2023
10(2), 537

Chemical Biology
LETTERS

In silico identification of small natural inhibitors against DNA methyl transferase 3-like protein by integrative molecular docking and molecular dynamics approach

Paratpar Sarkar[1], Niraj Niraj[2] and Vivek Srivastava[1*]

[1]Department of Chemistry & Biochemistry, Sharda School of Basic Sciences & Research, Sharda University, Greater Noida-201308, UP, India, [2]Dr. R.P.Centre, All India Institute of Medical Sciences, Delhi, India.

Supplementary Information file

Article cover image:

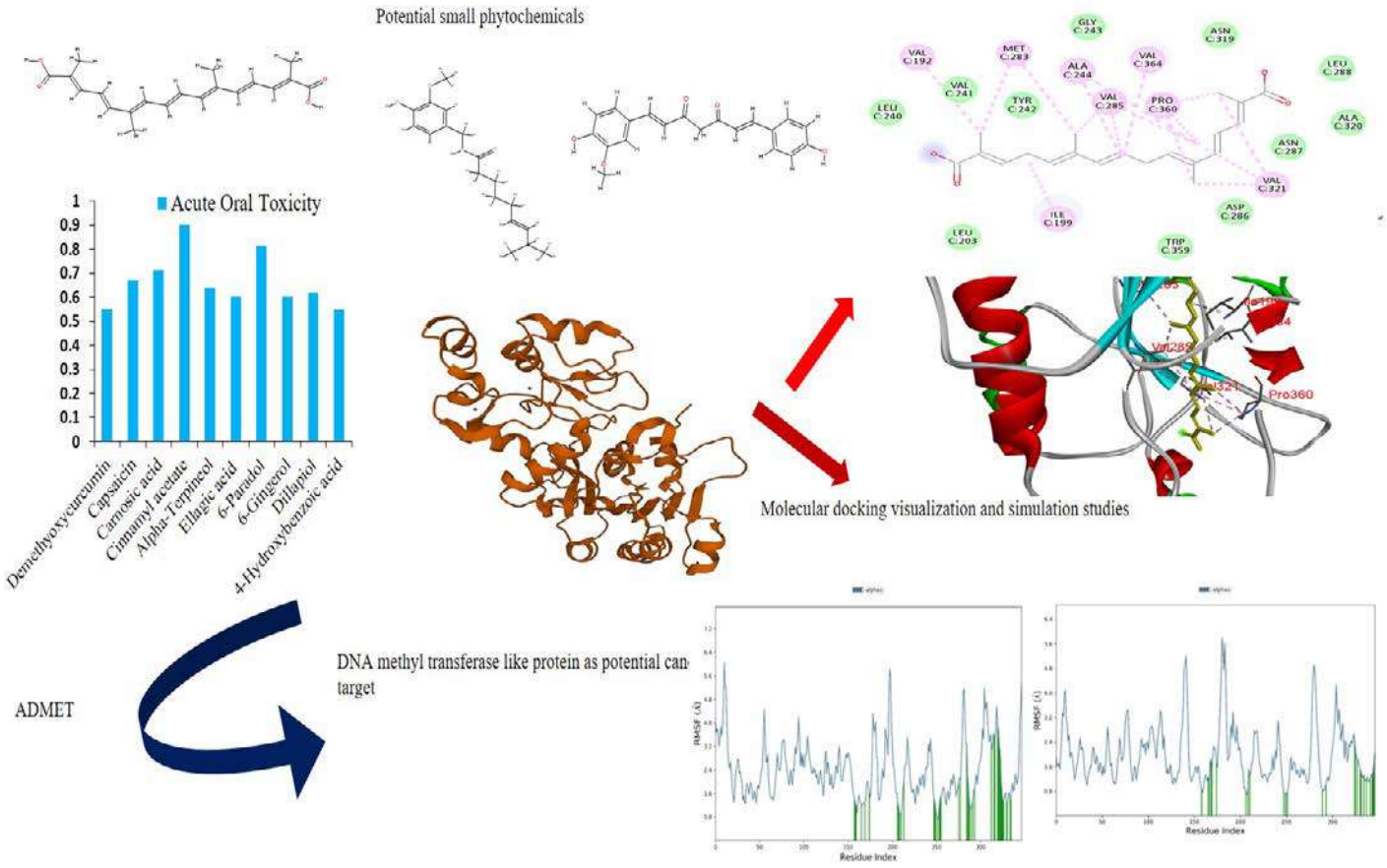
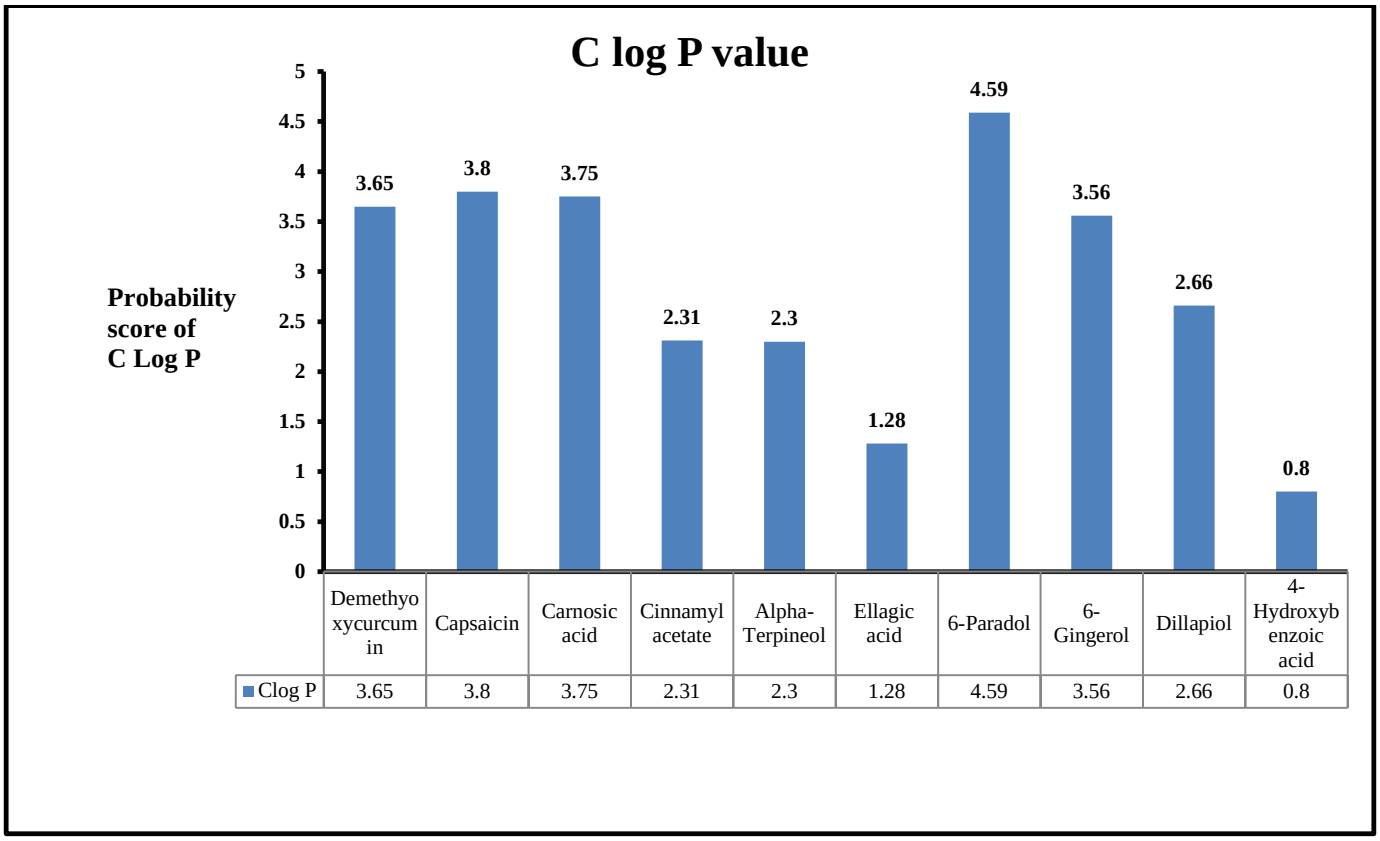


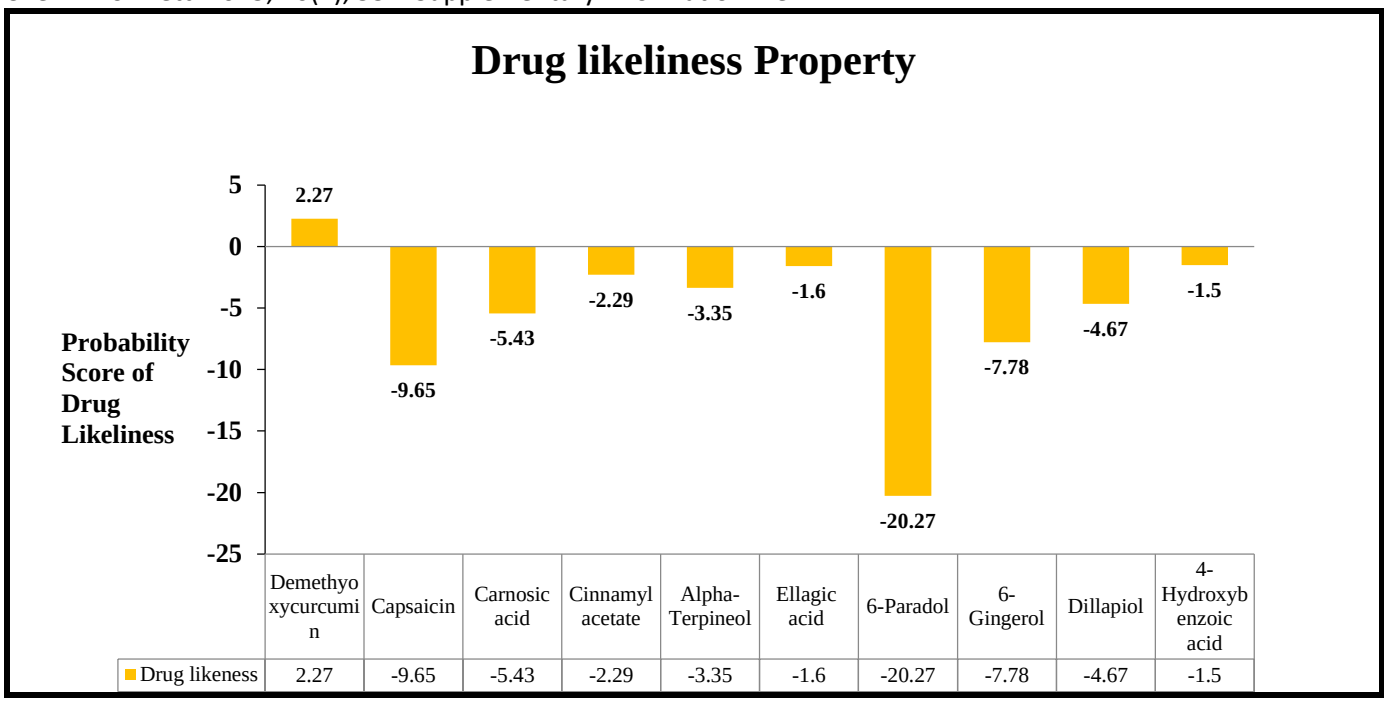
Table 1: Small phytochemicals from different plant species with their SMILE notation

Small phytochemicals	Comp (CID)	SMILE notation	Sources
Ferulic acid	445858	CC(C)(C)[Si](C)(C)OC1=C(C=C(C=C1)C=CC(=O)O[Si](C)(C)C(C)(C)C)OC	*Syzygium Aomaticum*
Crocetin	5281232	CC(=CC=CC=C(C)C=CC=C(C)C(=O)O)C=CC=C(C)C(=O)O	*Crocus sativus*
Cinnamic acid	444539	C1=CC=C(C=C1)C=CC(=O)O	*Cinnamon*
Eugenol	3314	COC1=C(C=CC(=C1)CC=C)O	*Cinnamon*
Cinnamaldehyde	637511	C1=CC=C(C=C1)C=CC=O	*Cinnamon*
Allicin	65036	C=CCSS(=O)CC=C	*Allium sativum*
Alpha tumerone	558173	CC1=CC=C(CC1)C(C)CC(=O)C=C(C)C	*Curcumin Longa*
Curcumin	10134351	CC1=C(C=C(C=C1)C=CC(=O)CC(=O)C=CC2=CC(=C(C=C2)C)[N+](=O)[O-])[N+](=O)[O-]	*Curcumin Longa*
Estragole	8815	COC1=CC=C(C=C1)CC=C	*Ocimum basilicum*
Shogaol	5281794	CCCCCC=CC(=O)CCC1=CC(=C(C=C1)O)OC	*Ginger officinale*
Demethoxy curcumin	5469424	COC1=C(C=CC(=C1)C=CC(=O)CC(=O)C=CC2=CC=C(C=C2)O)O	*Curcumin longa*
Capsaicin	1548943	CC(C)/C=C/CCCCC(=O)NCC1=CC(=C(C=C1)O)OC	*Piper nigrum*
Cinnamyl acetate	5282110	CC(=O)OC/C=C/C1=CC=CC=C1	*Curcumin Longa*
Alpha terpineol	17100	CC1=CCC(CC1)C(C)(C)O	*Elettaria cardamomum*
Ellagic acid	5281855	C1=C2C3=C(C(=C1O)O)OC(=O)C4=CC(=C(C(=C43)OC2=O)O)O	*Syzygium aromaticum*
6-Paradol	94378	CCCCCCCC(=O)CCC1=CC(=C(C=C1)O)OC	*Ginger officinale*
6-Gingerol	442793	CCCCCC(CC(=O)CCC1=CC(=C(C=C1)O)OC)O	*Ginger officinale*
Dillapiol(Di)	10231	COC1=C(C2=C(C=C1CC=C)OCO2)OC	*Foeniculum vulgare*
4-Hydroxybenzoic acid	135	C1=CC(=CC=C1C(=O)O)O	*Coriandrum sativum plant*

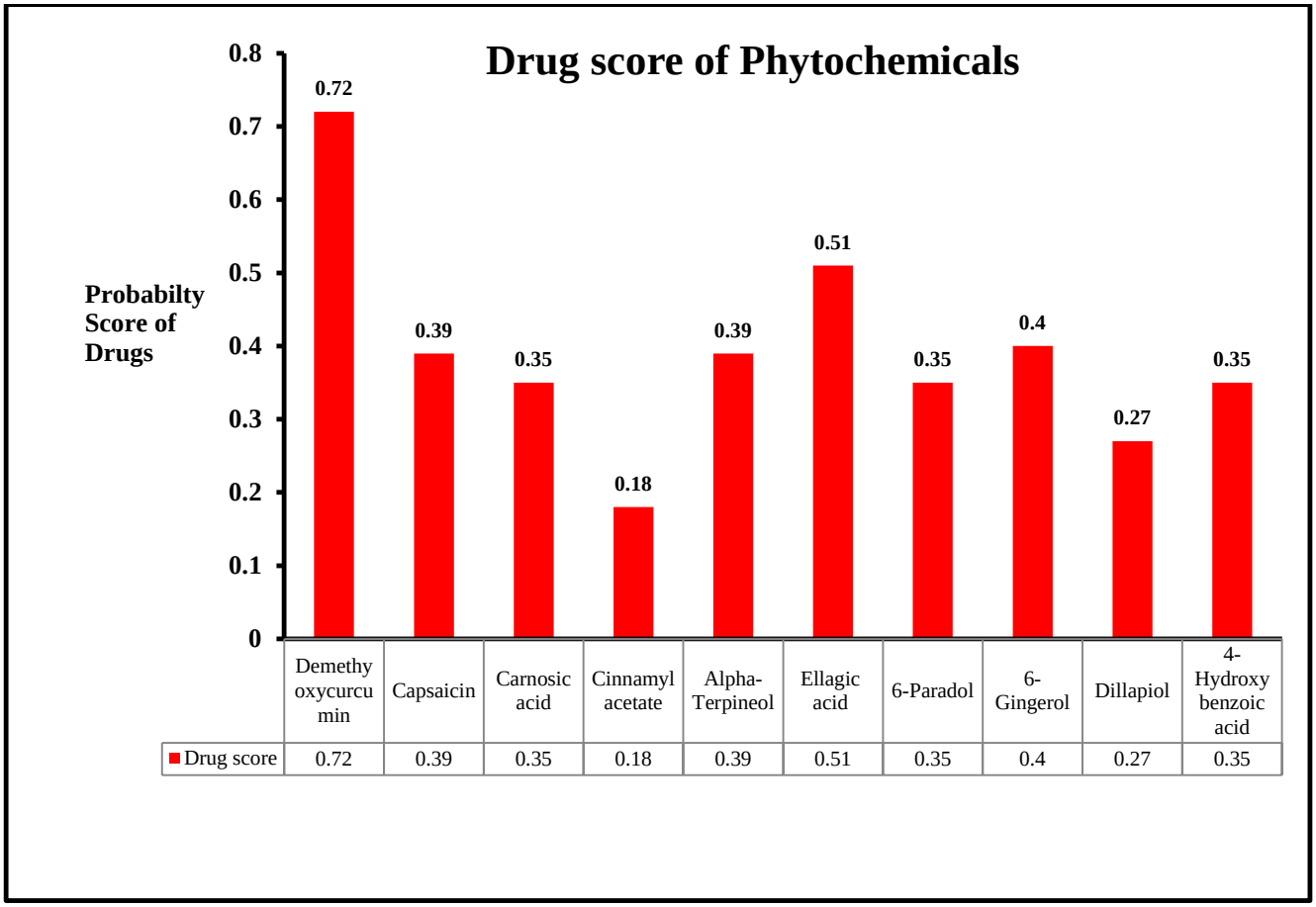
Carnosic acid	65126	CC(C)C1=C(C(=C2C(=C1)CCC3C2(CCCC3(C)C)C(=O)O)O)O	*Rosmarinus officinalis*



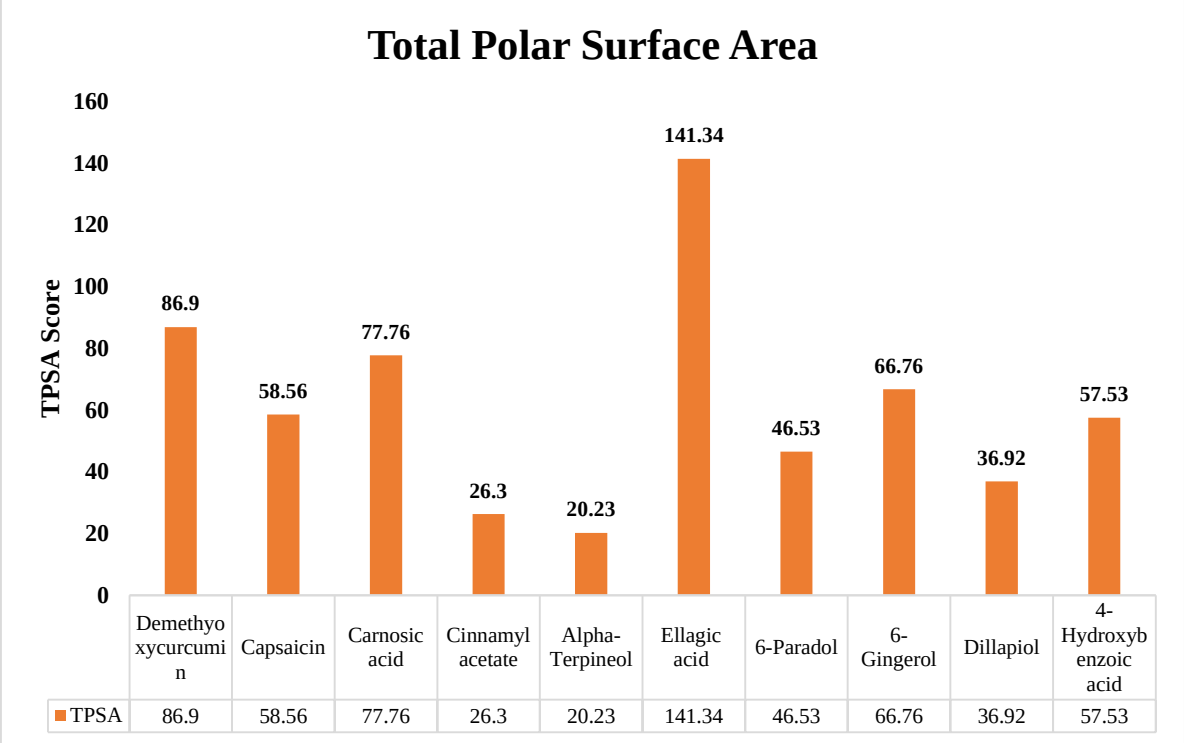
A


Drug likeliness Property
Probability Score of Drug Likeliness
5
0
-5
-10
-15
-20
-25
2.27
-9.65
-5.43
-2.29
-3.35
-1.6
-20.27
-7.78
-4.67
-1.5
Demethyoxycurcumin
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxybenzoic acid
Drug likeness
2.27
-9.65
-5.43
-2.29
-3.35
-1.6
-20.27
-7.78
-4.67
-1.5

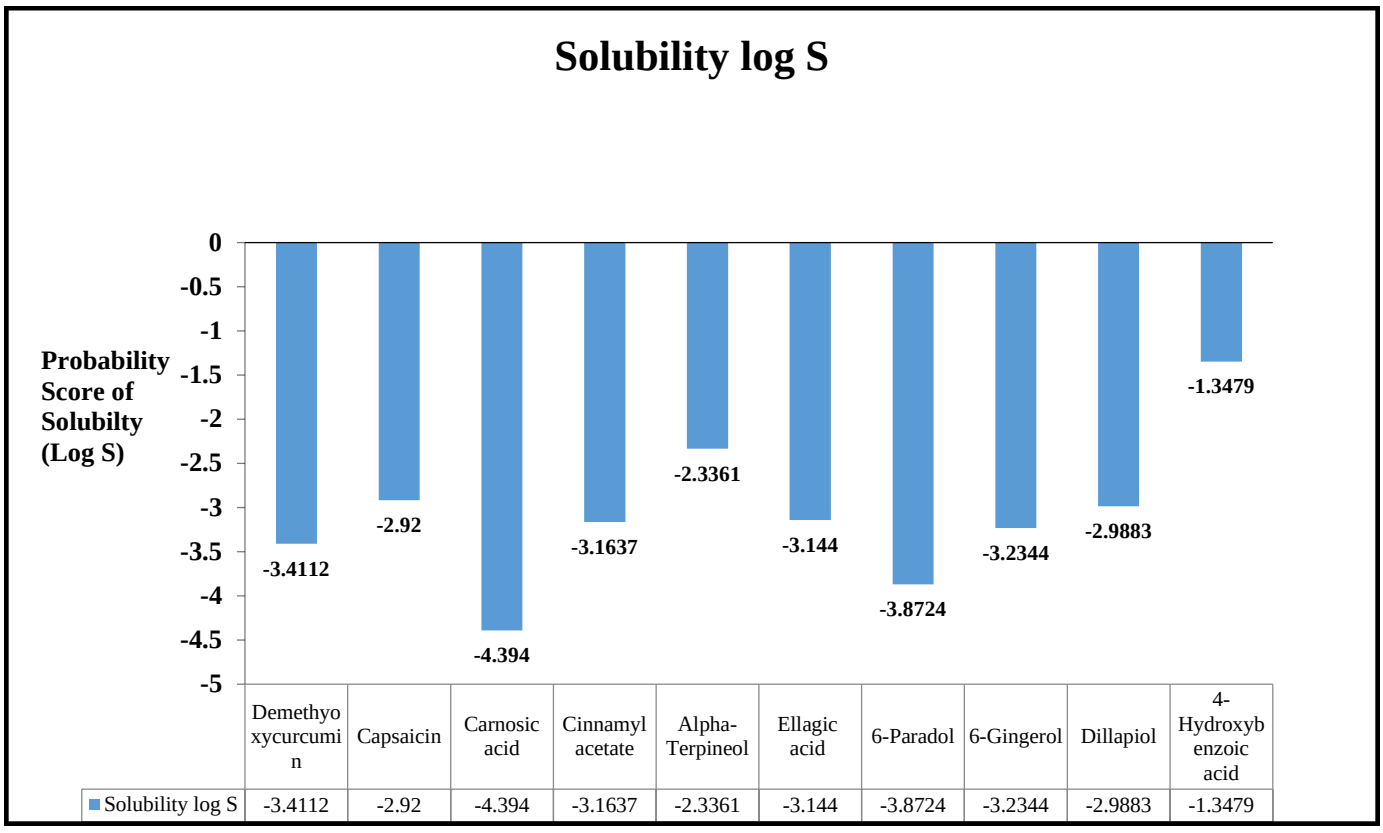
b


Drug score of Phytochemicals
Probabilty Score of Drugs
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.72
0.39
0.35
0.18
0.39
0.51
0.35
0.4
0.27
0.35
Demethy oxycurcumin
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxy benzoic acid
Drug score
0.72
0.39
0.35
0.18
0.39
0.51
0.35
0.4
0.27
0.35

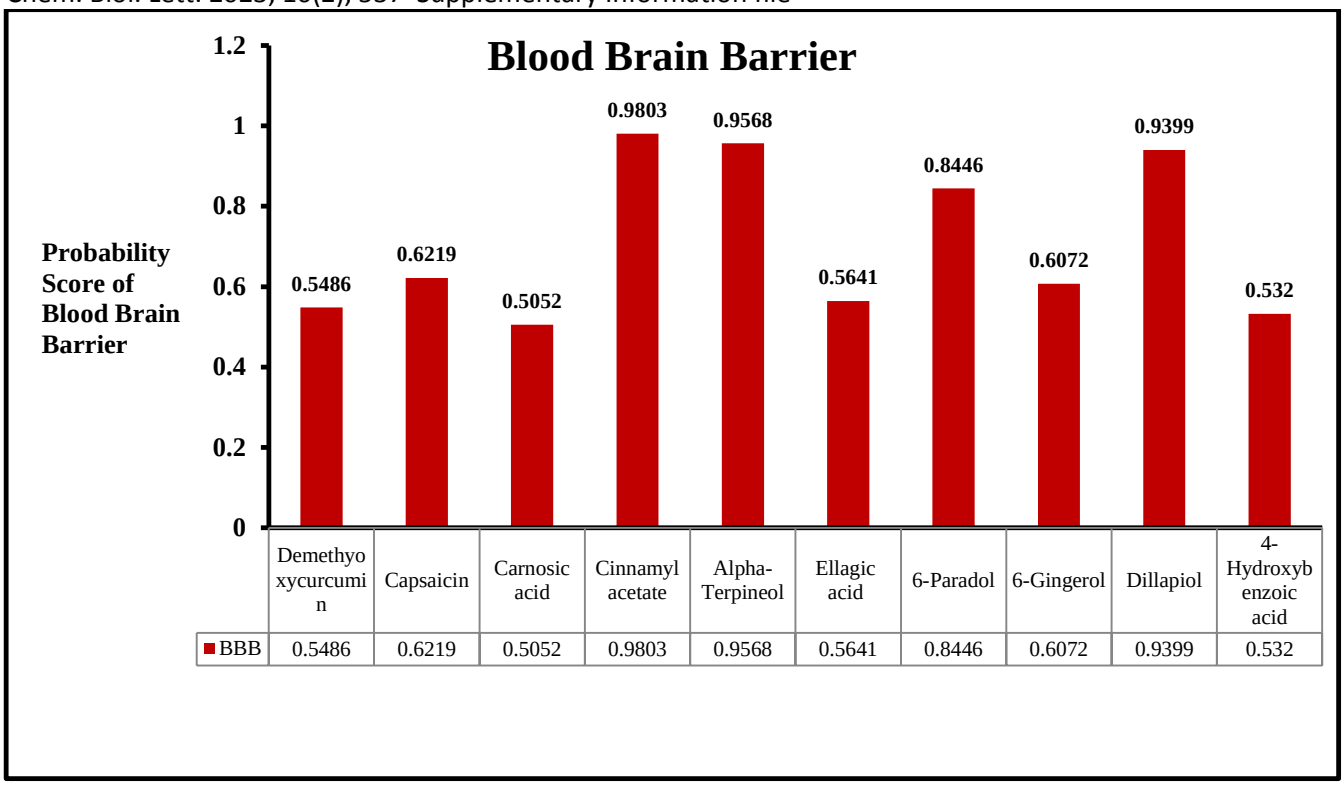
c


Total Polar Surface Area
TPSA Score
160
140
120
100
80
60
40
20
0
86.9
58.56
77.76
26.3
20.23
141.34
46.53
66.76
36.92
57.53
Demethyoxycurcumin
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxybenzoic acid
TPSA
86.9
58.56
77.76
26.3
20.23
141.34
46.53
66.76
36.92
57.53

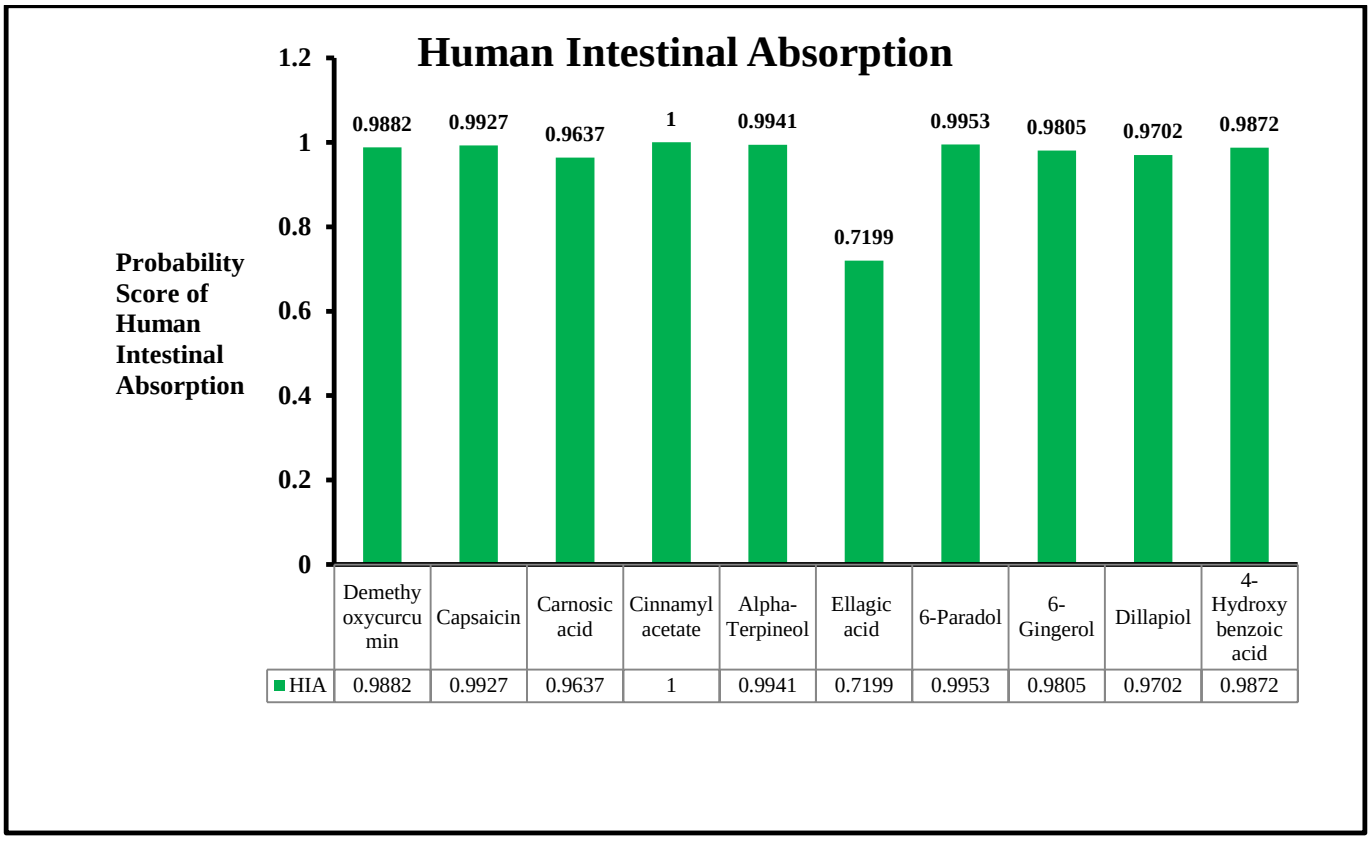
d


Solubility log S
Probability Score of Solubilty (Log S)
0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
-4
-4.5
-5
-3.4112
-2.92
-4.394
-3.1637
-2.3361
-3.144
-3.8724
-3.2344
-2.9883
-1.3479
Demethyoxycurcumin
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxybenzoic acid
Solubility log S
-3.4112
-2.92
-4.394
-3.1637
-2.3361
-3.144
-3.8724
-3.2344
-2.9883
-1.3479

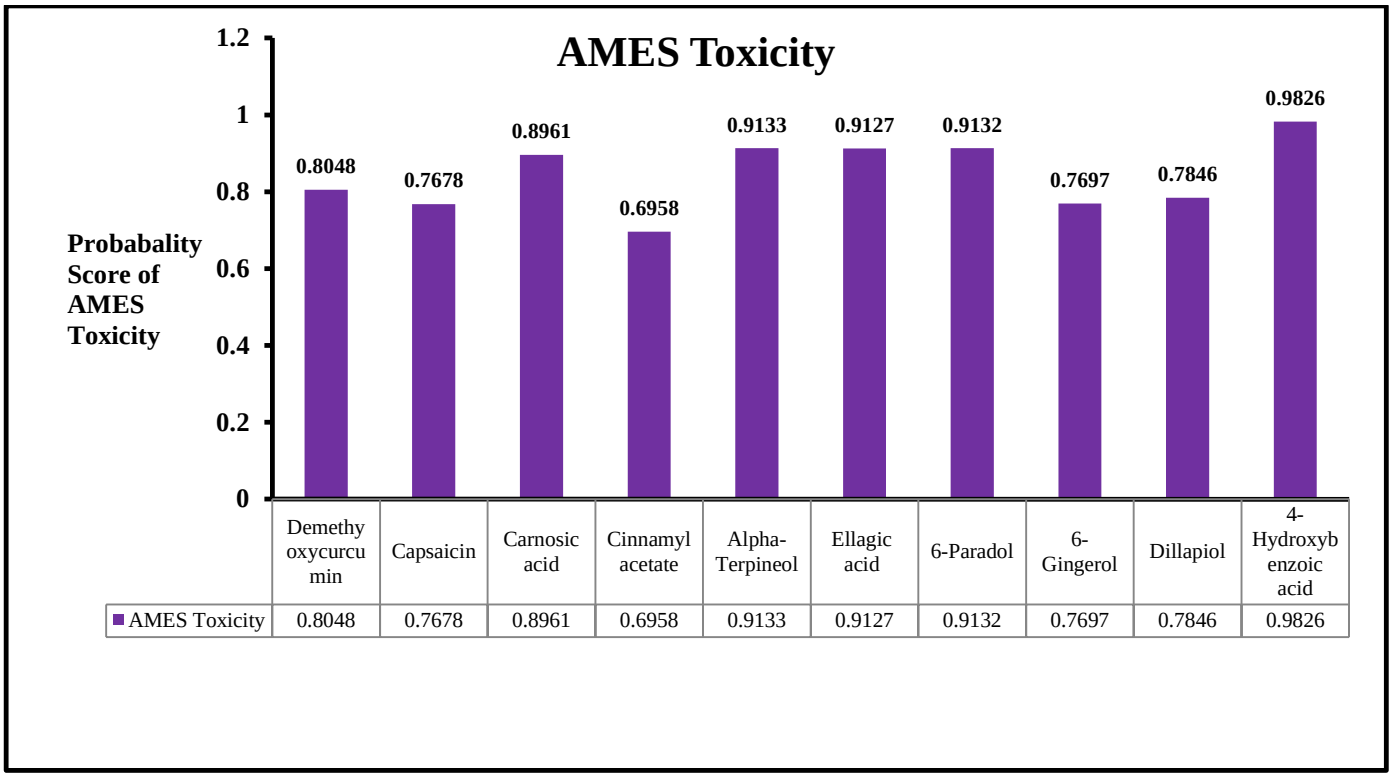
e


Blood Brain Barrier
Probability Score of Blood Brain Barrier
1.2
1
0.8
0.6
0.4
0.2
0
0.5486
0.6219
0.5052
0.9803
0.9568
0.5641
0.8446
0.6072
0.9399
0.532
Demethyoxycurcumin
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxybenzoic acid
BBB
0.5486
0.6219
0.5052
0.9803
0.9568
0.5641
0.8446
0.6072
0.9399
0.532

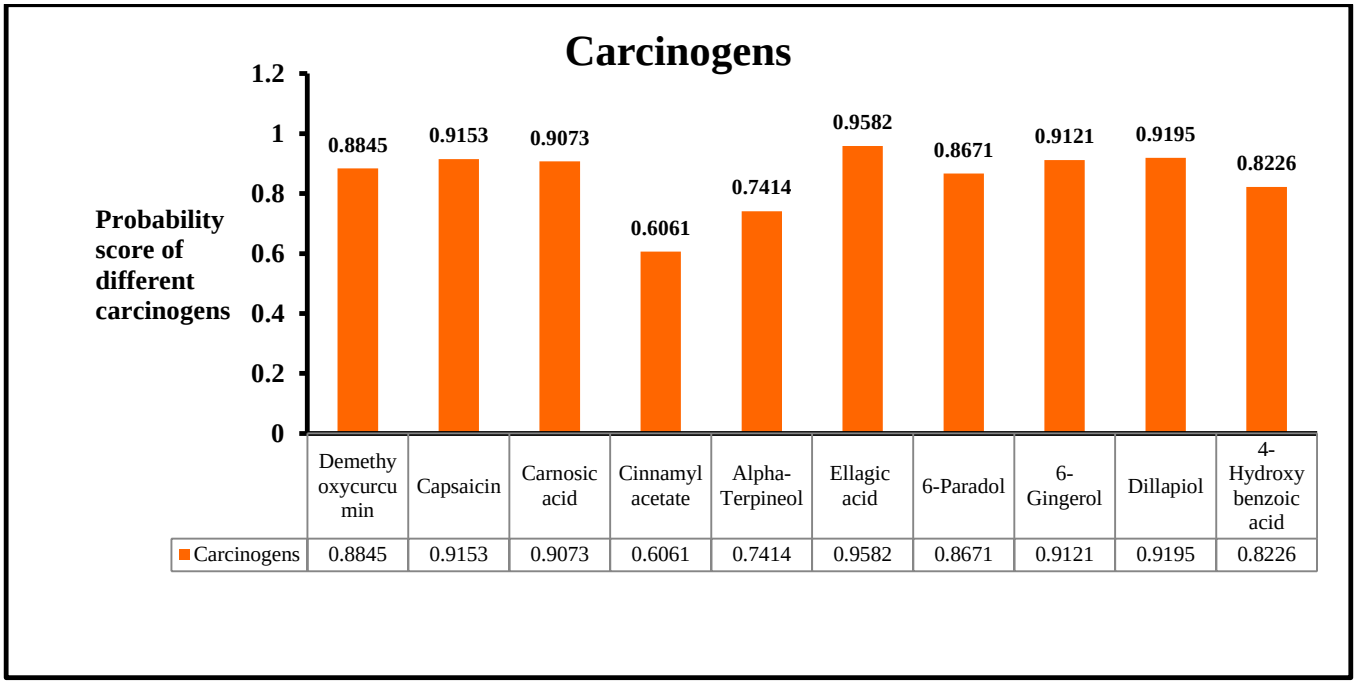
f


Human Intestinal Absorption
Probability Score of Human Intestinal Absorption
1.2
1
0.8
0.6
0.4
0.2
0
0.9882
0.9927
0.9637
1
0.9941
0.7199
0.9953
0.9805
0.9702
0.9872
Demethyoxycurcumin
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxybenzoic acid
HIA
0.9882
0.9927
0.9637
1
0.9941
0.7199
0.9953
0.9805
0.9702
0.9872

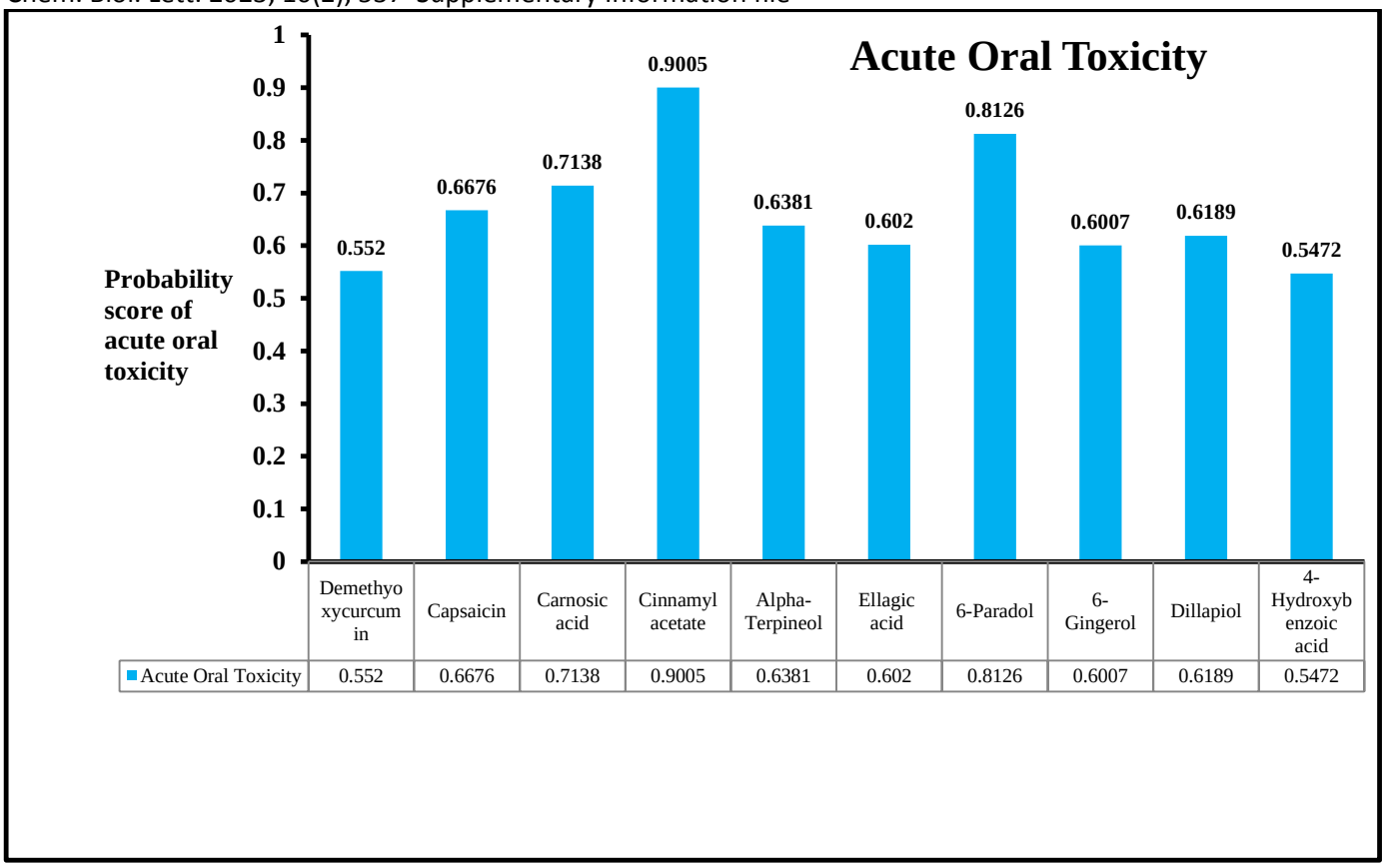
g


AMES Toxicity
Probability Score of AMES Toxicity
1.2
1
0.8
0.6
0.4
0.2
0
0.8048
0.7678
0.8961
0.6958
0.9133
0.9127
0.9132
0.7697
0.7846
0.9826
Demethy oxycurcu min
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxyb enzoic acid
AMES Toxicity
0.8048
0.7678
0.8961
0.6958
0.9133
0.9127
0.9132
0.7697
0.7846
0.9826

h


Carcinogens
Probability score of different carcinogens
1.2
1
0.8
0.6
0.4
0.2
0
0.8845
0.9153
0.9073
0.6061
0.7414
0.9582
0.8671
0.9121
0.9195
0.8226
Demethy oxycurcu min
Capsaicin
Carnosic acid
Cinnamyl acetate
Alpha-Terpineol
Ellagic acid
6-Paradol
6-Gingerol
Dillapiol
4-Hydroxy benzoic acid
Carcinogens
0.8845
0.9153
0.9073
0.6061
0.7414
0.9582
0.8671
0.9121
0.9195
0.8226

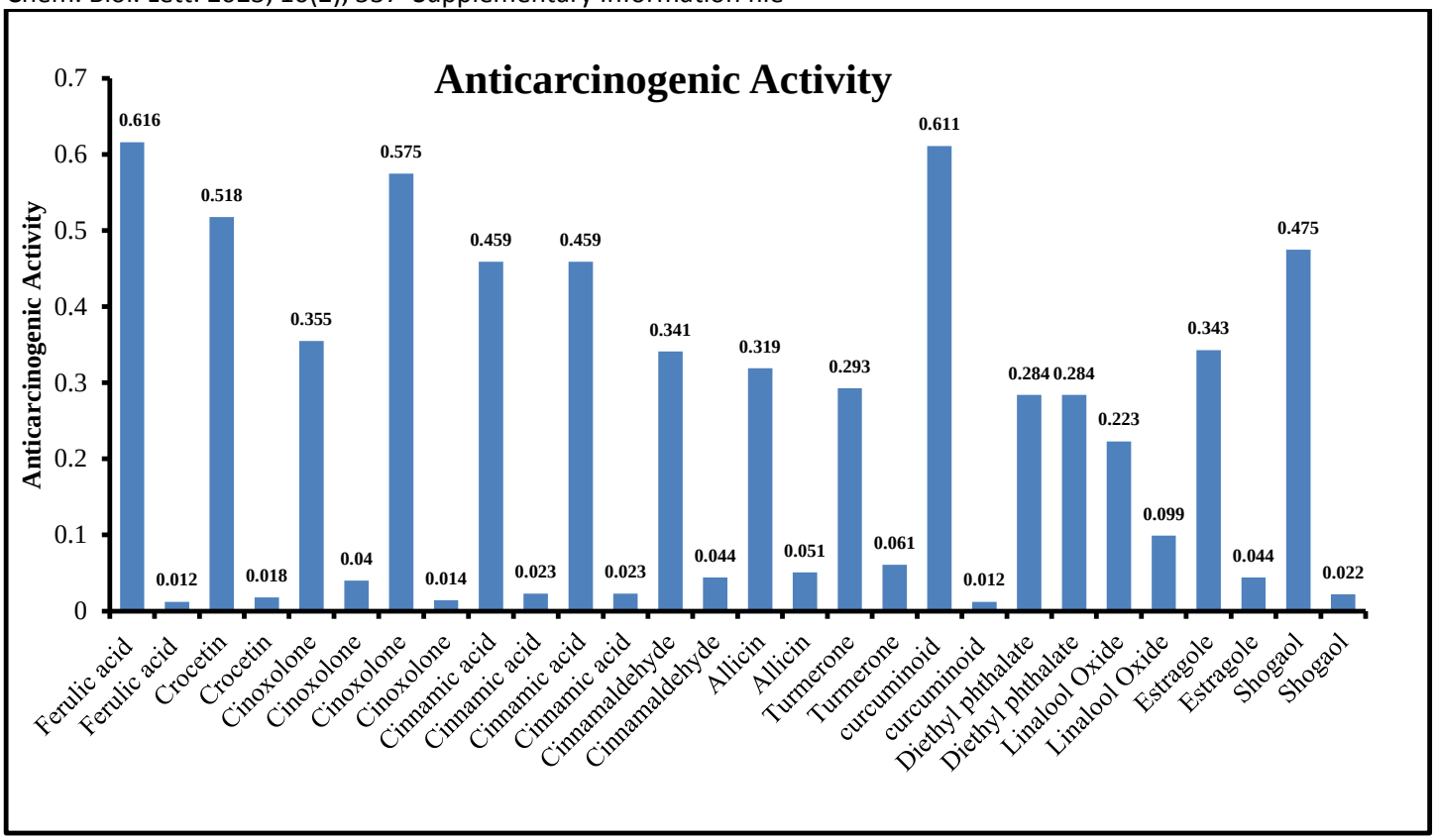
j



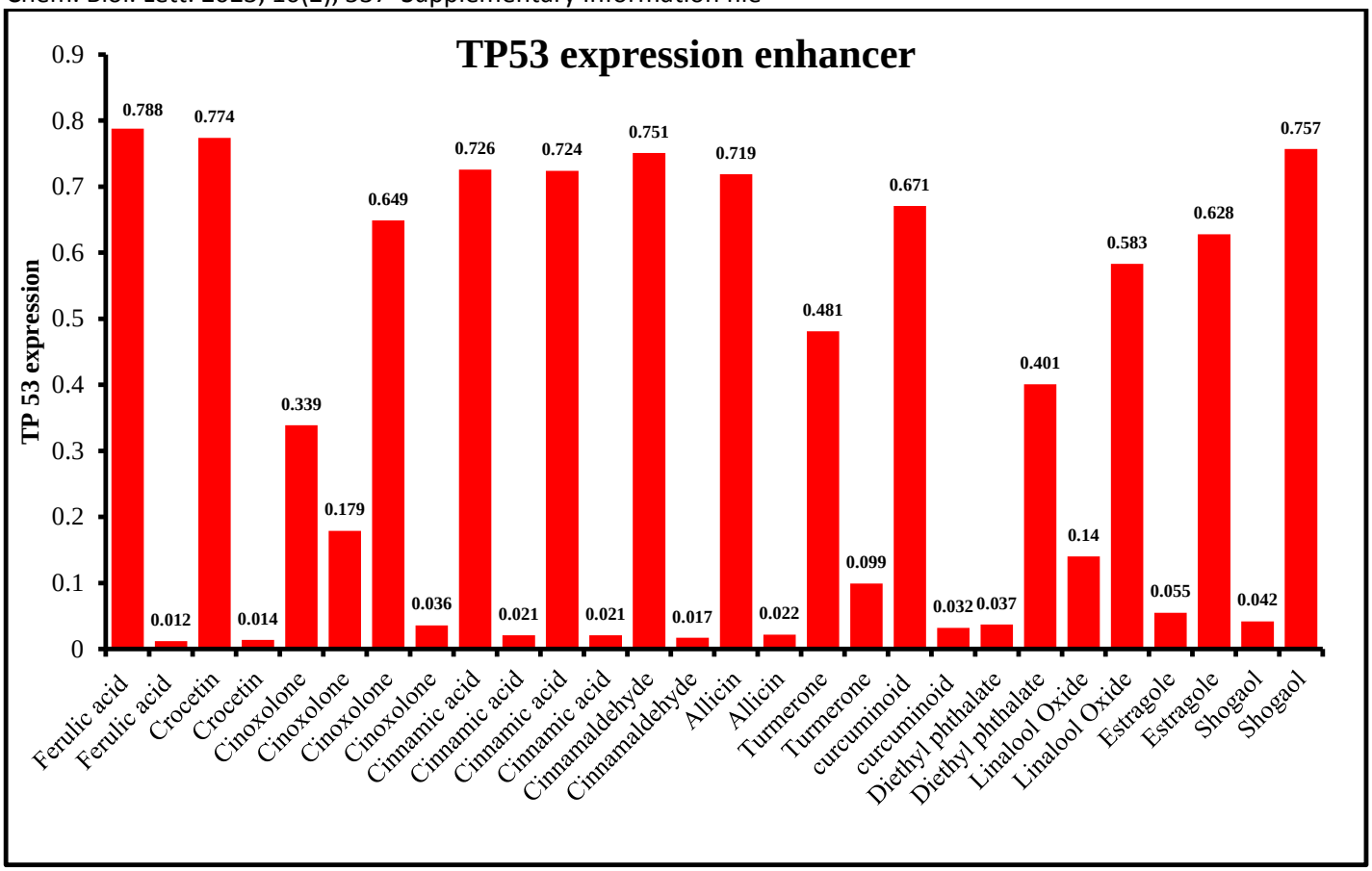
\

j

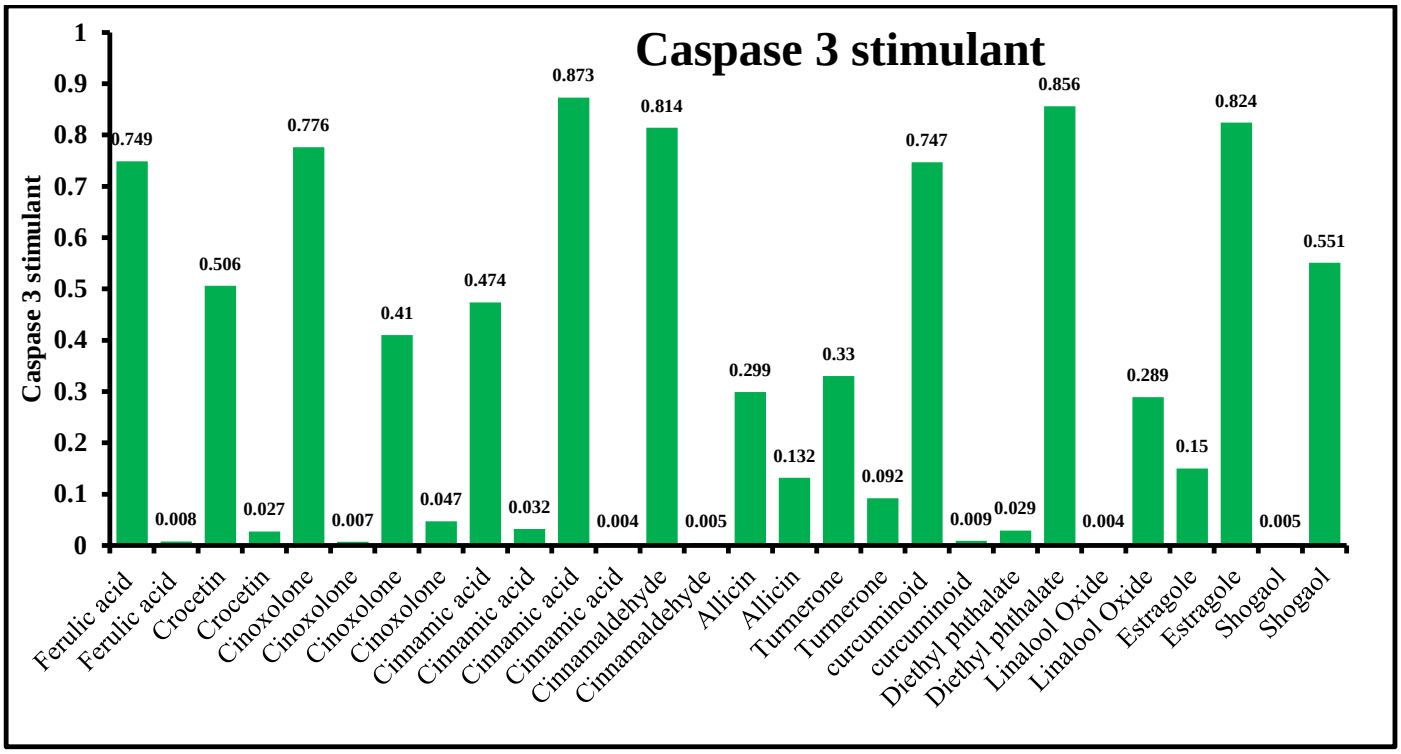
Figure 1 : a to j SWISS ADMET analysis of the studied phytochemicals using SWISS ADME software **a.** Calculated Log P prediction **b.** drug likeliness studies **c.** drug score **d.** total polar surface area **e.** aqueous solubility Log S **f.** Blood-brain barrier **g.** human intestinal absorption **h** AMES toxicity assay **i.** carcinogens **j.** acute oral toxicity


Anticarcinogenic Activity
Anticarcinogenic Activity
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.616
0.012
0.518
0.018
0.355
0.04
0.575
0.014
0.459
0.023
0.459
0.023
0.341
0.044
0.319
0.051
0.293
0.061
0.611
0.012
0.284
0.284
0.223
0.099
0.343
0.044
0.475
0.022
Ferulic acid
Ferulic acid
Crocetin
Crocetin
Cinoxolone
Cinoxolone
Cinoxolone
Cinoxolone
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamaldehyde
Cinnamaldehyde
Allicin
Allicin
Turmerone
Turmerone
curcuminoid
curcuminoid
Diethyl phthalate
Diethyl phthalate
Linalool Oxide
Linalool Oxide
Estragole
Estragole
Shogaol
Shogaol

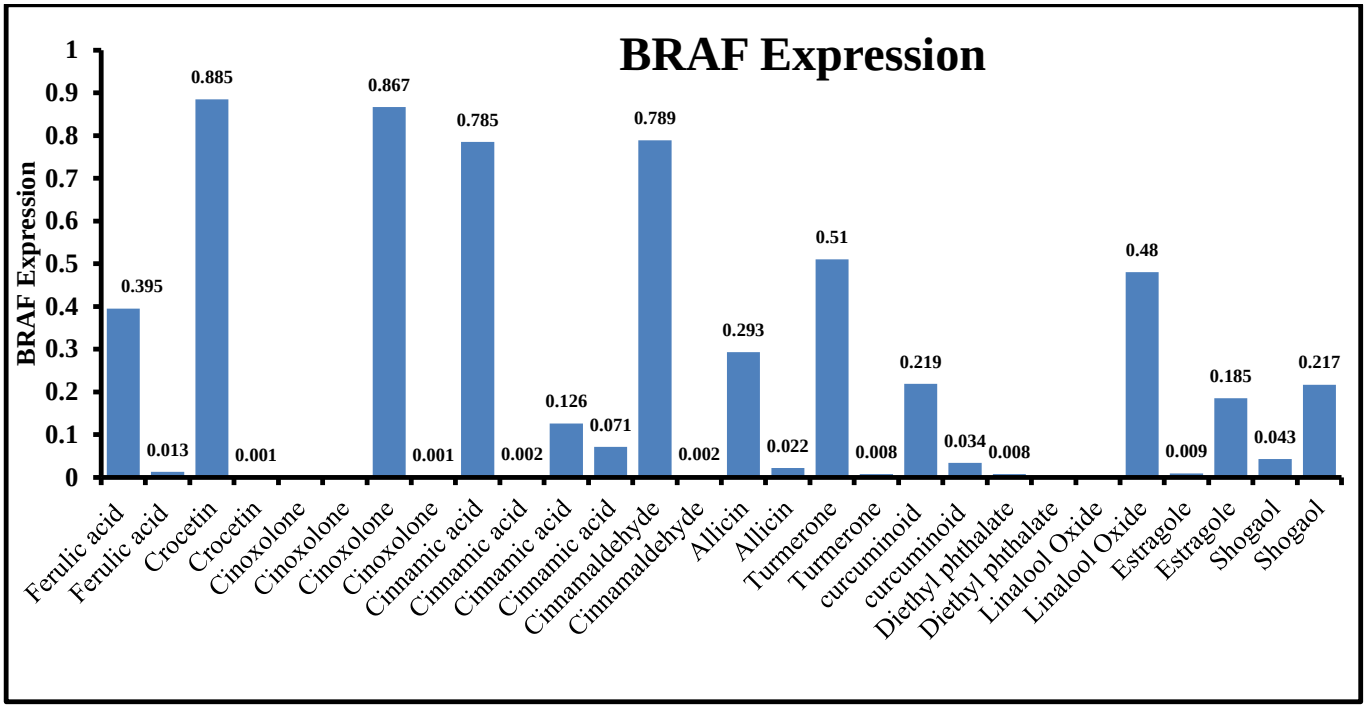
a


TP53 expression enhancer
TP 53 expression
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.788
0.012
0.774
0.014
0.339
0.179
0.649
0.036
0.726
0.021
0.724
0.021
0.751
0.017
0.719
0.022
0.481
0.099
0.671
0.032
0.037
0.401
0.14
0.583
0.055
0.628
0.042
0.757
Ferulic acid
Ferulic acid
Crocetin
Crocetin
Cinoxolone
Cinoxolone
Cinoxolone
Cinoxolone
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamaldehyde
Cinnamaldehyde
Allicin
Allicin
Turmerone
Turmerone
curcuminoid
curcuminoid
Diethyl phthalate
Diethyl phthalate
Linalool Oxide
Linalool Oxide
Estragole
Estragole
Shogaol
Shogaol

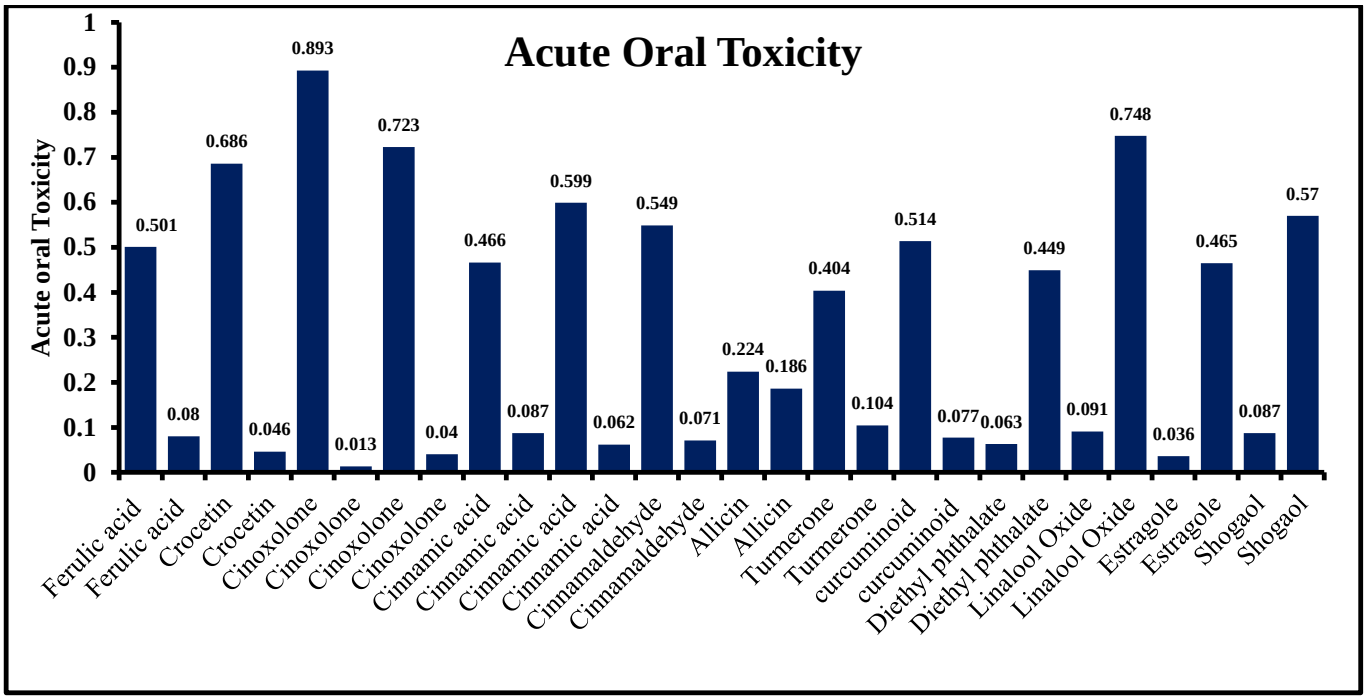
b


Caspase 3 stimulant
Caspase 3 stimulant
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.749
0.008
0.506
0.027
0.776
0.007
0.41
0.047
0.474
0.032
0.873
0.004
0.814
0.005
0.299
0.132
0.33
0.092
0.747
0.009
0.029
0.856
0.004
0.289
0.15
0.824
0.005
0.551
Ferulic acid
Ferulic acid
Crocetin
Crocetin
Cinoxolone
Cinoxolone
Cinoxolone
Cinoxolone
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamaldehyde
Cinnamaldehyde
Allicin
Allicin
Turmerone
Turmerone
curcuminoid
curcuminoid
Diethyl phthalate
Diethyl phthalate
Linalool Oxide
Linalool Oxide
Estragole
Estragole
Shogaol
Shogaol

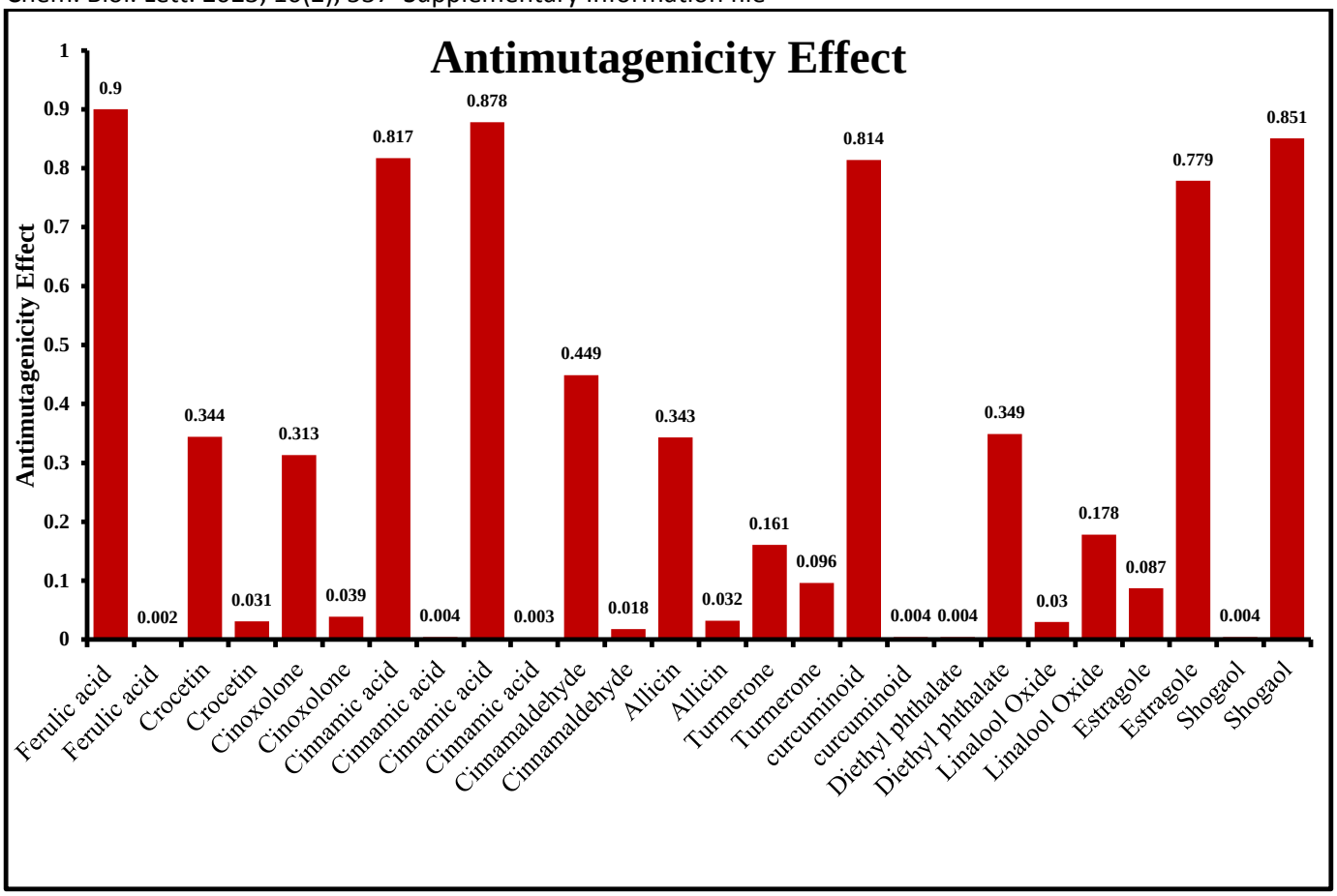
c


BRAF Expression
BRAF Expression
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.395
0.013
0.885
0.001
0.867
0.001
0.785
0.002
0.126
0.071
0.789
0.002
0.293
0.022
0.51
0.008
0.219
0.034
0.008
0.48
0.009
0.185
0.043
0.217
Ferulic acid
Ferulic acid
Crocetin
Crocetin
Cinoxolone
Cinoxolone
Cinoxolone
Cinoxolone
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamaldehyde
Cinnamaldehyde
Allicin
Allicin
Turmerone
Turmerone
curcuminoid
curcuminoid
Diethyl phthalate
Diethyl phthalate
Linalool Oxide
Linalool Oxide
Estragole
Estragole
Shogaol
Shogaol

d


Acute Oral Toxicity
Acute oral Toxicity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.501
0.08
0.686
0.046
0.893
0.013
0.723
0.04
0.466
0.087
0.599
0.062
0.549
0.071
0.224
0.186
0.404
0.104
0.514
0.077
0.063
0.449
0.091
0.748
0.036
0.465
0.087
0.57
Ferulic acid
Ferulic acid
Crocetin
Crocetin
Cinoxolone
Cinoxolone
Cinoxolone
Cinoxolone
Cinnamic acid
Cinnamic acid
Cinnamic acid
Cinnamaldehyde
Cinnamaldehyde
Allicin
Allicin
Turmerone
Turmerone
curcuminoid
curcuminoid
Diethyl phthalate
Diethyl phthalate
Linalool Oxide
Linalool Oxide
Estragole
Estragole
Shogaol
Shogaol

e



f

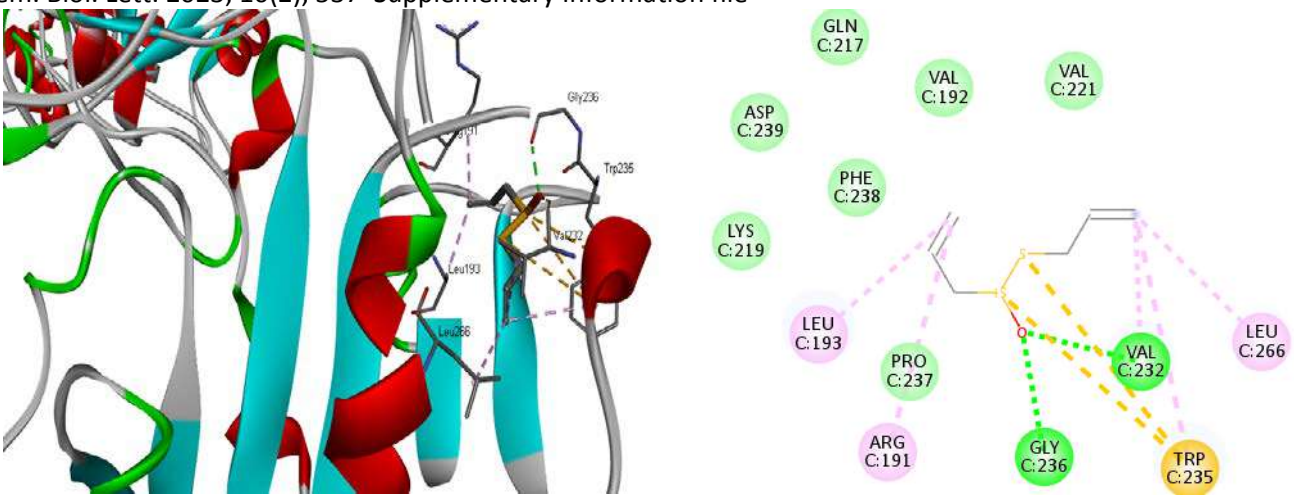
Figure 2 a-f: PASS (Prediction of Activity Spectra for Substances) analysis of studied phytochemicals

a. Anti-carcinogenic parameters **b.** TP53 expression pattern **c.** caspase 3 stimulant **d.** BRAF expression inhibitor **e.** toxicity studies **f.** antimutagenic studies

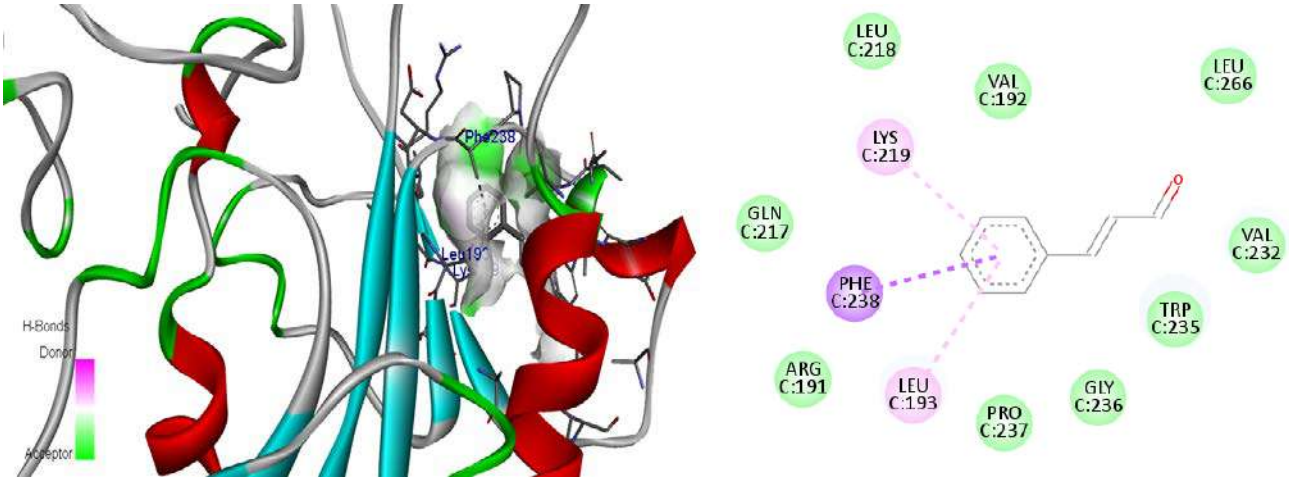
Table 2: Small phytochemicals with different binding energy released during interaction with 2PV0

Small phytochemicals	Pub Chem Comp ID (CID)	Binding energy (Global energy) after docking with 2PV0 (kcal/mol)	Sources of these phytochemicals
Ferulic acid	445858	-37.44	*Syzygium aomaticum*
Crocetin	5281232	**-49.27**	*Crocus sativus*
Cinnamic acid	444539	-30.99	*Cinnamon*
Eugenol	3314	-29.72	*Cinnamon*

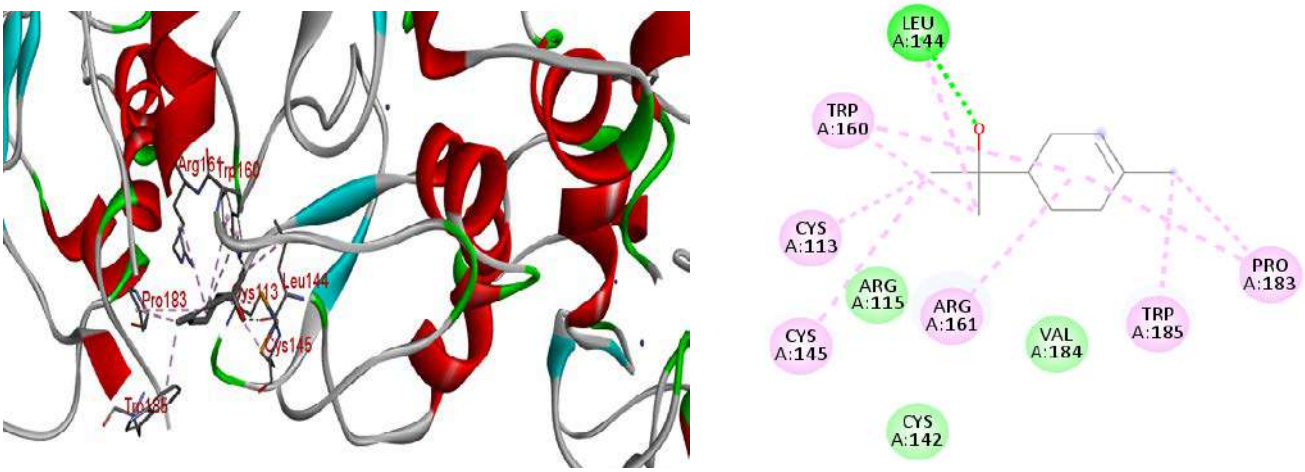
Cinnamaldehyde	637511	-27.60	*Cinnamon*
Allicin	65036	-33.27	*Allium sativum*
Curcumin	101341351	-38.19	*Curcumin longa*
Estragole	8815	-29.69	*Ocimum basilicum*
Demethoxy curcumin	5469424	**-48.27**	*Curcumin longa*
Capsaicin	1548943	**-42.85**	*Piper nigrum*
Cinnamyl acetate	5282110	-36.53	*Curcumin longa*
Alpha terpineol	17100	-27.65	*Elettaria cardamomum*
Carnosic acid	65126	**-44.72**	*Rosmarinus officinalis*


GLN C:217
VAL C:192
VAL C:221
ASP C:239
PHE C:238
LYS C:219
LEU C:193
PRO C:237
ARG C:191
GLY C:236
TRP C:235
VAL C:232
LEU C:266

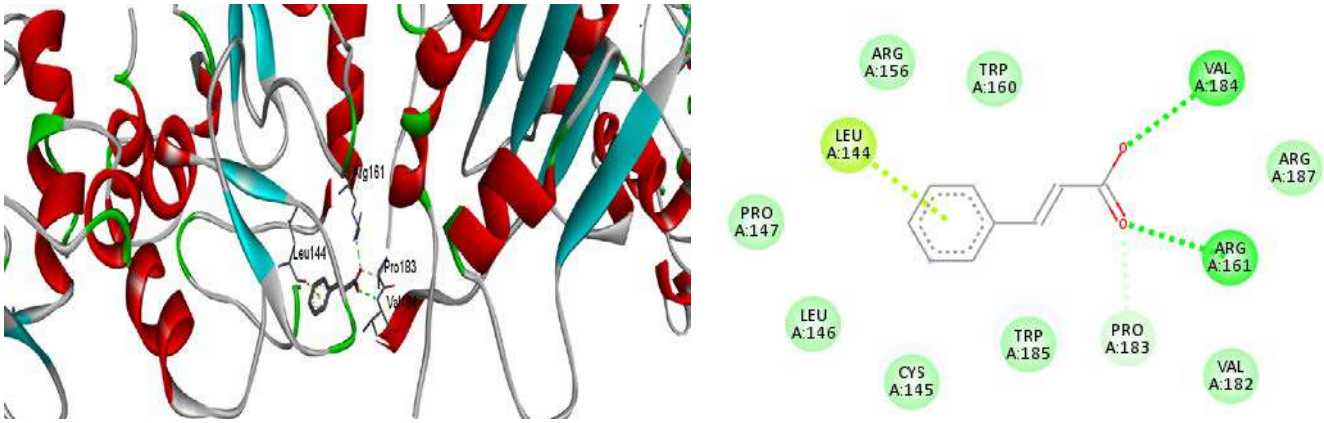
a


Phe238
H-Bonds
Donor
Acceptor
LEU C:218
VAL C:192
LEU C:266
LYS C:219
GLN C:217
VAL C:232
PHE C:238
TRP C:235
ARG C:191
LEU C:193
GLY C:236
PRO C:237

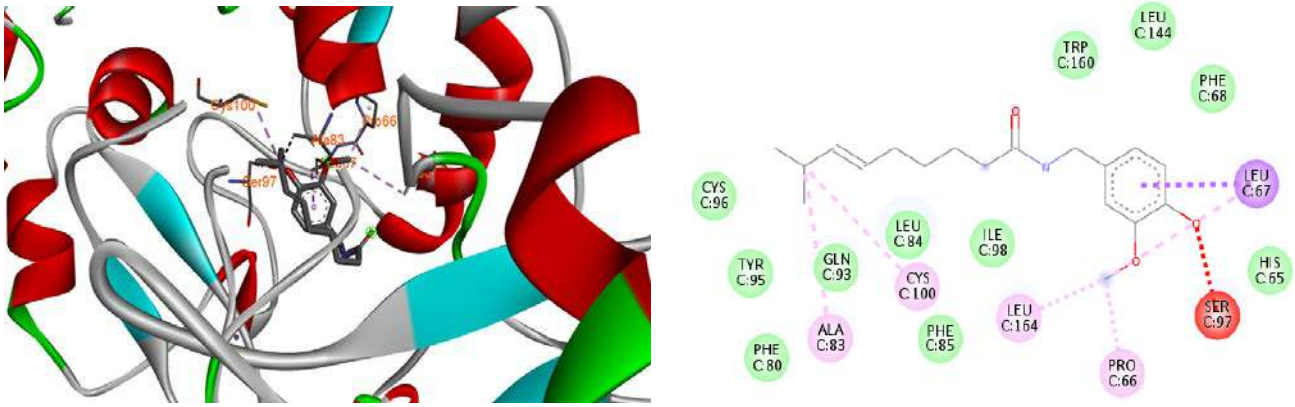
b


Arg161
Trp160
Pro183
Cys113
Leu144
Cys145
LEU A:144
TRP A:160
CYS A:113
ARG A:115
ARG A:161
PRO A:183
CYS A:145
VAL A:184
TRP A:185
CYS A:142

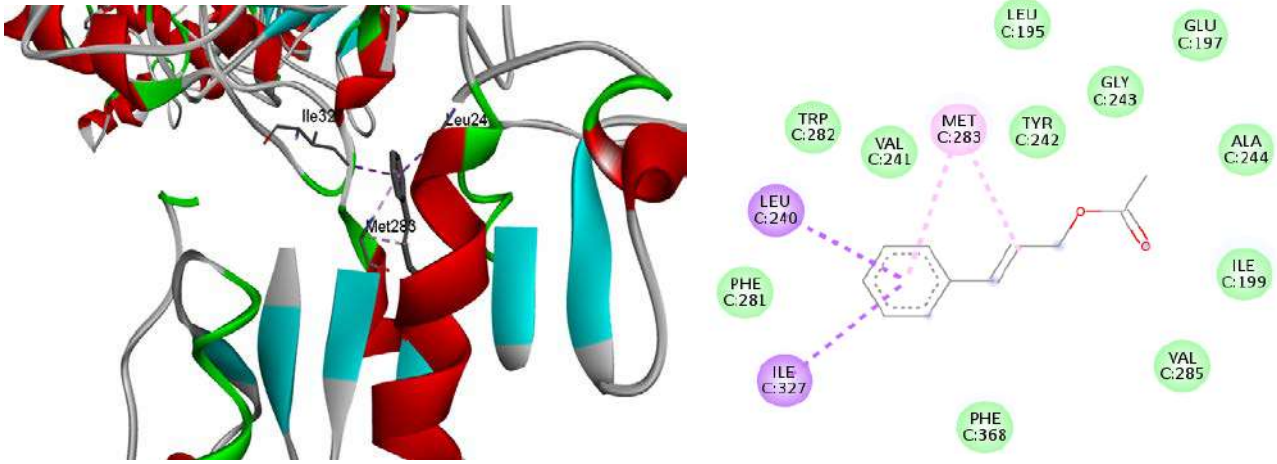
c


Leu144
Pro183
ARG A:156
TRP A:160
VAL A:184
LEU A:144
ARG A:187
PRO A:147
ARG A:161
LEU A:146
TRP A:185
PRO A:183
CYS A:145
VAL A:182

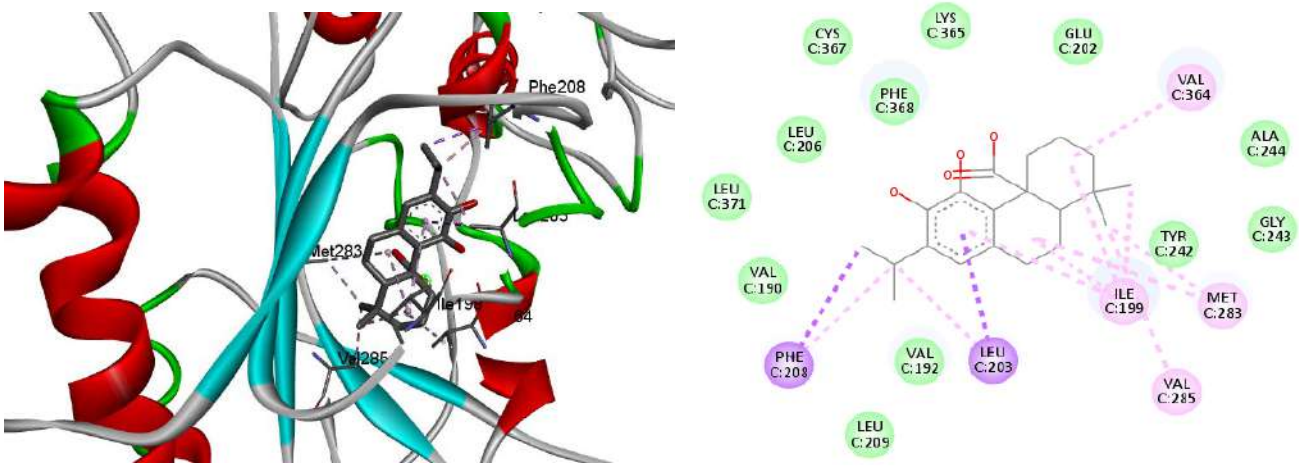
d


LEU C:144
TRP C:160
PHE C:68
CYS C:96
LEU C:67
LEU C:84
ILE C:98
TYR C:95
GLN C:93
CYS C:100
HIS C:65
SER C:97
LEU C:164
ALA C:83
PHE C:85
PHE C:80
PRO C:66

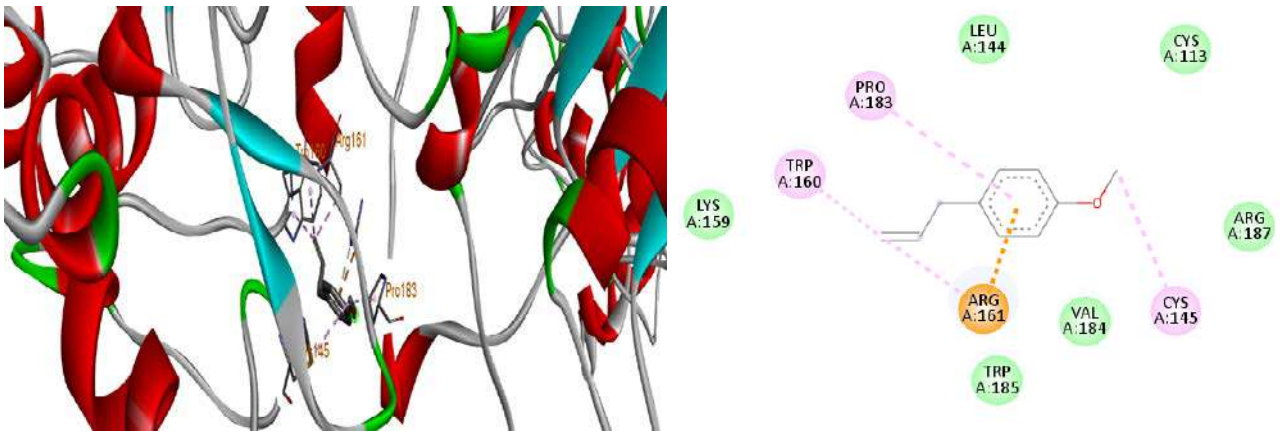
e


Ile32
Leu24
Met283
LEU C:195
GLU C:197
GLY C:243
TRP C:282
MET C:283
TYR C:242
VAL C:241
ALA C:244
LEU C:240
ILE C:199
PHE C:281
VAL C:285
ILE C:327
PHE C:368

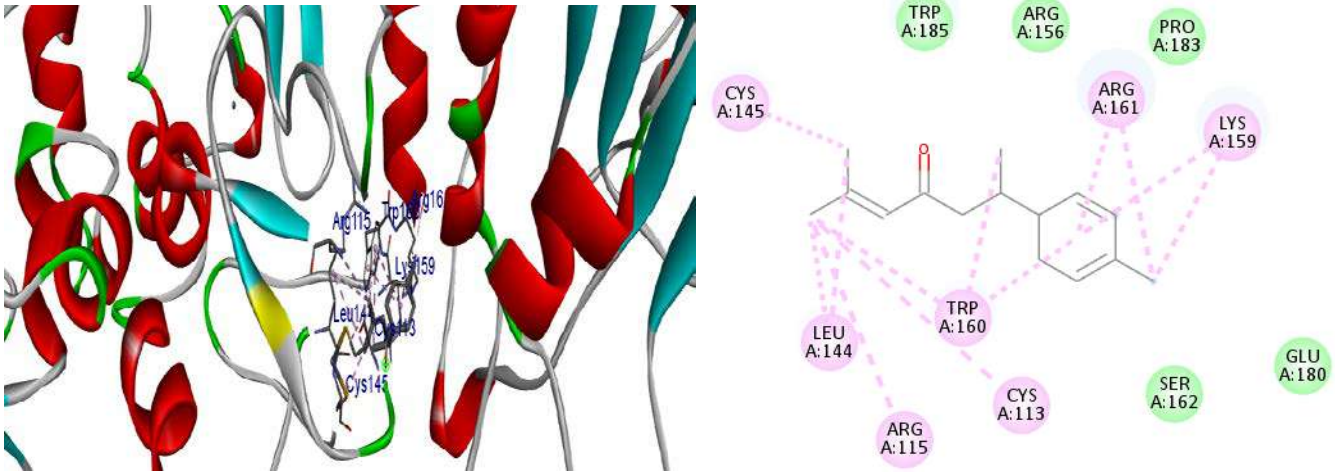
f


Phe208
Met283
CYS C:367
LYS C:365
GLU C:202
PHE C:368
VAL C:364
LEU C:206
ALA C:244
LEU C:371
GLY C:243
TYR C:242
VAL C:190
ILE C:199
MET C:283
PHE C:208
VAL C:192
LEU C:203
VAL C:285
LEU C:209

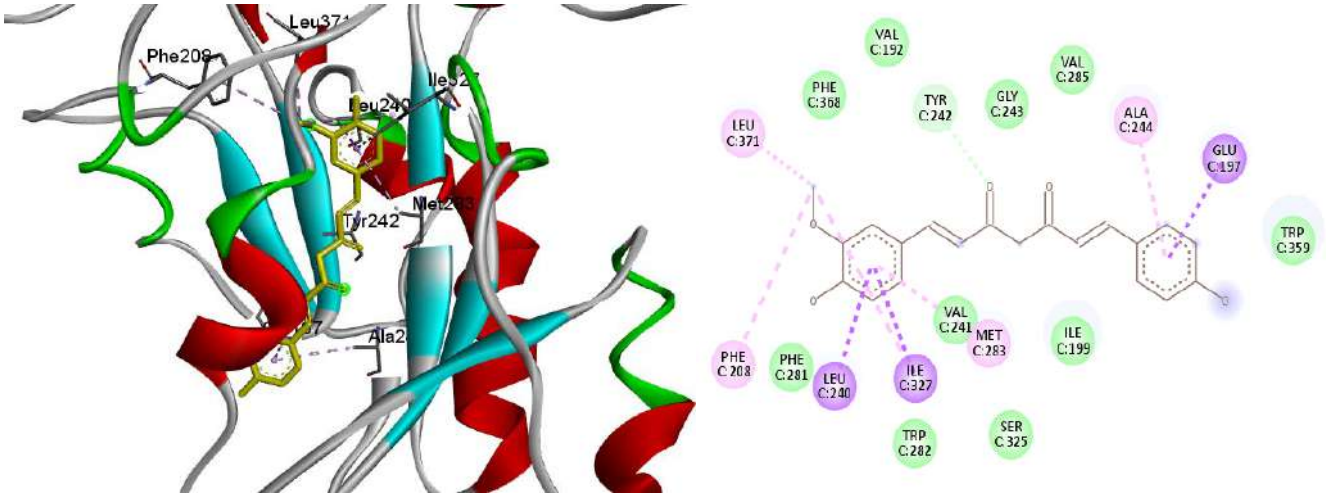
g


Pro183
LEU A:144
CYS A:113
PRO A:183
TRP A:160
LYS A:159
ARG A:187
ARG A:161
VAL A:184
CYS A:145
TRP A:185

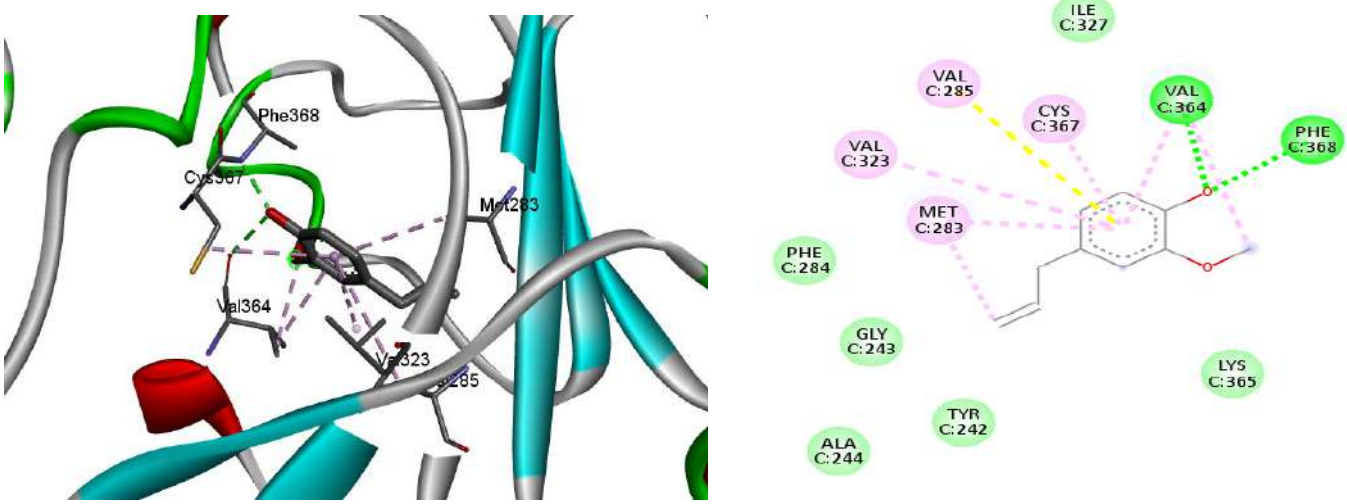
h


Arg115
Lys159
Cys145
TRP A:185
ARG A:156
PRO A:183
CYS A:145
ARG A:161
LYS A:159
TRP A:160
LEU A:144
GLU A:180
CYS A:113
SER A:162
ARG A:115

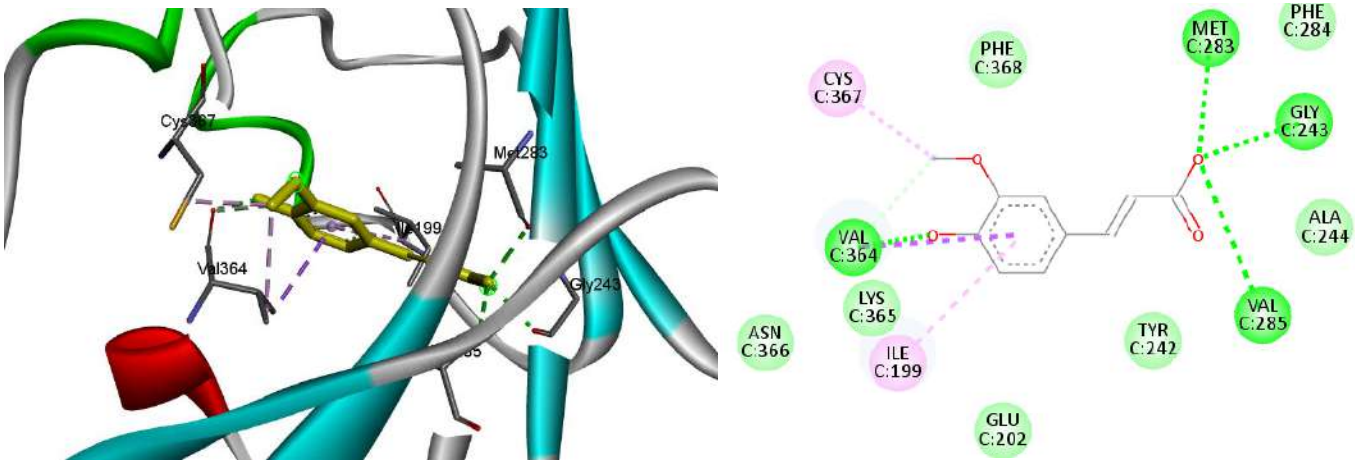
i


Phe208
Tyr242
Ala2
VAL C:192
PHE C:368
TYR C:242
GLY C:243
VAL C:285
ALA C:244
GLU C:197
LEU C:371
TRP C:359
VAL C:241
MET C:283
ILE C:199
PHE C:208
PHE C:281
LEU C:240
ILE C:327
TRP C:282
SER C:325

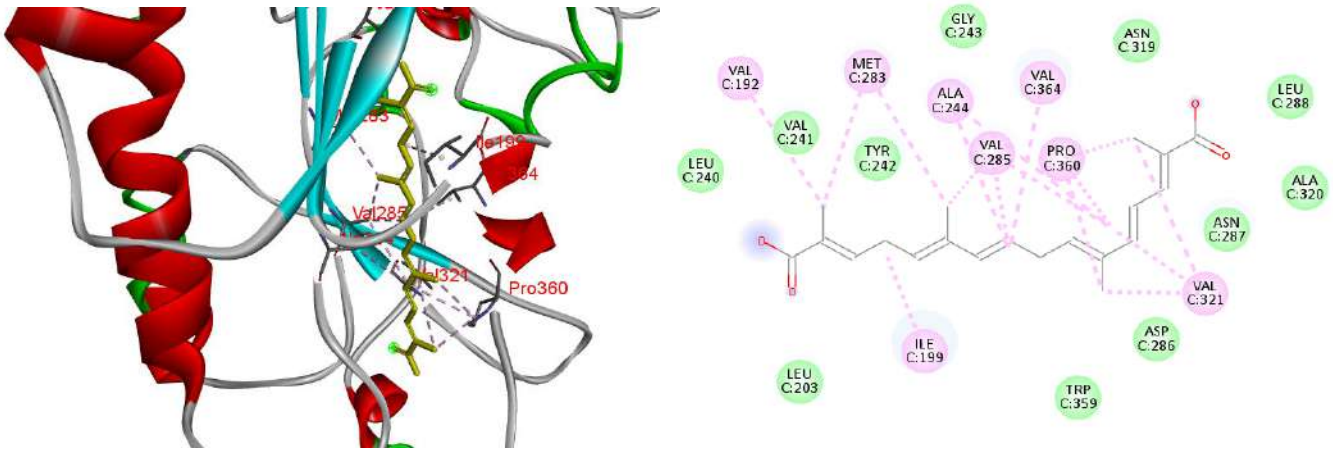
j


Phe368
Met283
Val364
ILE C:327
VAL C:285
CYS C:367
VAL C:364
PHE C:368
VAL C:323
MET C:283
PHE C:284
GLY C:243
LYS C:365
TYR C:242
ALA C:244

k


Met283
Val364
Gly243
PHE C:284
MET C:283
PHE C:368
CYS C:367
GLY C:243
ALA C:244
VAL C:364
LYS C:365
VAL C:285
ASN C:366
TYR C:242
ILE C:199
GLU C:202

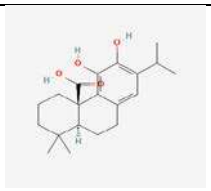
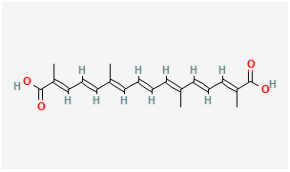
l

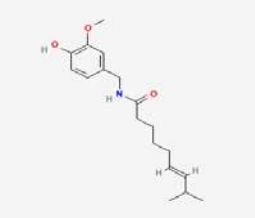
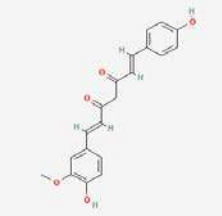


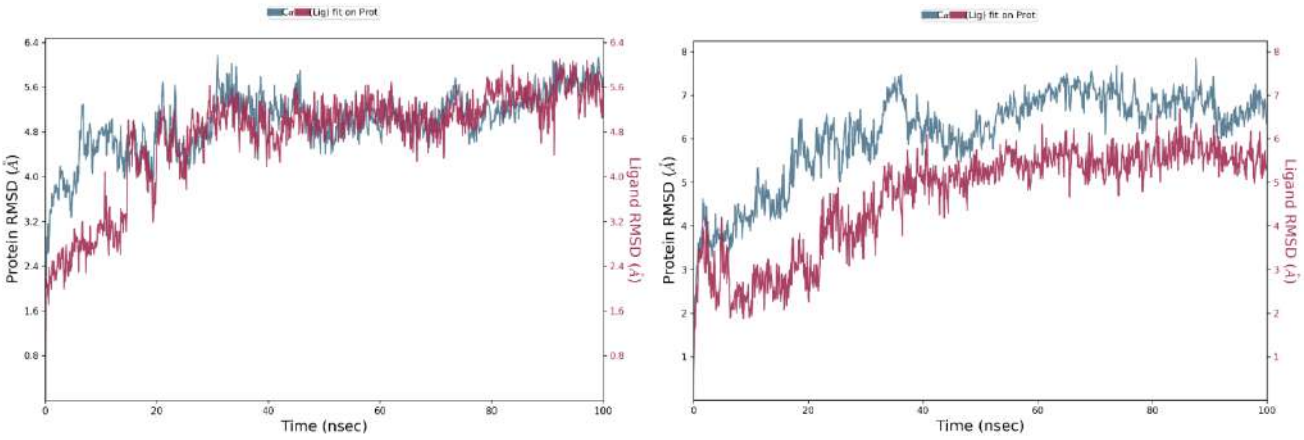
m

Figure 3 a –m 3D (Left panel) and 2D (right panel) representations of the binding modes of studied small phytochemicals against DNA methyl transferase 3-like protein (PDB ID: 2PV0). Also, hydrogen bonding interaction(shown in fluorescent green) and other types of non-covalent interaction(shown in dark and light purple) were also depicted throughout all figures **a**. allicin **b**. cinnamaldehyde **c**. alpha terpineol **d.** cinnamic acid **e.** capsaicin **f.** cinnamyl acetate **g.** carnosic acid **h.** estragole **i.** curcumin **j.** demethoxycurcumin **k.** eugenol **l.** ferulic acid **m.** crocetin

Table 3: Four best-optimized structure for MD simulation analysis

S.No.	Compound name	RMSD	Interacting residues during MD simulation	MMGBSA (ΔG-bind) post MD (kcal/mol)
1	 **Carnosic acid**	**0.5-1.0 Å**	Val192,Ile199,Leu203,Leu240,Tyr242,Thr361,Val364,and Tyr377	**-64.59±5**
2		**1.0-2.0 Å**	Val192,Leu193,Ile199,Leu203,Leu240,val241,Tyr242,Asn287,Lys358,Leu363,Val364,and Cys367	**-51.85±4**

	Crocetein			
3	 **Capsaicin**	**2.5-3.0 Å**	Leu84,Tyr87,Ser97,Leu164,Arg 161,Thr181 and Tyr374	**-48.93±5**
4	 **Demethoxycurcumin**	**3.0-4.0 Å**	Asp89,Gly243,Ile199,Phe208,Th r361,Asn287,Phe368 andThr380	**-41.85±4**



a **b**

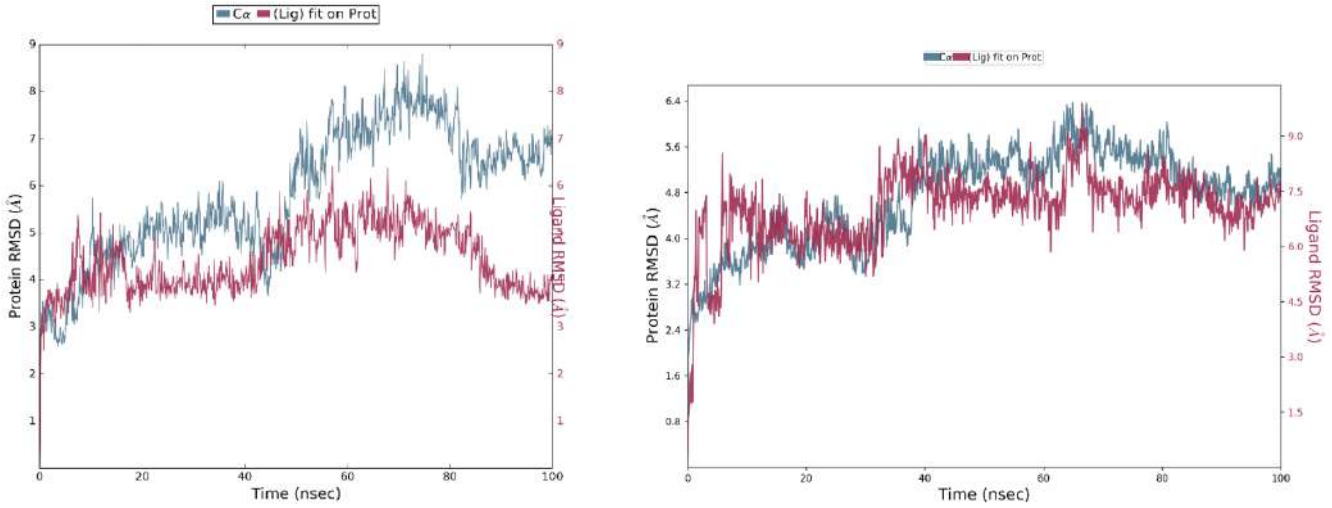
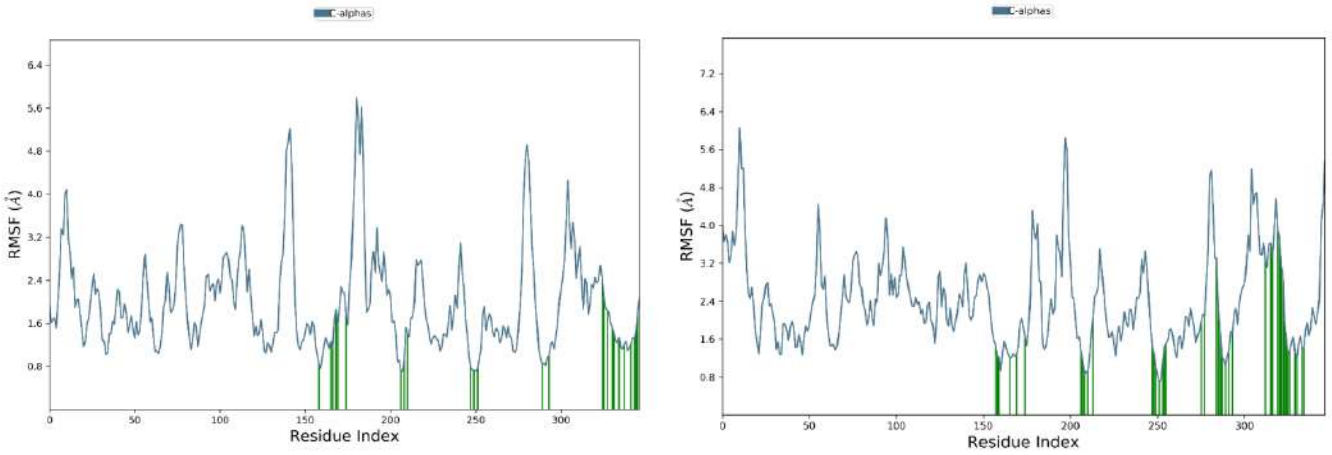
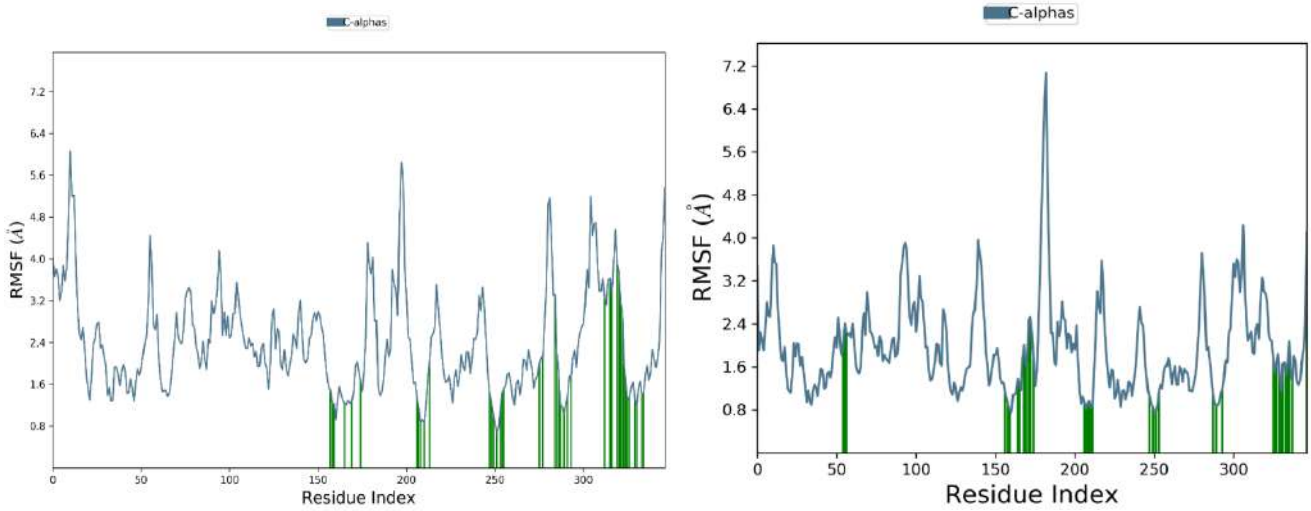


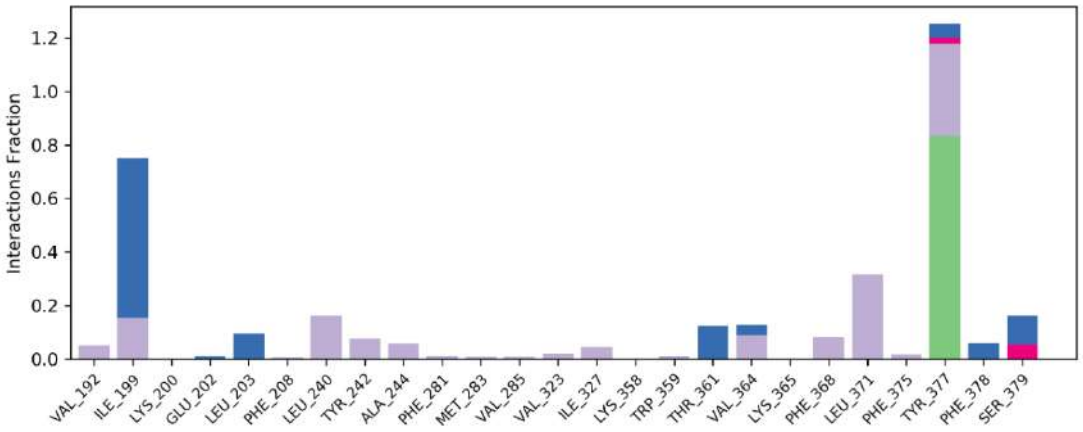
Figure 4. Representation of RMSD after MD simulation of the four top ranked compounds, X-axis represents simulation time in ns and the Y- axis represents the RMSD value. **a.** Carnosic acid **b.** Crocetin **c.** Capsaicin **d.** Demethoxycurcumin.



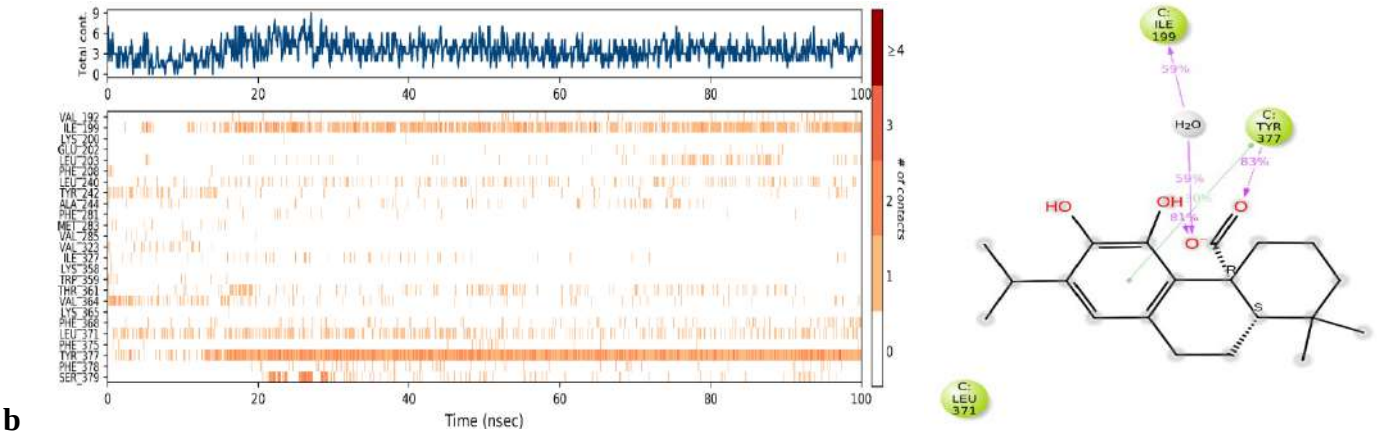


c **d**

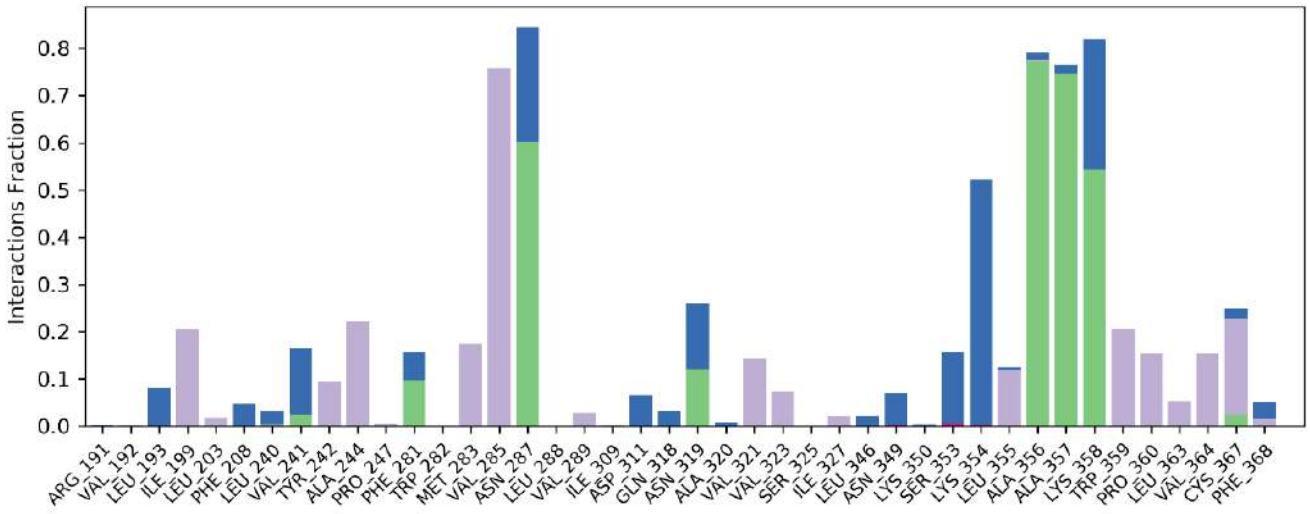
Figure 5. Representation of RMSF after MD simulation of the four top-ranked compounds. The X-axis represents residue numbers and the Y-axis represents RMSF values. **a.** Carnosic acid **b.** Crocetin **c.** Capsaicin **d.** Demethoxycurcumin. Blue color represents local changes along the protein chain, peaks indicate areas of the protein that fluctuate the most during the simulation, while green peaks show interactions of the residues with ligands.



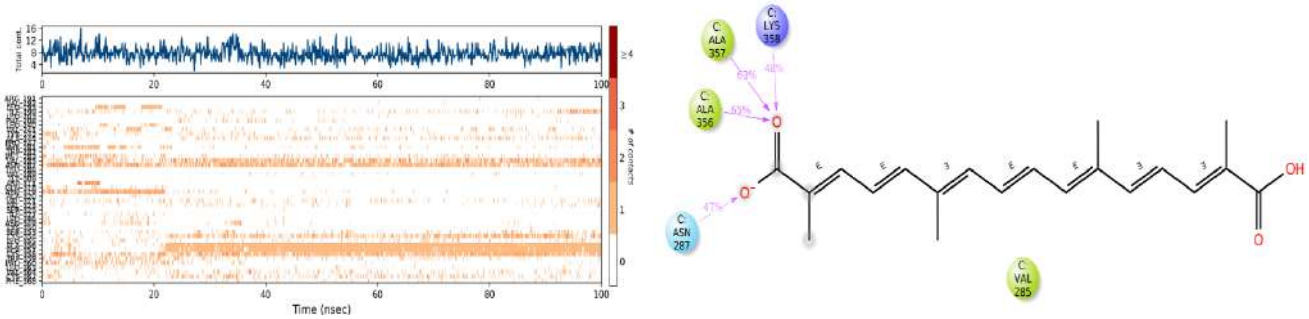
a



b c



d



e f

Figure. 6. Representation of protein-ligand contacts by MD simulation with compound Carnosic acid and Crocetin **(a)** and **(d)** stacked bar charts normalized over the course of the trajectory in which the X-axis shows residue name, Y-axis shows interaction fractions, **(b)** and **(e)** timeline of interactions and contacts (X-axis

simulation time in nanoseconds and Y-axis name of residue). **(c)** and **(f)** 2D interactions of protein in complex with Carnosic acid and Crocetin.