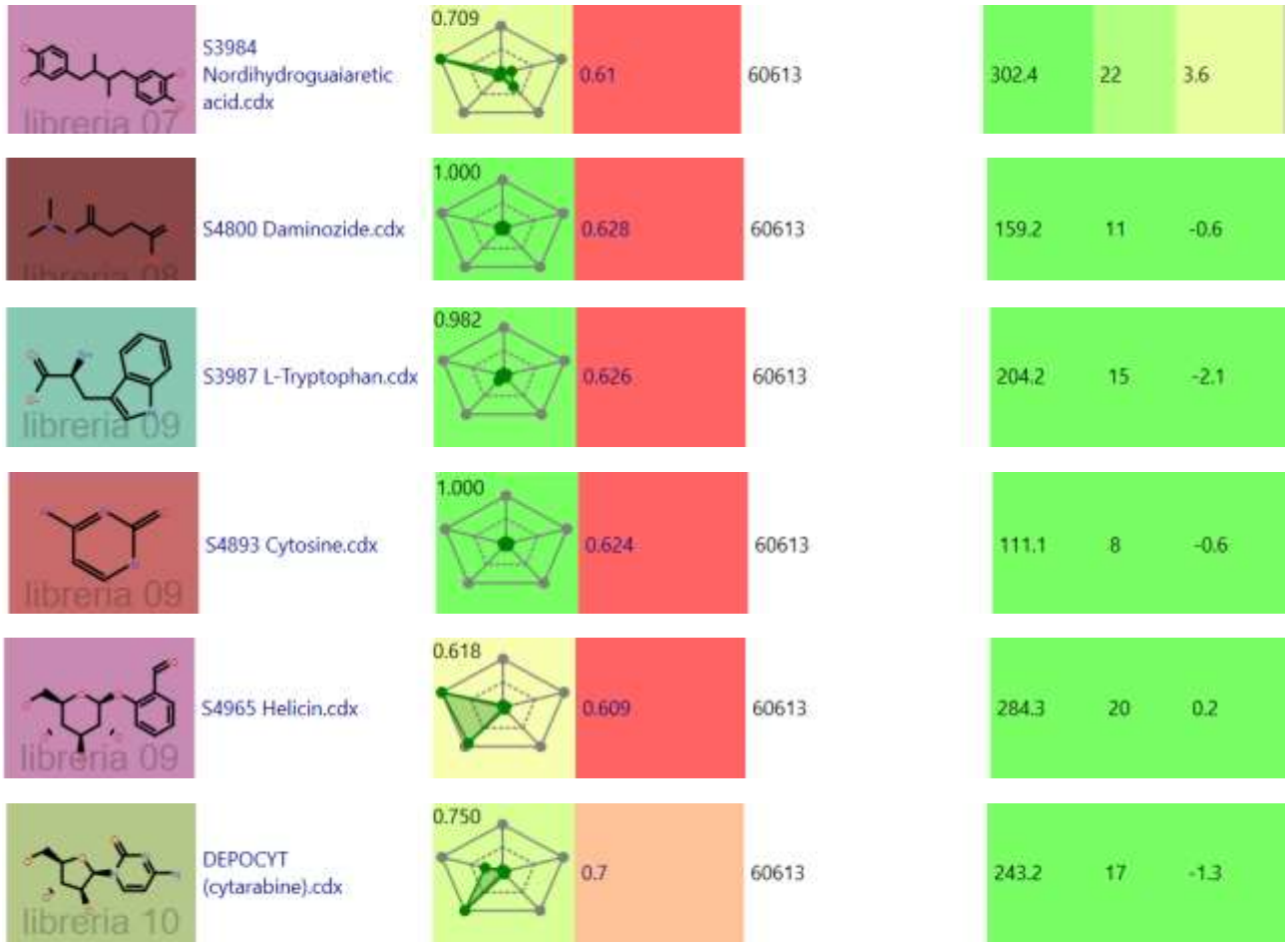


FIGURE 3 Cidofovir representation of electrostatic fields: positive red, negative blue, apolar brown

Molecules with a coefficient of "similarity" greater than 0.6 (considering the identity of 1) are taken into account.

Ligands

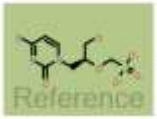
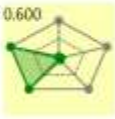
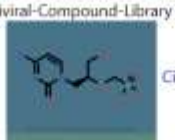
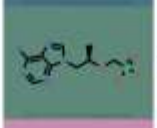
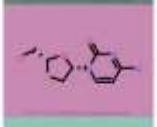
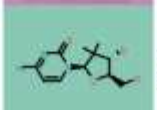
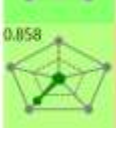
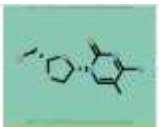
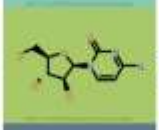
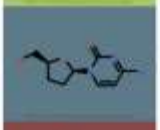
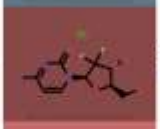
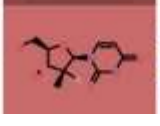
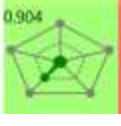
	Structure	Title	Radial Plot	Sim	Ref	Protein	MW	#Atoms	SlogP
		Cytidine.cdx	0.750	0.695	60613		243.2	17	-1.3
		Sulfamethoxypyridazin...	0.864	0.613	60613		280.3	19	0.9
		S4566 DL-Panthenol.cdx	0.751	0.603	60613		205.3	14	0.0
		S4558 Citiolone.cdx	0.975	0.605	60613		159.2	10	0.2

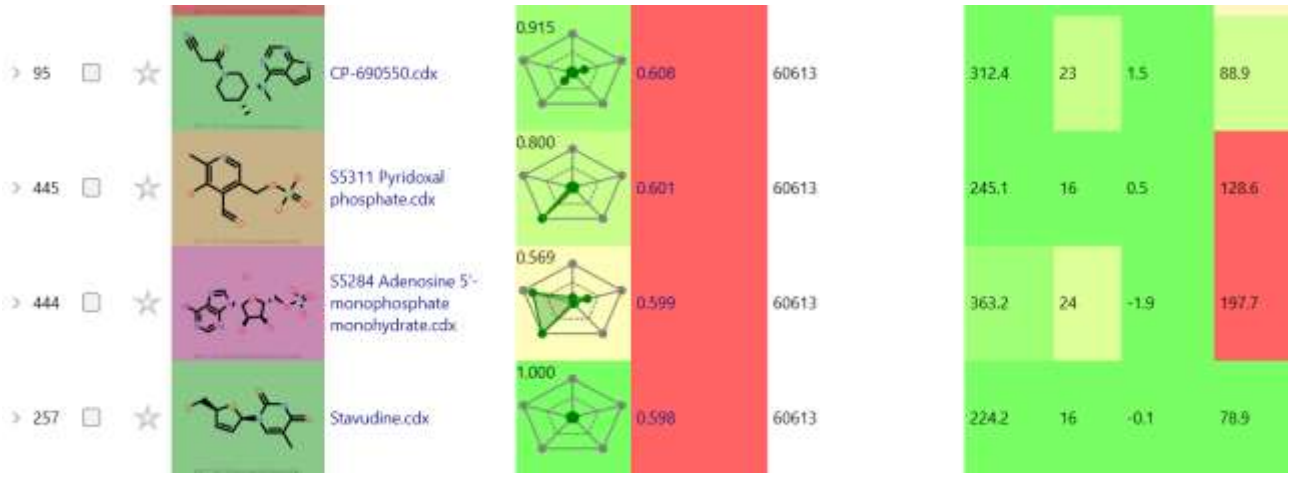


Screenshot 1

2. A screening (ligand-based) is performed on a library of 700 compounds with antiviral activity (ref. 2) with cidofovir as a reference (https://www.selleckchem.com/screening/antiviral-compound-library.html)

Ligands

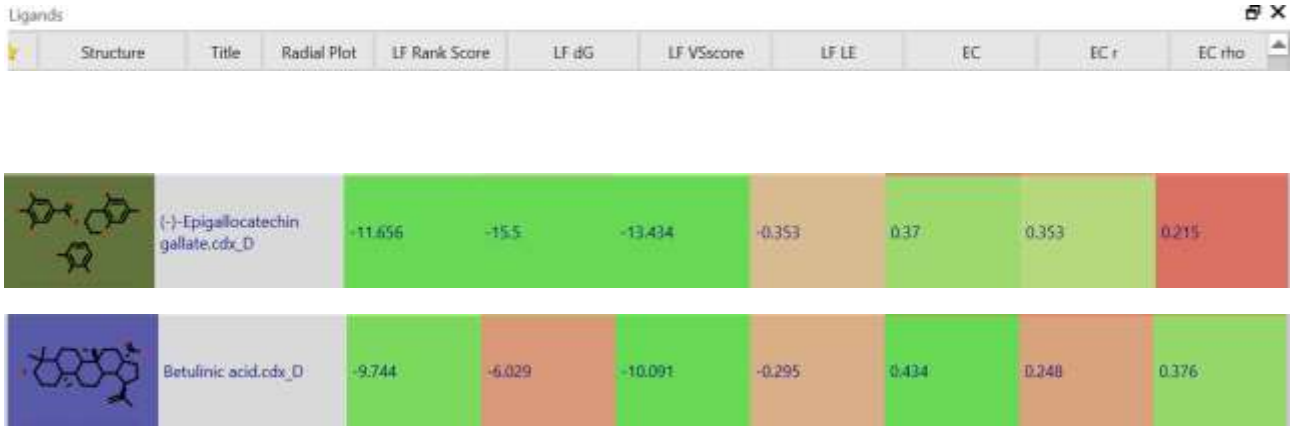
Index	Structure	Title	Radial Plot	Sim	Ref	Protein	MW	#Atoms	SlogP	TPSA
Reference (1)										
1		60613					277.2	18	-1	157.3
cidofovir (0)										
20211119-L7000-Antiviral-Compound-Library (700)										
339		Cidofovir .cdx		0.81	60613		277.2	18	-1	157.3
37		Tenofovir(PMPA).cdx		0.763	60613		285.2	19	-0.1	148
66		EPIVIR (lamivudine).cdx		0.738	60613		229.3	15	-0.1	88.2
64		GEMCITABINE (gemcitabine).cdx		0.706	60613		263.2	18	-0.4	108.4
226		Emtricitabine.cdx		0.698	60613		247.2	16	0.3	88.1
342		DEPOCYT (cytarabine).cdx		0.694	60613		243.2	17	-1.3	128.6
69		HIVID (zalcitabine).cdx		0.689	60613		211.2	15	0.0	88.1
332		Gemcitabine hydrochloride.cdx		0.665	60613		298.7	19	0.3	108.4
440		S5307 2'-deoxy-2'-fluoro-2'-C-methyluridine.cdx		0.629	60613		260.2	18	-0.5	99.1



Screenshots 2

Structure-based

There is a cocristallized in cidofovir protein data bank with PDB ID 5KM8 on natural compounds library[2]



Screenshot 3

FIGURE screenshot of the first classified betulinic acid, see legend screenshot

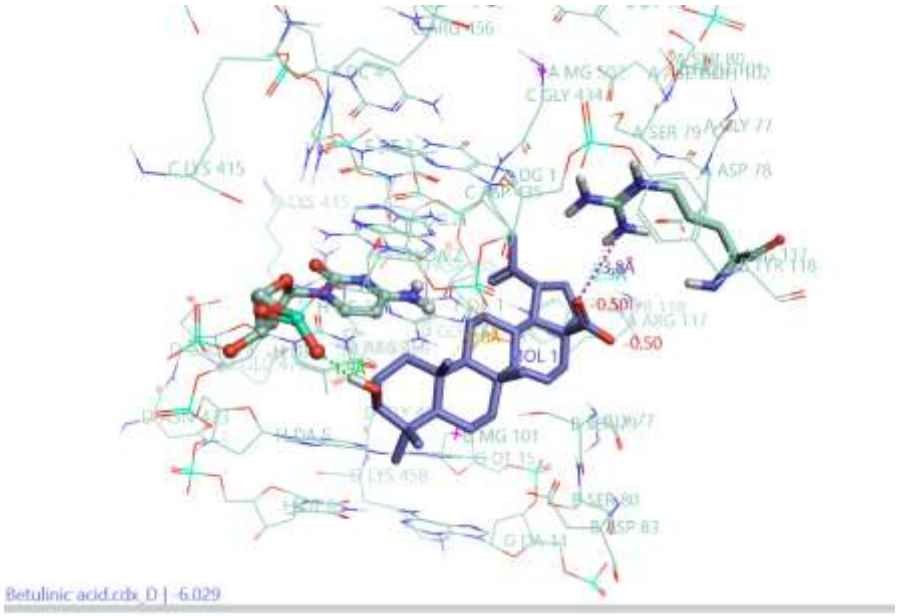


FIGURE 4 representation of betulinic acid in blue bound to the protein, the numbers indicate the binding distances with the individual amino acids, labeled

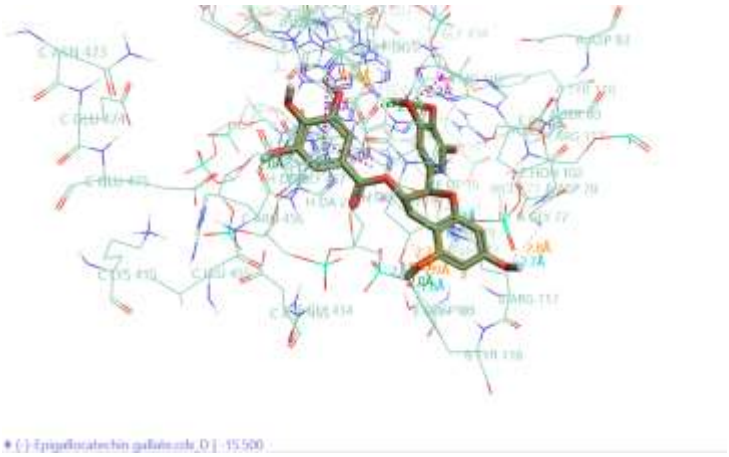


FIGURE 5 representation of the epigallocatechingallate in blue bound to the protein, the numbers indicate the binding distances with the individual amino acids, labeled

Protein alignment

The proteins topoisomerase of S. pneumonie (PDB ID 3RAE)[8 i] and topoisomerase of the smallpox virus (PDB ID 3IGC)[8 ii], are aligned, this is necessary, as in the crystallized of the smallpox protein there is no drug or reference molecule that highlight the possible binding pocket, while to the topoisomerase of S. pneumonia is bound levofloxacin.

Alignment of the two chains: they are equal to 15.56%

FIGURE 6 alignment of s topoisomerases. pneumoniae and smallpox

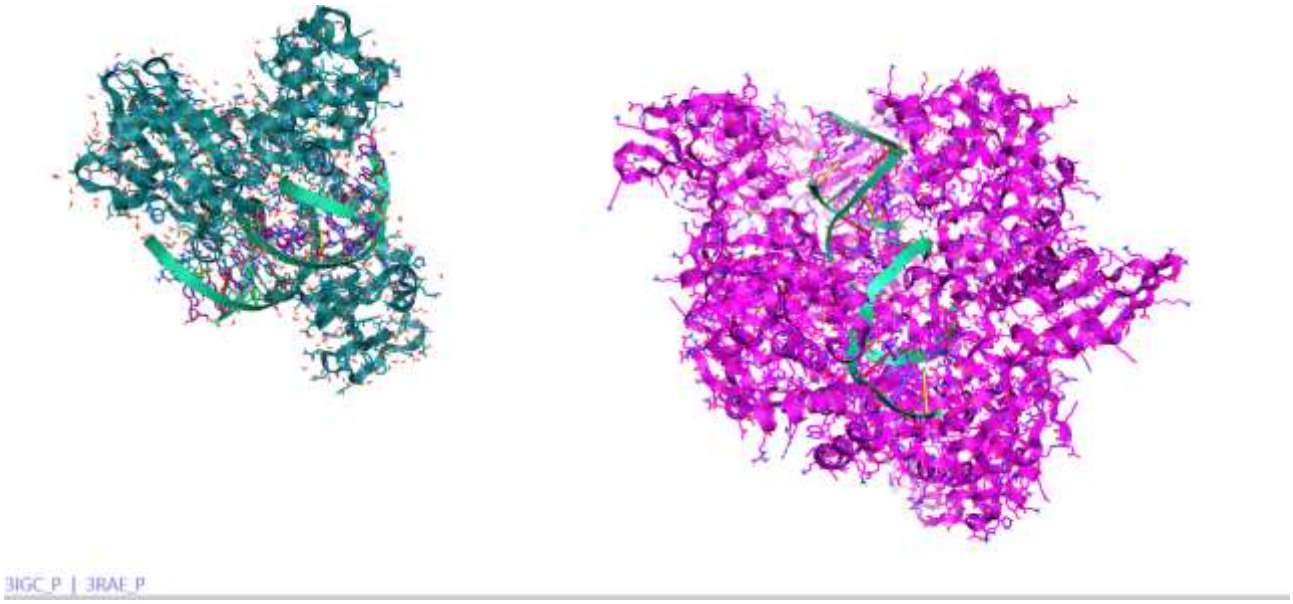


FIGURE 7 Topoisomerase of smallpox in green and S. pneumonie in pink (with DNA segment) separated.

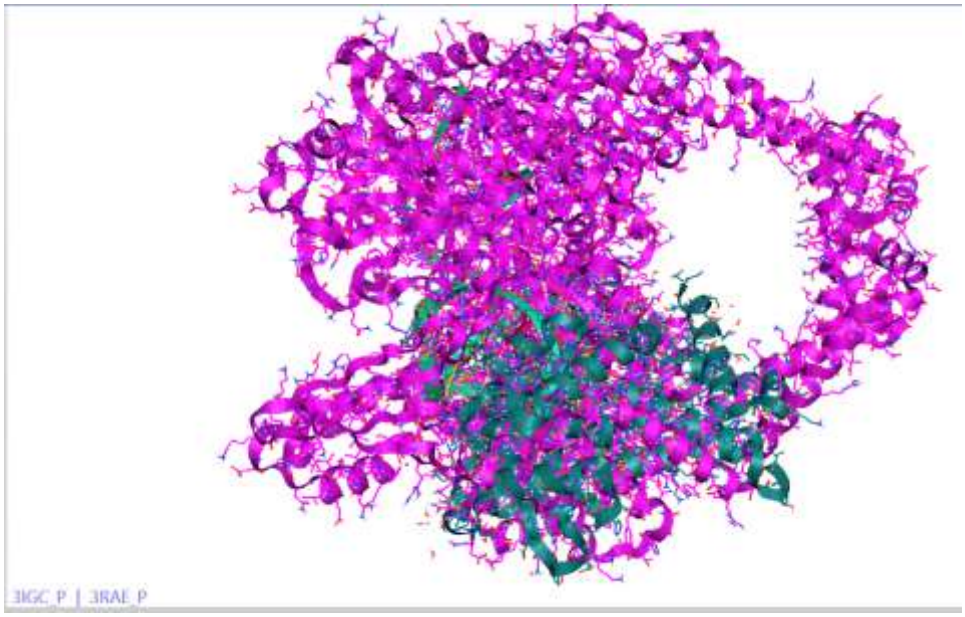


FIGURE 8 Topoisomerase of smallpox and S. pneumonie Overlapping,

unfortunately it is not possible to identify similarities between the two structures.

A docking is performed on the topoisomerase of smallpox by setting up a grid that contains almost entirely the entire protein with cocoa compounds (Tab 1).

The compounds appear to have affinities for two relatively circumscribed areas.

Tab 1 cocoa compound docking on S. pneumonie topoisomerase with grid containing all the protein (4000 atoms)

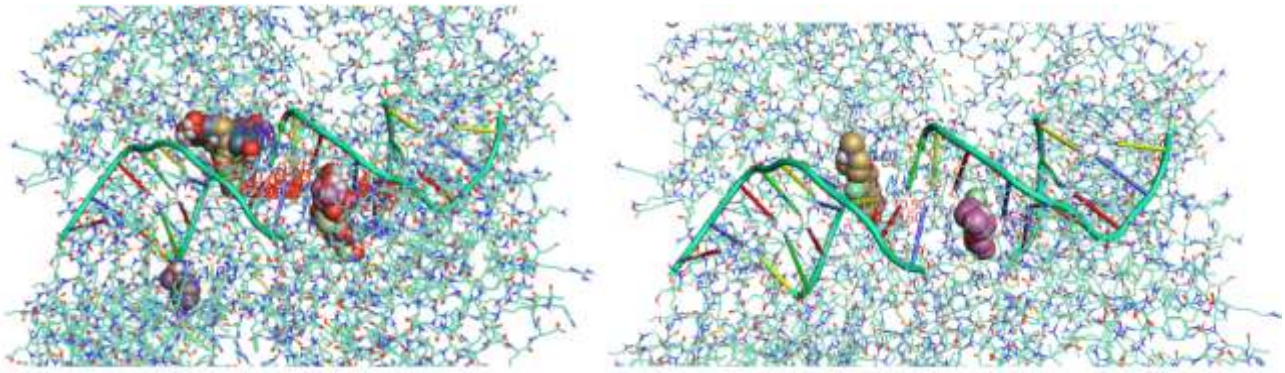
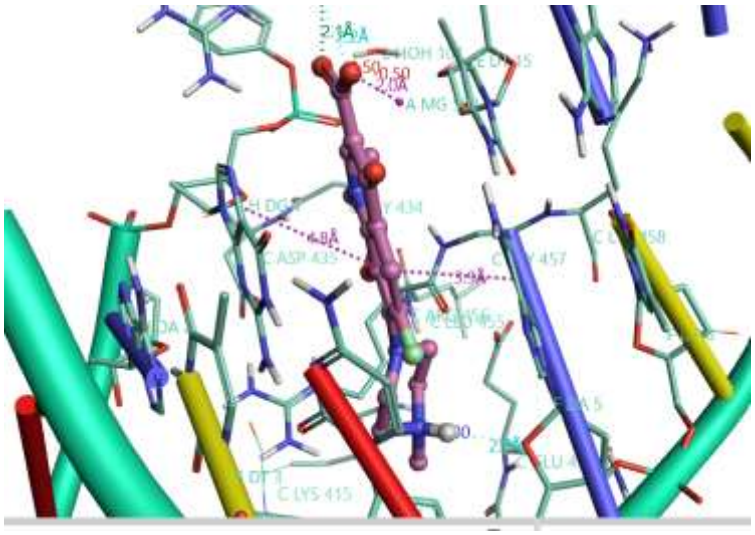


Figure 9 Representation of cocoa compounds (Table 1) bound to the protein simultaneously on the left and levofloxacin cocristallized to the same protein on the right

The compounds used as a probe are found to have a high affinity for two zones corresponding to the position of cocristallized levofloxacin (figure above)



Levofloxacin is inserted into S topoisomerase. pneumoniae results between two nucleotides in the chain:

the guanidine of the H chain (HDG1) and the adenine of the F chain (F DA 5), the nearest amino acid the arginine 456 asparagine 435 of the c chain and Tyr 118 of the b chain

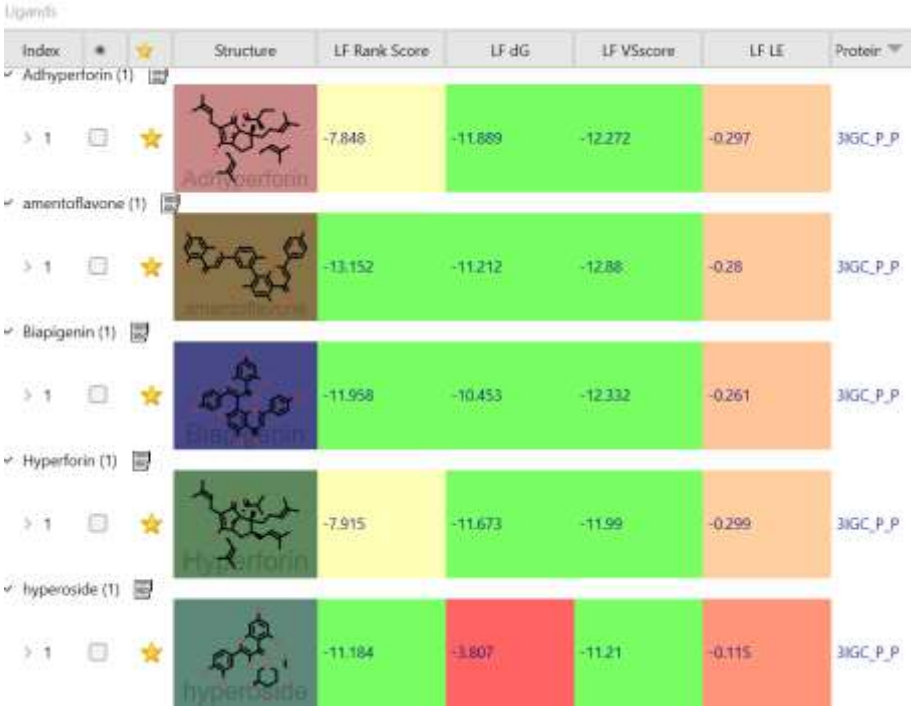
FIGURE 10 bound levofloxacin with relative bonding distances

By docking with levofloxacin on smallpox topoisomerase the residues with which the drug binds are different but the cocoa compounds used as a probe are found to have the same affinity for those two particular areas of the protein as it happens for S. pneumoniae

They have a high affinity for the two topoisomerases considered the Structures of catechin derivatives: catechin, epicatechin, gallocatechin (GC) epigallocatechin (EGC), catechin gallate, epicatechin gallate,- gallocatechin gallate, epigallocatechin gallate (EGCg), tea with scores similar to cocristallized . As confirmed by Yoscida et al. [5]

FIGURE 11 catechin gallate and quinine with relative scores and graphic representation of the bound epicatechin gallate (in pink)

A docking of The St. John's wort *compounds (Hypericum perforatum)* is performed on the topoisomerase of smallpox: smallpox infects the areas, and Hypericum oil with its principles could reach the infected subcutaneous areas. Hypericum compounds are active on topoisomerases as confirmed by Katherine et al. [6]



Screenshot n*

The 1,51-Å structure of the poxvirus L1 protein, a target of potent neutralizing antibodies

Hua-Poo Su , Scott C. Garman

1YPY Crystal structure of vaccine virus L1 protein

PDB DOI: 10.2210/pdb1YPY/pdb

Classification: VIRAL PROTEINS

Organism(s): vaccine virus

Expression system: Escherichia coli

Mutation(s): No (from Protein Data Bank)

The L1 structure reveals a hydrophobic cavity located adjacent to its terminus N. The cavity would be able to shield the myrist part, which is essential for the assembly of the virion.

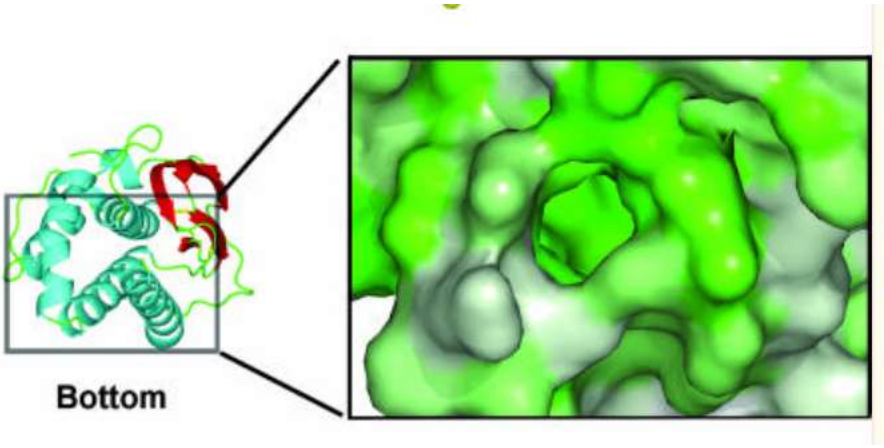


FIGURE 12

The cavity is coated with the 16 hydrophobic amino acids: Ile-7, Thr-10, Val-11, Leu-14, Ala-72, Thr-75, Tyr-76, Leu-79, Val-87, Met-90 , Phe-91, Val-104, Phe-108, Leu-163, Leu-166 and Ala-170

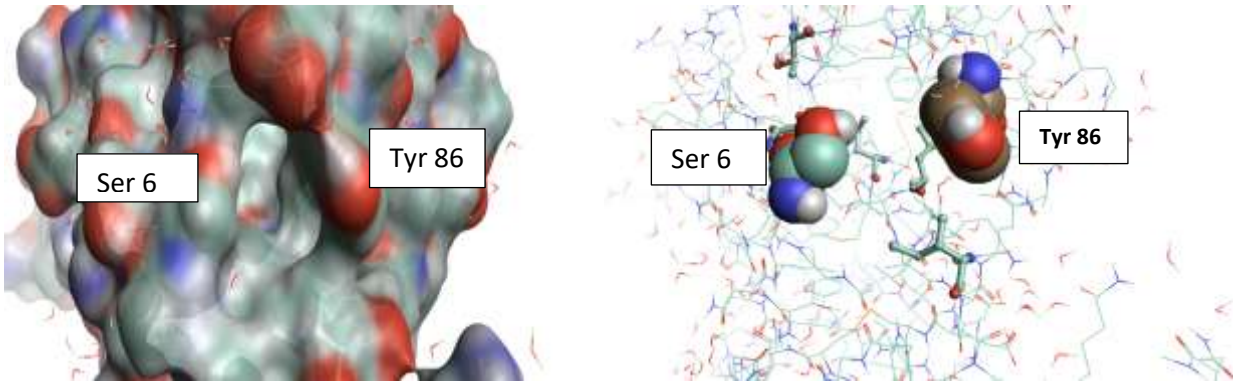


FIGURE 13

Ser 6 and tyr 86 are at the entrance of the hydrophobic bond pocket from which the myristoylated residue essential for the assembly of the viral membrane would enter and exit (ref 4)

Screening covalent docking with Flare (Cresset) using the tyr 86 residue as a residue to which compounds with nitrile residues would covalently bind.

the ligands to be attached must be designed to include an appropriate bond-forming functional group, also called a covalent head (Michael acceptors, nitriles, alkykinamides, alpha alo ketones, etc.). . These are electrophilic groups with low chemical reactivity that, after binding in a

non-covalent way to the target protein, are positioned near a specific nucleophilic residue (cis, lis, ser, tyr) in the active site to which they react quickly forming a bond

The scores of the compounds in tab 2 are calculated: nitrionic natural compounds

FIGURE 14

Menisdaurin scored the highest

Covalent docking is performed on drugs containing nitrile compounds (list in Table 3)

Vilazodone is the drug with the highest scores, binding covalently to serine at the entrance of the pocket and occupying it.

FIGURE 15

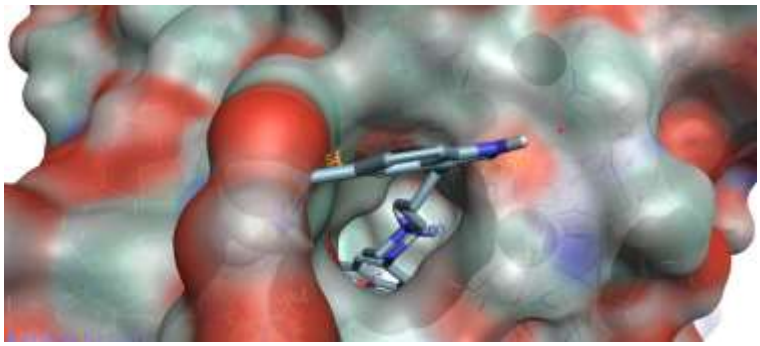


FIGURE 16

Vilazodone

From Wikipedia, the free encyclopedia

Vilazodone, sold among others under the brand name Viibryd, is a drug used to treat major depressive disorder. [1] While being studied for generalized anxiety disorder, that research had stopped starting in 2017. [3] It is taken orally. [1]

According to neratinib

FIG 17

Third cilomilast

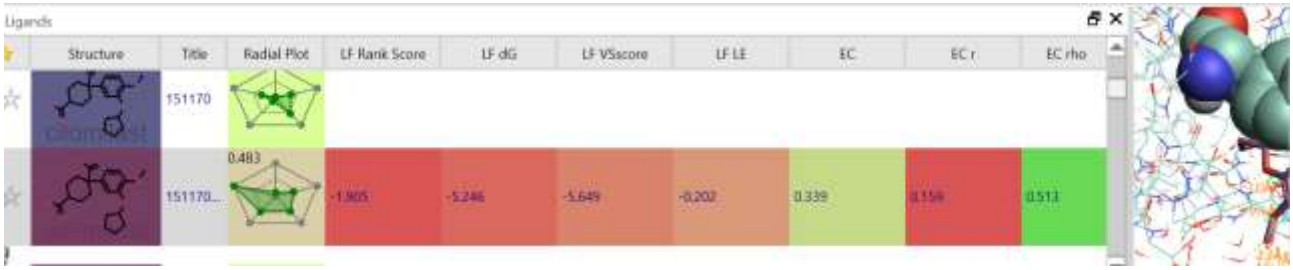


FIGURE 18

Cilomilast (Ariflo, SB-207,499) is a drug developed for the treatment of respiratory disorders such as asthma and chronic obstructive pulmonary disease (COPD). It is active orally and acts as a selective inhibitor of phosphodiesterase-4. After four clinical trials, the drug proved effective in treating COPD, however it was never marketed due to a poor side effect profile.

(from pubchem)

Covalent docking with natural compounds contained in hops, ivy (Tabs 4 and 5 list of individual compounds):

Falcarinone binds covalently to tyrosine, occupying the hydrophobic pocket with its tail.

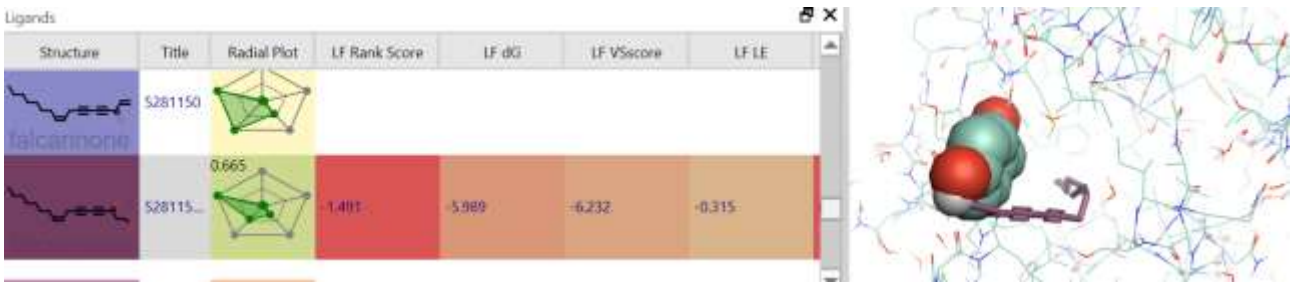


FIGURE 19

FIGURE 20

Rosmarinic acid and vitamin would bind to each other covalently, the vit C with a high electrostatic complementarity (EC rho: Spearmans rho rank)

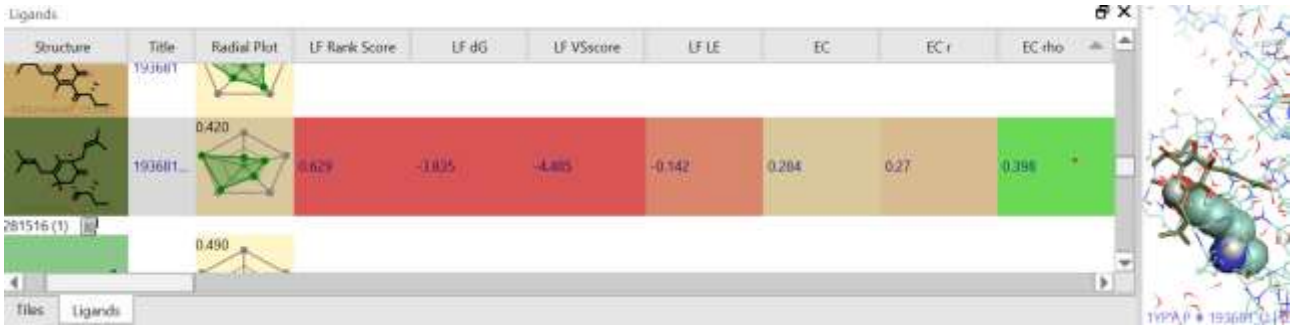


FIGURE 21

Adhumulone

By performing a covalent docking, with the cocoa compounds (see table 1 at the bottom) using serine 6 as a hook residue, slightly higher scores are obtained

FIGURE 22

Clovamide and chlorogenic cocoa acid, high score serine in low tyrosine.

CONCLUSIONS

A computational comparison was performed with Flare Cresset with docking techniques (ligand-based, structure-based, and covalent docking) on natural compounds and synthesis on some poxvirus targets.

Ligand based is used either cidifovir as a reference, as it is the only really effective drug in the clinic against systemic coxvirus infections.

Two unusual targets were used: the L1 protein of the poxvirus envelope that contains a hydrophobic pocket indispensable for viral replication, and smallpoisomerase topoisomerase.

The analyzed compounds derived from plant matrices and synthetic antiviral compounds were downloaded from a database[2]

The compounds of ivy, St. John's wort, cocoa, natural nitrile compounds and drugs were downloaded in sdf format from pubchem.

Some natural compounds have scores comparable to the reference drugs and there are some areas present on proteins with the same function in this case topoisomerases that have an affinity similar to the same molecules, despite the protein residues with which the ligands interact are totally different.

Curiosity, vitamin C would bind covalently according to the scores (not very high) with the residues of serine and tyrosine with a high electrostatic affinity at the mouth of the binding pocket of the L1 protein of smallpox, blocking its functionality, essential for the replication of the virion

CONFLICTS OF INTEREST

There are no conflicts of interest

FINANCING

The study is self-funded

THANKS

Thanks to **Cresset** for the academic license of **Flare**

REFERENCES:

1. **PHARMACEUTICAL MICROBIOLOGY**
   N. Carlone, R. Pompei (EdiSES)
2. **http://www.selleckchem.com**
3. **A Crystal Structure Based Guide to the Design of Human Histidine Triad Nucleotide Binding Protein 1 (hHint1) Activated ProTides**
   Kimberly M. Maize, Rachit Shah, Alex Strom, Sidath Kumarapperuma, Andrew Zhou, Carston R. Wagner, and Barry C. Finzel
4. **The 1,51-Å structure of the poxvirus L1 protein, a target of potent neutralizing antibodies**
   Hua-Poo Su , Scott C. Garman
5. **Inhibitory effects of catechin derivatives on mammalian DNA polymerase and topoisomerase activities and on the development of mouse unicellular zygote**
   Naoko Yoshida
6. **Catalytic inhibition of human DNA topoisomerase IIα by hypericin, a naphthodianthrone from St. John's wort (Hypericum perforatum)**

Katherine A Peebles

7. food chemistry

Cabras Martellli (PICCIN)

8. PROTEIN DATA BANK

https://www.rcsb.org/

i) 3RAE Quinolone(Levofloxacin)-DNA cleavage complex of type IV topoisomerase from S. pneumoniae
ii) 3IGC Smallpox virus topoisomerase-DNA transition state

CALCULATOR

Table 1 cocoa compounds[7] (*Theobroma cacao* L. , 1753)

(-)-Epicatechin	Epigallocatechin
(-)-Epigallocatechin gallate	Gallocatechin
+catechin	histamine_
1-Phenylethylamine_	luteolin
amandamideStructure2D	Naringenin_
apigenina	Oxalic acid
Caffeic acid	safrolo
caffeina	teofilline
-Catechin gallate_	tetraidrobetacarbolina
chinina	Theobromine
Chlorogenic acid	tryptamine
clovamide	Tyramine_
Epicatechin gallate	Vanillic acid

Table 2: Natural compounds with nitrile groups

5-methy tio esano solfossido nenitrile	epitionitrile	prunasin
5-methylthiopentanenitrile	epurpurins	rhodiocyanoside
acacipetalin	halimedin	ricinidine
Ambiguinine	heterodendrin	Saframycin
Bauhinin	Indoleacetonitrile	sarmentosin
b-cyanoglutamic acid	lithospermoside	simmonsin
borrelidin	lotaustralin	Sutherlandin
BURSATELLIN	malloapeltine	toyocamycin
calyculin Structure2D	mandellonitrile	valesamine n oxide
cianuro idrossibutene	menisdaurin	
Coclauril	multifidin	
Cyanopuupehenol	Mycalisine	
dnacin a1	Pre-kinamycin	

Table 3 : Drugs containing nitrile groups

alogliptin	cyanoguanidine	gallopamil	luliconazole	ravuconazole	Vildagliptin
anastrazole	dapivirine	IDX899 fosdevirine	milrinone	rilpivirine	zaleplon
Bicalutamide	diphenoxylate	isavuconazole	MIV-150	RS-8359	
BMS-191095	donitriptan	KR-31378	neratinib	RU-58841	
BMS-214662	EKB-569	kw-5092	nivaldipine	saxagliptin	
bosutinib	entacapone	L-778123	NO-1886 Ibrolipim	Strontium ranelate	
CC-5079	epanolol	lanoconazole	NVP-DPP-728	tanaproget	
CHS-828	escitalopram	lersivirine	olprinone	terbogrel	
cilomilast	etravirine	letrozole	pericyazine	trilostane	
cimetidine	fadrozole	levocabastine	PF-0998425	TYB-2285 Acreozast	
cromacalim	Febuxostat	levosimendan	pinacidil	verapamil	
cromakalim	finrozole	lodoxamide ethyl	piritramide	vidagliptin	
cyamemazine	FYX-051 Topiroxostat	lodoxamide	progesterone	vilazodone	

Tab 4: Hop compounds (*Humulus lupulus* L. , 1753)

- Adhumulone_193681
- alpha-Farnesene_CID_5281516
- beta-Caryophyllene_5281515
- Cohumulone_196915
- Humulinic acidCID_101739
- Humulone_442911
- Isoadhumulone_117231
- iso-cohumulone_25200613
- LupuloneCID_68051
- Myrcene_31253
- trans-Isohumulone_11175902
- Xanthohumol 639665

Table 5: Ivy compounds (*Hedera helix* L. , 1753)

2d ederacodide C	falcarinone	scopolin
2D Structure alpha ederin	hederagenin	stigmasterol
beta carotene	hidroxycinnamic acid	Structure2D_helixoside
beta sitosterol	kaempferol	vitamin c
caffeic acid	oleanolic acid	vitamin E
campesterol	oleic acid	
chlorogenic acid	palmitoleic acid	
cholesterol	panaxydol	
cis vaccenic acid	petroselinic acid	
emetin	quercetin	
falcarinol	rosmarinic acid	

Screenshot legend

1. Ligand-based

List of values of the chemical and physical properties of compounds: the molecules compared with cidofovir are listed, (if the box of individual values is colored green indicates a favorable score, if unfavorable red), whose name is written in the column "file name" with

- SIM = analogy score (values between zero and one, 1 = identity), radial graph = linear combination of several factors, scores close to one indicate a certain ease for the molecule to reach the receptor.
- # atom = number of atoms of the molecule
- 2D SIM = 2D analogy
- Slog P = octanol/water solubility coefficient, gives an indication of the hydrophilicity of the molecule
- TPSA = topological surface
- Flexibility indicates the possibility of the molecule to have more conformations

- Rol 5 = indicates how many of Lipinski's rules have been violated
- Radial plot, indicates a linear combination of different values, the dot if closer to the centeror has a better result.

2. Structure-based

SCREENSHOTS

Screenshot 1

The document was automatically translated by Microsoft Translator and written with Microsoft Word with an academic license

ADDITIONAL FIGURES

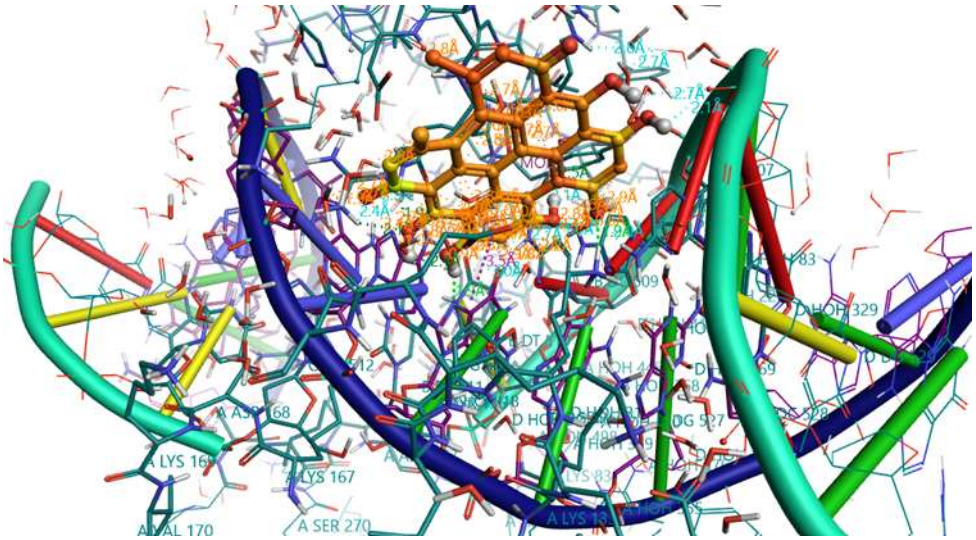


FIGURA 1 Supplemental Hypericin bound to smallpoisomerase