

**NGHIÊN CỨU HOẠT TÍNH KHÁNG UNG THƯ CỦA KAEMPFEROL-1,
DAIDZIN TỪ ZINGIBER ZERUMBET SM VÀ GLUCINE MAX L
SỬ DỤNG CÁC THAM SỐ MÔ TẢ PHÂN TỬ 2D VÀ 3D**

**Bùi Thị Phương Thúy¹, Phùng Văn Trung³, Hoàng Thị Kim Dung³,
Trần Dương², Phạm Văn Tất^{4*}**

¹Khoa Hóa học, Trường Đại học Khoa học, Đại học Huế

² Khoa Hóa học, Trường Đại học Sư phạm, Đại học Huế

³Viện Công nghệ Hóa học, TP. Hồ Chí Minh

^{4*}Khoa Khoa học và Công nghệ, Trường Đại học Hoa Sen, TP. Hồ Chí Minh

*E-mail: vantat@gmail.com

Ngày nhận bài: 24/5/2017; ngày hoàn thành phản biện: 14/6/2017; ngày duyệt đăng: 27/10/2017

TÓM TẮT

Trong nghiên cứu này, chúng tôi đã phân lập được hai dẫn xuất flavonoid từ *Zingiber zerumbet Sm* và *Glucine Max L* ở Việt Nam, với hoạt tính gây độc tế bào ung thư Hela tương đối mạnh. Các mô hình QSAR xây dựng được bao gồm: Mô hình QSAR_{MLR} (1), R^2_{train} là 0,955; R^2_{pred} là 0,745; mô hình QSAR_{MLR} (2), R^2_{train} là 0,854 và R^2_{pred} là 0,812; mô hình mạng thần kinh QSAR_{ANN} với kiến trúc mạng I(9)-HL(5)-O(1), R^2_{train} là 0,8963 và R^2_{pred} là 0,8883; Mô hình mạng thần kinh QSAR_{PCA-ANN} với kiến trúc mạng I(6)-HL(9)-O(1), R^2 là 0,8973 và R^2_{pred} là 0,8872. Hoạt tính kháng ung thư của các hợp chất flavonoid trong nhóm kiểm tra và các hợp chất phân lập được dự đoán từ những mô hình QSAR phù hợp với dữ liệu thực nghiệm từ các tài liệu tham khảo.

Từ khoá: QSAR_{MLR}, QSAR_{ANN}, QSAR_{PCA-ANN}, hoạt tính kháng ung thư Hela.

1. GIỚI THIỆU

Trong những năm gần đây, các tính toán hoá lượng tử được áp dụng rộng rãi trong các nghiên cứu không chỉ đối với tính chất hoá học và còn ứng dụng trong tìm kiếm các loại thuốc mới [4, 5]. Các nghiên cứu mối quan hệ định lượng giữa cấu trúc và hoạt tính (QSAR) của hợp chất tự nhiên là nguyên thủy cho các nghiên cứu thuốc mới và hỗ trợ trong sản xuất dược phẩm [3, 5]. Tại Việt Nam, có nhiều nhà khoa học từ các trường đại học và viện nghiên cứu đã nghiên cứu xây dựng các mô hình quan hệ định lượng giữa tính chất hóa lý và hoạt tính (QSSR), mô hình quan hệ định lượng

Nghiên cứu hoạt tính kháng ung thư của kaempferol-1, Daidzin từ Zingiber zerumbet Sm và ...

giữa tính chất phổ và hoạt tính (QSDAR) để dự đoán họa tính của các dẫn xuất flavonoid, trên cơ sở đó các hợp chất mới được thiết kế và có hoạt tính được cải thiện [3, 15]. Các nghiên cứu trước đây đối với nhóm dẫn xuất 3-aminoflavonoid đã dựa trên cơ sở các tính toán hoá lượng tử bán thực nghiệm [4]. Các nghiên cứu này đã chỉ ra phương thức thiết kế thuốc mới hiệu quả với sự hỗ trợ của máy tính. Việc xem xét các mối quan hệ định lượng giữa cấu trúc của flavonoid với hoạt tính kháng ung thư là nhiệm vụ quan trọng trong tìm kiếm, nghiên cứu các dẫn xuất mới [12, 15].

Các sản phẩm tự nhiên có nguồn gốc thực vật được quan tâm trong nghiên cứu thuốc chống ung thư mới và có thể tác động trực tiếp lên tế bào Hela và làm giảm tác dụng phụ tiêu cực của chúng [5, 16]. Trong nghiên cứu này, chúng tôi đã phân lập được hai dẫn xuất flavonoid từ Zingiber zerumbet Sm và Glucine Max L [13], hoạt tính kháng ung thư của chúng được thử nghiệm in vitro đã chứng minh được rằng chúng có hoạt tính gây độc tế bào ung thư Hela và một số loại ung thư khác tương đối mạnh [1, 9, 10, 16]. Các hoạt tính kháng ung thư của flavonoid nói chung được nghiên cứu rộng rãi [2, 11, 18].

Chúng tôi kết hợp các tính toán hoá lượng tử bán thực nghiệm và xây dựng mô hình QSAR từ 34 dẫn xuất flavone và isoflavone [5, 12]. Hoạt tính kháng ung thư $GI_{50}/\mu M$ của các dẫn xuất flavone và isoflavone trong nhóm kiểm tra, flavonoid chiết xuất được dự đoán từ các mô hình QSAR và so sánh với dữ liệu thực nghiệm [5, 12].

2. VẬT LIỆU VÀ PHƯƠNG PHÁP

2.1. Dữ liệu

Để đảm bảo xây dựng các mô hình QSAR chính xác, tập dữ liệu sử dụng để xây dựng mô hình QSAR bao gồm 34 hợp chất flavone và isoflavone và hoạt tính $GI_{50}/\mu M$ (nồng độ ức chế 50% sự phát triển của tế bào ung thư) được nghiên cứu bởi Wang và cộng sự [5, 12]. Giá trị GI_{50} được chuyển đổi sang giá trị pGI_{50} nhằm giảm biên độ biến thiên của tập dữ liệu đồng thời giảm sai số và tăng khả năng dự đoán của các mô hình QSAR [3, 15].

2.2. Tính toán các tham số mô tả phân tử

Cấu trúc thực nghiệm của các dẫn xuất flavonoid được xây dựng bằng chương trình HyperChem 8.0 [8]. Cấu trúc phân tử được tối ưu hoá bằng thuật toán Polak-Ribiere với gradient là $0,01 \text{ kcal.mol}^{-1}$ [8, 15]. Các tham số cấu trúc 2D, 3D được tính toán bằng phương pháp cơ học phân tử trên hệ thống QSARIS [6, 17].

2.3. Xây dựng mô hình QSAR

2.3.1. Mô hình QSAR tuyến tính

Mô hình QSAR tuyến tính là mối liên hệ tuyến tính định lượng giữa cấu trúc và hoạt tính, vì vậy hồi qui tuyến tính là một công cụ cần thiết để xây dựng mô hình QSAR [17]. Các giá trị thống kê R^2 , R^2_{pred} , sai số tương đối $ARE, \%$, trung bình sai số tương đối $MARE, \%$ được sử dụng để kiểm tra khả năng dự đoán của các mô hình [4, 17]. Sử dụng các mô hình này để dự đoán hoạt tính kháng ung thư của các hợp chất flavonoid trong nhóm kiểm tra, flavonoid chiết xuất và nhóm hợp chất mới thiết kế.

2.3.2. Mô hình mạng thần kinh QSAR_{ANN}

Mạng thần kinh thông minh nhân tạo là một công cụ xây dựng mối liên hệ phi tuyến giữa cấu trúc và hoạt tính [4, 15]. Mô hình QSAR_{ANN} được xây dựng bằng kỹ thuật neuro-fuzzy với thuật toán di truyền sử dụng chương trình INForm [19]. Mô hình mạng thần kinh nhân tạo có kiến trúc mạng $I(n)$ - $HL(m)$ - $O(k)$ được xây dựng. Lớp nhập $I(n)$ với n nơ ron tương ứng với n biến độc lập là các tham số cấu trúc 2D, 3D được chọn từ mô hình QSAR tuyến tính; lớp ẩn $HL(m)$ với m nơ ron và lớp xuất $O(k)$ với k nơ ron là biến phụ thuộc [19].

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Xây dựng mô hình QSAR_{MLR} để dự đoán hoạt tính của kaempferol-1

Dữ liệu cấu trúc và hoạt tính nhóm flavonoid được chia thành nhóm luyện và nhóm kiểm tra. Khả năng dự đoán của mô hình QSAR được đánh giá bằng cách so sánh kết quả dự đoán với hoạt tính pGI_{50} của các hợp chất trong nhóm kiểm tra. Sự biến thiên các giá trị R^2 , giá trị tương quan dự đoán R^2_{pred} và MSE (sai số chuẩn) trong các mô hình QSAR_{MLR} gồm các tham số mô tả 2D và 3D, tương ứng trong Bảng 1.

Bảng 1. Các mô hình QSAR_{MLR} (k từ 2 đến 10) cùng các giá trị R^2 , R^2_{pred} và MSE tương ứng

k	Tham số mô tả phân tử 2D, 3D trong các mô hình	MSE	R^2	R^2_{pred}
2	xvc3, Dipole	0,0628	0,6573	0,6299
3	xvp3, xvpc4, Dipole	0,0343	0,8202	0,7977
4	xvp6, xvpc4, Dipole, LogP	0,0267	0,8659	0,8425
5	xp8, Dipole, Volume, knotpv, LogP	0,0226	0,8913	0,8666
6	xp8, xvp8, Dipole, knotpv, SHBa, LogP	0,0188	0,9140	0,8895
7	xp8, xvp8, Dipole, Hmaxpos, knotpv, SHBa, LogP	0,0153	0,9331	0,9097
8	xp6, SaaCH, k2, nvx, knotp, knotpv, SHBa, LogP	0,0154	0,9361	0,9091
9	xp8, ABSQ, SsOH, Hmaxpos, phia, knotp, knotpv, SHBa, LogP	0,0115	0,9547	0,9321
10	xp8, ABSQ, SsOH, Hmaxpos, nvx, phia, knotp, knotpv, SHBa, LogP	0,0122	0,9549	0,9283

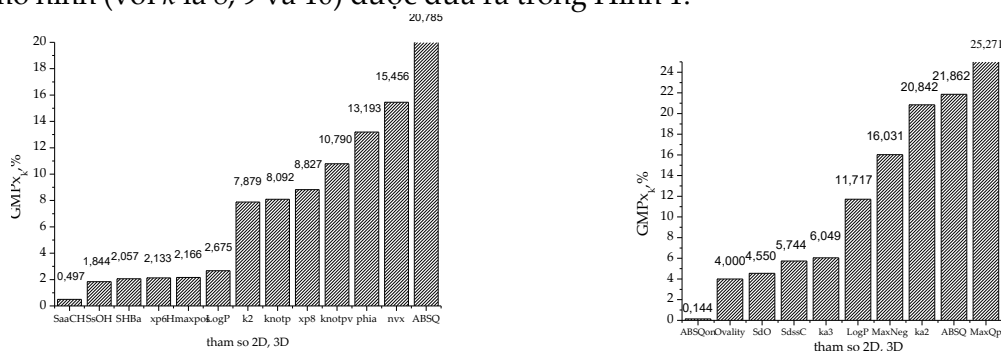
Để thành lập mô hình QSAR_{MLR}, các tham số mô tả 2D và 3D được chọn lựa bằng kỹ thuật hồi qui từng bước [17]. Các tham số mô tả 2D, 3D đưa vào mô hình trên cơ sở thay đổi các giá trị thống kê R^2 , SE và R^2_{pred} . Các mô hình được đánh giá chéo

Nghiên cứu hoạt tính kháng ung thư của kaempferol-1, Daidzin từ Zingiber zerumbet Sm và ...

bằng kỹ thuật loại bỏ dần từng trường hợp (LOO) để xác định giá trị R^2_{pred} [15]. Các mô hình phù hợp nhất được chỉ ra trong Bảng 1.

Các mô hình QSAR_{MLR} (với k từ 2 đến 10) được sắp xếp theo sự thay đổi các giá trị R^2 , SE và R^2_{pred} . Trong Bảng 1, các mô hình QSAR_{MLR} (với k từ 8 đến 10) là các mô hình có giá trị R^2_{pred} cao hơn cả. Mô hình QSAR_{MLR} với $k = 9$ có R^2 cao nhất là 0,955 và R^2_{pred} là 0,745. Vì vậy, ba mô hình (với k là 8, 9 và 10) được chọn để xác định giá trị phần trăm đóng góp của các tham số mô tả 2D, 3D đối với hoạt tính kháng ung thư.

Ảnh hưởng của các tham số cấu trúc đến hoạt tính kháng ung thư tính toán từ các mô hình (với k là 8, 9 và 10) được đưa ra trong Hình 1.



a) mô hình QSAR_{MLR} với $k = 8, 9, 10$ (Bảng 1)

b) mô hình QSAR_{MLR} với $k = 6, 7, 8$ (Bảng 2)

Hình 1. Ảnh hưởng của các tham số 2D, 3D đến hoạt tính kháng ung thư pGI₅₀

Phần trăm đóng góp trung bình toàn cục $GMP_{x_k}, \%$ [17] trong Hình 1a, đã mô tả ảnh hưởng quan trọng của các tham số mô tả 2D, 3D của các flavone và isoflavone. Sự sắp xếp mức độ ảnh hưởng quan trọng của các tham số 2D, 3D trong các mô hình QSAR_{MLR} theo $GMP_{m,x_k}, \%$: SaaCH < SsOH < SHBa < xp6 < Hmaxpos < LogP < k2 < knotp < xp8 < knotpv < phia < nvx < ABSQ. Các tham số mô tả ABSQ, nvx, phia, knotpv, xp8, knotp, k2 và LogP được xem là các tham số quan trọng nhất trong mỗi phân tử. Ngoài ra, các tham số mô tả phân tử đặc trưng cho nhóm carbonyl $C_4 = O_{11}$ là nguyên tố O_{11} . Các nguyên tố này có các cặp electron tự do với điện tử π của liên kết $C_2 = C_3$, và $C_4 = O_{11}$ tạo hệ liên hợp. Nhóm carbonyl $C_4 = O_{11}$ thể hiện toàn bộ phản ứng tự nhiên của nhóm carbonyl [5, 12]. Vì vậy, các tham số mô tả phân tử cũng chứng minh được định lượng ảnh hưởng trên toàn phân tử bởi giá trị GMP_{x_k} , và phù hợp với các giá trị thực nghiệm [5, 12].

Vị trí nguyên tử C_6 , C_2 và C_3 trong phân tử flavonoid còn trống, có thể gắn các nhóm thế mới khi thiết kế dẫn xuất mới. Các hợp chất mới được thiết kế và dự đoán hoạt tính sinh học nhằm sàng lọc các hợp chất mới có hoạt tính sinh học cao.

Mô hình QSAR_{MLR} được chọn với các tham số mô tả $k = 9$ là mô hình tốt nhất gồm các biến số độc lập bao gồm xp8, ABSQ, SsOH, Hmaxpos, phia, knotp, knotpv, SHBa và LogP. Kết quả dự đoán hoạt tính sinh học phù hợp với dữ liệu thực nghiệm

được minh chứng qua giá trị R^2 dự đoán R^2_{pred} . Mô hình QSAR_{MLR} với $k = 9$ trong Bảng 1 được sử dụng để dự đoán hoạt tính kháng ung thư Hela:

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

3.2. Xây dựng mô hình QSAR_{MLR} dự đoán hoạt tính của daidzin

Bằng cách xây dựng mô hình QSAR_{MLR} tương tự ở mục 3.1. khi thay đổi tập luyện và tập kiểm tra bằng cách chọn lựa ngẫu nhiên để tìm kiếm các mô hình QSAR có khả năng dự đoán hoạt tính của các hợp chất nghiên cứu [12, 15]. Các mô hình QSAR xây dựng được có k từ 2 – 10 với các giá trị thống kê tương ứng được đưa ra trong Bảng 2b. Các mô hình QSAR_{MLR} với 5, 6, 7 biến số có giá trị tương quan dự đoán R^2_{pred} cao hơn; đặc biệt mô hình QSAR_{MLR} có 6 biến số theo phương trình (2) là mô hình tốt nhất trong các mô hình ở Bảng 2, đã được chọn lựa để dự đoán hoạt tính kháng ung thư của các dẫn xuất flavonoid [12, 15].

$$pGI_{50} = 8,509 + 2,8540 \text{MaxQp} + 0,0247 \text{SdO} + 0,2192 \text{LogP} - 3,6969 \text{Ovality} + 0,2969 \text{SdssC} + 0,3653 \text{ka3} \quad (2)$$

Các mô hình QSAR_{MLR} với $k = 5, 6, 7$ biến số được chọn để tính toán khả năng ảnh hưởng của các tham số 2D, 3D. Thứ tự sắp xếp các tham số ảnh hưởng dựa vào kết quả tính toán theo giá trị $GMP_{X_k, \%}$: $ABSQ > \text{MaxQp} > \text{ka2} > \text{MaxNeg} > \text{LogP} > \text{ka3} > \text{SdssC} > \text{SdO} > \text{Ovality} > \text{ABSQon}$ (Hình 1b). Các tham số $ABSQ$; MaxQp ; ka2 ; MaxNeg ; LogP được sử dụng tính toán và xây dựng mô hình QSAR_{PCA} các thành phần chính $PC_1, PC_2, PC_3, PC_4, PC_5$, và PC_6 từ kỹ thuật phân tích thành phần chính.

Bảng 2. Mô hình QSAR_{MLR} (k từ 2-10) với các giá trị R^2 , R^2_{pred} và SE

k	Tham số mô tả phân tử 2D, 3D trong các mô hình	R^2	R^2_{pred}	SE
2	LogP, MaxNeg	0,756	0,731	0,43
3	LogP, MaxNeg, ka2	0,774	0,732	0,417
4	LogP, MaxNeg, ka2, SdO	0,805	0,772	0,39
5	LogP, MaxNeg, MaxQp, SdO, ka2	0,832	0,756	0,365
6	LogP, MaxQp, Ovality, SdO, SdssC, ka3	0,854	0,812	0,342
7	LogP, MaxNeg, MaxQp, ka2, SdO, ABSQ, ABSQon	0,836	0,721	0,365
8	LogP, MaxNeg, MaxQp, ka2, SdO, ka3, ABSQ, ABSQon	0,837	0,693	0,367
9	LogP, MaxNeg, MaxQp, ka2, SdO, ka3, ABSQ, ABSQon, Dipole	0,838	0,682	0,369
10	LogP, MaxNeg, MaxQp, ka2, SdO, ka3, ABSQ, ABSQon, Dipole, Ovality	0,841	0,65	0,368

3.3. Xây dựng mô hình QSAR_{ANN} và mô hình QSAR_{PCA-ANN}

Mô hình QSAR_{ANN} được xây dựng bằng kỹ thuật mạng thần kinh mờ với thuật toán di truyền trên hệ thống INForm [21]. Mô hình mạng thần kinh nhân tạo đã được xây dựng có kiến trúc mạng I(9)-HL(5)-O(1). Kiến trúc mạng này bao gồm lớp nhập I(9) với 9 nơ ron $xp8, ABSQ, SsOH, Hmaxpos, phia, knotp, knotpv, SHBa$ và LogP ; lớp ẩn HL(5) với 5 nơ ron và lớp xuất O(1) với 1 nơ ron là pGI_{50} . Thuật toán lan

Nghiên cứu hoạt tính kháng ung thư của kaempferol-1, Daidzin từ *Zingiber zerumbet* Sm và ...

truyền ngược được sử dụng để luyện mạng thần kinh. Mỗi tham số luyện mạng khác nhau được sử dụng cho quá trình luyện mạng với tốc độ luyện là 0,7 và tốc độ học là 0,01; tổng cộng sai số là 0,0003017, sai số dự đoán là 0,00001, số vòng lặp là 300000, R^2_{train} là 0,8963 và R^2_{pred} là 0,8883.

Mô hình QSAR_{PCA-ANN} được xây dựng tương tự mô hình QSAR_{ANN}, tốc độ luyện mạng là 0,7, tỉ lệ học 0,7, sai số MSE = 0,003447 với 5000 vòng lặp, giá trị R^2 là 0,897298, R^2_{pred} là 0,88718. Mô hình QSAR_{PCA-ANN} có kiến trúc mạng I(6)-HL(9)-O(1) với lớp nhập gồm 6 nơ ron là 6 thành phần chính từ PC₁ đến PC₆, lớp ẩn gồm 9 nơ ron, lớp đầu ra gồm 1 nơ ron là pGI₅₀.

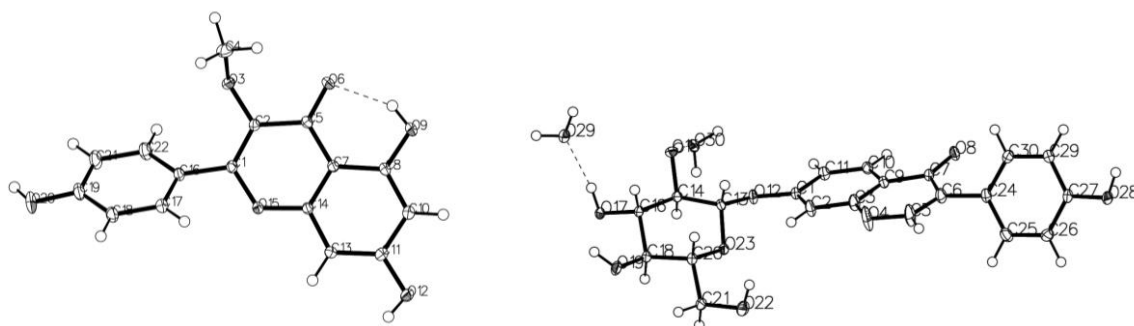
3.4. Phân lập Flavonoid

3.4.1. Phân lập flavonoid từ *Zingiber zerumbet* Sm.

Kaempferol-1 thu được dưới dạng tinh thể, màu vàng bằng phương pháp bay hơi chậm trong hệ dung môi CHCl₃ – MeOH (85:15). Dữ liệu nhiễu xạ đơn tinh thể của kaempferol-1 được đo trên máy Bruker D8 Quest (Khoa Hóa, Trường Đại học Khoa Học Tự Nhiên, Thành Phố Hồ Chí Minh) ở nhiệt độ 100 K, đối âm cực Mo với bước sóng K_α ($\lambda = 0,71073 \text{ \AA}$). Ảnh nhiễu xạ được ghi trên detector huỳnh quang dạng đĩa trong đường kính 34 cm, khoảng cách từ tinh thể đến detector là 10 cm. Cấu trúc kaempferol-1 được xác định bằng phương pháp đo nhiễu xạ tia X đơn tinh thể [7, 14]. Thông tin về cấu trúc tinh thể cũng như quá trình tính toán và tối ưu cấu trúc được tóm tắt trong Bảng 3. Cấu trúc phân tử *kaempferol* từ *Zingiber zerumbet* Sm được chỉ ra trong Hình 2a.

Bảng 3. Các thông tin về tinh thể kaempferol-1 và Daidzin

	kaempferol-3-O-methylether	Daidzin
Công thức phân tử	C ₁₆ H ₁₂ O ₆	C ₂₁ H ₁₅ O ₁₁
Hệ tinh thể	Trực thoi	Đơn tà
Nhóm không gian	Pna2 ₁	P2 ₁
Thông số mạng a, b c/Å	7,0060(4); 25,9070(14); 7,2062(4)	14,6876(11);-4,3568(3);-
$\alpha, \beta, \gamma /^\circ$	90; 90; 90	90; 111,658(3); 90
Thể tích (Volume) /Å ³	1307,96(13)	1029,32(13)
Tỉ khối tính toán	1,525	1,43
Kích thước tinh thể mm ³	0,1 × 0,08 × 0,04	0,1 × 0,05 × 0,05
Số phản xạ đo được	44181	25478
Số phản xạ độc lập	4114 [R _{int} = 0,0465, R _{sigma} =	6681 [R _{int} = 0,0399, R _{sigma} = 0,0506]
Hệ số phù hợp mô hình	1,092	1,022
Độ sai lệch [$I > 2\sigma(I)$]	R ₁ = 0,0387; wR ₂ = 0,0959	R ₁ = 0,0454; wR ₂ = 0,0962
Độ sai lệch tổng cộng (R)	R ₁ = 0,0427; wR ₂ = 0,0987	R ₁ = 0,0675; wR ₂ = 0,1049
Mật độ: đỉnh/lỗ / e Å ⁻³	0,44/-0,24	0,48/-0,39



a) Cấu trúc của Kaempferol-1 dưới dạng elipxoit với xác suất 50%; $GI_{50,exp}$ (μM) = 15.37 ± 0.401 b) Cấu trúc daidzin dưới dạng elipxoit với xác suất 50%; $GI_{50,exp}$ (μM) = $7,88 \pm 0.195$

Hình 2. Cấu trúc của phân tử: a) Kaempferol-1 và b) daidzin

3.4.2. Phân lập flavonoid từ *Glucine Max L.*

Daidzin thu được dưới dạng tinh thể, màu vàng trong hệ dung môi $CHCl_3$ – MeOH (75:5). Dữ liệu nhiễu xạ đơn tinh thể của daidzin được đo tương tự như đối với hợp chất Kaempferol-1 được đưa ra ở Bảng 3. Cấu trúc phân tử chỉ ra trong Hình 2b.

3.5. Khả năng dự đoán của mô hình QSAR

Khả năng dự đoán của các mô hình $QSAR_{MLR}$ và $QSAR_{ANN}$ được đánh giá cẩn thận bằng kỹ thuật loại bỏ từng trường hợp và giá trị R^2 dự đoán R^2_{pred} . Hoạt tính sinh học thực nghiệm *in vitro* $pGI_{50,exp}$ của hợp chất của kaempferol-1, daidzin được thực hiện tại phòng sinh học phân tử, Trường ĐH Tự Nhiên, TP. HCM. Hoạt tính sinh học dự báo $pGI_{50,pred}$ của kaempferol-1, daidzin và các hợp chất trong nhóm kiểm tra được dự đoán từ các mô hình tuyến tính $QSAR_{MLR}$ (1), $QSAR_{MLR}$ (2), mô hình $QSAR_{ANN}$ I(9)-HL(5)-O(1) và mô hình $QSAR_{PCA-ANN}$ I(6)-HL(9)-O(1) đưa ra ở Bảng 4.

Bảng 4. Hoạt tính sinh học pGI_{50} của nhóm kiểm tra và hai hợp chất phân lập kaempferol-1 và daidzin từ các mô hình $QSAR_{MLR}$ (1), $QSAR_{MLR}$ (2), $QSAR_{ANN}$ và $QSAR_{PCA-ANN}$

Hợp chất	Tham khảo	$pGI_{50,exp}$	$pGI_{50,pred}$		ARE,%	
			$QSAR_{MLR}$ (1)	$QSAR_{ANN}$	$QSAR_{MLR}$ (1)	$QSAR_{ANN}$
Fla-1	[5, 12]	5,6990	5,3879	5,7371	5,4589	0,6685
Fla-11	[5, 12]	5,6990	5,9188	6,0478	3,8568	6,1204
Fla-24	[5, 12]	5,6198	5,9316	5,7662	5,5482	2,6051
Fla-25	[5, 12]	5,6383	5,8627	5,6843	3,9799	0,8158
Fla-26	[5, 12]	5,6990	5,8841	5,7058	3,2479	0,1193
Fla-30	[5, 12]	5,0862	4,8677	4,9365	4,2959	2,9433
kaempferol-1	Phân lập	4,8130	4,8360	4,8061	0,4780	0,1430
MARE,%					3,838	1,917

Hợp chất	Tham khảo	pGI _{50,exp}	pGI _{50,pred}		ARE,%	
			QSAR _{MLR} (2)	QSAR _{PCA-}	QSAR _{MLR} (2)	QSAR _{PCA-}
				ANN		ANN
Fla-8	[5, 12]	5,9208	5,234	5,471	11,600	7,597
Fla-22	[5, 12]	5,7447	5,485	5,556	4,521	3,285
Fla-32	[5, 12]	6,0969	5,423	5,350	11,053	12,250
Fla-74	[5, 12]	5,6990	5,815	5,736	2,035	0,649
Fla-80	[5, 12]	5,6990	6,034	6,086	5,878	6,791
Fla-87	[5, 12]	5,298	4,415	5,106	16,667	3,624
daidzin	Phân lập	5,103	4,592	4,981	10,014	2,391
MARE,%					8,824	5,227

Hoạt tính sinh học dự đoán pGI_{50,pred} từ các mô hình được so sánh với hoạt tính sinh học thực nghiệm pGI_{50,exp} bằng giá trị sai số tương đối trung bình MARE,%. Từ các giá trị MARE,% trong Bảng 4 cho thấy rằng khả năng dự đoán của mô hình QSAR_{MLR} (1) thấp hơn mô hình QSAR_{ANN}. Tương tự khả năng dự đoán của mô hình QSAR_{PCA-ANN} cao hơn QSAR_{MLR} (2). Hoạt tính sinh học pGI₅₀ của sáu hợp chất trong nhóm kiểm tra và hai flavonoid kaempferol-1 và daidzin từ các mô hình QSAR_{MLR} (1), QSAR_{MLR} (2) QSAR_{ANN} và QSAR_{PCA-ANN} chỉ ra khả năng dự đoán chính xác với các sai số nằm trong khoảng không chắc chắn của phép đo thực nghiệm. Hợp chất phân lập là kaempferol-1 và daidzin với các vị trí còn trống C₆, C_{2'} và C_{3'} được chọn như hợp chất mẫu để thiết kế 5 hợp chất mới, bằng cách gắn các nhóm thế mới vào các vị trí C₆, C_{2'} và C_{3'}. Hoạt tính sinh học pGI₅₀ của các hợp chất mới thiết kế được dự đoán bằng mô hình QSAR_{ANN}, được đưa ra trong Bảng 5.

Kết quả dự đoán giá trị pGI_{50,pred} của các hợp chất mới được so sánh với giá trị thực nghiệm của kaempferol-1, daidzin (Bảng 5). Hoạt tính pGI_{50,pred} của 5 hợp chất mới cao hơn kaempferol-1 và daidzin. Từ đây, bằng kỹ thuật thiết kế phân tử hứa hẹn có thể tạo ra được phẩm mới từ hợp chất tự nhiên.

Bảng 5. Hoạt tính GI₅₀ (μM) của 5 hợp chất mới nhận được từ mô hình QSAR_{ANN}

Hợp chất mới	C ₆	C _{2'}	C _{3'}	pGI ₅₀	Phương pháp
kaempferol-1	H	H	H	4,813	<i>in vitro</i> test on HeLa
flav-1(new)	H	CH ₃	H	5,781	QSAR _{ANN}
flav-2(new)	CH ₃ CO-	H	OCH ₃	6,155	QSAR _{ANN}
flav-3(new)	CH ₃ CO-	OCH ₃	H	6,158	QSAR _{ANN}
flav-4(new)	CH ₃	OCH ₃	H	6,488	QSAR _{ANN}
flav-5(new)	CH ₃	H	OCH ₃	6,538	QSAR _{ANN}
Daidzin	H	H	H	5,103	<i>in vitro</i> test on HeLa
isoflav-1(new)	CH ₃	H	H	5,327	QSAR _{PCA-ANN}
isoflav-2(new)	NO ₂	H	NO ₂	5,740	QSAR _{PCA-ANN}

isoflav-3(new)	CH ₃ CO-	CH ₃	H	6,183	QSAR _{PCA-ANN}
isoflav-4(new)	OH	H	OH	6,216	QSAR _{PCA-ANN}
isoflav-5(new)	H	OH	OH	6,317	QSAR _{PCA-ANN}

4. KẾT LUẬN

Tính toán hoá lượng tử bán thực nghiệm, hồi qui đa biến, phân tích thành phần chính và mạng thần kinh được sử dụng để thành lập thành công quan hệ giữa các tham số mô tả phân tử 2D và 3D với hoạt tính kháng ung thư GI₅₀ (μM) của flavonoid. Mô hình QSAR_{MLR} (1) và QSAR_{MLR} (2) chỉ ra rằng các tham số mô tả phân tử xp8, ABSQ, SsOH, Hmaxpos, phia, knotp, knotpv, SHBa, LogP, MaxQp, SdO, Ka3, Ovality và SdssC của flavonoid, chúng ảnh hưởng quan trọng đối với hoạt tính gây độc tế bào ung thư Hela. Mô hình QSAR_{MLR} (1) và QSAR_{MLR} (2) cũng xác định được các vị trí quan trọng C₆, C_{2'} và C_{3'} để gắn nhóm thế mới nhằm thiết kế các flavonoid có hoạt tính cao hơn kaempferol-1, daidzin phân lập từ *Zingiber zerumbet Sm* và *Glucine Max L*. Mô hình QSAR_{ANN} và mô hình QSAR_{PCA-ANN} có khả năng dự báo tốt hơn cho hoạt tính của flavonoid. Hoạt tính sinh học nhận được từ các mô hình QSAR phù hợp với các kết quả nghiên cứu thực nghiệm. Mô hình QSAR trình bày trong nghiên cứu này hữu ích trong nghiên cứu *in vitro* hoạt tính gây độc tế bào ung thư Hela của các flavonoid khác nhau. Nghiên cứu thành lập các mô hình QSAR được ứng dụng cho các nghiên cứu, tìm kiếm các loại dược chất kháng ung thư.

TÀI LIỆU THAM KHẢO

- [1]. S. Ghahari, H. Alinezhad, G. A. Nematzadeh, M. Tajbakhsh, R. Baharfar (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.
- [2]. D. S. Jang, A. R. Han, G. Park, G. J. Jhon, E. K. Seo (2004), Flavonoids and aromatic compounds from the rhizomes of *Zingiber zerumbet*, *Arch Pharm Res*, 27(4), pp. 386-389.
- [3]. Bui Thi Phuong Thuy, Pham Van Tat (2012), QSSR models: prediction of physicochemical property and anticancer, *Viet Nam Journal of Chemistry*, 5(50), pp. 550-556.
- [4]. Pham Van Tat (2009), QSDAR model: Quantitative Spectrum Data ¹³C-NMR and ¹⁵O-NMR Activity Relationships (QSDAR) of a set of anti-cancer 3-aminoflavonoid, *Vietnamese Journal of Chemistry and Application*, 14, pp. 43-46.
- [5]. Tai-Chi Wang, etc. (2005), Synthesis, antiproliferative, and antiplatelet activities of oxime- and methyloxime-containing flavone and isoflavone derivatives, *Bioorganic & Medicinal Chemistry*, 13(21), pp. 6045-6053.
- [6]. QSARIS 1.0 (2001), "Statistical Solutions Ltd.", USA.
- [7]. XLSTAT Version 2014.4.08 (2014), "Copyright Addinsoft 1995-2014", U.S.A.

- [8]. I. Hypercube (2008), "HyperChem Release 8.03 for Windows", Hypercube. Inc. Gainesville, Florida, USA.
- [9]. Yob N. J., Mohd. Jofrry S., Meor. Mohd. Affandi M. M. R., Teh L. K., Salleh M. Z., Zakaria Z. A. (2011), Zingiber zerumbet (L.) Smith: A Review of Its Ethnomedicinal, Chemical, and Pharmacological Uses, *Evidence-Based Complementary and Alternative Medicine*, 11, pp. 1-12.
- [10]. E. H. Kim, S. H. Kim, J. I. Chung, H. Y. Chi, J. A. Kim, I. M. Chung (2005), Analysis of phenolic compounds and isoflavones in soybean seeds (*Glycine max* (L.) Merrill) and sprouts grown under different conditions, *European Food Research and Technology*, 222(1), pp. 201.
- [11]. Iris S. L. Lee, Mary C. Boyce, Michael C. Breadmore (2012), Extraction and on-line concentration of flavonoids in *Brassica oleracea* by capillary electrophoresis using large volume sample stacking, *Food Chemistry*, 133(1), pp. 205-211.
- [12]. Si Yan Liao, Jin Can Chen, Li Qianyong Shen, Kang Cheng Zheng (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.
- [13]. Do Tat Loi (2006), "Medicinal Plants and Drugs from Vietnam", Hanoi, Medicine.
- [14]. Bastien P., Sposito Vinzi V., Tenenhaus M. (2005), PLS Generalised Regression. *Computational Statistics and Data Analysis*, pp. 17-46.
- [15]. Bui Thi Phuong Thuy, Pham Van Tat (2012), Development of QSDAR models for anticancer flavone and isoflavone analogues Using ^{13}C -NMR and ^{15}O -NMR, 50(5A), pp. 203-208.
- [16]. Manjinder Singh, Maninder Kaur, Om Silakari (2014), Flavones: An important scaffold for medicinal chemistry, *European Journal of Medicinal Chemistry*, 84, pp. 206-239.
- [17]. Pham Van Tat (2009), "Development of Quantitative Structure-Activity Relationships (QSARs) and Quantitative Structure-Property Relationships (QSPRs)", Hanoi, Natural Science and Technology.
- [18]. S. H. Thilakarathna, H. P. V. Rupasinghe (2013), Flavonoid Bioavailability and Attempts for Bioavailability Enhancement, *Nutrients*, 5(9), pp. 3367-3387.
- [19]. INForm V2.0 (2000), "Intelligentsys Ltd.", UK.

ANTICANCER AGENTS OF KAEMPFEROL-3-O-METHYLEETHER AND DAIDZIN IN LEAF OF PLANTS *ZINGIBER ZERUMBET SM* AND *GLUCINE MAX LUSING* 2D, 3D DESCRIPTORS

Bui Thi Phuong Thuy¹, Phung Van Trung³, Hoang Thi Kim Dung³,
Tran Duong², Pham Van Tat^{4*}

¹Faculty of Chemistry, University of Sciences, Hue University

²Faculty of Chemistry, University of Education, Hue University

³Institute of Chemical Technology, Ho Chi Minh City

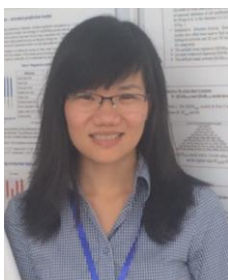
⁴Faculty of Science and Technology, Hoa Sen University, Ho Chi Minh City

*E-mail: vantat@gmail.com

ABSTRACT

In this study, we have isolated the two flavone and isoflavone in flavonoid group from *Zingiber zerumbet Sm* and *Glucine Max L* in Vietnam, with relatively strong cytotoxic activity in Hela cancer cells. In the QSAR study on cancer Hela cell line, the techniques of multiple linear regression, the artificial neural network and principal component analysis were integrated with artificial neural network to construct the different QSAR models such as: The best linear model QSAR_{MLR} (1), R^2_{train} of 0.955, R^2_{pred} of 0.745; QSAR_{MLR} (2), R^2_{train} of 0.854 and R^2_{pred} of 0.812; The artificial neural network QSAR_{ANN} with structural style I(9)-HL(5)-O(1) exhibited a better quality in statistical values R^2_{train} of 0.8963 and R^2_{pred} of 0.8883. The artificial neural network QSAR_{PCA-ANN} with structural style R^2 of 0.8973 and R^2_{pred} of 0.8872. The anti-cancer activities of the flavonoids in the test group and extraction of flavonoids resulting from those models turned out to give a good agreement with experimental data and those from literature.

Keywords: QSAR_{MLR}, QSAR_{ANN}, QSAR_{PCA-ANN}, anticancer activities Hela.



Bùi Thị Phương Thúy sinh ngày 10/08/1988 tại Đắk Lắk. Năm 2010, bà tốt nghiệp cử nhân chuyên ngành Sư phạm Hóa học tại Trường Đại học Tây Nguyên. Năm 2012, bà nhận bằng thạc sĩ chuyên ngành Hóa lý thuyết và hóa lý tại Trường Đại học Khoa học, Đại học Huế. Hiện nay bà đang công tác tại Trường Đại học Công nghệ Đồng Nai.

Lĩnh vực nghiên cứu: Hóa lý thuyết và hóa lý, hóa hữu cơ.



Hoàng Thị Kim Dung sinh ngày 27/3/1976 tại Thanh Hóa. Năm 1998, bà tốt nghiệp Đại học Bách khoa TP. HCM chuyên ngành Hóa hữu cơ. Năm 2003, bà nhận bằng thạc sĩ chuyên ngành Công nghệ Hóa học tại trường Đại học Bách khoa TP. HCM. Năm 2010, bà nhận bằng Tiến sĩ chuyên ngành Hóa hữu cơ tại Viện Công nghệ Hóa học – Viện Hàn lâm Khoa học và Công nghệ Việt Nam. Hiện nay bà đang công tác tại Viện Công nghệ Hóa học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Lĩnh vực nghiên cứu: hóa hữu cơ, hóa dược, hóa học các hợp chất thiên nhiên.



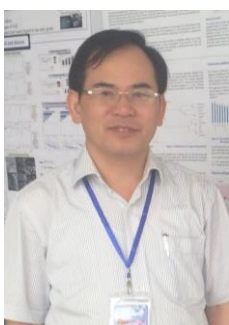
Phùng Văn Trung sinh ngày 01/05/1977 tại Thừa Thiên Huế. Năm 1998, ông tốt nghiệp cử nhân chuyên ngành Hoá học tại Trường ĐH Khoa học Tự nhiên Tp. Hồ chí Minh. Năm 2004, ông nhận bằng thạc sĩ chuyên ngành Hoá hữu cơ tại ĐH Khoa học Tự nhiên Tp. Hồ Chí Minh. Năm 2015, ông nhận bằng tiến sĩ Hoá học tại Viện Công nghệ Hoá học. Hiện nay, ông đang công tác tại Viện Công nghệ Hoá học.

Lĩnh vực nghiên cứu: Hoá học các hợp chất thiên nhiên.



Trần Dương sinh ngày 01/01/1960 tại Quảng Ngãi. Năm 1981, ông tốt nghiệp cử nhân ngành Hóa học tại Trường Đại học Sư phạm, Đại học Huế. Năm 1989, ông nhận bằng thạc sĩ chuyên ngành Hóa vô cơ tại Trường Đại học Sư phạm Hà Nội. Năm 2005, ông nhận bằng tiến sĩ chuyên ngành Hóa vô cơ tại Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội. Năm 2016, ông được công nhận chức danh phó giáo sư ngành Hóa học. Hiện nay, ông đang công tác tại Trường Đại học Sư phạm, Đại học Huế.

Lĩnh vực nghiên cứu: Vật liệu vô cơ, Hóa tin học



Phạm Văn Tất sinh ngày 30/11/1966 Tại Nam Định. Năm 1989 ông tốt nghiệp cử nhân Kỹ thuật Hóa học, Đại học Quốc gia Hà Nội. Năm 2002, ông nhận bằng thạc sĩ chuyên ngành Hoá học phân tích tại Trường Đại học Đà Lạt. Năm 2006 ông nhận bằng tiến sĩ chuyên ngành Hóa học tính toán - Lý thuyết tại Trường Đại học zu Köln, Deutschland. Năm 2011 ông được công nhận chức danh phó giáo sư ngành Hóa học. Hiện nay, ông đang công tác tại Trường Đại học Hoa Sen TP. HCM.

Lĩnh vực nghiên cứu: Hóa tính toán, Hóa môi trường, Hóa hữu cơ, hóa dược.