



DỰ BÁO HOẠT TÍNH KHÁNG UNG THƯ CỔ TỬ CUNG CỦA CÁC HỢP CHẤT FLAVONOID PHÂN LẬP TỪ *Cynara scolymus* L. VÀ *Artocarpus Incisa* L. SỬ DỤNG MÔ HÌNH QSAR 2D VÀ 3D

Bùi Thị Phương Thúy¹, Phùng Văn Trung², Hoàng Thị Kim Dung²,
Trần Dương³, Phạm Văn Tất^{4*}

¹Trường Đại học Khoa Học, Đại học Huế; Trường Đại học Công Nghệ Đồng Nai

²Viện Công nghệ Hóa học, TP. Hồ Chí Minh

³Trường Đại học Sư Phạm, Đại học Huế

⁴Trường Đại học Hoa Sen, TP. Hồ Chí Minh

Tóm tắt: Trong nghiên cứu này, hai hợp chất quercetin và cynaroside thuộc nhóm dẫn xuất flavonoid được phân lập từ actiso (*Cynara scolymus* L.) và xa kê (*Artocarpus Incisa* L.) với hoạt tính gây độc tế bào ung thư Hela. Đã xây dựng được các mô hình QSAR_{MLR} (với $k = 9$) với $R^2_{\text{train}} = 0,955$; $R^2_{\text{pred}} = 0,932$ và mô hình QSAR_{ANN} I(9)-HL(5)-O(1) có $R^2_{\text{train}} = 0,996$ và $R^2_{\text{pred}} = 0,988$ để xác định hoạt tính kháng ung thư Hela pGI₅₀ của các hợp chất flavonoid. Mô hình mạng nơ ron QSAR_{ANN} được sử dụng để dự báo hoạt tính pGI₅₀ của 2 hợp chất quercetin, cynaroside và 10 hợp chất mới thiết kế. Kết quả dự báo pGI₅₀ từ mô hình QSAR_{ANN} cho sai số $MARE, \% = 1,864$ nhỏ hơn mô hình QSAR_{MLR} với $MARE, \% = 3,906$ %. Hoạt tính kháng ung thư của quercetin, cynaroside và 10 hợp chất mới thiết kế mới dự đoán từ mô hình QSAR_{ANN} phù hợp tốt với thực nghiệm và cao hơn chất mẫu tương ứng.

Từ khóa: QSAR_{MLR}, QSAR_{ANN}, hoạt tính kháng ung thư Hela

1 Đặt vấn đề

Các phương pháp phòng và trị bệnh ung thư hiện nay vẫn còn nhiều hạn chế. Nhu cầu về dược chất kháng ung thư có khả năng phòng và trị bệnh đang rất lớn nhưng khả năng đáp ứng còn hạn chế [15, 16]. Các nhà khoa học, dược học đã và đang quan tâm nghiên cứu, tìm kiếm các loại dược chất mới. Trong đó nhóm flavonoid nói chung flavone, isoflavone nói riêng là nhóm dược chất có nhiều trong thực vật với hoạt tính kháng oxi hóa, kháng ung thư, kháng viêm, ... hiệu quả [10, 27].

Các nghiên cứu thực nghiệm trên thế giới và Việt Nam đã cung cấp một cơ sở dữ liệu quý giá về nguồn dược chất trong tự nhiên, tuy nhiên các nghiên cứu thực nghiệm thuần túy còn nhiều hạn chế để tạo ra hợp chất có hoạt tính kháng ung thư hiệu quả, nhanh chóng, kinh tế [7, 8]. Các nghiên cứu lý thuyết trên thế giới nói chung, trong nước nói riêng về nhóm flavone và

* Liên hệ: vantat@gmail.com

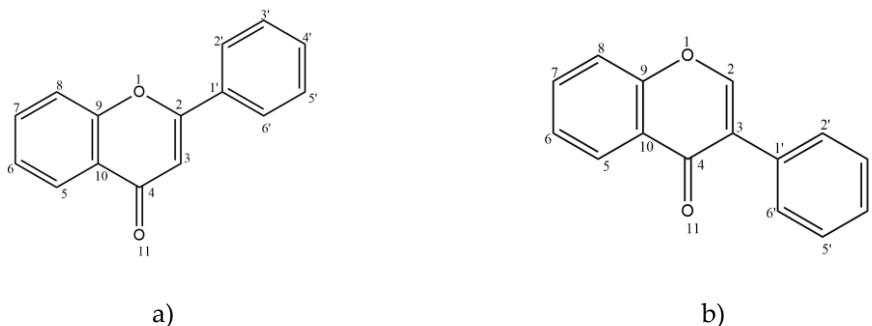
isoflavone có hoạt tính kháng ung thư cổ tử cung còn khá khiêm tốn [22, 26]. Nghiên cứu mối quan hệ cấu trúc – hoạt tính để khám phá những yêu cầu về cấu trúc cần thiết để thiết kế các dẫn xuất flavone, isoflavone mới có hoạt tính cải thiện; các nghiên cứu lý thuyết kết hợp thực nghiệm là rất cần thiết để thúc đẩy và làm tiền đề cho các nghiên cứu thực nghiệm, nhằm tìm kiếm các dược chất kháng ung thư hiệu quả [20, 22].

Trong nghiên cứu này, sử dụng các thông tin mô tả cấu trúc 2D và 3D của phân tử kết hợp các kỹ thuật phân tích hồi quy, mạng nơ ron để xây dựng các mối quan hệ định lượng cấu trúc - hoạt tính (QSAR) [23]. Các flavonoid được xây dựng và tối ưu hóa bằng các phương pháp cơ học phân tử MM+. Các tham số mô tả phân tử 2D, 3D được sử dụng để xây dựng các mô hình đa biến như hồi quy tuyến tính (MLR) và mạng nơ ron nhân tạo (ANN) [20, 22]. Xây dựng các mô hình QSAR nhằm xác định những yếu tố tham số mô tả phân tử ảnh hưởng đến tác dụng kháng ung thư cổ tử cung từ đó xác định hướng thiết kế phân tử mang lại hoạt tính cao hơn [22, 24, 26]. Trong đề tài này cũng đã tiến hành chiết tách và phân lập quercetin và cynaroside từ xa kê (*Artocarpus incisa* L.) và actiso (*Cynara Scolymus* L.), một vài kỹ thuật phân tích hóa lý cũng được sử dụng để xác định cấu trúc phân tử các dẫn xuất flavonoid. Các phân tử quercetin và cynaroside phân lập được sẽ được dự báo hoạt tính, và sử dụng làm chất mẫu để thiết kế hợp chất mới có hoạt tính cao hơn.

2 Vật liệu và phương pháp

2.1 Cơ sở dữ liệu

Tập dữ liệu mẫu sử dụng để xây dựng mô hình QSAR bao gồm 26 hợp chất flavone và isoflavone với hoạt tính $GI_{50}/\mu M$ (nồng độ ức chế 50% sự phát triển của tế bào ung thư) [20, 26], Hình 1. Giá trị GI_{50} được chuyển dạng dữ liệu sang giá trị pGI_{50} ($pGI_{50} = -\log(GI_{50})$) nhằm nâng cao chất lượng mô hình QSAR [24, 26].



Hình 1. Khung cấu trúc: a) flavone và b) isoflavone

2.2 Phương pháp phân lập và tinh chế quercetin, cynaroside

Hóa chất, thiết bị, nguyên liệu

Hóa chất

Các hóa chất dùng trong phân lập và tinh chế flavonoid từ xa kê (*Artocarpus incisa* L.) và actiso (*Cynara Scolymus* L.) bao gồm: hạt silica gel cỡ hạt 40-60 μm dùng cho pha thường và hạt silica gel pha đảo Rp18 cỡ hạt 30-50 μm ; sắc ký lớp mỏng (SKLM) được thực hiện trên bản nhôm silica gel 60F₂₅₄ (Merck) tráng sẵn có độ dày 0,2 mm; sắc ký cột (SKC) pha thường: sử dụng nhiều cột có kích cỡ khác nhau tùy thuộc vào số lượng mẫu, chất hấp phụ là silica gel cỡ hạt từ 0,0040-0,063 mm (240-400 mesh) của các hãng Himedia, Scharlau, Merck; dung môi dùng cho quá trình thí nghiệm gồm: hexane, chloroform (CHCl_3), methanol (MetOH), ethyl acetate (EtOAc), ethanol (EtOH), nước cất; thuốc thử hiện hình các vết chất hữu cơ trên bản mỏng là $\text{H}_2\text{SO}_4/\text{EtOH}$ của Merck [2, 6].

Thiết bị

Một số thiết bị sử dụng bao gồm: Đèn UV tử ngoại cầm tay, bước sóng 254 nm và 365 nm hiệu UVITEC dùng để hiện màu thuốc thử trên bản mỏng; cột sắc kí đường kính từ 2-5,5 cm dùng trong sắc kí cột; máy sắc ký trung áp [6, 10]. Cùng các thiết bị phân tích hóa lý $^1\text{H-NMR}$ (500 MHz), $^{13}\text{C-NMR}$ (125 MHz), COSY, DEPT, HSQC, HMBC trên máy Bruker AM500 FT-NMR Spectrometer.

Nguyên liệu

Lá actiso và lá xa kê dùng cho quá trình phân lập sau khi thu hái được loại bỏ phần đất cát, sâu bệnh, rửa sạch, để ráo sau đó phơi hoặc sấy khô, nghiền nhỏ để sử dụng cho quá trình chiết theo kỹ thuật lấy mẫu [14]. Mẫu thực vật được định danh và lưu tiêu bản tại Viện Công nghệ Sinh học và Môi trường, trường Đại học Tây Nguyên.

Phân lập quercetin và cynaroside

Lá actiso khô 2,2 kg đem ngâm trong dung môi EtOH 96%, lọc bỏ bã, lấy phần dịch chiết EtOH đem cô loại dung môi dưới áp suất kém thu được cao EtOH [5, 17]. Từ cao EtOH chiết lỏng lỏng với hệ dung môi có độ phân cực tăng dần: n-hexan, CHCl_3 , EtOAc, MetOH. Loại bỏ dung môi từ các dịch chiết thu được cao tương ứng [8, 14]. Cao EtOAc (190g) đem SKC silica gel pha thường với hệ dung môi CHCl_3 : MetOH có thể tích CHCl_3 thay đổi từ 100%-0%. Dựa vào kết quả SKLM để gom các phân đoạn giống nhau. Kết quả quá trình này thu được 12 phân đoạn. Lấy phân đoạn thứ 7 (17g) tiếp tục SKC silica gel pha thường lần 2 với hệ dung môi rửa

giải CHCl_3 : MetOH theo tỉ lệ thể tích CHCl_3 thay đổi từ 90% - 40% thu được 6 phân đoạn từ [14, 19]. Lấy phân đoạn thứ 2 (2,7g) tương ứng tỉ lệ thể tích CHCl_3 : MetOH là được tiếp tục SKC silica gel pha thường lần 3 trong hệ CHCl_3 : MetOH với tỉ lệ thể tích CHCl_3 giảm từ 80% đến 50% thu được 5 phân đoạn. Tại phân đoạn ứng với tỉ lệ CHCl_3 : MetOH là (60:40) thu được chất kết tinh khá sạch, màu trắng, đem rửa sạch và kiểm tra bằng SKLM. Lượng chất thu được trong phân đoạn này là 12 mg, kí hiệu CSL1.

Lá xa kê khô 1,7 kg được trích ly với EtOH 96% bằng phương pháp ngâm, lọc bỏ bã, phần dịch chiết sau khi loại dung môi thu được cao EtOH. Từ cao EtOH chiết lỏng lỏng với hệ dung môi n-hexan, CHCl_3 , EtOAc, MetOH [7, 14, 27], lấy phần dịch chiết đem cô loại dung môi thu được cao tương ứng. Dựa vào kết quả SKLM, lấy cao EtOAc (120g) SKC silica gel pha thường với hệ dung môi CHCl_3 : MetOH có tỉ lệ thể tích CHCl_3 thay đổi từ 100%-0% [7, 27]. Dựa vào kết quả SKLM để gom các phân đoạn giống nhau. Kết quả quá trình này thu được 7 phân đoạn. Lấy phân đoạn thứ 5 (10g) tiếp tục SKC silica gel pha thường lần 2 với hệ dung môi CHCl_3 : MetOH với các tỉ lệ thể tích 95 : 5; 90 : 10; 85 : 15; 80 : 20; 75 : 27; 70 : 30; 65 : 35; 60 : 40 thu được 8 phân đoạn (B1-B8) [14, 19]. Tại phân đoạn thứ 4 với tỉ lệ CHCl_3 : MetOH (80 : 20) thu được 7mg chất kết tinh khá sạch, màu trắng, đem rửa sạch và kiểm tra bằng SKLM, ký hiệu hợp chất là AIL1.

Thử hoạt tính gây độc tế bào ung thư - Sulforhodarmine – B

Các hợp chất CSL1 và AIL1 được thử hoạt tính kháng ung thư cổ tử cung tại Phòng Sinh học phân tử, trường đại học Khoa học Tự nhiên, thành phố Hồ Chí Minh. Quá trình thử nghiệm được thực hiện theo phương pháp so màu (SRB). Tế bào ung thư Hela được nuôi cấy trong môi trường E'MEM (Sigma), bổ sung phenol red 0,4%, L-glutamine 200 mM, NaHCO_3 7,5%, HEPES 1M, Penicillin - Streptomycin 200 IU/mL, Amphotericin - B 0,1% và huyết thanh phôi thai bò 5%, trên đĩa 96 giếng với mật độ 104 tế bào/giếng, trong điều kiện vô trùng, ở 37°C, 5% CO_2 . Sau 24 giờ, tế bào nuôi cấy được xử lý với dung dịch flavonoid 100 $\mu\text{g/mL}$. Sau 48h, tế bào Hela được nhuộm SRB, đo mật độ quang (optical density – OD) ở bước sóng 490nm và 620nm. Chất đối chứng dương sử dụng là Camptothecin 0,01 $\mu\text{g/mL}$, chất đối chứng âm là nước cất vô trùng 10% pha trong môi trường và DMSO 0,25%. Tỉ lệ ức chế tăng trưởng tế bào tính theo công thức (1).

$$I(\%) = (1 - (\text{ODThử nghiệm}/\text{ODChứng})) \times 100\% \quad (1)$$

Tất cả các thí nghiệm được lặp lại 3 lần. Số liệu được xử lý bằng phần mềm Microsoft Office Excel 2007.

2.3 Xây dựng mô hình QSAR

Tính toán các tham số 2D, 3D

Sử dụng phương pháp cơ học phân tử MM+ để tối ưu hóa cấu trúc phân tử flavonoid. Các tham số mô tả cấu trúc 2D, 3D được tính toán bằng phương pháp cơ học phân tử [12], phương pháp Grid [4, 11], phương pháp Gasteiger trên hệ thống QSARIS [4].

Xây dựng mô hình tuyến tính QSARMLR

Hồi quy tuyến tính đa biến MLR [13] là thường được sử dụng xây dựng mô hình QSAR. Biểu thức tổng quát của một phương trình MLR: $y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n$. Với y là biến hoạt tính pGI₅₀; các tham số x_1, x_2, \dots, x_n là các tham số mô tả phân tử 2D và 3D; hệ số hồi quy tương ứng a_1, a_2, \dots, a_n , tương ứng, a_0 là hệ số tự do của mô hình. Chất lượng của mô hình MLR được xác định bằng các giá trị hệ số xác định bội R^2 luyện (R^2_{train}) (2), giá trị R^2 dự đoán (R^2_{pred}), sai số (SE) [21, 22].

$$R^2_{train} = 1 - \frac{\sum (y_{tn} - y_{tt})^2}{\sum (y_{tn} - \bar{y}_m)^2} \quad (2)$$

Trong các phương trình trên, y_{tn} và y_{tt} là giá trị pGI₅₀ thực nghiệm và tính toán từ mô hình; \bar{y}_m là giá trị trung bình của pGI₅₀ thực nghiệm. Giá trị R^2_{pred} được tính toán tương tự bằng công thức (2) [22, 28].

Xây dựng mô hình mạng nơ ron QSARANN

Mạng nơ ron là một công cụ xây dựng mối liên hệ phi tuyến tính giữa cấu trúc và hoạt tính [23, 25]. Mô hình QSAR_{ANN} được xây dựng bằng kỹ thuật neuro-fuzzy với thuật toán di truyền trong hệ thống INForm [13]. Mạng nơ ron cấu tạo: I(j) – HL(m) – O(n). Trong đó: I là lớp nhập với j là số tham số mô tả phân tử trong mô hình tuyến tính QSAR_{MLR}; HL là lớp ẩn, m là số nút của lớp ẩn, O là lớp xuất, n là số nơ ron lớp xuất, tức là hoạt tính pGI₅₀ [13].

3 Kết quả và thảo luận

3.1 Xây dựng mô hình QSAR_{MLR}

Các phân tử flavonoid được tối ưu hoá bằng thuật toán Polak-Ribiere với gradient là 0,01 kcal.mol⁻¹. Các mô hình tuyến tính QSAR_{MLR} gồm các tham số cấu trúc 2D, 3D đưa ra trong Bảng 1, được tính toán trên hệ thống QSARIS [4]. Trong các mô hình QSAR_{MLR} đã xây dựng từ

nhóm luyện (26 hợp chất) xem mục 2.1. Khả năng dự đoán của các mô hình QSAR_{MLR} được đánh giá chất lượng giữa các mô hình bằng cách so sánh giá trị thống kê R^2_{train} , R^2_{pred} và MSE [12, 23]. Sự thay đổi các tham số mô tả 2D và 3D tương ứng trong các mô hình QSAR_{MLR} trong Bảng 1 bằng kỹ thuật phân tích hồi quy từng bước [12]. Các tham số mô tả 2D, 3D được chọn dựa trên cơ sở thay đổi các giá trị thống kê R^2_{train} và giá trị trung bình sai số MSE [25]. Trong Bảng 1 đưa ra 9 mô hình phù hợp nhất. Trong đó mô hình QSAR_{MLR} với $k = 9$ tốt nhất được chọn dự đoán.

Bảng 1. Các mô hình QSAR_{MLR} ($k = 2$ đến 10) với các giá trị R^2_{train} , R^2_{pred} và MSE

1	Tham số mô tả phân tử 2D, 3D trong các mô hình	MSE	R^2_{train}	R^2_{pred}
2	xvc3, Dipole	0.0628	0.657	0.630
3	xvp3, xvpc4, Dipole	0,0343	0.820	0,798
4	xvp6, xvpc4, Dipole, LogP	0.0267	0.866	0.843
5	xp8, Dipole, Volume, knotpv, LogP	0,0226	0.891	0,867
6	xp8, xvp8, Dipole, knotpv, SHBa, LogP	0,0188	0,914	0,890
7	xp8, xvp8, Dipole, Hmaxpos, knotpv, SHBa, LogP	0,0153	0,933	0,910
8	xp6, SaaCH, k2, nvx, knotp, knotpv, SHBa, LogP	0,0154	0,936	0,910
9	xp8, ABSQ, SsOH, Hmaxpos, phia, knotp, knotpv, SHBa, LogP	0,0115	0,955	0,932
10	xp8, ABSQ, SsOH, Hmaxpos, nvx, phia, knotp, knotpv, SHBa, LogP	0,0122	0,955	0,928

Các mô hình QSAR_{MLR} (với $k = 2$ đến 10) được sắp xếp một cách trật tự với sự thay đổi các giá trị R^2_{train} , R^2_{pred} và MSE . Trong Bảng 1 các mô hình QSAR_{MLR} (với $k = 8$ đến 10) là các mô hình có giá trị R^2_{pred} cao hơn cả. Cụ thể, trong mô hình QSAR_{MLR} với $k = 9$ có R^2 cao nhất là 0,954 và R^2_{pred} là 0,932. Vì vậy, lựa chọn ba mô hình (với k là 8, 9 và 10) để xác định giá trị phần trăm đóng góp của các tham số mô tả 2D, 3D đối với hoạt tính kháng ung thư. Các giá trị phần trăm đóng góp trung bình toàn cục $GMP_{m,x_k},\%$ (3) [22] và các giá trị thống kê khác đối với các mô hình (với k là 8, 9 và 10), được đưa ra trong Bảng 2.

$$GMP_{m,x_k},\% = \frac{1}{n} \sum_{n=1}^3 MP_{m,x_k}$$

$$\text{với } MP_{m,x_k},\% = \frac{1}{m} \sum_{m=1}^{26} P_{m,x_k}; P_{m,x_k},\% = 100 \frac{|b_k x_k|}{C_{\text{total}}} \text{ và } C_{\text{total}} = \sum_{i=1}^k |b_{m,k} x_{m,k}| \quad (3)$$

trong đó: C_{total} là tổng phần trăm đóng góp của các tham số mô tả trong một hợp chất. P_{mXk} , % là phần trăm đóng góp trung bình của mỗi tham số mô tả đối với m hợp chất. GMP_{mXk} , % là phần trăm đóng góp trung bình toàn cục của mỗi tham số mô tả phân tử đối với các mô hình $n = 3$.

Giá trị phần trăm đóng góp trung bình toàn cục GMP_{mXk} , % trong **Bảng 2**, đã mô tả mức độ ảnh hưởng của các tham số mô tả 2D, 3D của các flavonoid. Sự sắp xếp mức độ ảnh hưởng quan trọng của các tham số 2D, 3D trong các mô hình QSAR_{MLR} theo thứ tự là GMP_{mXk} , % là SaaCH < SsOH < SHBa < xp6 < Hmaxpos < LogP < k2 < knotp < xp8 < knotpv < phia < nvx < ABSQ.

Bảng 2. Các giá trị **thống kê** của mô hình QSAR_{MLR} (với $k = 8, 9, 10$) và các giá trị MP_{mXk} , %; GMP_{mXk} , %

Biến số X_k	QSAR _{MLR}			MP_{mXk} , %			GMP_{mXk} , %
	$k = 8$	$k = 9$	$k = 10$	$k = 8$	$k = 9$	$k = 10$	
R^2_{train}	0.936	0.955	0.955				
R^2_{pred}	0.910	0.932	0.928				
SE	0,124	0,107	0,110				
Hằng số	-13,133	3,044	2,946				
k2	-4,926	-	-	23,638			7,879
xp6	-2,899	-	-	6,399			2,133
xp8	-	4,473	4,282		14,301	12,181	8,827
nvx	3,487	-	0,036	45,229		1,138	15,456
knotp	-8,350	-2,673	-2,794	9,072	7,879	7,325	8,092
knotpv	34,40	15,54	15,917	9,681	11,878	10,813	10,790
SHBa	-0,225	0,019	0,016	4,398	1,004	0,770	2,057
Hmaxpos	-	-1,203	-1,201		3,439	3,060	2,166
SaaCH	-0,166	-	-	1,490			0,497
ABSQ	-	-5,982	-5,939		33,101	29,255	20,785
SsOH	-	0,359	0,358		2,924	2,609	1,844
Phia	-	3,067	2,999		21,167	18,413	13,193
LogP	0,050	0,839	0,793	0,093	4,308	3,624	2,675

Từ Bảng 2, chúng tôi thấy các tham số mô tả ảnh hưởng chính đến hoạt tính pGI₅₀ là tổng điện tích nguyên tử trong phân tử (ABSQ), chỉ số đỉnh của khung phân tử (nvx), giá trị (phia) và các tham số hình học khác như knotpv, xp8, knotp, chỉ số Kappa 2 (k2) cũng như logP đã thể hiện mức độ ảnh hưởng có ý nghĩa nhất trong mỗi phân tử. Ngoài ra, các tham số này cũng đặc trưng cho nhóm carbonyl C₄ = O₁₁ và nguyên tố O₁. Các nguyên tố này có các cặp electron tự do với electron π của liên kết C₂ = C₃ và C₄ = O₁₁ tạo hệ liên hợp. Nhóm carbonyl C₄ = O₁₁ thể hiện toàn bộ phản ứng tự nhiên của nhóm của nhóm cacbonyl [20, 26]. Vì vậy, các tham số mô tả phân tử được mô tả ở đây đã chứng minh được về mặt định lượng tổng điện tích trên phân tử bởi giá trị $GMP_{m\chi_k, \%}$ đã phù hợp với nhận định trong nghiên cứu của Wang [26].

Tuy nhiên để thiết kế một phân tử mới từ các chất mẫu quercetin và cynaroside được phân lập từ tự nhiên là nhiệm vụ của bài báo này. Từ kết quả xây dựng các mô hình QSAR_{MLR} đã giúp định hướng thiết kế dẫn xuất mới có hoạt tính cao hơn hoạt tính tự nhiên bằng cách gắn nhóm thế vào phân tử để làm tăng giá trị ABSQ và giảm nvx và knotpv. Điều này sẽ dẫn đến pGI₅₀ sẽ giảm, có nghĩa là hoạt tính kháng ung thư sẽ tăng. Chúng tôi chọn lựa các vị trí thuận lợi để gắn thêm các nhóm thế mới tạo ra các dẫn xuất mới là các vị trí còn trống tại C₆, C_{3'} sẽ dễ dàng là tăng điện tích toàn phân tử dễ dàng do đặc tính cấu trúc của phân tử. Điều này làm cơ sở định hướng cho các nhà nghiên cứu thực nghiệm tổng hợp được chất. Hai hợp chất quercetin và cynaroside phân lập được từ lá xa kê và actiso có thể được dự đoán pGI₅₀ từ mô hình (4):

$$pGI_{50} = 3,044 + 4,473 \times xp8 - 5,982 \times ABSQ + 0,359 \times SsOH - 1,203 \times Hmaxpos + 3,067 \times phia - 2,673 \times knotp + 15,540 \times knotpv + 0,019 \times SHBa + 0,839 \times LogP \quad (4)$$

Với $n = 26$; $R^2_{train} = 0,955$; $R^2_{pred} = 0,932$; $SE = 0,107$; $F = 42,182$.

Khả năng dự đoán hoạt tính pGI₅₀ của mô hình QSAR_{MLR} (4) này được minh chứng qua giá trị thống kê R^2_{train} , R^2_{pred} lớn và sai số SE nhỏ, đáp ứng khả năng dự đoán hoạt tính kháng ung thư của mô hình.

3.2 Xây dựng mô hình QSAR_{ANN}

Mô hình mạng nơ ron được xây dựng bằng kỹ thuật neuro-fuzzy với thuật toán di truyền bằng mô hình INForm [13]. Mô hình QSAR_{ANN} có kiểu kiến trúc mạng tối ưu là I(9)-HL(5)-O(1). Cấu trúc mạng nơ ron này được xây dựng từ lớp nhập I(9) với 9 nơ ron tương ứng với 9 tham số mô tả phân tử xp8, ABSQ, SsOH, Hmaxpos, phia, knotp, knotpv, SHBa và logP; lớp ẩn HL(5) với 5 nơ ron và lớp xuất O(1) với 1 nơ ron là hoạt tính sinh học pGI₅₀. Thuật toán sử dụng cho quá trình nhận thức của mạng là lan truyền ngược sai số. Hàm truyền từ mỗi nơ ron là hàm

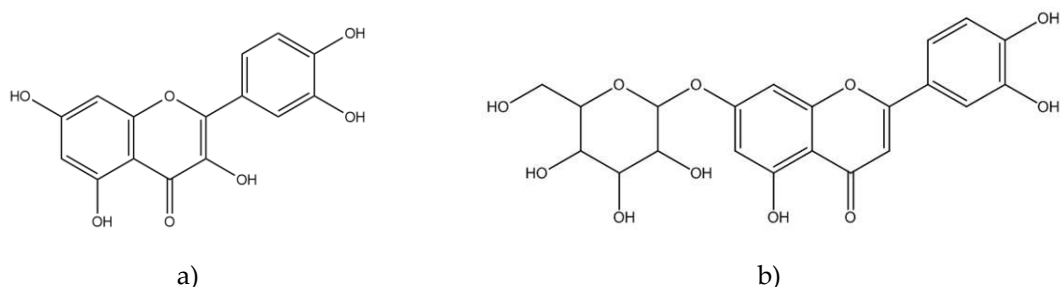
hyperbol; tham số luyện mạng nơ ron được sử dụng cho quá trình luyện mạng là tham số tốc độ luyện 0.7 và tốc độ học 0,01; sau khi luyện mạng nơ ron kết quả sai số tổng cộng là 0,0003017 ; sai số dự đoán là 0,00001 với 300000 vòng lặp, $R^2_{\text{train}} = 0,996$ và $R^2_{\text{pred}} = 0,988$.

3.3 Thực nghiệm xác định cấu trúc flavonoid tự nhiên

Các hợp chất flavonoid trong phân đoạn CSL1 và AIL1 được phân lập từ các kỹ thuật tách chiết và chạy sắc ký giấy, sắc ký bản mỏng theo mục 2.2.2 , được xác định cấu trúc bằng phương pháp đo phổ $^1\text{H-NMR}$ (500 MHz), $^{13}\text{C-NMR}$ (125 MHz), COSY, DEPT, HSQC, HMBC được ghi trên máy Bruker AM500 FT-NMR Spectrometer [1, 3] tại Viện Hóa học, Viện Khoa học và Công nghệ Việt Nam. Độ dịch chuyển hoá học δ (ppm), hằng số tương tác J tính bằng Hz, chất nội chuẩn là TMS (Tetrametyl Silan).

AIL1 phân lập từ *Artocarpus Incisa* L. theo mục 2.2.2, điểm nóng chảy 314 – 316°C. Phổ $^1\text{H-NMR}$ (DMSO- d_6 , 500 MHz, δ ppm) kết hợp với phổ HSQC, HMBC: δ 6,26 (1H, d, J = 1,5 Hz, H₆); δ 6,52 (1H, s (t), H₈); δ 7,82 (1H, d, J = 1,5 Hz, H₂); δ 7,68 (1H, dd, J = 8,5 và 2 Hz, H₅); δ 6,98 (1H, d, J = 8,5 Hz, H₆); δ 12,16 (1H, s). Phổ $^{13}\text{C-NMR}$ (DMSO- d_6 , 125 Hz) kết hợp với phổ DEPT, HSQC, HMBC: δ C 146,9 (C₂); δ 136,6 (C₃); δ C 176,5, (C₄); δ 162,2 (C₅); δ 99,1 (C₆); δ 164,9 (C₇); δ 94,4 (C₈); δ 157,7 (C₉); δ 104,0 (C₁₀); δ 121,4 (C₁); δ 116,1 (C₂); 145,7 (C₃); δ 148,2 (C₄); δ 115,6 (C₅); δ 123,7 (C₆) [5]. So sánh với dữ liệu phổ chuẩn [9] xác định được cấu trúc AIL1 là quercetin, Hình 2.

CSL1 phân lập và tinh chế từ *Cynara scolymus* L. theo mục 2.2.2. điểm nóng chảy 256 – 257°C. Phổ $^1\text{H-NMR}$ (500 MHz, DMSO, δ ppm): δ 6,39 (s, H₃); 6.44 (d, J = 2 Hz, H₆); 6,78 (d, J = 2 Hz, H₈); 5,08 (1H, d, J = 7,5 Hz, H_{1''}); 7,41 (d, J = 8.5 Hz); 6,90 (d, J = 8,5 Hz, H₃); 7,44 (dd, J = 8,5 Hz, J = 2, H₂); 3,489 (1H, m, H_{2''}); 3,476 (1H, m, H_{4''}); 3,466(1H, m, H_{3''}); 3,500(1H, m, H_{5''}); 3,725 (1H, dd, J = 12,3; 2,5 HZ, H_{6''}); 3,702 (1H, dd, J = 12,4; 6,3 HZ, H_{6''}); 12,8 (C₅OH); 3,466–3,725 (7H, m, glucose protons). Phổ $^{13}\text{C-NMR}$ (DMSO, δ ppm, 125 MHz), DEPT: δ 164,5(C₂); 105,3(C₃); 181,785 (C₄); 161,2 (C₅); 100,2 (C₆); 162,9 (C₇); 95 (C₈); 156,9 (C₉); 103,1 (C₁₀); 121,4 (C₁); 113,6 (C₂); 145,7 (C₃); 149,9 (C₄); 115,9 (C₅); 119,2 (C₆); 100,0 (C_{1''}); 76,4 (C_{2''}); 77,2 (C_{3''}); 69,5 (C_{4''}); 78 (C_{5''}); 60,6 (C_{6''}). Tương tác giữa C và H trong phổ HMBC và HSQC thể hiện tương tác: H₆-C₅-C₇-C₈-C₁₀; H₈-C₆-C₇-C₉-C₁₀; H₂-C₂-C₁-C₃-C₄-C₆; H₅-C₁-C₃-C₄-C₆; H₆-C₂-C₁-C₂-C₄-C₅. So sánh với dữ liệu phổ chuẩn [18] xác định được cấu trúc **CSL1** là cynaroside. Hình 2 mô tả cấu trúc của hợp chất quercetin và cynaroside trong phân đoạn AIL1 và CSL1 tương ứng đã được thử nghiệm hoạt tính *in vitro* gây độc tế bào ung thư Hela tại phòng sinh học phân tử, trường đại học Khoa học Tự nhiên theo kỹ thuật mô tả ở mục 2.2.3. Kết quả thử nghiệm cho ở Bảng 3.



Hình 2. Cấu trúc hợp chất a) quercetin, b) cynaroside

3.4 Dự đoán hoạt tính của dẫn xuất mới

Các hợp chất quercetin và cynaroside được phân lập từ *Artocarpus Incisa* L và *Cynara scolymus* L cùng với 6 hợp chất flavonoid (fla) trong nhóm kiểm tra được tách ngẫu nhiên từ dữ liệu mẫu [20, 26] đã được dự đoán pGI_{50} từ mô hình $QSAR_{MLR}$ (4) và $QSAR_{ANN}$ I(9)-HL(5)-O(1), được đưa ra ở Bảng 3.

Table 3. Bảng 3. So sánh hoạt tính pGI_{50} thực nghiệm và dự đoán của các chất trong nhóm kiểm tra và hai hợp chất quercetin và cynaroside từ các mô hình $QSAR_{MLR}$ (4) và $QSAR_{ANN}$ I(9)-HL(5)-O(1)

Hợp chất	$pGI_{50,tn}$	$pGI_{50,pred}$		ARE,%	
		$QSAR_{MLR}$	$QSAR_{ANN}$	$QSAR_{MLR}$	$QSAR_{ANN}$
Fla-1	5,6990 [20, 26]	5,3879	5,7371	5,4589	0,6685
Fla-11	5,6990 [20, 26]	5,9188	6,0478	3,8568	6,1204
Fla-24	5,6198 [20, 26]	5,9316	5,7662	5,5482	2,6051
Fla-25	5,6383 [20, 26]	5,8627	5,6843	3,9799	0,8158
Fla-26	5,6990 [20, 26]	5,8841	5,7058	3,2479	0,1193
Fla-30	5,0862 [20, 26]	4,8677	4,9365	4,2959	2,9433
Quercetin	4,859 [nghiên cứu này]4,9010		4,8400	0,8644	0,3910
Cynaroside	4,977 [nghiên cứu này]5,1760		4,9150	3,9984	1,2457
			MARE,%	3,9063	1,8636

Hoạt tính pGI₅₀ dự đoán từ các mô hình QSAR_{MLR} được so sánh với hoạt tính pGI_{50,tn} thực nghiệm bằng giá trị tuyệt đối sai số tương đối $ARE, \%$ và giá trị trung bình tuyệt đối của sai số tương đối $MARE, \%$. Qua kết quả so sánh đưa ra ở Bảng 3, chúng ta thấy rằng khả năng dự đoán của mô hình QSAR_{MLR} có giá trị $MARE, \% = 3,906\%$ cao hơn $MARE, \% = 1,864$ của mô hình QSAR_{ANN} I(9)-HL(5)-O(1). Như vậy, chất lượng mô hình QSAR_{ANN} cho khả năng dự báo tốt hơn. Trong công trình này, hai hợp chất quercetin và cynaroside là các chất tách ra từ tự nhiên đã được chúng tôi thử hoạt tính pGI₅₀ *in vitro* như ở Bảng 3, có thể xem như là chất mẫu để thiết kế các chất mới nhằm cải thiện hoạt tính tự nhiên của chúng.

Kết quả dự đoán được đánh giá và so sánh bằng sai số $MARE, \%$

$$MARE, \% = \frac{\sum_{i=1}^n ARE_i, \%}{n} \quad \text{với} \quad ARE, \% = \frac{|y_m - y_{tt}|}{y_m} \cdot 100\% \quad (5)$$

Ở đây n là số hợp chất flavonoid pGI₅₀ kiểm tra; y_{tn} và y_{tt} là giá trị pGI₅₀ thực nghiệm và tính toán.

Từ hai hợp chất quercetin và cynaroside, tiến hành thiết kế 10 hợp chất mới bằng cách gắn thêm nhóm thế khác nhau vào các vị trí C₆ và C_{3'}, cấu trúc của 10 hợp chất được đưa ra trong Bảng 4.

Hoạt tính kháng ung thư pGI₅₀ của 6 hợp chất flavonoid trong nhóm kiểm tra và 2 hợp chất quercetin và cynaroside dự đoán từ mô hình QSAR_{MLR} (4) và QSAR_{ANN} I(9)-HL(5)-O(1) được so sánh với hoạt tính ung thư thực nghiệm, kết quả đưa ra ở Bảng 3.

Sau khi sử dụng mô hình QSAR_{ANN} I(9)-HL(5)-O(1) để dự đoán hoạt tính pGI₅₀ của 10 hợp chất mới thiết kế từ các vị trí còn trống C₆ và C_{3'} của quercetin và cynaroside, đưa ra ở Bảng 4. Hoạt tính pGI₅₀ của 10 hợp chất mới cao hơn hoạt tính quercetin (với pGI_{50,tn} = 4,859) và hoạt tính cynaroside (với pGI_{50,tn} = 4,977) tương ứng. Như vậy các hợp chất mới thiết kế thử nghiệm gợi ý tốt có thể giúp định hướng phương pháp phát triển và tổng hợp được chất mới từ hợp chất tự nhiên hiệu quả hơn.

Bảng 4. Hoạt tính pGI₅₀ của 10 hợp chất mới dự đoán được từ mô hình QSAR_{ANN} I(9)-HL(5)-O(1)

Hợp chất mới	Vị trí C ₆	Vị trí C _{3'}	pGI ₅₀	Phương pháp Xác định
Quercetin	-H	-H	4,859	Thử <i>in vitro</i>
flav-1(n)	-OCH ₂ CONHCH ₃	-OH	5,382	QSAR _{ANN}
flav-2(n)	-OCH ₂ CONHC ₆ H ₄ F	-H	5,461	QSAR _{ANN}
flav-3(n)	-OH	-OCH ₂ CONHCH ₃	5,463	QSAR _{ANN}
flav-4(n)	-OCH ₂ (CH ₃)C=NOH	-OH	5,637	QSAR _{ANN}
flav-5(n)	-OCH ₂ CONHC ₆ H ₄ OCH ₃	-H	5,525	QSAR _{ANN}
Hợp chất mới	Vị trí C ₆	Vị trí C _{3'}	pGI ₅₀	Phương pháp Xác định
cynaroside	-H	-H	4,977	<i>in vitro</i>
flav-6(n)	-OCH ₂ CONHCH ₃	-OH	5,456	QSAR _{ANN}
flav-7(n)	-OH	-OCH ₂ CONHCH ₃	5,597	QSAR _{ANN}
flav-8(n)	-OCH ₂ CONHC ₆ H ₄ F	-H	5,597	QSAR _{ANN}
flav-9(n)	-OCH ₂ CONHC ₆ H ₄ OCH ₃	-H	5,610	QSAR _{ANN}
flav-10(n)	-OCH ₂ (CH ₃)C=NOH	-OH	5,731	QSAR _{ANN}

4 Kết luận

Công trình này đã tính toán và sàng lọc được các tham số mô tả phân tử 2D và 3D ảnh hưởng chính đến hoạt tính của các hợp chất flavonoid; đã xây dựng và đánh giá thành công khả năng dự báo của các mô hình QSAR_{MLR} và QSAR_{ANN} I(9)-HL(5)-O(1). Đã xác định được cấu trúc và thử hoạt tính pGI₅₀ *in vitro* đối với 2 hợp chất quercetin và cynaroside phân lập từ *Artocarpus Incisa* L và *Cynara scolymus* L.

Đã dự đoán hoạt tính pGI₅₀ của 10 hợp chất mới được thiết kế từ các chất mẫu quercetin và cynaroside. Hoạt tính kháng ung thư pGI₅₀ của 10 hợp chất mới tốt hơn hoạt tính kháng ung thư của chất mẫu. Như vậy, việc tiến hành nghiên cứu xây dựng mô hình QSAR 2D và 3D trong

nghiên cứu này là một định hướng hữu ích trong nghiên cứu tìm kiếm và tổng hợp các flavonoid khác nhau từ tự nhiên.

Tài liệu tham khảo

1. Nguyễn Thanh Hồng (2007), *Các phương pháp phổ trong hóa hữu cơ*, NXB Khoa học và kỹ thuật, Hà Nội
2. Nguyễn Kim Phi Phụng (2007), *Phương pháp cô lập hợp chất hữu cơ ĐHQG-HCM*
3. Nguyễn Kim Phi Phụng (2005), *Phổ NMR sử dụng trong phân tích hữu cơ*, NXB ĐH Quốc gia TP. HCM.
4. Gasteiger, J., Marsili, M. (1980), "Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges", *Tetrahedron*, 36(22), pp. 3219-3228.
5. Ghahari, S., Alinezhad, H., Nematzadeh, G.A., Tajbakhsh, M., Baharfar, R. (2017), "Chemical Composition, Antioxidant and Biological Activities of the Essential Oil and Extract of the Seeds of Glycine max (Soybean) from North Iran", *Curr Microbiol*, 74(4), pp. 522-531.
6. Hossain, M.A., Mizanur, R.S.M. (2015), "Isolation and characterisation of flavonoids from the leaves of medicinal plant *Orthosiphon stamineus*", *Arabian Journal of Chemistry*, 8(2), pp. 218-221.
7. Jalal, T.K., Ahmed, I.A., Mikail, M., Momand, L., Draman, S., Isa, M.L., Abdull Rasad, M.S., Nor Omar, M., Ibrahim, M., Abdul Wahab, R. (2015), "Evaluation of antioxidant, total phenol and flavonoid content and antimicrobial activities of *Artocarpus altilis* (breadfruit) of underutilized tropical fruit extracts", *Appl Biochem Biotechnol*, 175(7), pp. 3231-3243.
8. Jang, D.S., Han, A.R., Park, G., Jhon, G.J., Seo, E.K. (2004), "Flavonoids and aromatic compounds from the rhizomes of *Zingiber zerumbet*", *Arch Pharm Res*, 27(4), pp. 386-389.
9. Kalailingam, P., Balasubramanian, K., Kannaiyan, B., Mohammed, A.K.N., Meenakshisundram, K., Tamilmani, E., R., K. (2013), "Isolation and quantification of flavonoids from ethanol extract of *Costus igneus* rhizome (CiREE) and impact of CiREE on hypoglycaemic, electron microscopic studies of pancreas in streptozotocin (STZ)-induced diabetic rats", *Biomedicine & Preventive Nutrition*, 3(3), pp. 285-297.
10. Kamdem, W.A.F., Coombes, P.H., Mulholland, D.A., Nkengfack, A.E., Fomum, Z.T. (2006), "Flavones and isoflavones from the west African Fabaceae *Erythrina vogelii*", *Phytochemistry*, 67(5), pp. 459-463.
11. Kier, L., Hall, H., Lowell (1991), *The Molecular Connectivity Chi Indices and Kappa Shape Indices in Structure-Property Modeling*. Vol. 2. VCH publishers New York
12. Kunal, R., Supratik, K., Rudra, N.D. (2015), *Chapter 1 QSAR/QSPR Methods*, in *A primer on QSAR/QSPR modeling - 2015 SpringerBriefs in Molecular Science*, pp. 61-103.
13. Kunal, R., Supratik, K., Rudra, N.D. (2015), *Chapter 2 Statistical Methods in QSAR/QSPR*, in *A primer on QSAR/QSPR modeling - 2015 SpringerBriefs in Molecular Science*, pp. 37-59.
14. Markham, K.R. (1975), *Chapter 1 Isolation Techniques for Flavonoids*, in *The Flavonoids* Springer US, pp. 1-44.
15. Rebecca, L.S., Kimberly, D.M., Ahmedin, J. (2015), "Cancer Statistic, 2015", *CA: A Cancer journal for Clinicians*, 65(1), pp. 5-29.
16. Rebecca, S., Deepa, N., Ahmedin, J. (2012), "Cancer statistics, 2012", *CA: A Cancer journal for Clinicians*, 62(1), pp. 10-29.

17. Sarawek, S., Feistel, B., Pischel, I., Butterweck, V. (2008), "Flavonoids of *Cynara scolymus* possess potent xanthinoxidase inhibitory activity in vitro but are devoid of hypouricemic effects in rats after oral application", *Planta Med*, 74(3), pp. 221-227.
18. Saxena, V.K., Aprajita, G. (2008), "Isolation and study of the flavone glycoside; luteolin-7- O- β -d-glucopyranoside from the seeds of the *capparis decidua* (forsk)", *International journal chemical sciences*, 6(1), pp. 7-10.
19. Senousy, A.S.E., Farag, M.A., Al-Mahdy, D.A., Wessjohann, L.A. (2014), "Developmental changes in leaf phenolics composition from three artichoke cvs. (*Cynara scolymus*) as determined via UHPLC-MS and chemometrics", *Phytochemistry*, 108, pp. 67-76.
20. Si, Y.L., Jin, C.C., Li, Q.Y.S., Kang, C.Z. (2008), "QSAR, action mechanism and molecular design of flavone and isoflavone derivatives with cytotoxicity against HeLa", *European Journal of Medicinal Chemistry*, 43, pp. 2159-2170.
21. Snedecor, G., Cochran, W. (1967), *Statistical methods*, Oxford and IBH, New Delhi.
22. Tat, P.V. (2017), *Development of new anticancer agents from Leaf of plants in Viet Nam*, LAP Lambert Academic Publishing. 01-80 trang.
23. Tat, P.V. (2009), "QSDAR model: Quantitative Spectrum Data ^{13}C -NMR and ^{15}O -NMR Activity Relationships (QSDAR) of a set of anti-cancer 3-aminoflavonoid", *Vietnamese Journal of Chemistry and Application*, 14, pp. 43-46.
24. Tat, P.V. (2009), *Development of Quantitative Structure-Activity Relationships (QSARs) and Quantitative Structure-Property Relationships (QSPRs)*, Hanoi, NXB Natural Science and Technology.
25. Thuy, B.T.P., Tat, P.V. (2012), "Development of QSDAR models for anticancer flavone and isoflavone analogues Using ^{13}C -NMR and ^{15}O -NMR", *Viet Nam Journal of Chemistry*, 50(5A), pp. 203-208.
26. Wang, T.C., Chen, I.L., Lu, P., Wong, C., Liao, C., Tsiao, K., Chang, K., Chen, Y., Tzeng, C. (2005), "Synthesis, antiproliferative, and antiplatelet activities of oxime- and methyloxime-containing flavone and isoflavone derivatives", *Bioorganic & Medicinal Chemistry*, 13(21), pp. 6045-6053.
27. Wang, Y., Xu, K., Lin, L., Pan, Y., Zheng, X. (2007), "Geranyl flavonoids from the leaves of *Artocarpus altilis*", *Phytochemistry*, 68(9), pp. 1300-1306.
28. Yousefinejad, S., Hemmateenejad, B. (2015), "Chemometrics tools in QSAR/QSPR studies: A historical perspective", *Chemometrics and Intelligent Laboratory Systems*, 149, Part B, pp. 177-204.

PREDICTION OF ANTICANCER ACTIVITY HELA OF FLAVONOID COMPOUNDS ISOLATING FROM *Cynara scolymus* L. AND *Artocarpus Incisa* L. USING MODELS QSAR 2D VÀ 3D

Bui Thi Phuong Thuy¹, Phung Van Trung², Hoang Thi Kim Dung², Tran Duong³, Pham Van Tat^{4*}

¹Science University of Hue, Hue City; Dong Nai University of Technology

²Institute of Chemical Technology, Ho Chi Minh City, Viet Nam

³Hue University of Education, Hue City

⁴Hoa Sen University, Ho Chi Minh City

Abstract: In this work, the compounds quercetin and cynaroside of flavonoid derivatives were isolated from *Cynara scolymus* L. and *Artocarpus Incisa* L. with toxigenicity activity for Hela cell line. The models QSAR_{MLR} ($k = 9$) with R^2_{train} of 0.955, R^2_{pred} of 0.932 and QSAR_{ANN} I(9)-HL(5)-O(1) with R^2_{train} of 0.996 and R^2_{pred} of 0.988 were constructed to define the activity pGI₅₀ of flavonoids for cell line Hela. The model QSAR_{ANN} was used to predict the pGI₅₀ of 10 new designed compounds from lead compounds quercetin and cynaroside. The value $MARE, \%$ of 1.864 for the predicted results pGI₅₀ from QSAR_{ANN} was smaller than the value $MARE, \%$ = 3.906% of QSAR_{MLR} model. The anticancer activities of quercetin and cynaroside, and 10 new designed compounds resulting from model QSAR_{ANN} were a good agreement with those from experiment and higher those from lead compounds, respectively.

Keywords: QSAR_{MLR}, QSAR_{ANN}, anticancer activities Hela