

Article



## In Silico Mining of Terpenes from Red-Sea Invertebrates for SARS-CoV-2 Main Protease (M<sup>pro</sup>) Inhibitors

Mahmoud A. A. Ibrahim <sup>1,\*</sup>, Alaa H. M. Abdelrahman <sup>1</sup>, Tarik A. Mohamed <sup>2</sup>, Mohamed A. M. Atia <sup>3</sup>, Montaser A. M. Al-Hammady <sup>4</sup>, Khlood A. A. Abdeljawaad <sup>1</sup>, Eman M. Elkady <sup>4</sup>, Mahmoud F. Moustafa <sup>5,6</sup>, Faris Alrumaihi <sup>7</sup>, Khaled S. Allemailem <sup>7</sup>, Hesham R. El-Seedi <sup>8,9,10,\*</sup>, Paul W. Paré <sup>11</sup>, Thomas Efferth <sup>12</sup> and Mohamed-Elamir F. Hegazy <sup>2,12,\*</sup>

- <sup>1</sup> Computational Chemistry Laboratory, Chemistry Department, Faculty of Science, Minia University, Minia 61519, Egypt; a.abdelrahman@compchem.net (A.H.M.A.); kh.abdeljawaad@compchem.net (K.A.A.A.)
- <sup>2</sup> Chemistry of Medicinal Plants Department, National Research Centre, 33 El-Bohouth St., Dokki, Giza 12622, Egypt; tarik.nrc83@yahoo.com
- <sup>3</sup> Molecular Genetics and Genome Mapping Laboratory, Genome Mapping Department, Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), Giza 12619, Egypt; matia@ageri.sci.eg
- <sup>4</sup> National Institute of Oceanography & Fisheries, NIOF, 11516 Cairo, Egypt; coralreef\_noif1@yahoo.com (M.A.M.A-H); emelkady@yahoo.com (E.M.E.).
- <sup>5</sup> Department of Biology, College of Science, King Khalid University, Abha 9004, Saudi Arabia; hamdony@yahoo.com
- <sup>6</sup> Department of Botany & Microbiology, Faculty of Science, South Valley University, Qena, Egypt
- <sup>7</sup> Department of Medical Laboratories, College of Applied Medical Sciences, Qassim University, Buraydah, Saudi Arabia; f\_alrumaihi@qu.edu.sa (F.A.); k.allemailem@qu.edu.sa (K.S.A.)
- <sup>8</sup> Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, S-106 91 Stockholm, Sweden; hesham.elseedi@su.se
- <sup>9</sup> Department of Chemistry, Faculty of Science, El-Menoufia University, Shebin El-Kom 32512, Egypt
- <sup>10</sup> International Research Center for Food Nutrition and Safety, Jiangsu University, Zhenjiang 212013, China
- <sup>11</sup> Department of Chemistry & Biochemistry Texas Tech University, Lubbock, TX 79409 USA; paul.pare@ttu.edu
- <sup>12</sup> Department of Pharmaceutical Biology, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University, Staudinger Weg 5, 55128 Mainz, Germany; efferth@uni-mainz.de
- \* Correspondence: m.ibrahim@compchem.net (M.A.A.I.); hesham.elseedi@su.se (H.R.E.-S.); elamir77@live.com (M.E.F.H.); Tel.: +2-010-241-61-444 (M.A.A.I.); +46-73-566-8234 (H.R.E.-S.); +2-033-371-635 (M.E.F.H.)



**Figure S1.** 2D representations of interactions of lopinavir and the top 27 potent marine natural products (MNPs) with the proximal amino acid residues of SARS-CoV-2 main protease (M<sup>pro</sup>).

















224

Figure S1. Continued.











197



Figure S1. Continued.



Figure S1. Continued.



Figure S1. Continued.



Figure S2. 3D representations of predicted binding modes of (i) 190, (ii) 178, (iii) 226 and (iv) lopinavir towards SARS-CoV-2 main protease (M<sup>pro</sup>).



**Figure S3.** 3D representations of binding modes of (i) erylosides B (226)- and (ii) lopinavir-M<sup>pro</sup> complexes according to an average structure over a 100 ns MD simulation.



**Figure S4.** A genome-wide Reactome hierarchy map of the pathways influenced by the top 20 gene targets in response to erylosides B (**226**) in term of SARS-CoV-2 infection. Reactome pathways are arranged in a hierarchy. Each step away from the Center represents the next level lower in the pathway hierarchy. The color code denotes the over-representation of that pathway in the input dataset. Light grey signifies pathways that are not significantly over-represented.

Table S1.	Evaluated	docking score	(in kcal/mol) for	lopinavir and	all investigated	l marine natura	l products (MNPs	) against SARS-
CoV-2 ma	in protease	(M <sup>pro</sup> ).						

No.	Plant Compound Name Source		Chemical Structure	Docking Score (kcal/mol)
1	Lopinavir			-9.8
2	Depresosterol <b>(190)</b>	L. depressum	HOWING OF THE REPORT OF	-12.3
3	3β-25-Dihydroxy-4-methyl- 5α,8α-epidioxy-2- ketoergost-9-ene <b>(178)</b>	Sinularia candidula	HO HO UNIT O O O O O O O O O O O O O O O O O O O	-12.2
4	Erylosides B <b>(226)</b>	E. lendenfeldi		-12.1
5	Sipholenol H <b>(157)</b>	S. siphonella	HO HO OH	-12.0













39	Sipholenol G (154)	S. siphonella	HO <sub>IIIII</sub> HO <sub>IIIII</sub> HO <sub>IIIII</sub>	-9.5
40	Sigmosceptrellin B methyl ester <b>(139)</b>	D. erythraeanus		-9.5
41	3β-Hydroxycholest-5-en-7- one <b>(205)</b>	A. dichotoma	HO	-9.5
42	Nuapapuin A methyl ester (132)	Diacarnus erythraeanus		-9.4
43	Sipholenol A (150)	S. siphonella	HOMMIN HOMMIN HOMMIN	-9.4
44	Siphonellinol E <b>(170)</b>	S. siphonella		-9.4

























111	Erylosides B <b>(226)</b>	E. lendenfeldi	-8.3
112	Sipholenoside A <b>(159)</b>	S. siphonella	-8.3
113	Sipholenoside B (160)	S. siphonella	-8.3
114	Thunbergol <b>(57)</b>	L. pauciflorum	 -8.3
115	14(15)-Epoxyxeniaphyllene (77)	X. lilielae	-8.2
116	Bilosespens A <b>(131)</b>	D. cinerea	-8.2







134	Eryloside A <b>(197)</b>	Genus Erylus		-7.9
135	3-Deoxy-20- acetylpresinularolide B <b>(64)</b>	L. crassum		-7.9
136	Sarcophine <b>(30)</b>	S. glaucum		-7.9
137	Xenicin <b>(73)</b>	Xenia macrosoiculata		-7.8
138	Sarcophytol M <b>(65)</b>	Litophyton arboreum	HO	-7.8
139	Smenotronic acid <b>(26)</b>	Smenospongia sp.	O O O O O O O O O O O O O O O O O O O	-7.8

140	Zaatirin <b>(109)</b>	C. erecta		-7.8
141	Cholest-5-en-3β-yl-formate <b>(204)</b>	A. dichotoma		-7.8
142	Pachycladin B <b>(85)</b>	C. pachyclados	HONING UNIT OF	-7.8
143	Sclerophytin A <b>(93)</b>	C. pachyclados	HOILING	-7.8
144	epi-Sigmosceptrellin B <b>(147)</b>	D. erythraeanus		-7.8
145	Pachycladin D <b>(91)</b>	C. pachyclados	OH IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	-7.7































		Autodock	MM/GBSA
No.	Compound Name	<b>Binding Score</b>	Binding Energy
		(kcal/mol)	(kcal/mol)
1	Lopinavir	-9.8	-39.4
2	Erylosides B (226)	-12.1	-50.8
3	Erylosides K (224)	-11.1	-50.4
4	3β-Hexadecanoylcholest-5-en-7-one <b>(202)</b>	-10.0	-49.3
5	Eryloside A (197)	-10.7	-47.8
6	SipholenolA-4-O-3',4'-dichlorobenzoate (151)	-10.5	-42.6
7	Sipholenone E (163)	-9.9	-40.9
8	Sipholenone A (175)	-11.0	-37.7
9	3β-25-Dihydroxy-4-methyl-5α,8α-epidioxy-2- ketoergost-9-ene <b>(178)</b>	-12.2	-36.1
10	Sipholenol D (176)	-11.0	-35.9
11	(22R,24E,28E)-5β,6β-Ероху-22,28-oxido-24-	-11.4	-35.2
	methyl-5 $\alpha$ cholestan-3 $\beta$ ,25,28-triol <b>(191)</b>		
12	Sipholenol I (174)	-11.8	-35.1
13	Sipholenol H (157)	-12.0	-34.7
14	Sipholenone D (155)	-10.7	-33.4
15	Siphonellinol C (172)	-11.3	-33.0
16	Neviotine B <b>(158)</b>	-10.9	-32.2
17	Tasnemoxide A (144)	-11.4	-32.2
18	Siphonellinol-C-23-hydroperoxide (171)	-11.2	-31.8
19	Brassicasterol (222)	-10.1	-31.4
20	Cholest-5-en-3β,7β-diol <b>(206)</b>	-10.3	-31.1
21	Dahabinone A (162)	-11.9	-30.2
22	Clionasterol (219)	-10.3	-29.9
23	Depresosterol (190)	-12.3	-28.1
24	Campesterol (221)	-10.3	-27.8
25	Cholesterol (184)	-10.3	-27.4
26	Lobophytosterol (188)	-11.5	-27.4
27	24-Methylcholestane-5-en-3β,25-diol (187)	-10.6	-27.3
28	Stigmasterol (220)	-10.5	-26.4

**Table S2.** Computed Autodock and MM/GBSA binding energies (in kcal/mol) for the top 27 potent marine natural products (MNPs) against SARS-CoV-2 main protease (M<sup>pro</sup>) over 250 ps implicit solvent MD simulations<sup>a</sup>.

<sup>a</sup>Data sorted according to the calculated MM/GBSA binding energies.

	10,	1 0 ,	. ,	
Name	BetweennessCentrality	ClosenessCentrality	Degree	Number Of Undirected Edges
VEGFA	0.16167593622379026	0.543859649122807	30	30
DRD2	0.09010641139487129	0.4946808510638298	27	27
STAT3	0.08946266709776997	0.51666666666666666	24	24
JUN	0.0981884572759623	0.510989010989011	23	23
ADRA2A	0.024166127699934013	0.45365853658536587	23	23
ADRA2C	0.017619738652461147	0.44285714285714284	22	22
F2	0.06340725344246699	0.469696969696969697	21	21
ADRA2B	0.016564892030468053	0.4407582938388625	21	21
SLC6A3	0.02851326445210035	0.44497607655502397	20	20
ADRA1A	0.008093604972074064	0.41150442477876104	19	19
ADRA1B	0.013961744615366147	0.4246575342465753	19	19
REN	0.060652471264093046	0.484375	19	19
PRKCA	0.06205750483643582	0.48186528497409326	18	18
DRD3	0.008419520597214505	0.43457943925233644	18	18
OPRK1	0.019575984856832866	0.42081447963800905	18	18
ADRA1D	0.005035012045256654	0.40611353711790393	18	18
EDNRB	0.040084345951539736	0.44497607655502397	18	18
IL2	0.0567559135447914	0.47692307692307695	17	17
HTR2A	0.006128365201270778	0.4096916299559471	17	17
FGF2	0.023107182275248352	0.484375	16	16
PRKCB	0.030044228440275655	0.469696969696969697	16	16
HTR2C	0.004745649753776972	0.40086206896551724	16	16
SLC6A2	0.012665535612403297	0.39914163090128757	16	16
ACHE	0.017816808762018056	0.4133333333333333333	16	16

Table S3. Network topological analysis for the predicted targets for erylosides B (226).

Pathway name	Entities found	Entities total	Interactors found	Interactors total	Entities ratio	Entities pValue	Entities FDR	Reactions found	Reactions total	Reactions ratio
Signaling by GPCR	15	1497	9	1666	0.1016	0.00157934	0.0728	45	445	0.03363822
GPCR										
downstream	14	1355	3	1157	0.092	5.72E-04	0.0353	31	260	0.01965379
signalling										
GPCR ligand	10		,	501	0.0451	1 105 07	1.405.04	14	105	0.01000440
binding	13	665	6	581	0.0451	1.10E-06	1.48E-04	14	185	0.01398443
Class A/1										
(Rhodopsin-like	13	475	6	552	0.03226	1.21E-07	2.69E-05	14	158	0.01194346
receptors)										
Amine ligand-	10	00	4	FO	0.0050	2 20E 12	2 27E 10	(	22	0.00166201
binding receptors	10	88	4	50	0.0059	3.39E-13	2.27E-10	6	22	0.00166301
G alpha (q)	0	202	2	4/1	0.0102	1 705 04	0.01/1	7	25	0.000(457
signalling events	8	283	3	461	0.0192	1.70E-04	0.0161	7	35	0.0026457
Platelet										
activation,	0	20(	2	((0)	0.0201	7.02E.04	0.0426	15	115	0.009(0202
signaling and	8	290	2	007	0.0201	7.93E-04	0.0436	15	115	0.00869302
aggregation										
Adrenoceptors	6	48	1	35	0.0032	2.97E-08	9.94E-06	3	7	5.29E-04
G alpha (z)	5	62	0	48	0.0042	4.02E-06	4.46E-04	6	13	9.83E-04
signalling events										
Interleukin-4 and										
Interleukin-13	5	216	1	153	0.0146	0.00110401	0.0563	21	47	0.0035528
signaling										
POU5F1 (OCT4),										
SOX2, NANOG										
activate genes	4	21	2	132	0.0014	3.29E-04	0.0273	4	16	0.00120946
related to										
proliferation										
Response to										
elevated platelet	4	144	0	104	0.0097	0.00206465	0.0802	4	14	0.00105828
cytosolic Ca2+										
Transcriptional										
regulation of	4	45	2	231	0.003	0 00300441	0 1051	6	35	0 0026457
pluripotent stem	I	10	2	201	0.000	0.00000111	0.1001	0	00	0.0020107
cells										

**Table S4**. Top 20 most relevant pathways for erylosides B (**226**) targets resulted from Pathway Enrichment Analysis (PEA). PEA was performed using a binomial test and p-values were False Discovery Rate (FDR)-corrected for multiple testing.

Platelet Aggregation (Plug Formation)	4	53	1	315	0.0036	0.00778789	0.1999	4	27	0.00204097
Adrenaline										
signalling										
through Alpha-2	3	5	0	0	3.40E-04	2.05E-07	3.42E-05	1	1	7.56E-05
adrenergic										
receptor										
G alpha (12/13)				6			0.4000			
signalling events	3	87	0	63	0.0059	0.00402801	0.1329	3	15	0.00113387
VEGFR2										
mediated cell	3	31	0	132	0.0021	0.00556703	0.1725	7	12	9.07E-04
proliferation										
Dopamine	2	(	2	0	4.095.04	4 525 04	0.0225	2	2	2 27E 04
receptors	Z	0	3	9	4.08E-04	4.53E-04	0.0355	2	3	2.27 E-04
Depolymerisation										
of the Nuclear	2	23	0	8	0.0015	0.00216767	0.0802	2	6	4.54E-04
Lamina										
RUNX1 and										
FOXP3 control										
the development	2	17	2	39	0.0011	0.00637006	0.1771	6	20	0.00151183
of regulatory T										
lymphocytes										