Chapter 5

Isolation, Cytotoxic Evaluation, and *In silico* Screening of Coumarins from *Psoralea corylifolia* (L.)

5 Isolation, Cytotoxic Evaluation, and *In silico* Screening of Coumarins from *Psoralea corylifolia* (L.)

5.1 Introduction

Chemically coumarins are characterized as 2*H*-chromen-2-one (1,2-benzopyrone, or 2H-1-benzopyran-2-one) oxo-heterocycle. Classical examples include anticoagulants warfarin, and its analogs as vitamin K antagonists [143]. Other examples are umbelliferone and armillarisin A with choleretic activity and antibiotic novobiocin with bacterial DNA gyrase inhibitory activity [144]. *Psoralea corylifolia* L. (syn. *Cullen corylifolium* (L.) Medik., common name Bawachi) holds significant importance in Ayurvedic medicine, traditional Chinese medicine, and South African medicine because of its photosensitizing, estrogenic, antimicrobial, and cytotoxic properties [128, 145]. The entire plant, particularly seeds, possesses notable medicinal properties thus, used to treat a variety of illnesses. Apart from its use in the treatment of psoriasis, leucoderma, leprosy, and inflammatory skin illnesses, bawachi is also used as an aphrodisiac, laxative, anthelmintic, and diuretic in traditional medicine [146]. The seed oil of *P. corylifolia* is extensively investigated at preclinical and clinical levels for the treatment of psoriasis [128]. The major active constituents like psoralen, isopsoralen, bakuchiol, bavachalcone, neobavaisolflavones, and corylifols are predominantly found in seeds [147]. These phytoconstituents exhibit various pharmacological properties like antioxidant, antibacterial, antifungal, antitumor, estrogenic, anti-inflammatory, osteoblastic, and immunomodulatory [125]. The aerial parts of plant are not much investigated for phytochemical screening. The current study is focused to create the library of coumarins from the aerial parts and seeds of *P. corylifolia.* Here, various coumarins were isolated and characterized. The isolated coumarins were also evaluated for cytotoxicity against MDA-MB-231 and A549 lines. Later, to determine their possible mechanism, molecular

modelling studies of isolated coumarins with EGFR protein was also performed as they are highly overexpressed in both MDA-MB-231 and A549 cell lines.

5.2 Experimental section

5.2.1 General experimental procedures

As mentioned in Section 4.2.7.1.

5.2.2 Plant material

The authentication of *P. corylifolia* seeds is mentioned in Section 4.2.7.2. The *P. corylifolia* leaves were collected and authenticated by Dr. Bikarma Singh, Department of Botanical Garden, Plant Conservation and Agrotechnologies, CSIR-NBRI, Lucknow, India, from the Jammu region (J&K, India). The specimen sample of leaves was preserved (Voucher no. RRLH54225 and RRLH9475) in Janaki Ammal Herbarium-IIIM Jammu, J&K, India.

5.2.3 Extraction and isolation of plant material

The procedure for extraction and isolation of leaves (500 g) is same as mentioned for that of seeds in Section 4.2.7.3. The obtained compounds were characterized via NMR spectroscopy, mass spectrometry, and the structures were confirmed by comparing the NMR data with the reported literature (Appendix, Figure A.19-A.34) [148].

1. **Marmin** (**1**). $C_{19}H_{24}O_5$ 1H NMR (500 MHz, $CDCl_3$) δ 7.63 (d, J = *9.5* Hz, 1H), 7.35 (d, J = *8.5* Hz, 1H), 6.84 (dd, J = *8.5*, *2.4* Hz, 1H), 6.81 (d, J = *2.4* Hz, 1H), 6.23 (d, J = *9.5* Hz, 1H), 5.54 – 5.49 (m, 1H), 4.60 (d, J = *6.6* Hz, 2H), 3.35 (dd, J = *10.6*, *2.1* Hz, 1H), 2.42 – 2.32 (m, 1H), 2.21 – 2.11 (m, 1H), 1.77 (s, 3H), 1.68 – 1.58 (m, 1H), 1.52 – 1.41 (m, 1H), 1.23 (s, 3H), 1.16 (s, 3H). 13C NMR (126 MHz, $CDCl_3$) δ 162.1, 161.4, 155.8, 143.5, 142.2, 128.7, 118.9, 113.3, 113.0, 112.5, 101.6, 77.9, 73.1, 65.4, 36.6, 29.4, 26.5, 23.3, 16.8. HRMS *m/z*: $[M+Na]^+$ calc. 355.1516, obs. 355.1487 [149].

2. **Seselin** (**2**). $C_{14}H_{12}O_3$ 1H NMR (500 MHz, $CDCl_3$) δ 7.58 (d, *J = 9.5* Hz, 1H), 7.19 (d, *J = 8.4* Hz, 1H), 6.86 (dd, *J = 10.1, 0.8* Hz, 1H), 6.70 (dd, *J = 8.4, 0.8* Hz, 1H), 6.21 (d, *J = 9.4* Hz, 1H), 5.71 (d, *J = 10.1* Hz, 1H), 1.46 (s, 6H). 13C NMR (126 MHz, $CDCl_3$) δ 161.0, 156.3, 150.1, 143.9, 130.8, 127.9, 115.0, 113.5, 112.6, 109.3, 77.6, 28.1. HRMS *m/z*: $[M+H]^+$ calc. 229.0859, obs. 229.0827 [150].

3. **Xanthotoxin** (**3**). $C_{12}H_8O_4$ 1H NMR (500 MHz, $CDCl_3$) δ 7.78 (d, *J = 9.6* Hz, 1H), 7.71 (d, *J = 2.2* Hz, 1H), 7.37 (s, 1H), 6.84 (d, *J = 2.2* Hz, 1H), 6.38 (d, *J = 9.6* Hz, 1H), 4.31 (s, 3H). 13C NMR (126 MHz, $CDCl_3$) δ 160.5, 147.7, 146.7, 144.4, 142.9, 132.8, 126.1, 116.5, 114.7, 112.9, 106.7, 61.3. HRMS *m/z*: $[M+H]^+$ calc. 217.0496, obs. 217.0491 [151].

4. **Apterin** (**4**). $C_{20}H_{24}O_{10}$ 1H NMR (500 MHz, CD_3OD) δ 7.92 (d, *J = 9.6* Hz, 1H), 7.57 (d, *J* = 8.5 Hz, 1H), 6.90 (d, *J = 8.5* Hz, 1H), 6.26 (d, *J = 9.6* Hz, 1H), 5.63 (d, *J = 6.4* Hz, 1H), 4.80 (d, *J = 7.8* Hz, 1H), 4.56 (d, *J = 6.4* Hz, 1H), 3.51 (dd, *J = 11.9, 4.6* Hz, 1H), 3.43 – 3.33 (m, 2H), 3.24 (dd, *J = 11.9, 2.4* Hz, 1H), 3.20 – 3.12 (m, 2H), 1.63 (d, *J = 1.8* Hz, 6H). 13C NMR (126 MHz, CD_3OD) δ 163.3, 161.4, 151.9, 144.8, 130.9, 116.6, 113.4, 111.6, 107.5, 97.7, 92.0, 77.8, 76.6, 76.2, 73.6, 69.6, 68.9, 60.4, 23.7, 23.0. HRMS *m/z*: $[M+NH_4]^+$ calc. 442.1707, obs. 442.1701 [152].

5. **Isoimperatorin** (**5**). $C_{16}H_{14}O_4$ 1H NMR (500 MHz, $CDCl_3$) δ 8.16 (d, *J = 9.8* Hz, 1H), 7.60 (d, *J = 2.3* Hz, 1H), 7.15 (s, 1H), 6.97 (d, *J = 3.4* Hz, 1H), 6.27 (d, *J = 9.8* Hz, 1H), 5.55 (t, *J = 6.3* Hz, 1H), 4.93 (d, *J = 7.0* Hz, 2H), 1.82 (s, 3H), 1.71 (s, 3H). 13C NMR (126 MHz, $CDCl_3$) δ 161.3, 158.1, 152.7, 148.9, 144.9, 139.8, 139.6, 119.1, 114.2, 112.5, 107.5, 105.1, 94.2, 69.7, 25.8, 18.2. HRMS *m/z*: $[M+H]^+$ calc. 271.0965, obs. 271.0955 [153].

6. **Bergapten** (**6**). $C_{12}H_8O_4$ 1H NMR (500 MHz, $CDCl_3$) δ 8.18 (d, *J = 10.5* Hz, 1H), 7.62 (d, *J = 2.4* Hz, 1H), 7.16 (s, 1H), 7.04 (d, *J = 3.4* Hz, 1H), 6.30 (d, *J = 9.7* Hz, 1H), 4.29 (s, 3H). 13C NMR (126 MHz, $CDCl_3$) δ 161.3, 158.4, 152.7, 149.6, 144.8, 139.3, 112.7, 112.6, 106.4, 105.0, 93.9, 60.1. HRMS *m/z*: $[M+H]^+$ calc. 217.0496. obs. 217.0451 [153].

7. **Isobergapten** (**7**). $C_{12}H_8O_4$ 1H NMR (500 MHz, $CDCl_3$) δ 8.18 (d, *J = 9.8* Hz, 1H), 7.59 (d, *J = 2.3* Hz, 1H), 7.05 (d, *J = 3.1* Hz, 1H), 6.91 (s, 1H), 6.33 (d, *J = 9.8* Hz, 1H), 3.99 (s, 3H). 13C NMR (126 MHz, $CDCl_3$) δ 160.9, 157.9, 154.3, 148.8, 144.3, 139.9, 112.2, 110.1, 105.8, 103.9, 90.5, 56.3. HRMS *m/z*: $[M+H]^+$ calc. 217.0496. obs. 217.0487 [154].

8. **Heratomine** (**8**). $C_{16}H_{14}O_4$ 1H NMR (500 MHz, $CDCl_3$) δ 7.74 (d, *J = 9.6* Hz, 1H), 7.69 (d, *J = 2.1* Hz, 1H), 7.12 (d, *J = 1.2* Hz, 1H), 6.80 (s, 1H), 6.39 (d, *J = 9.5* Hz, 1H), 5.58 (t, *J = 6.8* Hz, 1H), 4.74 (d, *J = 6.9* Hz, 2H), 1.80 (d, *J = 16.8* Hz, 6H). 13C NMR (126 MHz, $CDCl_3$) δ 161.1, 147.3, 145.9, 144.5, 143.1, 142.2, 139.4, 118.9, 118.6, 114.4, 113.6, 105.1, 104.5, 66.3, 25.9, 18.3. HRMS *m/z*: $[M+H]^+$ calc. 271.0965, obs. 271.0955 [155].

9. **Marmesin** (**9**). $C_{14}H_{14}O_4$ 1H NMR (500 MHz, DMSO-d_6) δ 7.93 (d, *J = 9.5* Hz, 1H), 7.48 (s, 1H), 6.79 (s, 1H), 6.22 (d, *J = 9.5* Hz, 1H), 4.71 (s, 1H), 3.19 (d, *J = 7.8* Hz, 2H), 1.15 (d, *J = 9.6* Hz, 6H). 13C NMR (126 MHz, DMSO- d_6) δ 163.8, 161.0, 155.5, 145.2, 126.0, 124.4, 112.6, 111.7, 97.2, 91.4, 70.5, 29.1, 26.3, 25.4. HRMS *m/z*: $[M+H]^+$ calc. 247.0965, sobs. 247.0961 [156].

10. **Xanthotoxol** (**10**). $C_{11}H_6O_4$ 1H NMR (500 MHz, CD_3OD) δ 8.01 (d, *J = 9.6* Hz, 1H), 7.85 (d, *J = 2.0* Hz, 1H), 7.36 (s, 1H), 6.91 (d, J = 2.2 Hz, 1H), 6.35 (d, *J = 9.6* Hz,

1H). 13C NMR (126 MHz, CD_3OD) δ 161.6, 146.9, 145.7, 145.6, 139.6, 130.3, 125.9, 116.4, 113.2, 109.8, 106.5. HRMS *m/z*: $[M+H]^+$ calc. 203.0339, obs. 203.0329 [151].

11. **Heratomol benzoate** (**11**). $C_{18}H_{10}O_5$ 1H NMR (500 MHz, $CDCl_3$) δ 8.28 (dd, *J = 8.5, 1.3* Hz, 2H), 7.79 (d, *J = 9.6* Hz, 1H), 7.73 – 7.66 (m, 2H), 7.57 (t, *J = 7.6* Hz, 2H), 7.33 (s, 1H), 7.19 (d, *J = 2.1* Hz, 1H), 6.44 (d, *J = 9.6* Hz, 1H). 13C NMR (126 MHz, $CDCl_3$) δ 165.0, 161.3, 149.4, 147.2, 146.9, 144.8, 134.9, 133.6, 131.2, 129.5, 129.2, 119.8, 116.8, 115.7, 114.4, 105.5. HRMS *m/z*: $[M+H]^+$ calc. 307.0601, obs. 307.0557 [155].

12. **Isopsoralen** (**12**). $C_{11}H_6O_3$ 1H NMR (500 MHz, $CDCl_3$) δ 7.82 (d, *J = 9.6* Hz, 1H), 7.70 (d, *J = 2.2* Hz, 1H), 7.44 (dd, *J = 8.5, 0.9* Hz, 1H), 7.38 (d, *J = 8.5* Hz, 1H), 7.14 (dd, *J = 2.2, 0.9* Hz, 1H), 6.40 (d, *J = 9.6* Hz, 1H), 1.63 (s, 1H). 13C NMR (126 MHz, $CDCl_3$) δ 160.9, 157.4, 148.5, 145.9, 144.5, 123.8, 116.9, 114.1, 113.5, 108.8, 104.1. HRMS *m/z*: $[M+H]^+$ calc. 187.0390, obs. 187.0381 [157].

13. **Oxypeucedanin** (**13**). $C_{16}H_{14}O_5$ 1H NMR (500 MHz, $CDCl_3$) δ 8.07 (s, 1H), 7.47 (d, *J = 12.8* Hz, 1H), 6.98 (d, *J = 13.7* Hz, 1H), 6.80 (d, *J = 16.0* Hz, 1H), 6.11 (d, *J = 9.6* Hz, 1H), 4.28 (d, *J = 14.2* Hz, 2H), 3.07 (td, *J = 7.3, 4.0* Hz, 1H), 1.18 (d, *J = 31.0* Hz, 6H). 13C NMR (126 MHz, $CDCl_3$) δ 161.9, 158.1, 152.2, 148.4, 145.6, 139.8, 114.4, 112.4, 107.3, 104.4, 94.5, 72.2, 61.4, 58.6, 24.2, 18.6. HRMS *m/z*: $[M+H]^+$ cal. 287.0914, obs. 287.0867 [153].

14. **Psoralen** (**14**). $C_{11}H_6O_3$ 1H NMR (500 MHz, $CDCl_3$) δ 7.82 (d, *J = 9.6* Hz, 1H), 7.71 (d, *J = 2.3* Hz, 1H), 7.70 (s, 1H), 7.48 (s, 1H), 6.85 (dd, *J = 2.2, 0.8* Hz, 1H), 6.39 (d, *J = 9.6* Hz, 1H). 13C NMR (126 MHz, $CDCl_3$) δ 161.1, 156.4, 152.0, 146.9, 144.1, 124.9, 119.9, 115.4, 114.6, 106.4, 99.9. HRMS *m/z*: $[M+H]^+$ calc. 187.0390, obs. 187.0381 [157].

5.2.4 Cytotoxicity screening

5.2.4.1 Cell culture

As mentioned in Section 4.2.8.1.

5.2.4.2 Cell viability assay

As mentioned in Section 4.2.8.2.

5.2.5 Molecular docking studies

As mentioned in Section 4.2.3.

5.2.6 MD simulation

As mentioned in Section 3.2.9.

5.2.7 Binding free energy calculation

As mentioned in Section 3.2.10.

5.3 Result & Discussion

The seeds and shade-dried leaves of *P. corylifolia* were coarsely grounded and extracted with a mixture of chloroform: methanol (in 1:1 ratio) via cold maceration with intermittent stirring. The cycle was repeated three times to ensure complete extraction. Both the extracts were dried and the resultant extracts were weighed to get final yield of approximately 113 g (22.6 %) for seed and approximately 48 g (9.6 %) for leaves. The extracts were then subjected to silica-gel column chromatography for isolation of phytoconstituents using *n*-hexane: ethyl acetate as mobile phase in a step gradient manner starting from 100% hexane to 100% ethyl acetate. Different fractions were collected at each increasing polarity and were pooled based on TLC profiling. The obtained fractions were subjected to repeated column chromatography to obtain pure compounds, which resulted in isolation of fourteen coumarins (Figure 5.1, Figure 5.2). The coumarins isolated from the seed and leaf extract of *P. corylifolia* include marmin (**1**, $C_{19}H_{24}O_5$), seselin (**2**, $C_{14}H_{12}O_3$), xanthotoxin (**3**, $C_{12}H_8O_4$), apterin (**4**, $C_{20}H_{24}O_{10}$), isoimperatorin

(**5**, $C_{16}H_{14}O_4$), bergapten (**6**, $C_{12}H_8O_4$), isobergapten (**7**, $C_{12}H_8O_4$), heratomine (**8**, $C_{16}H_{14}O_4$), marmesin (**9**, $C_{14}H_{14}O_4$), xanthotoxol (**10**, $C_{11}H_6O_4$), heratomol benzoate (**11**, $C_{18}H_{10}O_5$), isopsoralen (**12**, $C_{11}H_6O_3$), oxypeucedanin (**13**, $C_{16}H_{14}O_5$), and psoralen (**14**, $C_{11}H_6O_3$) (Figure 5.3). Out of fourteen isolated coumarins, compound **3**, **6**, **7**, **10**, **12**, and **14** were isolated from seeds while rest were isolated from leaves.

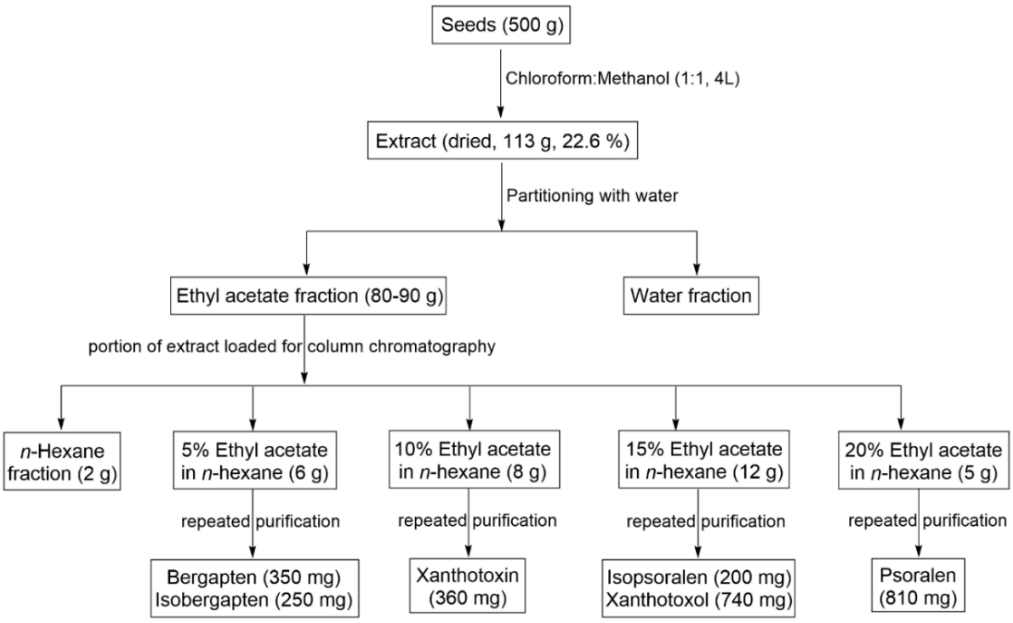


Figure 5.1 Flowchart of methodology for isolation of coumarins from seeds of *P. corylifolia*

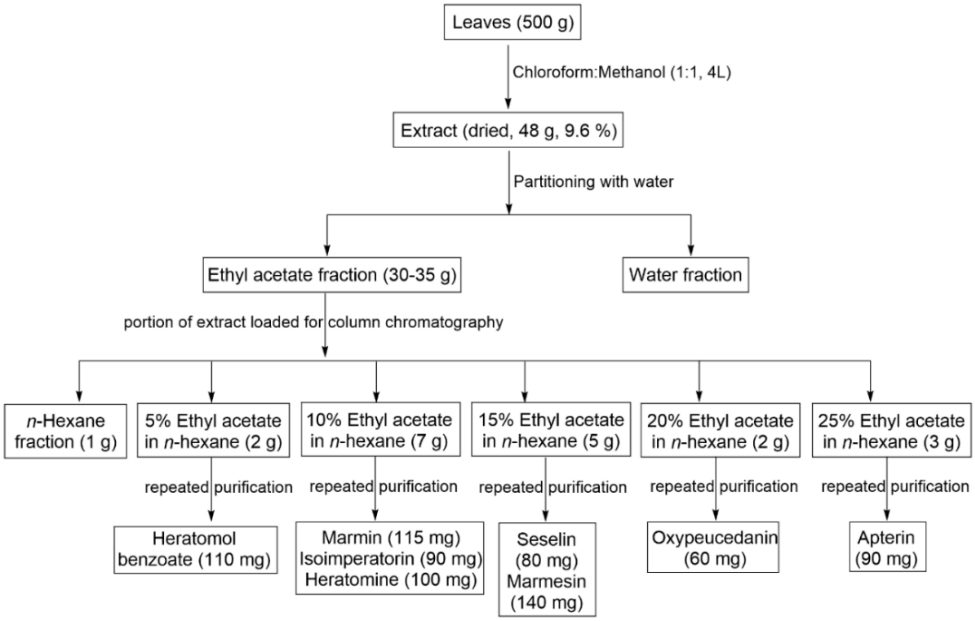


Figure 5.2 Flowchart of methodology for isolation of coumarins from leaves of *P. corylifolia*.

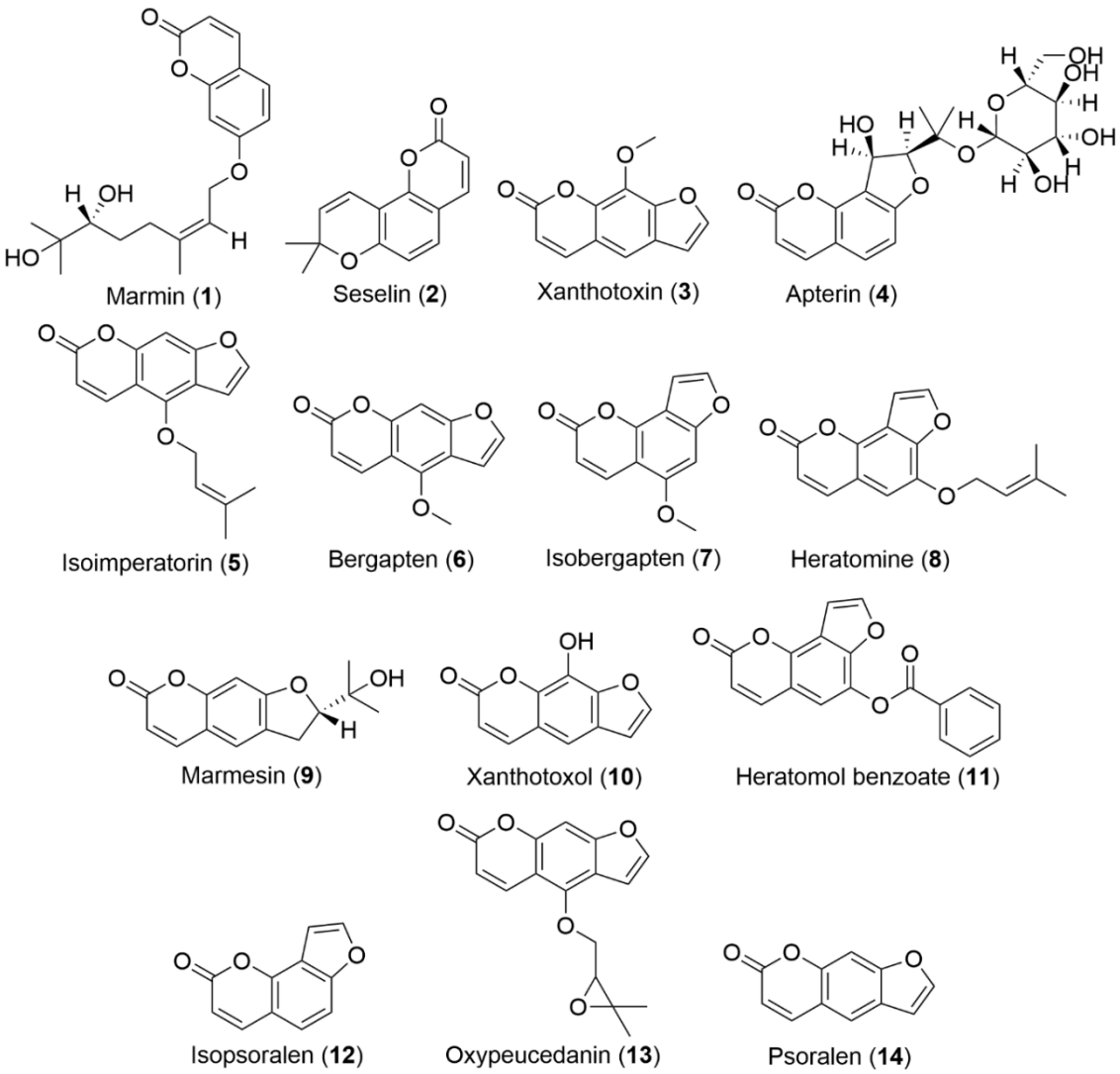


Figure 5.3 Chemical structures of isolated coumarins from *P. corylifolia*.

The presence of coumarin in *P. corylifolia* was revealed by characteristic fluorescence of crude extract and isolated compounds on TLC at a long UV wavelength of 350 nm. The common structural feature of the isolated compounds observed in 13C NMR was the presence of nine aromatic carbons with a carbonyl carbon peak of lactone at approximately δ 160 ppm. The other structure-determining feature visible in the NMR of all compounds was double-bonded carbons of the lactone ring of coumarin at approximately δ 114 and 144 ppm and corresponding H at approximately δ 6-6.5 ppm and δ 7.5-8 ppm as a doublet (J = 9.6 Hz), respectively in all the compounds. In case of furanocoumarins, the additional common features among the compounds of this class

were the presence of two peaks at δ 106 ppm and δ 146 ppm and corresponding H at δ 7-8 ppm and δ 7.5-8 ppm as a doublet (J = 9.6 Hz), respectively of the furan ring. Further, the DEPT-135 spectra was used to identify the substitution pattern of the compounds and get an idea about the -CH, -CH_2, and -CH_3 groups. The common peak pattern that was observed in the DEPT-135 spectra was the presence of peaks at approximately δ 106 ppm, δ 114 ppm, and δ 144 ppm. Moreover, for the structure determination of the isolated compounds, the HRMS spectra was recorded to get the mass-to-charge (*m/z*) ratio. After observing the basic structural scaffold, possible substitution and relating it to the m/z ratio, the possible deduced structure was confirmed by comparing the NMR spectra with the reported literature [149-157].

Here the isolation of compounds **1**, **2**, **8**, **9**, **11**, and **13** was reported for the first time from *P. corylifolia*. These furanocoumarins are reported earlier from other plant species but in this study, their presence in *P. corylifolia* was found, but in small amounts only. Interestingly, a new derivative of heratomol; benzoate ester of heratomol was obtained and confirmed by the mass and NMR spectra. Earlier studies have reported the isolation of heratomol acetate from *Platytaenia lasiocarpa* [155]. All the isolated compounds were subjected for cytotoxicity screening against the cancer cell lines.

5.3.1 Cytotoxicity screening of compounds 1-14 against cancer cells

The isolated compounds (**1-14**) were evaluated against two human cancer cell lines viz. MDA-MB-231 and A549 for their cytotoxicity screening using MTT assay to identify the active compounds (Table 5.1). The isolated coumarins showed significant cytotoxicity on the cancer cell lines. Compound **1**, **2**, and **3** unveiled their high cytotoxic potential against MDA-MB-231 cells with IC_{50} values of 0.49, 0.56, and 0.84 µM, respectively while they showed negligible or no cytotoxicity against A549 cells. Compound **10** showed high cytotoxicity with IC_{50} 0.9 µM against A549 cells while it showed moderate cytotoxicity

in MDA-MB-231 cells. Compound **6** was the only compound that displayed good cytotoxic potential in both cell lines. Whereas, psoralen possessed no cytotoxicity in both the cell lines. Overall, in MBD-MB-231 cell lines, compound **1**, **2**, **3**, **4**, **5**, **6**, and **7**, showed high cytotoxic potential with IC_{50} less than 10 µM. While in A549 cell lines, only compound **10** and **5** possessed cytotoxicity less than 10 µM, but few compounds viz. **7**, **11**, and **13**, showed cytotoxicity in the range of 10-15 µM. Based on cytotoxicity screening compounds **1**, **2**, **3**, and **10** were identified as lead molecules.

Table 5.1 Cytotoxicity screening of isolated phytoconstituents.

Ligands	**IC_{50} (µM)**	
	MDA-MB-231	**A549**
1	0.49 ± 0.06	66.93 ± 10.62
2	0.56 ± 0.14	-
3	0.84 ± 0.10	36.65 ± 1.31
4	2.04 ± 0.13	> 100
5	6.90 ± 0.78	8.67 ± 1.10
6	9.00 ± 0.52	39.44 ± 5.05
7	9.63 ± 0.37	13.25 ± 2.16
8	15.13 ± 1.34	43.31 ± 2.61
9	15.14 ± 0.48	> 100
10	24.87 ± 0.87	0.9 ± 0.001
11	25.72 ± 1.21	15.96 ± 1.83
12	31.05 ± 1.66	32.87 ± 1.33
13	37.35 ± 0.59	13.56 ± 1.15
14	-	> 100
Doxorubicin	9.84 ± 0.263	16.37 ±0.233

*Data represented in Mean ± SEM. Doxorubicin was used as positive control. DMSO (vehicle) was used as negative control/untreated control.

Further, to determine the possible mechanism of action of isolated coumarins, the molecular docking and molecular simulation studies were performed against EGFR protein. The EGFR was the target of choice as they are overexpressed in both MDA-MB-231 and A549 cells lines [158-161].

5.3.2 Molecular docking studies

Further, the molecular docking studies of compound **1**, **2**, **3**, and **10** with EGFR protein (PDB: 6DUK) was performed to study their possible binding modes with the active site of protein. The binding energy of isolated compounds was compared against Gefitinib, an FDA-approved EGFR inhibitor. Compound **10** displayed highest binding energy which was -10.62 kcal/mol and it was more gefitinib (−9.47 kcal/mol). The binding energy of rest of the compound **1**, **2**, and **3** was approximately equivalent to Gefitinib (Table 5.2). All the compounds showed important interactions with binding site residues like H-bond interaction with Lys 745 and Arg 776, charged interactions with Thr 854 and Asp 855, and hydrophobic interactions with Leu 858. The interactions of the compounds were similar to that of Gefitinib. The interaction table and diagram are presented in Table 5.2 and Figure 5.4. The docking interaction table and interaction diagram of the remaining compounds are provided in Appendix (Table A.5 and Figure A.35).

Table 5.2 The binding energy, ligand efficiency, and interactions of isolated coumarins with EGFR protein (PDB: 6DUK)

Ligands	**Binding Energy (kcal/mol)**	**Ligand Efficiency (kcal/mol)**	**Interactions with EGFR (PDB Id: 6DUK)**
10	-10.62	-0.708	A:Lys 745 (H-Bond), A:Arg 776 (H-Bond), A:Ile 789 (Hydrophobic), A:Met 790 (Hydrophobic), A:Thr 854 (Charged-positive), A:Asp 855 (Charged-negative), A:Phe 856 (H-Bond, Hydrophobic)
1	-9.75	-0.406	A:Lys 745 (H-Bond), A:Arg 776 (H-Bond), A:Leu 777 (Hydrophobic), A:Thr 854 (H-Bond, Charged-positive), A:Asp 855 (Charged-negative), A:Leu 858 (Hydrophobic)
2	-9.42	-0.554	A:Lys 745 (H-Bond, Pi-Cation), A:Ile 789 (Hydrophobic), A:Met 790 (Hydrophobic), A:Thr 854 (Charged-positive), A:Asp 855 (Charged-negative), A:Leu 858 (Hydrophobic)
3	-9.36	-0.585	A:Lys 745 (H-Bond), A:Arg 776 (H-Bond), A:Thr 854 (Charged-positive), A:Asp 855 (Charged-negative), A:Phe 856 (H-Bond, Hydrophobic), A:Leu 858 (Hydrophobic)
Gefitinib	-9.47	-0.305	A:Lys 745 (H-Bond), A:Arg 776 (H-Bond), A:Arg 841 (H-Bond), A:Thr 854 (Charged-positive), A:Asp 855 (H-Bond, Charged-negative), A:Leu 858 (Hydrophobic)

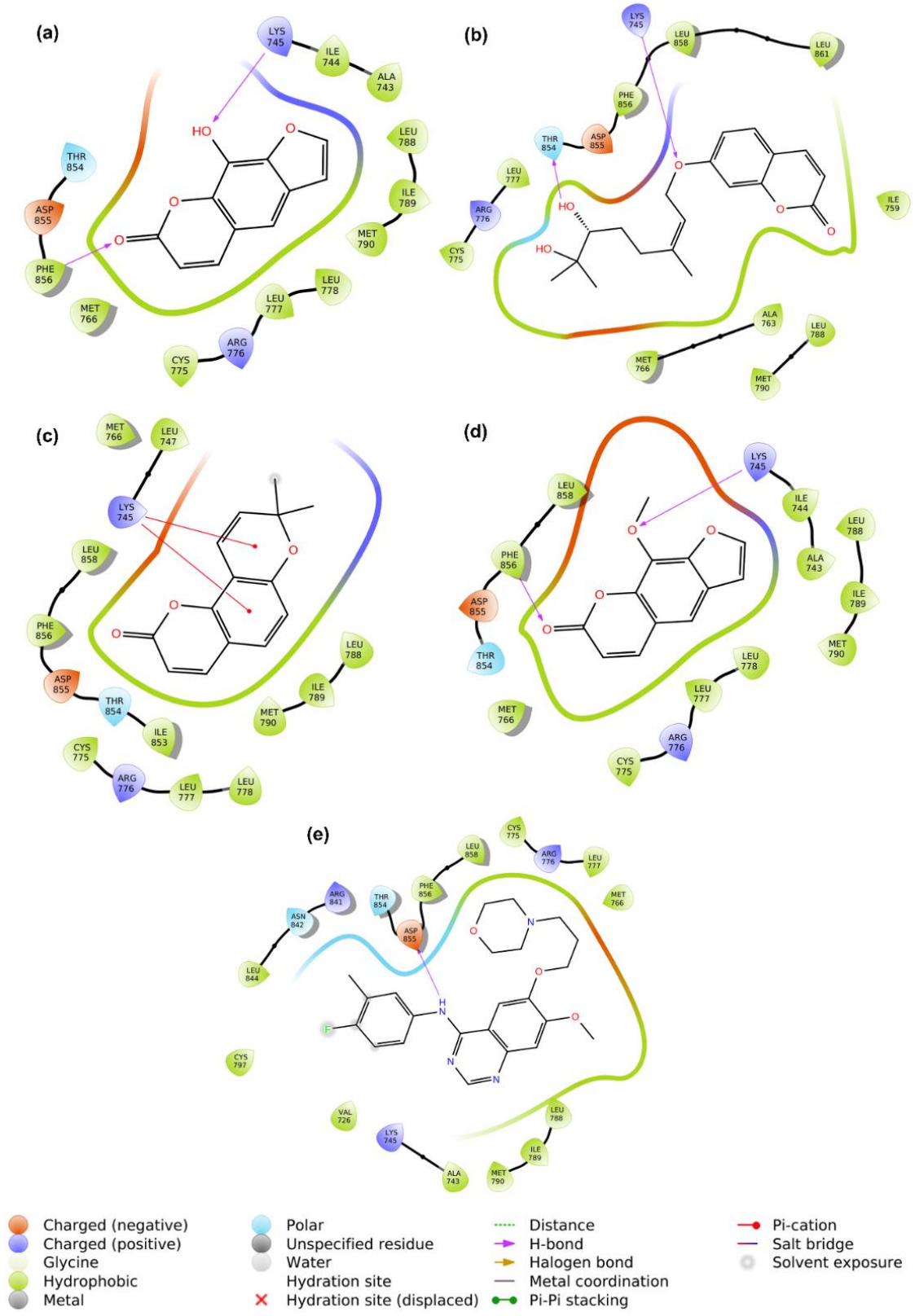


Figure 5.4 Ligand-interaction diagrams of (a) Compound 10, (b) Compound 1, (c) Compound 2, (d) Compound 3, and (e) Gefitinib with EGFR protein (PDB: 6DUK).

5.3.3 MD Studies

MD simulation was employed to ascertain the stability of four complexes: Compound **1**-EGFR complex, Compound **2**-EGFR complex, Compound **3**-EGFR complex, and Compound **10**-EGFR complex.

5.3.3.1 RMSD analysis

Figure 5.5 illustrates that the protein maintained a stable RMSD throughout the simulation run. The average protein backbone RMSD for compound **1**-EGFR complex, compound **2**-EGFR complex, compound **3**-EGFR complex, and compound **10**-EGFR complex was found to be 2.89 Å, 2.87 Å, 2.87 Å, and 2.38 Å, respectively. These values are considered to be perfectly acceptable for small and globular proteins. For the gefitinib-EGFR complex, the average protein backbone RMSD was 2.87 Å, which aligned well with that of the isolated compounds (Figure 5.5). Regarding ligand RMSD, it was observed that the values were not significantly higher than the protein RMSD, indicating that the ligand remained close to its initial binding site without diffusing away. The average ligand RMSD for compound **1**-EGFR complex, compound **2**-EGFR complex, compound **3**-EGFR complex, and compound **10**-EGFR complex was found to be 1.36 Å, 1.32 Å, 1.33 Å, and 0.84 Å, respectively. Similarly, the average ligand RMSD for the gefitinib-EGFR complex was 1.35 Å (Figure 5.6).

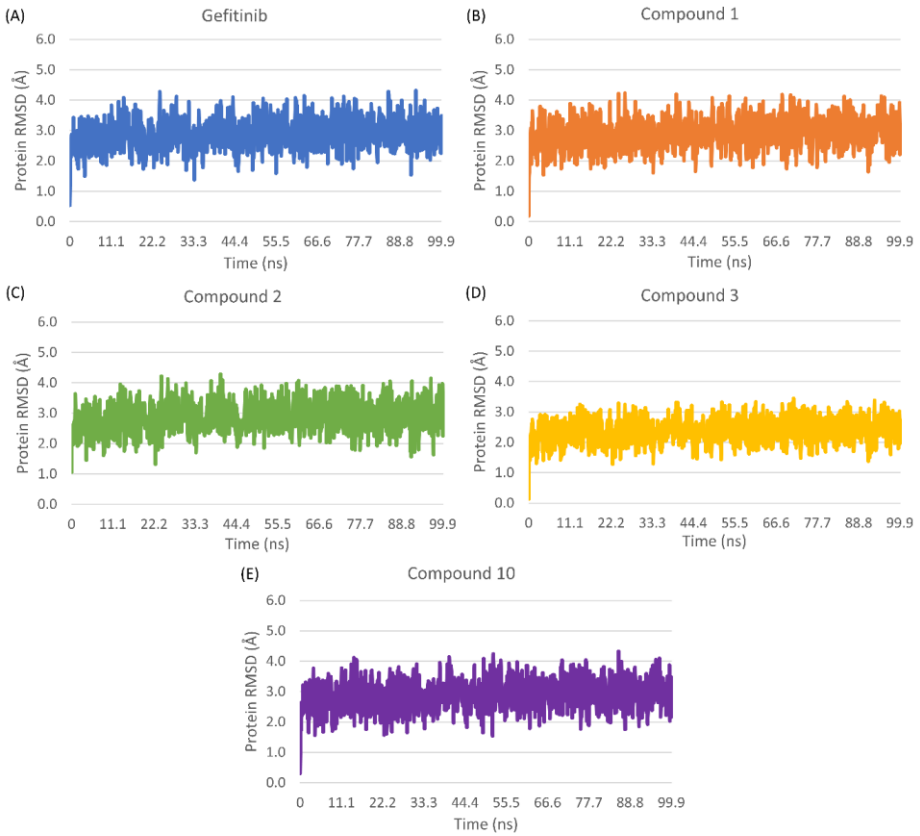


Figure 5.5 Protein RMSD of the protein-ligand complexes.

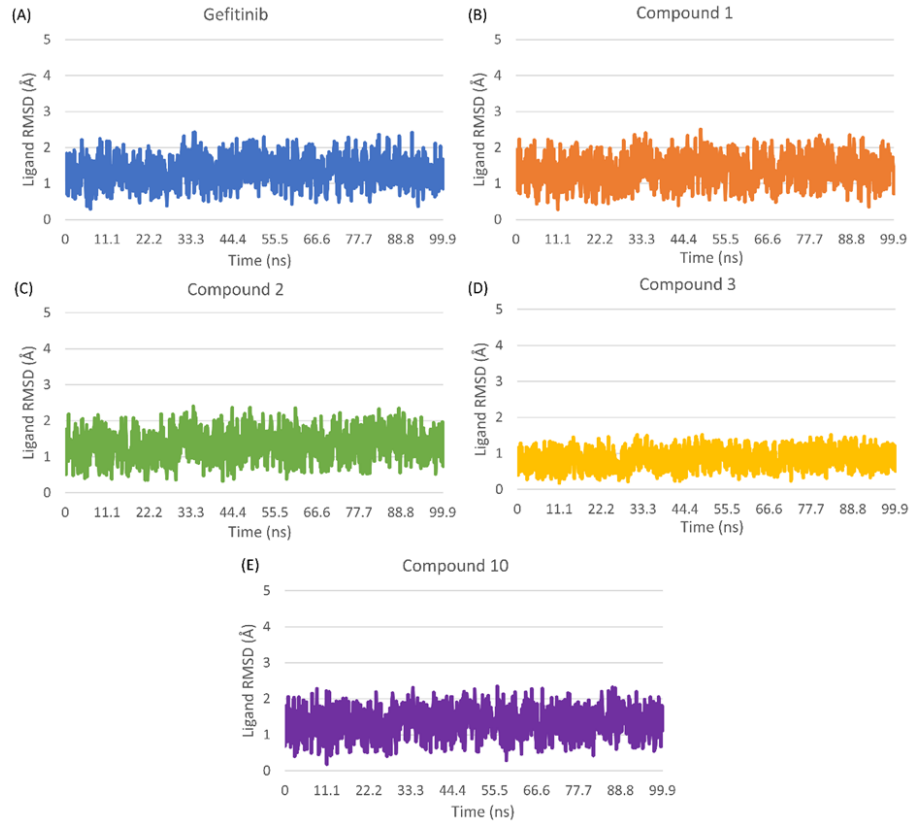


Figure 5.6 Ligand RMSD of the protein-ligand complexes.

5.3.3.2 RMSF analysis

The average RMSF values for the compound **1**-EGFR complex, compound **2**-EGFR complex, compound **3**-EGFR complex, and compound **10**-EGFR complex, and gefitinib-EGFR complex were 1.03 Å, 0.98 Å, 1.01 Å, 1.11 Å, and 0.95 Å, respectively. The residues engaged in interactions with ligands, namely Lys 745, Met 790, Thr 854, and Asp 855, exhibited higher fluctuations, while the remaining protein residues remained stable throughout the analysis (Figure 5.7).

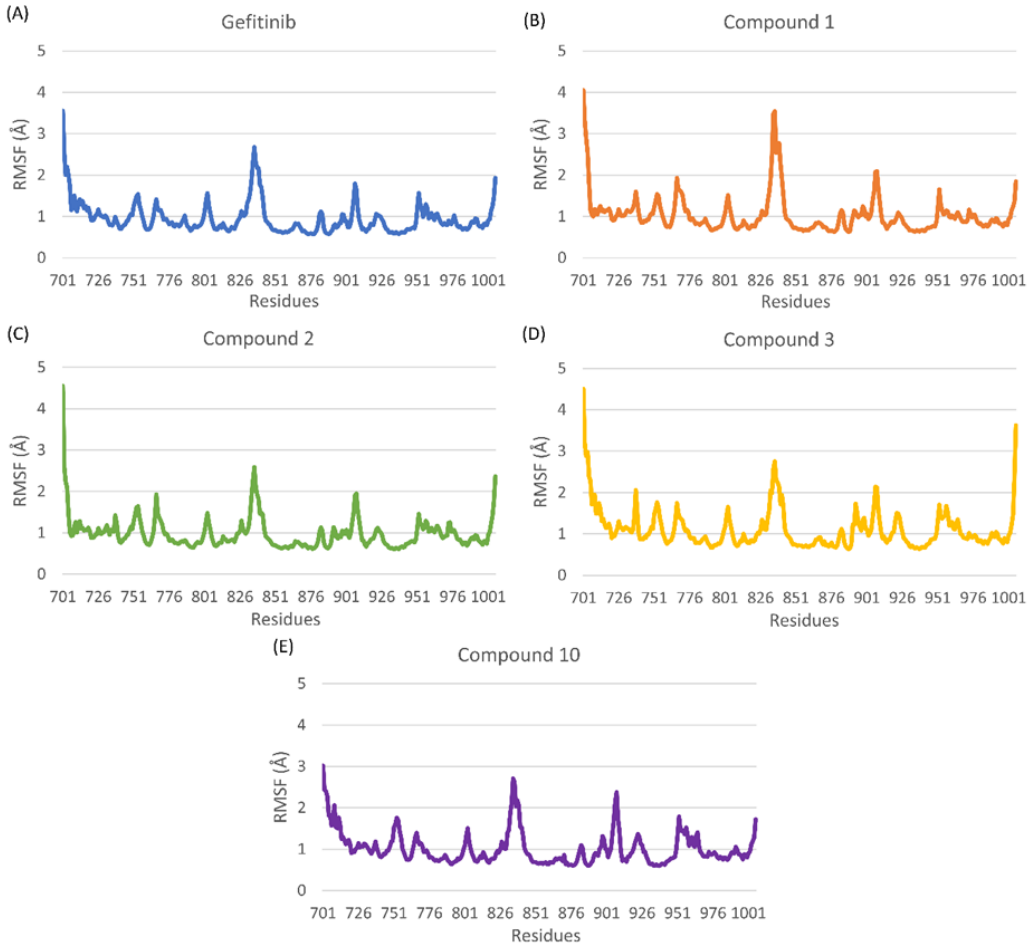


Figure 5.7 Residue wise RMSF deviations in the protein-ligand complexes.

5.3.3.3 H-Bond interaction analysis

The compounds primarily exhibited H-bond interactions with Leu 745 and Arg 776, with occasional interactions observed with Arg 841, Thr 854, and Phe 856 (Figure 5.8). These interactions were similarly observed in the docking studies.

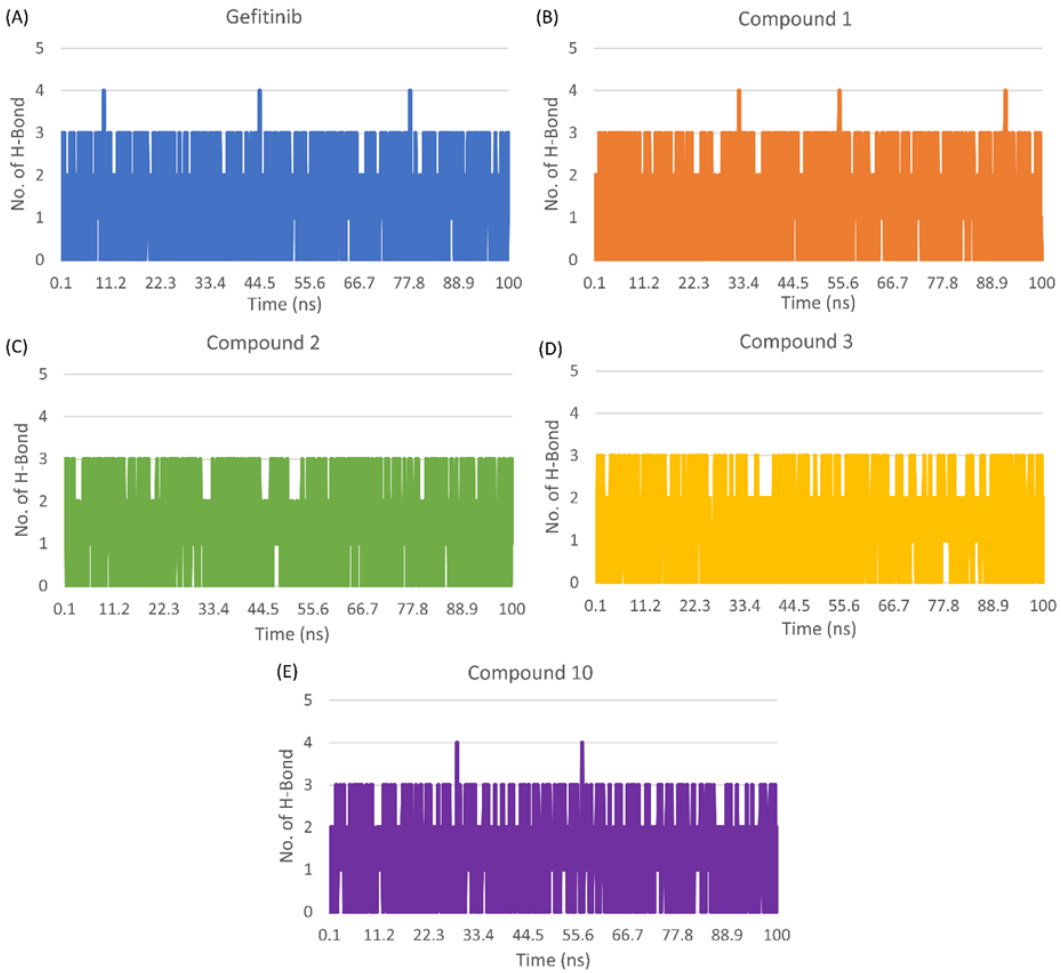


Figure 5.8 H-bond interactions between the ligand and protein during MD simulation run.

5.3.4 Binding free energy calculation

The binding free energy of complexes was calculated using, MM-GBSA and MM-PBSA methods. The results showed that all five complexes exhibited high stability, indicated by low net binding free energies obtained from both methods. However, the gas phase energy contributions (ΔE_{vdw} and ΔE_{ele}) were high, suggesting that the complex's stability primarily arose from the ligand's conformation in relation to the receptor. Compound **10**, **1**, and **3** demonstrated superior binding free energy compared to gefitinib. While, compound **2** exhibited nearly equal binding free energy to gefitinib in the GB solvation model. In the PB solvation model, all complexes exhibited favorable binding free energy

compared to Gefitinib. A comprehensive summary of the energy contributions in both the MM-GBSA and MM-PBSA assay is provided in Table 5.3 and Table 5.4.

Table 5.3 Energy contributions of complexes obtained from MM-GBSA.

Ligand	$\Delta E_{vdw(kcal/mol)}$	$\Delta E_{ele(kcal/mol)}$	$\Delta G_{GB(kcal/mol)}$	$\Delta G_{SA(kcal/mol)}$	$\Delta G_{MMGBSA(kcal/mol)}$
Gefitinib	-51.67 ± 4.93	-15.02 ± 7.17	24.29 ± 4.97	-5.47 ± 0.14	-47.88 ± 2.74
1	-51.27±2.28	-13.81± 4.51	12.61 ± 5.62	-4.72 ± 0.84	-57.19 ± 4.87
2	-55.69 ± 2.31	-28.46 ± 3.64	39.61 ± 4.01	-3.22 ± 0.39	-47.76 ± 3.72
10	-53.13± 2.38	-27.89 ± 1.85	28.43 ± 2.33	-6.21 ± 0.26	-58.80 ± 2.84
3	-51.23 ± 1.15	-35.38 ± 6.92	35.87 ± 6.09	-5.29 ± 0.16	-56.05 ± 1.87

Data expressed in Mean±S.D. (*n*=5)

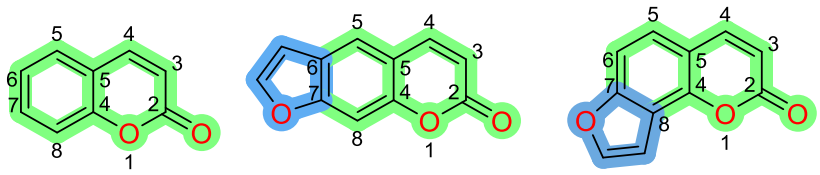
Table 5.4 Energy contributions of complexes obtained from MM-PBSA.

Ligand	$\Delta E_{vdw(kcal/mol)}$	$\Delta E_{ele(kcal/mol)}$	ΔG_{PB} (kcal/mol)	$\Delta G_{Non-Polar(kcal/mol)}$	$\Delta G_{Dispersion}$ (kcal/mol)	ΔG_{MMPBSA} (kcal/mol)
Gefitinib	-51.67±4.93	-15.02±7.17	27.69±2.90	-28.14 ± 0.35	53.87 ± 0.99	-13.27 ± 3.13
1	-51.27±2.28	-13.81± 4.51	24.73±1.42	-24.05 ± 1.29	47.49 ±4.18	-16.91 ± 2.77
2	-55.69±2.31	-28.46±3.64	40.18±0.93	-20.32 ± 3.69	49.62 ± 2.18	-14.67 ± 0.97
10	-53.13± 2.38	-27.89 ± 1.85	40.36±2.92	-40.28 ± 0.74	66.24 ± 2.96	-14.70 ± 3.58
3	-51.23±1.15	-35.38±6.92	45.02±5.81	-27.39 ± 0.66	53.56 ± 0.97	-15.44 ± 3.56

Data expressed in Mean±S.D. (*n*=5)

5.3.5 Effect of substitution

The coumarins are well known to possess cytotoxic and anticancer properties. The arrangements of rings, type of substitution, and the substituted group affects their activity. From the cytotoxicity screening of isolated coumarins (majority of them were furanocoumarin), it was observed that some groups played vital role in altering the cytotoxicity. Although there is not enough data to draw a clear picture of SAR however, some points can be noticed in general.



Coumarin | Linear Furanocoumarin | Angular Furanocoumarin

The isolated furanocoumarins could be classified as linear or angular where the cytotoxicity of the angular ones i.e., compound **12** and compound **7** is higher than the linear ones i.e., compound **14** and compound **6**. In case of linear furanocoumarins, it could be inferred that C5 and C8 substitution makes molecule more potent compared to those with no substitution. The hydroxyl substitution at C8 showed potent activity as observed in case of compound **10** while further substitution on hydroxyl group like methoxy attenuated the activity as seen in compound **3**. For C5 substituted furanocoumarins a contrasting pattern was noticed, as here O-alkyl substitution is tolerable for the activity, but higher alkyl chain substitution such as O-prenyl (compound **5**) or O-substituted prenyl (compound **13**) displayed more cytotoxicity than O-methyl substitution (compound **6**). In case of angular furanocoumarins, the substitution at C5 and C6 position is tolerable. The O-alkyl substitution with methyl (compound **7**), prenyl (compound **8**), and benzoyl groups (compound **11**) displayed similar activity.

As the compounds were limited and also their activities varied depending on the type of cancer cell lines, we weren't able to predict exact SAR but we tried to postulate a generalized scenario. Overall, it could be summarized that linear furanocoumarin with substitution at C5 and C8 and long alkyl chains tend to be more cytotoxic compared to others in their category.

5.4 Conclusion

The study resulted in isolation of fourteen coumarins from the seeds and leaves of *P. corylifolia*, including six (Marmin, Seselin, Marmesin, Heratomine, Heratomol benzoate,

and Oxypeucedanin) previously unreported coumarins from this plant. The study provides valuable insights into the cytotoxic and *in silico* molecular interactions of isolated coumarins from *P. corylifolia*. These coumarins exhibited significant cytotoxicity against cancer cell lines, particularly MDA-MB-231 cell lines. Xanthotoxol displayed the highest cytotoxicity against A549 cells, while Marmesin, Seselin, and Xanthotoxin showed potent cytotoxicity against MDA-MB-231 cells. The molecular docking studies conducted to determine the possible mechanism of action, revealed strong binding affinities of Marmin, Seselin, Xanthotoxin, and Xanthotoxol with the EGFR protein. The MD simulations confirmed the stability of coumarin and EGFR complexes, further supporting their cytotoxic potential. The *in silico* findings suggested EGFR inhibition as the mode of action, but rigorous *in vitro* and *in vivo* studies are needed to validate this mechanism. Overall, these findings highlight the potential of *P. corylifolia*-derived coumarins as promising leads for developing novel cancer therapies, especially targeting EGFR-driven malignancies.