

Hình 3.8. Đồ thị phân tán của các tập dữ liệu liên quan cho mục đích sàng lọc *in silico* chất ức chế và chất không ức chế P-gp, dựa trên 02 thành phần chính đầu tiên.

Trong quá trình sàng lọc học máy, 10/95 chalcon nội bộ: F7, F33, F36, F37, F59, F60, F89, F90, F91, F95 (**Bảng 3.14**, tham khảo cấu trúc tương ứng trong **Phụ lục 1**) và 738/6676 hợp chất Ngân hàng Thuốc được xác định là các chất ức chế P-gp từ kết quả dự đoán của cả tám mô hình đơn lẻ là C5.0, mạng nơron, SVM, hồi quy logistic, CHAID, C&R Tree, mạng Bayesian, QUEST và mô hình kết hợp. Chỉ có một chalcon nội bộ duy nhất là F62 (tham khảo cấu trúc tương ứng trong **Phụ lục 1**) thu được từ kết quả dự đoán của mô hình mặt nghiêng quyết định, cùng với bốn mô hình khác là C5.0, SVM, hồi quy logistic và mạng Bayesian (**Bảng 3.14**). Tất cả mười mô hình được tạo ra, bao gồm mặt nghiêng quyết định cùng thống nhất dự đoán 211/738 hợp chất Ngân hàng Thuốc là các chất dương tính về khả năng ức chế P-gp. Kết quả sàng lọc được trình bày chi tiết trong **tập tin TLBS.xlsx, Sheet10**.

Bảng 3.14. Tóm tắt kết quả sàng lọc *in silico* của 95 chalcon nội bộ.

Chất	Số mô hình đồng thuận là chất ức chế P-gp		Dự đoán bằng mô hình kết hợp		Sàng lọc Ph4		Điểm số docking (kJ/mol)
	Trong số 9 mô hình ^a	Trong số 10 mô hình ^b	pIC ₅₀ dự đoán	IC ₅₀ dự đoán (μM)	HHHaV ^c	RRHdV ^d	
F1	0	0	2,95	1117,19			-15,30
F2	0	0	3,41	385,29			-16,00
F3	0	0	3,51	307,15			-15,29
F4	1	1	4,87	13,55			-13,81
F5	5	5	5,01	9,83			-12,75
F6	4	4	4,84	14,35			-14,48
F7	9	9	5,27	5,33			-12,74
F8	0	0	4,69	20,38			-14,00
F9	0	0	4,64	22,91			-14,50
F10	0	0	4,81	15,48			-12,67
F11	1	1	4,98	10,39			-13,48
F12	0	0	4,81	15,55			-11,48
F13	0	0	4,74	18,13			-12,04
F14	2	2	5,21	6,21			-12,39
F15	0	0	3,47	337,09			-14,09
F16	7	7	5,15	7,09			-12,71
F17	2	2	5,52	3,00			-12,61
F18	2	2	5,11	7,82			-16,00
F19	0	0	3,60	251,79			-16,20
F20	0	0	3,65	225,23			-19,15
F21	0	0	3,31	493,47			-18,57
F22	0	0	2,93	1166,20			-17,32
F23	0	0	3,45	358,77			-19,42
F24	1	1	3,85	140,02			-18,93
F25	2	2	4,48	33,33			-16,43

Chất	Số mô hình đồng thuận là chất ức chế P-gp		Dự đoán bằng mô hình kết hợp		Sàng lọc Ph4		Điểm số docking (kJ/mol)
	Trong số 9 mô hình ^a	Trong số 10 mô hình ^b	pIC ₅₀ dự đoán	IC ₅₀ dự đoán (μ M)	HHHaV ^c	RRHdV ^d	
F26	5	5	4,16	69,12			-18,23
F27	5	5	4,56	27,75			-16,94
F28	3	3	4,21	61,37			-18,18
F29	8	8	4,83	14,94			-15,38
F30	7	7	4,95	11,17			-14,19
F31	6	6	4,74	18,20			-14,72
F32	8	8	4,92	12,13			-15,73
F33	9	9	5,14	7,25			-14,59
F34	8	8	5,00	9,97			-15,75
F35	8	8	5,22	6,07			-14,18
F36	9	9	5,32	4,76			-15,13
F37	9	9	5,42	3,80			-11,86
F38	0	0	3,76	175,34			-18,50
F39	0	0	3,79	162,43			-18,80
F40	0	0	4,33	47,03			-18,93
F41	0	0	4,31	48,72			-18,24
F42	0	0	4,34	45,72			-18,86
F43	0	0	3,48	333,83			-16,52
F44	0	0	5,43	3,76			-16,91
F45	0	0	5,53	2,92			-17,12
F46	0	0	2,99	1026,83			-19,72
F47	1	1	4,43	36,76			-15,24
F48	7	7	4,59	25,69			-15,38
F49	1	1	4,81	15,35			-12,37
F50	0	0	4,76	17,56			-11,44
F51	0	0	4,57	26,86			-14,71
F52	0	0	3,68	209,82			-20,75
F53	0	0	3,75	176,37			-21,75
F54	0	0	4,00	99,57			-21,08
F55	0	0	3,92	119,88			-21,48
F56	0	0	3,83	148,37			-20,49
F57	0	0	3,04	920,89			-23,85
F58	8	8	4,56	27,28	x		-16,08
F59	9	9	4,22	59,86	x		-16,20
F60	9	9	4,15	71,39			-15,94
F61	1	1	5,00	9,95			-13,63
F62	4	5	3,97	106,93			-12,15
F63	7	7	5,16	6,97			-15,87
F64	1	1	4,20	63,02			-12,84
F65	1	1	4,99	10,33			-14,93
F66	0	0	3,25	567,41			-14,81
F67	0	0	3,86	139,45			-13,78
F68	0	0	3,86	138,81			-15,83
F69	0	0	4,31	48,67			-12,14
F70	0	0	3,82	150,09			-12,98
F71	0	0	4,73	18,49			-13,36

Chất	Số mô hình đồng thuận là chất ức chế P-gp		Dự đoán bằng mô hình kết hợp		Sàng lọc Ph4		Điểm số docking (kJ/mol)
	Trong số 9 mô hình ^a	Trong số 10 mô hình ^b	pIC ₅₀ dự đoán	IC ₅₀ dự đoán (μM)	HHHaV ^c	RRHdV ^d	
F72	0	0	3,53	293,83			-14,13
F73	0	0	4,55	28,43			-11,57
F74	0	0	4,09	80,81			-11,30
F75	0	0	4,53	29,54			-12,48
F76	0	0	3,09	820,90			-13,87
F77	0	0	3,42	380,30			-12,26
F78	0	0	3,70	198,46			-10,83
F79	0	0	3,82	151,21			-11,38
F80	0	0	3,40	400,33			-16,22
F81	0	0	3,70	200,07			-11,17
F82	0	0	3,21	616,48			-16,92
F83	0	0	3,61	247,44			-18,70
F84	0	0	3,20	627,32			-17,13
F85	0	0	3,56	278,04			-18,40
F86	0	0	3,57	268,00			-18,68
F87	0	0	3,61	247,44			-18,70
F88	8	8	5,57	2,67		x	-19,66
F89	9	9	6,19	0,64	x	x	-20,66
F90	9	9	6,10	0,80	x	x	-20,41
F91	9	9	6,06	0,87	x	x	-21,52
F92	1	1	3,25	564,75			-16,56
F93	0	0	3,78	167,21			-14,40
F94	0	0	3,69	204,79			-16,02
F95	9	9	5,16	6,98			-17,05

^aKhông kể mô hình mặt nghiêng quyết định; ^bBao gồm mô hình mặt nghiêng quyết định; ^cPharmacophore chất ức chế P-gp mạnh; ^dPharmacophore chất ức chế NorA nhưng không ức chế P-gp.

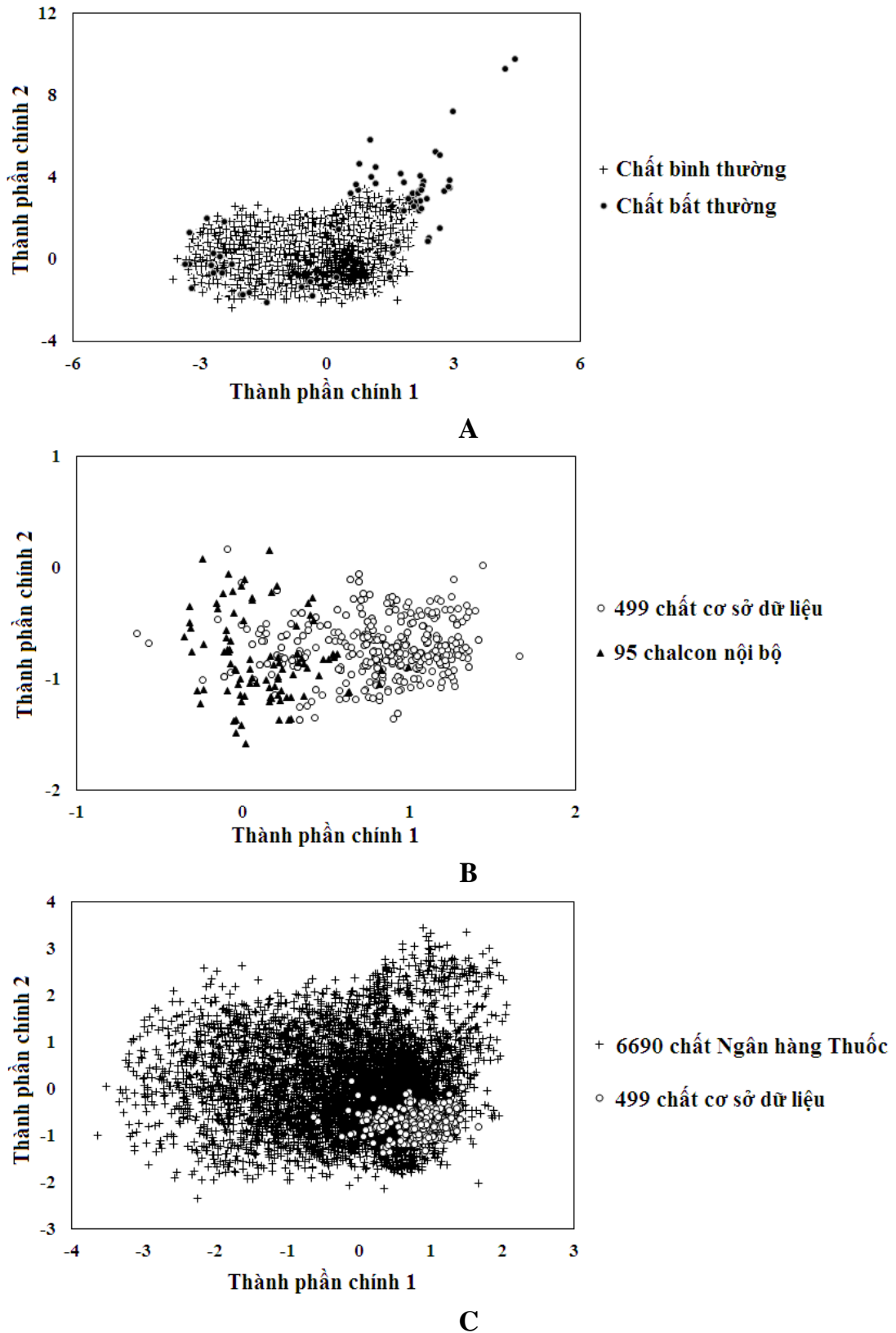
3.3.2. Dự đoán hoạt tính ức chế P-gp

Việc dự đoán hoạt tính sinh học trên P-gp (giá trị IC₅₀ hoặc pIC₅₀) của các tập dữ liệu nghiên cứu nhằm mục đích xác định các chất ức chế tiềm năng với IC₅₀ ≤ 15μM (giá trị ngưỡng để được xem là chất ức chế P-gp) [147] của bơm ngược đề kháng đa thuốc này. Dựa trên 34 thông số mô tả được trình bày trong **Phụ lục 4**, các chất có đầy đủ giá trị thuộc tính cùng với 499 chất của toàn bộ tập dữ liệu dùng trong việc huấn luyện và đánh giá các mô hình dự đoán được kiểm tra nhanh bằng hạch “Anomaly Detection” trong Clementine để phát hiện các chất bất thường. Tỷ lệ của các chất bất thường vẫn được thiết lập ở mức 1 %. Ngoài ra, phân tích thành phần

chính (PCA) của tất cả các chất này cũng được thực hiện với số lượng tối đa là năm thành phần.

Sau quá trình sàng lọc sơ bộ, 73 chất bất thường được phát hiện từ 6763 hợp chất Ngân hàng Thuốc không bị thiếu thông số mô tả nào (**tập tin TLBS.xlsx, Sheet9**). Sự phân bố của các chất thuộc các cơ sở dữ liệu khác nhau theo hai thành phần chính đầu tiên được thể hiện trong **Hình 3.9**. Đồ thị phân tán đầu tiên (**Hình 3.9A**) cho thấy tính hợp lý của việc loại bỏ các chất bất thường là những chất có phần tách biệt khỏi các chất khác trong không gian hai chiều được chọn. Trong khi đồ thị thứ hai (**Hình 3.9B**) cho thấy sự phân tán của 95 chalcon nội bộ nằm trong không gian hóa học của cơ sở dữ liệu 499 chất dùng cho việc phát triển các mô hình dự đoán, thì trong đồ thị thứ ba (**Hình 3.9C**) nhiều chất thuộc tập Ngân hàng Thuốc lại phân tán nằm ngoài khu vực này. Nói cách khác, các mô hình QSAR có nhiều khả năng dự đoán đúng hoạt tính sinh học của các chalcon trong tập dữ liệu nội bộ hơn so với các chất trong tập Ngân hàng Thuốc.

Kết quả dự đoán hoạt tính sinh học bằng các mô hình QSAR trên hai tập dữ liệu đề cập (95 chalcon nội bộ và 6690 hợp chất Ngân hàng Thuốc) được trình bày chi tiết trong **tập tin TLBS.xlsx, Sheet11**. Trong đó, mô hình tốt nhất là mô hình kết hợp đã dự đoán 27 chalcon nội bộ (**Bảng 3.14**, tham khảo cấu trúc tương ứng trong **Phụ lục 1**) và 2373 hợp chất Ngân hàng Thuốc có giá trị IC_{50} thấp hơn giá trị ngưỡng là 15 μM . Dựa trên các nhóm thế, các chalcon này có thể được phân loại thành năm nhóm là: 2'-hydroxy (F29, F30, F32, F33, F34, F35, F36, F37, F44, F45); 4'-bromo (F11, F14, F16, F17); methoxy (F4, F5, F6, F7, F61, F63); amino (F18, F95) và dị vòng A (F65, F88, F89, F90, F91). Trong số 47 hợp chất Ngân hàng Thuốc có giá trị pIC_{50} dự đoán ≥ 7 (**Bảng 3.15**), amiodaron, tacrolimus, indinavir, ritonavir, cyclosporin và saquinavir là các chất ức chế P-gp đã biết, theo Cơ quan Quản lý Thực phẩm và Dược phẩm Hoa Kỳ (U.S. FDA) [187].



Hình 3.9. Đồ thị phân tán của các tập dữ liệu liên quan cho mục đích dự đoán *in silico* hoạt tính ức chế P-gp, dựa trên 02 thành phần chính đầu tiên.

Bảng 3.15. Tóm tắt kết quả sàng lọc *in silico* của 47 chất từ Ngân hàng Thuốc với các giá trị pIC₅₀ trên P-gp được dự đoán bởi mô hình kết hợp ≥ 7 .

Chất	Nhóm thuốc	Tên generic	Số mô hình đồng thuận là chất ức chế P-gp		Dự đoán bằng mô hình kết hợp		Sàng lọc Ph4		Điểm số docking (kJ/mol)
			Trong số 9 mô hình ^a	Trong số 10 mô hình ^b	pIC ₅₀ dự đoán	IC ₅₀ dự đoán (μM)	HHHaV ^c	RRHdV ^d	
DB01249	A	Iodixanol	2	2	9,2657	0,0005			-17,78
DB01118	A; I	Amiodaron	9	9	8,5249	0,0030			-7,50
DB04869	I	Olcegepant	9	9	8,3513	0,0045			-20,06
DB01578	A	Metrizamid	1	1	7,9275	0,0118			-17,95
DB01362	A	Iohecol	0	0	7,6472	0,0225			-13,48
DB04775	E	Reidispongiolid C	9	9	7,5282	0,0296			2,17
DB02872	E	-	9	10	7,5173	0,0304			-14,82
DB02860	E	Calyculin A	6	6	7,4863	0,0326			-6,19
DB04774	E	Reidispongiolid A	9	9	7,4600	0,0347			Không dock được
DB06748	N	Ginsenosid C	2	2	7,4478	0,0357	Không có chất nào thỏa mãn các mô hình Ph4		Không dock được
DB06749	N	Ginsenosid Rb1	2	2	7,4242	0,0377			1,30
DB04783	E	Sphinxolid B	9	9	7,4125	0,0387			2,19
DB04933	I	Eritoran	6	6	7,4077	0,0391			Không dock được
DB03621	E	L-709,587	9	9	7,3917	0,0406			Không dock được
DB04711	A	Iodipamid	0	0	7,3042	0,0496			-12,19
DB06439	A; I	Tyloxapol	8	9	7,2874	0,0516			-0,01
DB00864	A; I	Tacrolimus	9	9	7,2809	0,0524			0,49

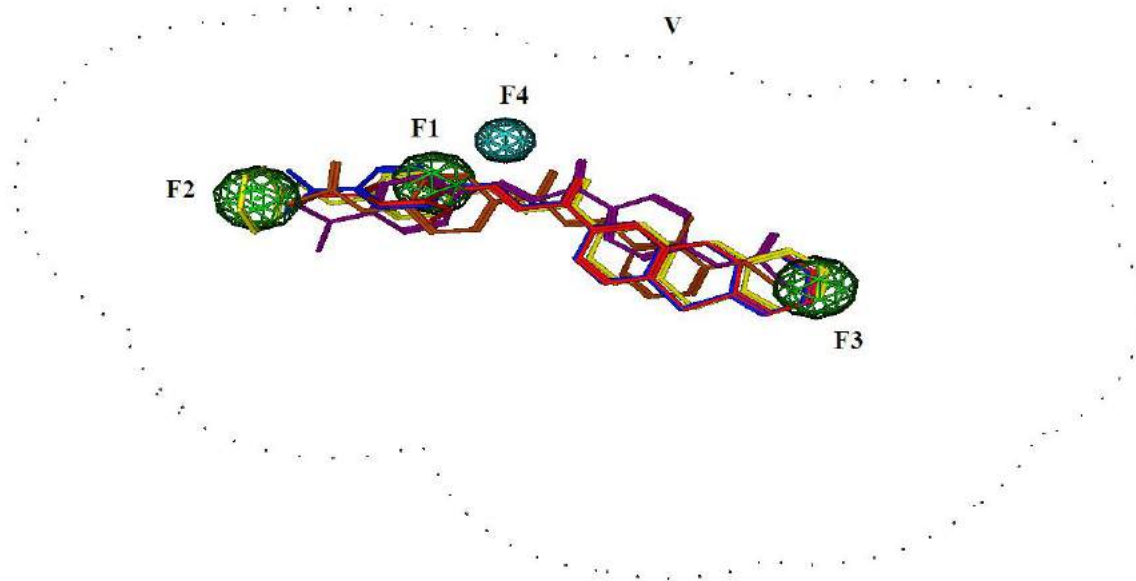
Chất	Nhóm thuốc	Tên generic	Số mô hình đồng thuận là chất ức chế P-gp		Dự đoán bằng mô hình kết hợp		Sàng lọc Ph4		Điểm số docking (kJ/mol)
			Trong số 9 mô hình ^a	Trong số 10 mô hình ^b	pIC ₅₀ dự đoán	IC ₅₀ dự đoán (μM)	HHHaV ^c	RRHdV ^d	
DB01590	A	Everolimus	9	9	7,2567	0,0554			Không dock được
DB02169	E	9,10-Deepithio-9,10-Didehydroacanthifolicin	1	1	7,2470	0,0566			-7,99
DB00877	A; I	Sirolimus	9	9	7,2203	0,0602			-0,27
DB03933	E	Chromophor thom hóa C-1027	6	7	7,2201	0,0602			-11,04
DB00224	A	Indinavir	9	9	7,2121	0,0614			-11,67
DB01721	E	Đồng đẳng của Indinavir	9	10	7,1863	0,0651			-12,64
DB01321	A	Josamycin	7	7	7,1804	0,0660			Không dock được
DB00050	A; I	Cetrorelix	9	10	7,1644	0,0685	Không có chất nào thỏa mãn các mô hình Ph4		Không dock được
DB08947	A	Iopamidol	0	0	7,1543	0,0701			-10,18
DB00503	A; I	Ritonavir	9	10	7,1521	0,0705			-7,43
DB04748	E	Oximinoarylsulfonamid	9	9	7,1511	0,0706			-10,54
DB02253	E	-	9	10	7,1479	0,0711			-10,09
DB04934	I	Benzoxazinorifamycin	Chất bất thường bị loại		7,1385	0,0727			-5,44
DB08889	A	Carfilzomib	9	9	7,1330	0,0736			-9,50
DB02785	E	-	9	9	7,1320	0,0738			-18,06
DB01723	E	-	1	2	7,1055	0,0784			-9,95
DB02888	E	FKB-001	9	10	7,0942	0,0805			-3,91
DB02009	E	L-756,423	9	10	7,0694	0,0852			-17,07
DB00403	A	Ceruletid	0	1	7,0560	0,0879			Không dock được

Chất	Nhóm thuốc	Tên generic	Số mô hình đồng thuận là chất ức chế P-gp		Dự đoán bằng mô hình kết hợp		Sàng lọc Ph4		Điểm số docking (kJ/mol)
			Trong số 9 mô hình ^a	Trong số 10 mô hình ^b	pIC ₅₀ dự đoán	IC ₅₀ dự đoán (μM)	HHHaV ^c	RRHdV ^d	
DB01226	A	Mivacurium	Chất bất thường bị loại		7,0407	0,0911			5,69
DB01152	W	Candididin	3	3	7,0395	0,0913			Không dock được
DB08874	A	Fidaxomicin	6	6	7,0379	0,0916			Không dock được
DB01180	A	Rescinnamin	9	10	7,0376	0,0917			-13,65
DB04378	E	-	9	10	7,0273	0,0939	Không có chất nào thỏa mãn các mô hình Ph4		-16,80
DB00954	A	Dirithromycin	8	8	7,0172	0,0961			1,65
DB04318	E	-	9	10	7,0135	0,0969			-14,07
DB00091	A; I	Cyclosporin	9	9	7,0119	0,0973			Không dock được
DB01232	A; I	Saquinavir	9	10	7,0079	0,0982			-19,31
DB01135	A	Doxacurium	Chất bất thường bị loại		7,0054	0,0988			6,88
DB00778	A; W	Roxithromycin	6	6	7,0026	0,0994			-4,83

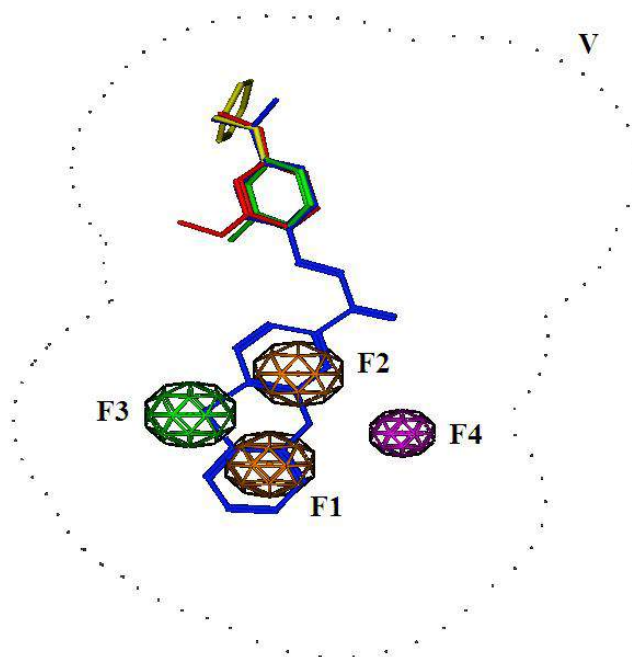
^aKhông kể mô hình mặt nghiêng quyết định; ^bBao gồm mô hình mặt nghiêng quyết định; ^cPharmacophore chất ức chế P-gp mạnh; ^dPharmacophore chất ức chế NorA nhưng không ức chế P-gp; A: Đã được phê duyệt; E: Đang thử nghiệm; I: Đang nghiên cứu; N: Dinh dưỡng; W: Đã bị thu hồi.

3.3.3. Sàng lọc bằng pharmacophore

Các hình thể ba chiều (3D) được tạo ra cho 95 chalcon nội bộ và 4456 chất từ cơ sở dữ liệu Ngân hàng Thuốc gồm 6874 chất. Trong quá trình sàng lọc bằng pharmacophore, có 5/95 chalcon nội bộ là F58, F59, F89, F90, F91 (**Bảng 3.14, Hình 3.10**, tham khảo cấu trúc tương ứng trong **Phụ lục 1**) và 69/4456 hợp chất Ngân hàng Thuốc thỏa mãn mô hình pharmacophore của các chất ức chế P-gp mạnh (HHHaV); đồng thời có 4/95 chalcon nội bộ là F88, F89, F90, F91 (**Bảng 3.14, Hình 3.11**, tham khảo cấu trúc tương ứng trong **Phụ lục 1**) và 611/4456 hợp chất Ngân hàng Thuốc thỏa mãn mô hình pharmacophore của các chất ức chế NorA mà không ức chế P-gp (RRHdV). Kết quả sàng lọc được trình bày chi tiết trong **tập tin TLBS.xlsx, Sheet12**.



Hình 3.10. Năm chalcon thỏa pharmacophore chất ức chế P-gp mạnh (F1, F2, F3: Nhóm kỵ nước; F4: Nhóm nhận liên kết hydro; V: Giới hạn thể tích): F58 (tím); F59 (cam); F89 (vàng); F90 (đỏ); F91 (xanh dương).



Hình 3.11. Bốn chalcone thỏa pharmacophore chất ức chế NorA mà không ức chế P-gp (F1, F2: Yếu tố vòng thơm/vòng Pi; F3: Nhóm kỵ nước; F4: Nhóm cho liên kết hydro; V: Giới hạn thể tích): F88 (xanh lá); F89 (vàng); F90 (đỏ); F91 (xanh dương).

3.3.4. Docking phân tử

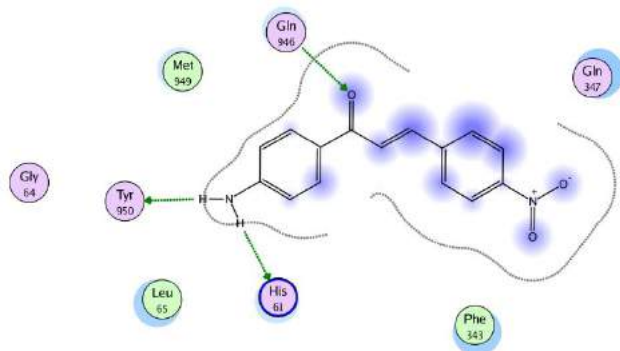
Nghiên cứu docking được thực hiện với 95 chalcone nội bộ (**Bảng 3.14**) và 47 hợp chất Ngân hàng Thuốc có giá trị IC_{50} dự đoán thấp nhất (**Bảng 3.15**) để xác định các mô hình gắn kết có thể và ái lực gắn kết của các chất này với mô hình tương đồng tốt nhất của P-gp bằng gói FlexX trong LeadIT. Ngoài ra, ba chất ức chế phân tử nhỏ của P-gp là reserpin, tariquidar và elacridar [173] cũng được tiến hành docking để so sánh. Tổng cộng 134 phân tử bao gồm 95 chalcone nội bộ, 36 hợp chất Ngân hàng Thuốc và 3 chất so sánh đã được dock thành công vào túi gắn kết của protein mục tiêu, được xác định bao gồm các acid amin trong bán kính 6,5 Å tính từ phối tử tham khảo QZ59-RRR. Tuy nhiên, thuật toán docking đã thất bại với 11 hợp chất Ngân hàng Thuốc còn lại do cấu trúc phức tạp và công kênh của chúng. Kết quả docking được trình bày chi tiết trong **tập tin TLBS.xlsx, Sheet13, Bảng 3.14** và **Bảng 3.15**.

Với điểm số docking nằm trong khoảng -23,85 đến -10,83 kJ/mol, các chalcone nội bộ cho thấy khả năng gắn kết tốt vào khoang bên trong của mô hình tương đồng

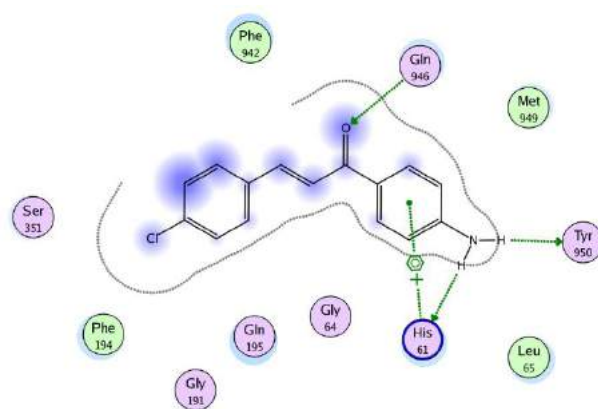
P-gp. Trong đó, ba chalcon F57, F53 và F91 có điểm số docking thấp nhất, lần lượt là -23,85; -21,75 và -21,52 kJ/mol (**Hình 3.12**). Mặc dù là những dẫn xuất được dự đoán có ái lực gắn kết mạnh nhất với P-gp trong số 95 chalcon, F57 và F53 lại có các giá trị IC_{50} dự đoán $> 15 \mu M$ (xấp xỉ 921 và 176 μM), không được bất cứ mô hình phân loại nào xác định là chất ức chế P-gp và không thỏa bất kỳ pharmacophore nào được xây dựng.

Điểm số docking cao hơn về tổng thể (trong khoảng -20,06 đến 6,88 kJ/mol) cho thấy các hợp chất Ngân hàng Thuốc có khả năng gắn kết vào mục tiêu tác động kém hơn so với các dẫn xuất chalcon, mặc dù chúng có các giá trị IC_{50} dự đoán thấp hơn đáng kể. Không kể các chất không dock được, 7/36 hợp chất Ngân hàng Thuốc có điểm số docking dương. Trong số 29 chất còn lại, 3 chất có điểm số docking thấp nhất là DB04869 (olcegepant; -20,06 kJ/mol); DB01232 (saquinavir; -19,31 kJ/mol) và DB02785 (-18,06 kJ/mol) (**Hình 3.12**). Ngoài ra, cả ba chất này cũng có các giá trị IC_{50} dự đoán $< 0,1 \mu M$, được ít nhất 9 mô hình phân loại xác định là chất ức chế P-gp, dù không thỏa bất kỳ pharmacophore nào được xây dựng.

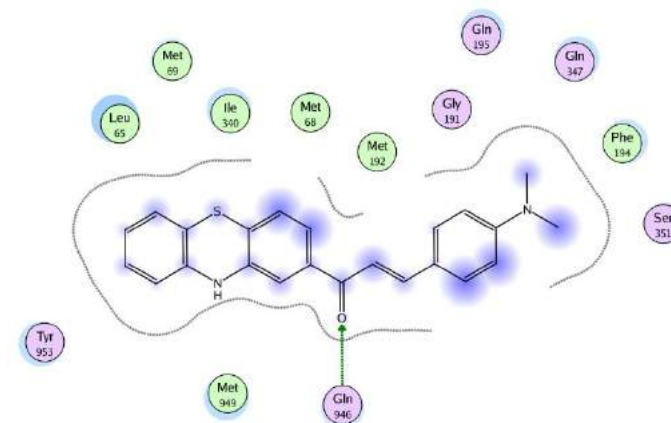
Tất cả các chất thử nghiệm trong **Hình 3.12** đều có điểm số docking tương đương elacridar (-22,86 kJ/mol) và tốt hơn so với reserpin và tariquidar (-3,47 kJ/mol và -13,26 kJ/mol). Qua cách thức gắn kết của các chất này, các acid amin chịu trách nhiệm chính cho các tương tác gắn kết phối tử - protein (liên kết hydro, aren-cation và kỵ nước) được xác định là His61, Gln946, Tyr950, Leu65, Met949, Phe194, Gln195 trong trường hợp các chalcon và Gln347, Phe942, Phe343, Phe194, Gln195, Ser344, Gln946 trong trường hợp các hợp chất Ngân hàng Thuốc. Trong đó, các acid amin Phe194, Gln195 ở vùng xuyên màn TM3 và Gln946, Tyr950 ở vùng xuyên màn TM11 giúp tái khẳng định túi gắn kết thuốc đặt tại mặt phân giới TM3/TM11 đã được Chiba và cộng sự [25] chỉ ra trước đó. Các phương pháp khác như đánh dấu ái lực quang học (photoaffinity labeling), đột biến điểm định hướng (site-directed mutagenesis), đầu dò thiol và nghiên cứu cấu trúc đồng kết tinh của P-gp với các phối tử nói trên có thể được sử dụng để đánh giá vai trò của các acid amin được chỉ ra trong nghiên cứu này.



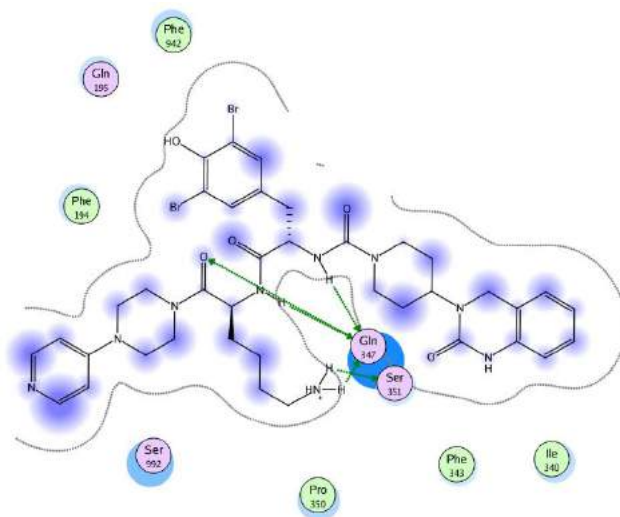
F57 (-23,85 kJ/mol)



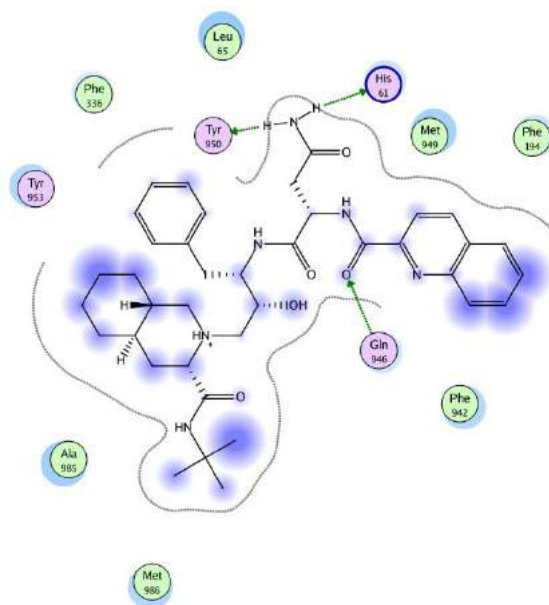
F53 (-21,75 kJ/mol)



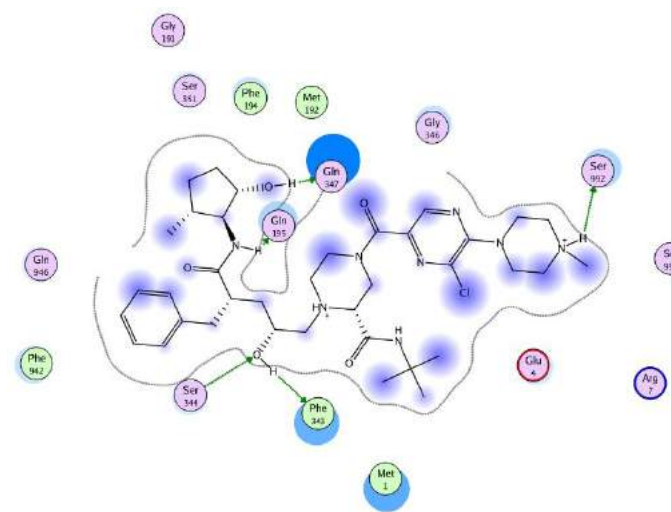
F91 (-21,52 kJ/mol)



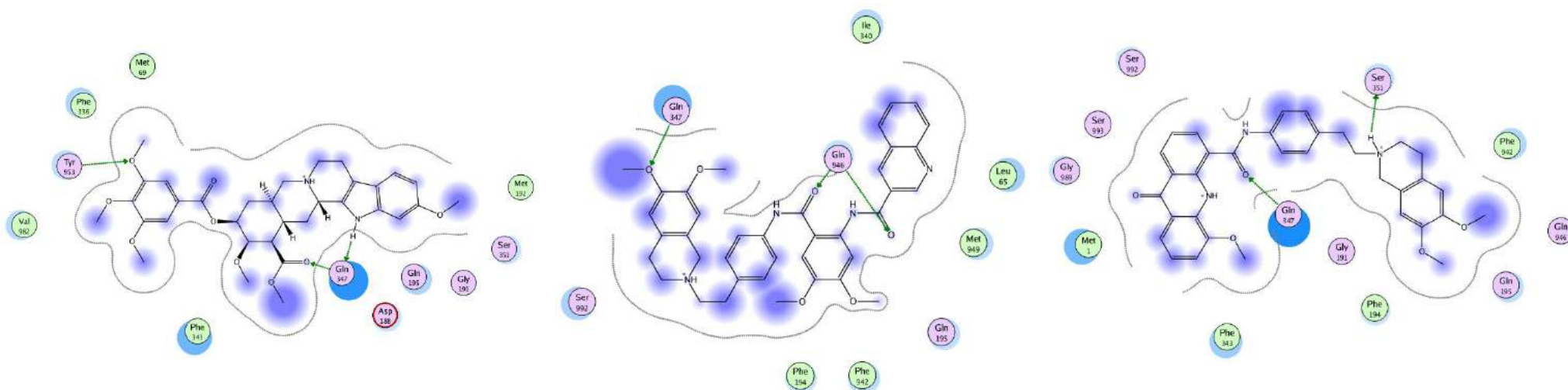
DB04869 (Olcegepant; -20,06 kJ/mol)



DB01232 (Saquinavir; -19,31 kJ/mol)



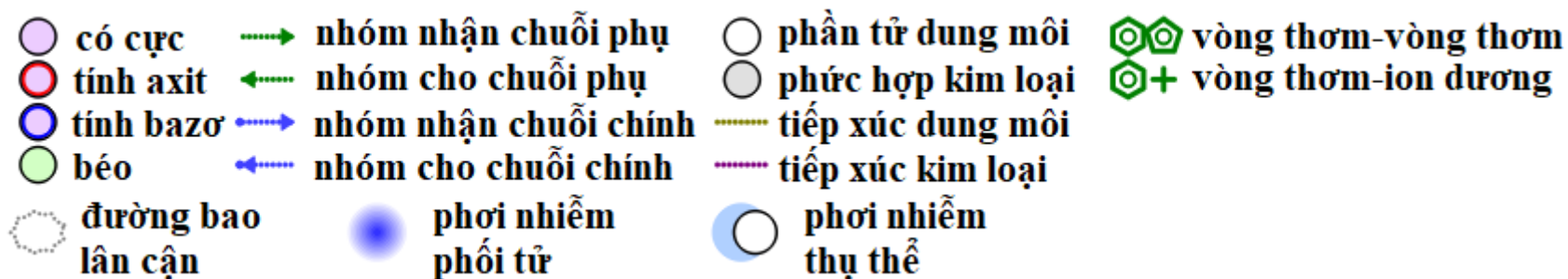
DB02785 (-18,06 kJ/mol)



Reserpin (-3,47 kJ/mol)

Tariquidar (-13,26 kJ/mol)

Elacridar (-22,86 kJ/mol)



Hình 3.12. Hình ảnh docking vào mô hình tương đồng của P-gp của ba chalcon và ba hợp chất Ngân hàng Thuốc có điểm số docking tốt nhất, cùng với ba chất ức chế P-gp đã biết là reserpin, tariquidar và elacridar.

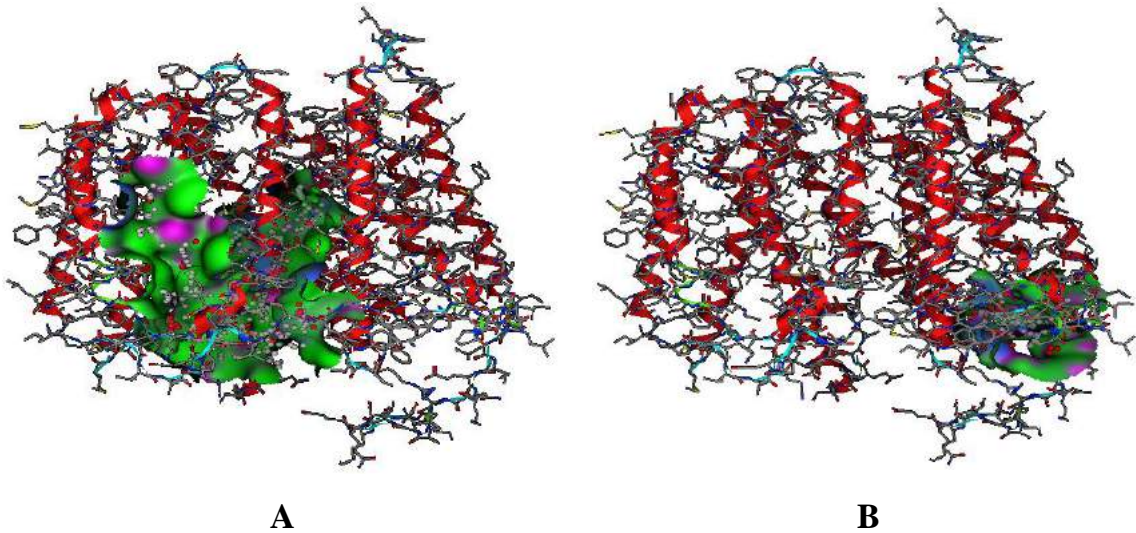
3.4. Sàng lọc *in silico* và thử nghiệm *in vitro* đánh giá tác dụng ức chế bơm ngược NorA trên *S. aureus* của một số chalcon nội bộ

Dựa vào kết quả sàng lọc *in silico* trên P-gp (**Bảng 3.14**) và tính sẵn có của các chalcon tổng hợp, bốn dẫn xuất F29, F88, F90 và F91 được chọn để kiểm tra hoạt tính ức chế bơm NorA của *S. aureus* trên *in silico* và *in vitro*, do giữa hai bơm ngược này có sự chia sẻ hay chồng phủ phối tử nhất định [17]. Ngoài pharmacophore chất ức chế NorA nhưng không ức chế P-gp được xây dựng trong nghiên cứu này, các công cụ máy tính khác bao gồm mô hình tốt nhất dự đoán IC₅₀ trên NorA (mô hình D) và mô hình tương đồng của bơm ngược này được tạo ra trong một nghiên cứu khác của nhóm chúng tôi [181] cũng được sử dụng để dự đoán máy tính cho các chalcon được chọn trước khi tiến hành thử nghiệm thực nghiệm tác dụng ức chế bơm ngược trên vi khuẩn.

3.4.1. Đánh giá *in silico* trên NorA

Mô hình hồi quy tuyến tính D (12) được thiết lập bằng kỹ thuật bình phương tối thiểu từng phần (Partial Least Square) sử dụng ba thông số mô tả là rings, balabanJ và logS, từ một tập dữ liệu 45 chất ($R^2 = 0,80$) và được đánh giá bằng một tập ngoại 15 chất ($R^2 = 0,60$) [181]. Trong khi đó, mô hình tương đồng tốt nhất của NorA (C-score = 1,35; TM-score = $0,9 \pm 0,06$; RMSD = $4,0 \pm 2,7$; mật độ đám = 0,91) cũng được mô hình hóa bằng server tự động I-TASSER, sử dụng đĩa là chuỗi A của protein chuyên chở đa thuốc EmrD của *E. coli* (mã định danh: 2gfp), thuộc cùng Liên họ trợ giúp chính (Major Facilitator Superfamily) với NorA. Hai vị trí gắn kết có xếp hạng cao nhất của mô hình này là khoang trung tâm và Walker B (**Hình 3.13**), được xác định bằng Site Finder trong MOE để phục vụ docking [181].

$$pIC_{50} = 5,91838 - 0,11195*rings - 1,38078*balabanJ - 0,24425*logS \quad (12)$$

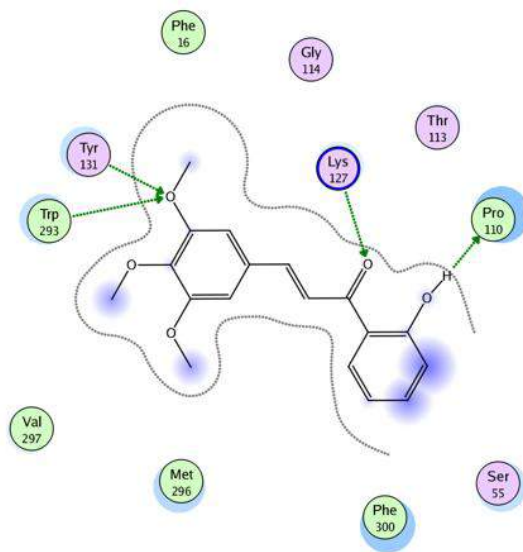


Hình 3.13. Mô hình tương đồng tốt nhất của NorA với 02 vị trí gắn kết phối tử được dự đoán: (A) Khoảng trung tâm; (B) Walker B.

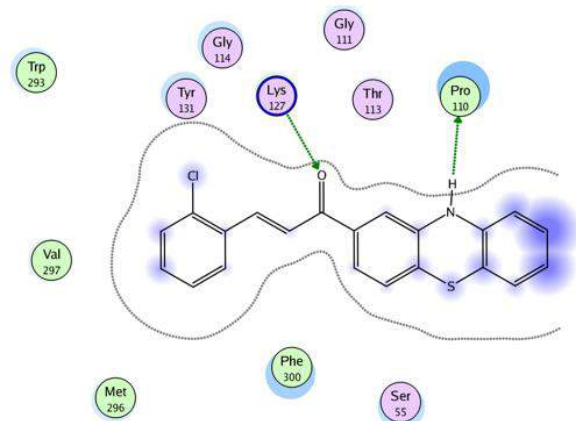
Kết quả đánh giá hoạt tính *in silico* trên NorA của các chalcon nội bộ được trình bày trong **Bảng 3.16** và **Hình 3.14**. Theo mô hình D, chalcon F88 là dẫn xuất tiềm năng nhất do sở hữu giá trị IC_{50} dự đoán thấp nhất ($3,84 \mu\text{M}$), tiếp theo là F90 và F91 (xấp xỉ $5 \mu\text{M}$), sau cùng là F29 ($88,27 \mu\text{M}$). Tất cả chalcon đều được dock thành công vào khoang trung tâm và Walker B của mô hình tương đồng NorA với điểm số docking tốt (từ $-30,33$ đến $-22,65 \text{ kJ/mol}$), đồng nghĩa với việc các chất này có ái lực gắn kết mạnh với protein. Các acid amin đóng vai trò quan trọng cho sự gắn kết phối tử ở khoang trung tâm là Lys127, Tyr131, Pro110, Trp293, Met296, Phe300, Val297, Ser55, Gln51 và ở Walker B là Glu376, His379, Lys377, Leu374, Arg380, Gln378, Phe306, Phe259, Asn200. Hiệu lực ức chế bơm ngược NorA của chúng tiếp tục được kiểm chứng bằng thử nghiệm kháng khuẩn.

Bảng 3.16. Kết quả dự đoán hoạt tính ức chế NorA bằng mô hình D và docking vào mô hình tương đồng của protein này của các chalcon “hit”.

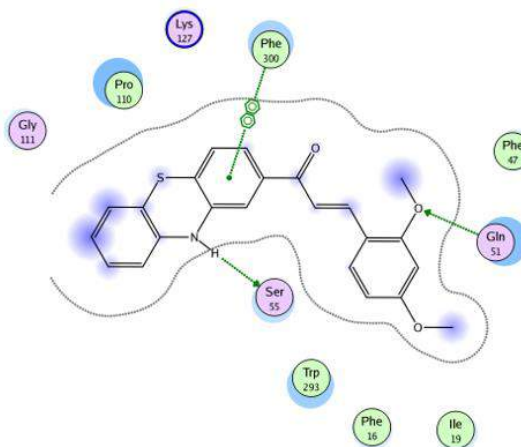
Chất	rings	balabanJ	logS	pIC ₅₀ dự đoán	IC ₅₀ dự đoán (μ M)	Điểm số docking (kJ/mol)	
						Khoang trung tâm	Walker B
F29	2	1,85	-3,75	4,05	88,27	-23,67	-26,69
F88	4	1,26	-6,92	5,42	3,84	-22,65	-29,06
F90	4	1,25	-6,29	5,28	5,24	-23,76	-30,33
F91	4	1,21	-6,12	5,30	5,07	-24,96	-27,78



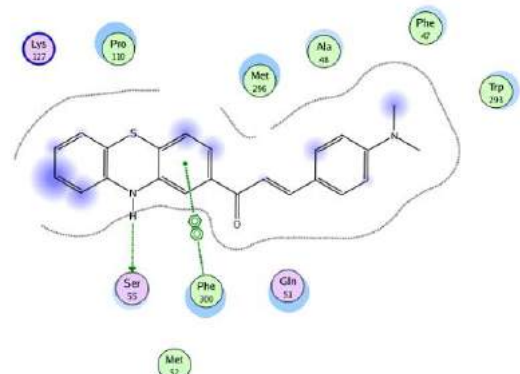
F29 (-23,67 kJ/mol)



F88 (-22,65 kJ/mol)

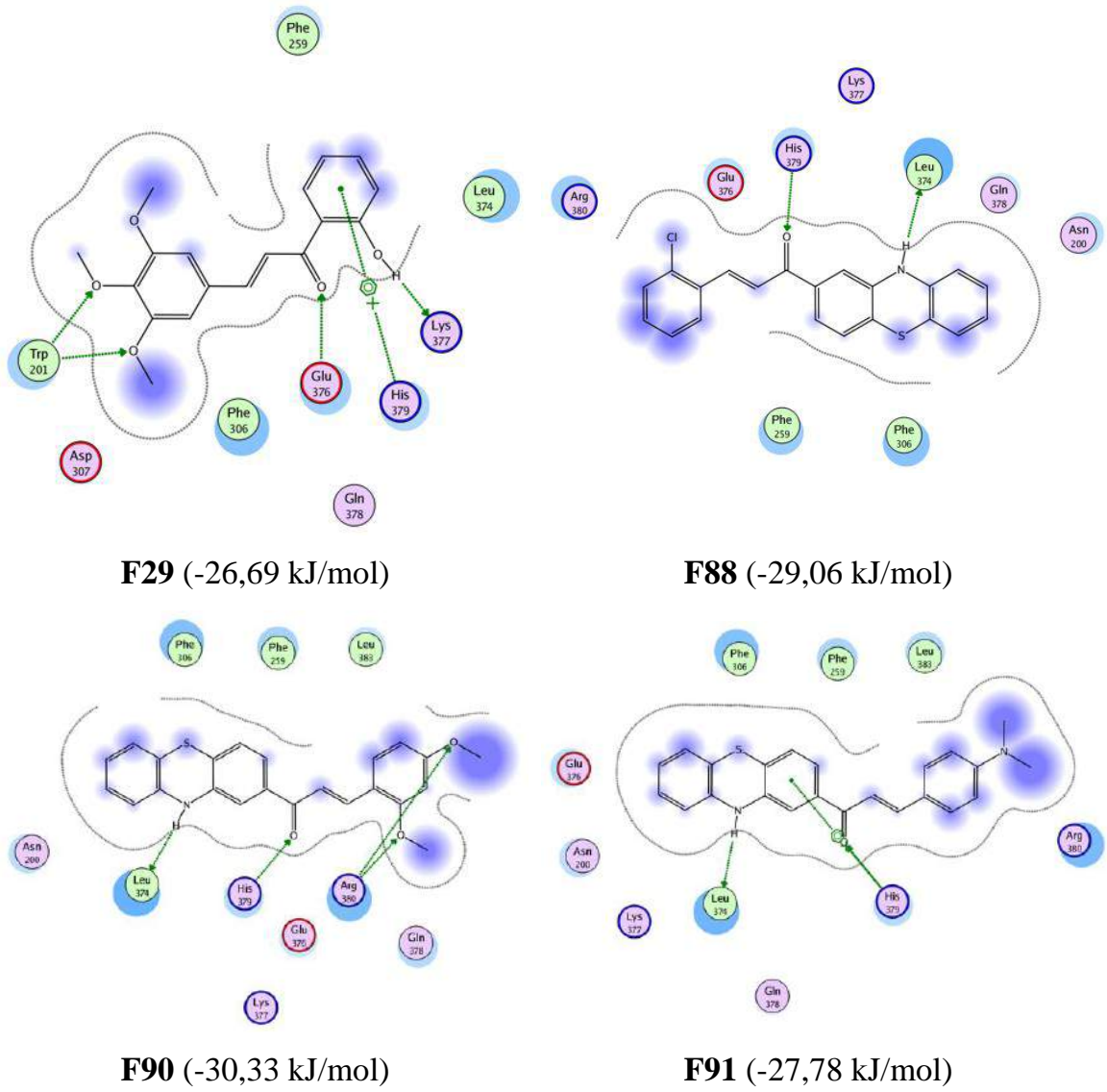


F90 (-23,76 kJ/mol)



F91 (-24,96 kJ/mol)

A. Khoang trung tâm



B. Walker B

- | | | | |
|---------------------|-------------------------|-----------------------|--------------------------|
| ○ có cực | → nhóm nhận chuỗi phụ | ○ phân tử dung môi | ⊗ ⊗ vòng thơm-vòng thơm |
| ● tính axit | ← nhóm cho chuỗi phụ | ○ phức hợp kim loại | ⊗ ⊗+ vòng thơm-ion dương |
| ● tính bazơ | → nhóm nhận chuỗi chính | --- tiếp xúc dung môi | |
| ● béo | ← nhóm cho chuỗi chính | ⋯ tiếp xúc kim loại | |
| ○ đường bao lân cận | ● phơi nhiễm phối tử | ○ phơi nhiễm thụ thể | |

Hình 3.14. Hình ảnh docking vào mô hình tương đồng của NorA của bốn chalcone “hit”: (A) Vào khoang trung tâm; (B) Vào Walker B.

3.4.2. Các thử nghiệm *in vitro*

3.4.2.1. Thử nghiệm tác dụng ức chế bơm ngược NorA trên các chủng vi khuẩn *S. aureus* SA-1199 và SA-1199B

Kết quả thử nghiệm (**Bảng 3.17**) cho thấy:

- Trên chủng tự nhiên SA-1199: Kháng sinh ciprofloxacin vẫn nhạy cảm (MIC < 0,125 µg/mL) khi không có mặt hay có mặt các chalcon.
- Trên chủng đột biến SA-1199B (sản xuất quá mức NorA):
 - Hiệu quả của ciprofloxacin giảm khi được sử dụng riêng lẻ (MIC = 4 µg/mL).
 - So với khi sử dụng riêng lẻ, mức độ nhạy cảm của ciprofloxacin được cải thiện rõ rệt (MIC = 2 µg/mL) khi có sự hiện diện của chalcon F29 ở nồng độ 100 µg/mL và các chalcon F88, F90 ở nồng độ 50 µg/mL và 100 µg/mL; nhưng lại không thay đổi (MIC = 4 µg/mL) khi có mặt chalcon F91 ở cả hai nồng độ thử nghiệm.

Bảng 3.17. Giá trị MIC (µg/mL) của ciprofloxacin trên các chủng *S. aureus* SA-1199 và SA-1199B khi vắng mặt và khi có mặt các chalcon nghiên cứu.

MIC	SA-1199	SA-1199B
Ciprofloxacin	< 0,125 µg/mL	4 µg/mL
Ciprofloxacin + F29 ^a	< 0,125 µg/mL	4 µg/mL
Ciprofloxacin + F29 ^b	< 0,125 µg/mL	2 µg/mL
Ciprofloxacin + F88 ^a	< 0,125 µg/mL	2 µg/mL
Ciprofloxacin + F88 ^b	< 0,125 µg/mL	2 µg/mL
Ciprofloxacin + F90 ^a	< 0,125 µg/mL	2 µg/mL
Ciprofloxacin + F90 ^b	< 0,125 µg/mL	2 µg/mL
Ciprofloxacin + F91 ^a	< 0,125 µg/mL	4 µg/mL
Ciprofloxacin + F91 ^b	< 0,125 µg/mL	4 µg/mL

^a50 µg/mL; ^b100 µg/mL.

3.4.2.2. Thử nghiệm sàng lọc các chủng vi khuẩn *S. aureus* phân lập từ lâm sàng đề kháng ciprofloxacin qua trung gian bơm ngược

Kết quả thử nghiệm (**Bảng 3.18**) cho thấy:

- Trên tất cả 156 chủng SA lâm sàng, MIC của ciprofloxacin được xác định thấp nhất là 0,06 µg/mL và cao nhất là 256 µg/mL.
- Lần lượt có 36 chủng và 96 chủng với giá trị MIC của ciprofloxacin giảm 4 lần và 2 lần khi có sự hiện diện của chất ức chế bơm ngược đã biết là PaβN ở nồng độ 20 µg/mL. Trong khi đó, giá trị MIC của kháng sinh này vẫn giữ nguyên trên 24 chủng còn lại khi được sử dụng kết hợp với chất ức chế bơm chuẩn ở nồng độ đã cho.
- Trong số 36 chủng tăng nhạy cảm với ciprofloxacin tốt nhất (MIC giảm 4 lần) khi kết hợp với PaβN, 10 chủng (I16.1412, I16.1421, I16.1461, I16.1505, I16.1562, I16.1604, I16.1617, I16.1635, I16.1672 và I16.1773) được chọn ngẫu nhiên để tiếp tục làm thử nghiệm ức chế bơm ngược với các dẫn xuất chalcon.

Bảng 3.18. Giá trị MIC (µg/mL) của ciprofloxacin trên các chủng *S. aureus* lâm sàng khi vắng mặt và khi có mặt chất ức chế bơm PaβN.

STT	Chủng SA	MRSA	MIC ciprofloxacin (µg/mL) (X)	MIC ciprofloxacin (µg/mL) + PaβN (Y)	Tỷ số X/Y
1	I16.1412	+	32	8	4
2	I16.1421	-	1	0,25	4
3	I16.1430	-	0,25	0,06	4
4	I16.1446	-	0,25	0,06	4
5	I16.1447	-	0,25	0,06	4
6	I16.1461	+	64	16	4
7	I16.1474	-	0,25	0,06	4
8	I16.1477	-	0,25	0,06	4
9	I16.1478	-	0,25	0,06	4
10	I16.1479	-	0,25	0,06	4
11	I16.1480	-	0,25	0,06	4
12	I16.1481	-	0,25	0,06	4
13	I16.1493	+	0,25	0,06	4
14	I16.1496	-	0,25	0,06	4
15	I16.1505	+	64	16	4
16	I16.1510	-	0,25	0,06	4
17	I16.1528	-	0,25	0,06	4
18	I16.1531	-	0,25	0,06	4

STT	Chủng SA	MRSA	MIC ciprofloxacin ($\mu\text{g/mL}$) (X)	MIC ciprofloxacin ($\mu\text{g/mL}$) + Pa β N (Y)	Tỷ số X/Y
19	I16.1537	-	0,25	0,06	4
20	I16.1538	-	0,25	0,06	4
21	I16.1540	+	0,25	0,06	4
22	I16.1562	+	64	16	4
23	I16.1563	-	0,25	0,06	4
24	I16.1567	+	0,25	0,06	4
25	I16.1604	+	16	4	4
26	I16.1615	+	0,25	0,06	4
27	I16.1617	+	32	8	4
28	I16.1635	+	32	8	4
29	I16.1672	+	32	8	4
30	I16.1690	+	0,25	0,06	4
31	I16.1701	+	0,25	0,06	4
32	I16.1702	-	0,25	0,06	4
33	I16.1773	+	32	8	4
34	I16.1778	+	0,25	0,06	4
35	I16.1801	+	0,25	0,06	4
36	I16.1806	-	0,25	0,06	4
37	I16.1407	+	16	8	2
38	I16.1411	-	0,5	0,25	2
39	I16.1413	+	0,5	0,25	2
40	I16.1416	+	64	32	2
41	I16.1420	-	32	16	2
42	I16.1425	+	0,5	0,25	2
43	I16.1432	+	32	16	2
44	I16.1436	+	32	16	2
45	I16.1437	+	32	16	2
46	I16.1454	-	16	8	2
47	I16.1456	+	16	8	2
48	I16.1460	+	0,5	0,25	2
49	I16.1464	+	16	8	2
50	I16.1465	+	16	8	2
51	I16.1466	+	0,5	0,25	2
52	I16.1469	+	256	128	2
53	I16.1471	+	32	16	2
54	I16.1485	+	32	16	2
55	I16.1491	+	32	16	2
56	I16.1499	+	64	32	2
57	I16.1503	-	32	16	2
58	I16.1517	+	32	16	2

STT	Chủng SA	MRSA	MIC ciprofloxacin ($\mu\text{g/mL}$) (X)	MIC ciprofloxacin ($\mu\text{g/mL}$) + Pa β N (Y)	Tỷ số X/Y
59	I16.1520	+	32	16	2
60	I16.1527	+	16	8	2
61	I16.1546	+	32	16	2
62	I16.1547	+	32	16	2
63	I16.1549	+	32	16	2
64	I16.1550	+	32	16	2
65	I16.1559	+	32	16	2
66	I16.1564	+	32	16	2
67	I16.1569	+	16	8	2
68	I16.1570	+	16	8	2
69	I16.1571	+	32	16	2
70	I16.1576	+	32	16	2
71	I16.1581	+	16	8	2
72	I16.1582	+	32	16	2
73	I16.1585	+	16	8	2
74	I16.1591	+	16	8	2
75	I16.1592	-	8	4	2
76	I16.1603	+	32	16	2
77	I16.1606	+	16	8	2
78	I16.1610	+	32	16	2
79	I16.1611	+	16	8	2
80	I16.1613	+	32	16	2
81	I16.1614	+	32	16	2
82	I16.1616	-	32	16	2
83	I16.1618	+	0,5	0,25	2
84	I16.1619	+	256	128	2
85	I16.1620	-	256	128	2
86	I16.1621	+	32	16	2
87	I16.1622	+	0,5	0,25	2
88	I16.1623	+	0,5	0,25	2
89	I16.1629	-	32	16	2
90	I16.1630	-	32	16	2
91	I16.1632	+	32	16	2
92	I16.1633	-	64	32	2
93	I16.1634	+	32	16	2
94	I16.1638	+	16	8	2
95	I16.1639	+	32	16	2
96	I16.1641	+	16	8	2
97	I16.1645	+	32	16	2
98	I16.1648	+	0,5	0,25	2

STT	Chủng SA	MRSA	MIC ciprofloxacin ($\mu\text{g/mL}$) (X)	MIC ciprofloxacin ($\mu\text{g/mL}$) + Pa β N (Y)	Tỷ số X/Y
99	I16.1650	+	32	16	2
100	I16.1651	+	0,5	0,25	2
101	I16.1655	+	32	16	2
102	I16.1656	+	0,5	0,25	2
103	I16.1657	+	32	16	2
104	I16.1658	+	32	16	2
105	I16.1659	+	32	16	2
106	I16.1660	+	256	128	2
107	I16.1661	+	32	16	2
108	I16.1666	+	256	128	2
109	I16.1667	-	0,5	0,25	2
110	I16.1671	+	16	8	2
111	I16.1673	+	32	16	2
112	I16.1675	+	32	16	2
113	I16.1680	+	0,5	0,25	2
114	I16.1681	+	0,5	0,25	2
115	I16.1685	+	32	16	2
116	i16.1688	+	16	8	2
117	I16.1704	+	32	16	2
118	I16.1708	+	0,5	0,25	2
119	I16.1713	+	64	32	2
120	I16.1716	+	0,5	0,25	2
121	I16.1721	+	32	16	2
122	I16.1727	+	32	16	2
123	I16.1728	+	32	16	2
124	I16.1756	+	0,5	0,25	2
125	I16.1775	+	0,5	0,25	2
126	I16.1776	+	0,5	0,25	2
127	I16.1782	+	0,5	0,25	2
128	I16.1783	+	32	16	2
129	I16.1784	+	16	8	2
130	I16.1791	+	256	128	2
131	I16.1796	+	32	16	2
132	I16.1809	+	32	16	2
133	I16.1404	-	0,5	0,5	1
134	I16.1414	-	32	32	1
135	I16.1422	+	32	32	1
136	I16.1433	+	32	32	1
137	I16.1449	+	32	32	1
138	I16.1467	+	0,25	0,25	1

STT	Chủng SA	MRSA	MIC ciprofloxacin ($\mu\text{g/mL}$) (X)	MIC ciprofloxacin ($\mu\text{g/mL}$) + Pa β N (Y)	Tỷ số X/Y
139	I16.1489	+	32	32	1
140	I16.1490	+	32	32	1
141	I16.1494	+	16	16	1
142	I16.1501	+	16	16	1
143	I16.1522	+	16	16	1
144	I16.1525	+	16	16	1
145	I16.1530	+	32	32	1
146	I16.1534	+	32	32	1
147	I16.1536	-	0,06	0,06	1
148	I16.1572	+	16	16	1
149	I16.1578	+	16	16	1
150	I16.1580	+	16	16	1
151	I16.1605	+	8	8	1
152	I16.1679	-	0,06	0,06	1
153	I16.1772	+	0,06	0,06	1
154	I16.1797	+	32	32	1
155	I16.1798	+	32	32	1
156	I16.1808	+	0,25	0,25	1

3.4.2.3. Thử nghiệm tác dụng ức chế bơm ngược trên các chủng vi khuẩn *S. aureus* lâm sàng đề kháng ciprofloxacin bằng bơm ngược

Kết quả thử nghiệm (**Bảng 3.19**) cho thấy:

- Kháng sinh ciprofloxacin có giá trị MIC được xác định thấp nhất là 0,5 $\mu\text{g/mL}$ và cao nhất là 32 $\mu\text{g/mL}$ khi được sử dụng riêng lẻ trên các chủng SA lâm sàng.
- Các chalcon thử nghiệm ở nồng độ 20 $\mu\text{g/mL}$ giúp làm giảm MIC của ciprofloxacin trong một số trường hợp. Trong đó, chalcon F88 cho thấy khả năng làm giảm hiện tượng đề kháng với ciprofloxacin trên nhiều chủng vi khuẩn lâm sàng nhất (5/10 chủng); tiếp đến là F90 (3/10 chủng) và F29, F91 (2/10 chủng).

Bảng 3.19. Giá trị MIC ($\mu\text{g/mL}$) của ciprofloxacin (Ci) trên các chủng *S. aureus* lâm sàng khi vắng mặt và khi có mặt các chalcon nghiên cứu.

Chủng SA	MIC Ci ($\mu\text{g/mL}$)	MIC Ci ($\mu\text{g/mL}$) + F29	MIC Ci ($\mu\text{g/mL}$) + F88	MIC Ci ($\mu\text{g/mL}$) + F90	MIC Ci ($\mu\text{g/mL}$) + F91
I16.1412	16	16	16	16	16
I16.1421	0,5	0,125	0,125	0,125	0,5
I16.1461	32	32	32	32	32
I16.1505	32	32	16	32	32
I16.1562	32	32	16	32	32
I16.1604	16	16	16	16	16
I16.1617	16	16	16	16	16
I16.1635	32	32	16	16	16
I16.1672	32	16	16	16	16
I16.1773	16	16	16	16	16

CHƯƠNG 4. BÀN LUẬN

4.1. Các mô hình máy tính dựa trên phối tử

4.1.1. Các mô hình phân loại chất ức chế và chất không ức chế P-gp

Cho đến nay, nhiều chiến lược nghiên cứu khác nhau bao gồm các phương pháp dựa trên phối tử (SAR, QSAR, pharmacophore) và các phương pháp dựa trên cấu trúc (pharmacophore, mô hình hóa tương đồng, docking) được sử dụng rộng rãi để dự đoán sự ức chế P-gp [137]. Do thiếu các cấu trúc tinh thể ba chiều (3D) ở độ phân giải cao của protein này nên chỉ một số lượng nhỏ các nghiên cứu dựa vào cấu trúc được thực hiện trong quá khứ [22], [45], [83], [170]. Ngược lại, nhiều công trình mô hình hóa dựa vào phối tử lại được công bố cho mục đích xác định các chất ức chế P-gp [23], [34], [50], [83], [142], [143], [144], [146], [164], [174], [180], [193], [195] sử dụng cả các phương pháp học máy có giám sát như SVM, mạng nơron, cây quyết định, rừng ngẫu nhiên, Kappa/biến-láng giềng gần nhất, Naive Bayesian, QSAR nhị phân, ... và không giám sát như sơ đồ tự tổ chức. Các mô hình máy tính thu được trước đó từ nhiều loại thông số mô tả và dấu vân tay với các phương pháp học máy khác nhau được mô tả tóm tắt trong **Bảng 4.1**.

Bảng 4.1. Tóm tắt các mô hình phân loại chất ức chế và chất không ức chế P-gp được công bố trong các nghiên cứu trước và trong nghiên cứu này.

Nghiên cứu	Năm	Phương pháp	Thông số mô tả	Cơ sở dữ liệu		Kết quả
				Tập huấn luyện	Tập đánh giá	
Sun [174]	2005	Naive Bayesian	Thông số mô tả loại nguyên tử và dấu vân tay	424 chất	185 chất	Tập đánh giá: Độ đúng = 82,2 %
Yang và cộng sự [195]	2005	Sơ đồ tự tổ chức Kohonen	Thông số mô tả phân tử Molconn-Z	206 chất		Độ đúng trung bình = 82,3 %
		Mạng nơron lan truyền phía sau		174 chất (tập huấn luyện 50 %, tập đánh giá nội 25 % và tập đánh giá ngoại 25 %)	Tập đánh giá: Độ đúng (chất nền) = 46,6 %; Độ đúng (chất ức chế) = 29,0 %	

Nghiên cứu	Năm	Phương pháp	Thông số mô tả	Cơ sở dữ liệu		Kết quả
				Tập huấn luyện	Tập đánh giá	
Crivori và cộng sự [34]	2006	Phân tích phân biệt bình phương tối thiểu từng phần	Thông số mô tả dựa vào GRIND-pharmacophore	23 chất	125 chất	Tập đánh giá: Độ đúng = 82 %
Cruciani và cộng sự [18]	2011	Lĩnh vực tương tác phân tử	Volsurf+ và FLAP	772 chất	Nội: 85 chất Ngoại: 418 chất	Tập đánh giá ngoại: Độ đúng = 86 %
Hou và cộng sự [23]	2011	Cây quyết định; phân chia đệ quy; Naive Bayesian	Dấu vân tay và tính chất phân tử	973 chất	300 chất	Bộ phân loại Bayesian: Tập huấn luyện: Độ đúng = 81,7 %; Tập đánh giá: Độ đúng = 81,2 %
Bianucci và cộng sự [146]	2012	Cây ngẫu nhiên; C4.5; consensus	Thông số mô tả phân tử và dấu vân tay (PaDEL)	39 chất	20 chất	Đánh giá chéo bỏ một: Độ đúng 1 = 84,6 %; Độ đúng 2 = 76,9 %; Độ đúng 3 = 76,9 % Tập đánh giá nội: Độ đúng 1 = 80 %; Độ đúng 2 = 75 %; Độ đúng 3 = 75 % Tập ngoại: Độ đúng 1 = 77,8 %; Độ đúng 2 = 72,3 %; Độ đúng 3 = 76,6 %
Ecker và cộng sự [142]	2012	Rừng ngẫu nhiên; Kappa láng giềng gần nhất; máy vector hỗ trợ	Dấu vân tay Checkmol	1935 chất		Toàn tập dữ liệu: Độ đúng (Rừng ngẫu nhiên) = 81 % Tập đánh giá: Độ đúng (Rừng ngẫu nhiên) = 75 %
Ecker và cộng sự [83]	2014	Kappa láng giềng gần nhất; máy vector hỗ trợ; rừng ngẫu nhiên; cây quyết định; QSAR nhị phân	Thông số mô tả MOE, dấu vân tay MACCS và dưới cấu trúc	1201 chất	Nội: 407 chất Ngoại: 346 chất	Tập đánh giá ngoại: Độ đúng (máy vector hỗ trợ) = 75 %; Độ đúng (rừng ngẫu nhiên) = 73 %
Éric và cộng sự [50]	2014	Mạng nơron nhân tạo; máy vector hỗ trợ; kết hợp	Thông số mô tả ADMET Modeler	101 chất	34 chất	Mô hình tốt nhất: Tập huấn luyện: Độ đúng trung bình = 96 %; Tập đánh giá: Độ đúng trung bình = 89 %

Nghiên cứu	Năm	Phương pháp	Thông số mô tả	Cơ sở dữ liệu		Kết quả
				Tập huấn luyện	Tập đánh giá	
Thái Khắc Minh và cộng sự [180]	2015	Mạng nơron lan truyền ngược	Thông số mô tả MOE và PaDEL	106 chất	27 chất	Tập đánh giá: Độ đúng tổng thể (SIN) = 82 %; Độ đúng chất ức chế (SIN) = 100 %
Nantasenamat và cộng sự [143]	2015	Cây quyết định; mạng nơron nhân tạo; máy vector hỗ trợ	Thông số mô tả hóa lượng tử và thông số mô tả phân tử Dragon	1301 chất	215 chất	Tập huấn luyện: Độ đúng 1 = 92 %; Độ đúng 2 = 91 %; Độ đúng 3 = 89 % Đánh giá chéo 10 lần: Độ đúng 1 = 87 %; Độ đúng 2 = 88 %; Độ đúng 3 = 88 % Đánh giá ngoại: Độ đúng 1 = 87 %; Độ đúng 2 = 86 %; Độ đúng 3 = 84 %
Schyman và cộng sự [164]	2016	Biến-láng giềng gần nhất (v-NN)	Dấu vân tay ECFP4	2076 chất	200 chất	Mô hình độ đúng cao: Độ đúng (tập huấn luyện) = 87 %; Độ đúng (tập đánh giá) = 81 % Mô hình độ phủ cao: Độ đúng (tập huấn luyện) = 83 %; Độ đúng (tập đánh giá) = 77 %
Worachartcheewan và cộng sự [144]	2017	CORrelation And Logic (CORAL)	Dựa trên SMILES	2254 chất		Các giá trị độ đúng > 80 %
Tseng và cộng sự [193]	2019	C&R Tree (CART)	Dấu vân tay PubChem 2D 881 bit (PaDEL) và 196 thông số mô tả (RDKit Release)	2056 chất		Mô hình tốt nhất: Độ đúng = 83 %; Độ nhạy = 93 %; Độ đặc hiệu = 69 % F-score = 0,87
Nghiên cứu này	2016	Mạng nơron; C5.0; C&R Tree; QUEST; CHAID; hồi quy logistic; mặt nghiêng quyết định; mạng Bayesian; phân tích phân biệt; máy vector hỗ trợ; kết hợp	Thông số mô tả MOE và PaDEL; dấu vân tay MACCS, Pubchem và dưới cấu trúc	1690 chất	Nội: 419 chất Ngoại: 22 chất	Mô hình kết hợp (Ensemble): Độ đúng (tập huấn luyện đa dạng) = 84 %; Độ đúng (tập đánh giá nội) = 92 %; Độ đúng (tập đánh giá ngoại) = 100 %

Trong một nghiên cứu trước đó [180], các mô hình ba lớp và bốn lớp giúp phân loại các chất ức chế và chất nền của P-gp được tạo ra sử dụng mạng nơron lan truyền ngược; và trong số đó, mô hình SIN là mô hình tốt nhất khi dự đoán đúng 82 % số chất trong tập kiểm tra và 100 % chất ức chế mà không phải là chất nền của protein chuyên chở này. Mặc dù có các giá trị thống kê khá tốt, nhưng công trình nghiên cứu này [180] cũng như các công trình khác của Sun [174], Yang và cộng sự [195], Crivori và cộng sự [34], Bianucci và cộng sự [146], Éric và cộng sự [50] lại chỉ được phát triển từ các tập dữ liệu tương đối nhỏ và vì vậy khó xác định đúng một chất bất kỳ có khả năng ức chế P-gp hay không bởi vì tính hỗn tạp của protein. Được xây dựng từ các tập dữ liệu lớn với hơn 1000 - 2000 phân tử đại diện cho một không gian hóa học lớn hơn, nghiên cứu này cùng với các công trình của Cruciani và cộng sự [18], Hou và cộng sự [23], Ecker và cộng sự [83], [142], Nantasenamat và cộng sự [143], Schyman và cộng sự [164], Worachartcheewan và cộng sự [144], Tseng và cộng sự [193] có thể phù hợp hơn trong việc giải quyết các vấn đề liên quan đến tính đặc hiệu chất nền rộng của các loại bơm ngược, bao gồm P-gp.

Các nghiên cứu của Cruciani và cộng sự [18] và Hou và cộng sự [23] đã công bố hai bộ cơ sở dữ liệu gốc bao gồm lần lượt 1275 chất và 1273 chất; và đều được sử dụng trong các nghiên cứu sau đó của Ecker và cộng sự [83], [142], Schyman và cộng sự [164], Tseng và cộng sự [193], cũng như trong nghiên cứu này. Trong khi đó, hai nghiên cứu của Nantasenamat và cộng sự [143] và Worachartcheewan và cộng sự [144] lại sử dụng tập dữ liệu các chất tương tác với P-gp từ cơ sở dữ liệu admetSAR do Tang và cộng sự [24] tạo ra. Mặc dù đã được báo cáo trước đó vào năm 2016, các giá trị độ đúng dự đoán của mô hình kết hợp thu được từ nghiên cứu này là có thể so sánh với các mô hình tốt nhất hiện tại (**Bảng 4.1**).

Trong các công trình được đề cập, các mô hình máy tính được tạo ra sử dụng các kỹ thuật tính toán thông số mô tả, phân chia dữ liệu và mô hình hóa khác nhau, nhưng hầu hết đều là các mô hình học máy dựa trên các thuật toán đơn lẻ với những hạn chế nhất định như đã được chứng minh ở trên. Hiện có rất ít mô hình kết hợp được xây dựng, chẳng hạn như các mô hình dựa trên sự kết hợp giữa mạng nơron

nhân tạo và SVM mà Éric và cộng sự đã công bố gần đây cho mục đích dự đoán sự ức chế P-gp, với trọng số bằng nhau cho các dự đoán của mỗi mô hình đơn lẻ [50]. So với phương pháp bầu chọn đơn giản (simple voting) áp dụng nguyên tắc số đông được nhóm nghiên cứu của Éric sử dụng [50], phương pháp bầu chọn dựa trên trọng số độ tin cậy (confidence-weighted voting) được sử dụng trong nghiên cứu này là hợp lý hơn và được ưa chuộng bởi vì dự đoán với độ tin cậy cao hơn sẽ được lựa chọn trong mỗi trường hợp. Các kết quả khả quan thu được từ cả hai nhóm nghiên cứu góp phần củng cố quan điểm đã được đặt ra trước đó về việc kết hợp nhiều mô hình nhằm tăng cường khả năng phân loại [180].

Phần lớn các chất ức chế P-gp có những tính chất lý hóa và cấu trúc chung, bao gồm tính thân dầu, kích thước hay trọng lượng phân tử, tính cồng kềnh, tính thơm và nhóm nhận liên kết hydro [144]. Trong khi tính thân dầu là yếu tố giúp cho khả năng thấm qua màng, sự hiện diện của nhóm nhận liên kết hydro, nguyên tử N (bậc ba), vòng thơm và các vùng kỵ nước lần lượt cần thiết cho các tương tác gắn kết hydro, tĩnh điện, π -stacking và kỵ nước [102]. Từ tập hợp 24 thuộc tính được chọn làm biến độc lập trong các mô hình phân loại (**Phụ lục 2**), nghiên cứu này cũng đã khẳng định vai trò của tính cồng kềnh về mặt cấu trúc (số vòng 8), khả năng tích điện (các thông số mô tả điện tích), nguyên tử N (các dấu vân tay Pubchem), các nhóm nhận và cho hydro (aldehyd, carboxylic, diol, ...) trong việc xác định các chất ức chế P-gp.

4.1.2. Các mô hình dự đoán hoạt tính ức chế P-gp

Trong nhiều thập kỉ, các phương pháp dựa vào phối tử (SAR, 2D-QSAR, 3D-QSAR và mô hình hóa pharmacophore) được sử dụng chủ yếu cho việc khám phá và thiết kế các chất ức chế P-gp mới và hiệu quả, bởi vì cấu trúc của protein này vẫn chưa được phân giải [137]. Các cơ chế phân tử của sự điều hòa MDR nói chung và sự ức chế P-gp nói riêng cho đến nay vẫn chưa được hiểu rõ, gây khó khăn cho các nghiên cứu QSAR về chất ức chế P-gp [9]. Bên cạnh các mô hình phân loại nhị phân dự đoán sự ức chế P-gp (1 là chất ức chế; 0 là chất không ức chế), nhiều nghiên cứu 2D-QSAR khác [40], [57], [58], [59], [72], [87], [101], [128], [139], [165], [167], [171], [191], [194], [196], [205], [207] với hoạt tính ức chế P-gp là biến liên tục (pIC_{50} ,

pK_d) cũng được báo cáo trong thời gian qua và được tóm tắt trong **Bảng 4.2**. Sử dụng các phương pháp khác nhau từ đơn giản (hồi quy đa tuyến tính) đến phức tạp (các dạng cây, mạng nơron, máy vector hỗ trợ, phương pháp lai, ...), các nghiên cứu này đã thu được các kết quả thống kê tốt và góp phần xác định các yếu tố cấu trúc cần thiết cho sự đảo ngược MDR qua trung gian P-gp. Tuy nhiên, khả năng ứng dụng của các mô hình được công bố trên các tập dữ liệu khác nhau vẫn còn phải bàn cãi bởi vì chúng chỉ được tạo ra từ một số lượng phân tử nghiên cứu hạn chế và các phân tử này lại thuộc cùng một nhóm cấu trúc trong nhiều trường hợp (phenothiazin, methylxanthin, flavonoid, diterpen, imidazol, polyphenol, ...). Trong số đó, chỉ có các công trình của Dearden và cộng sự [40], Ghafourian và cộng sự [165] là được thực hiện với trên 200 phân tử.

Bảng 4.2. Tóm tắt các mô hình QSAR hai chiều dự đoán hoạt tính ức chế P-gp (biến liên tục) được công bố trong các nghiên cứu trước và trong nghiên cứu này.

Nghiên cứu	Năm	Cơ sở dữ liệu và phương pháp	Cơ chế tác động mục tiêu	Kết quả và thông số mô tả phân tử/tính chất liên quan đến sự điều hòa P-gp
Litman và cộng sự [101]	1997	- 34 thuốc khác nhau - Hồi quy tuyến tính	Ức chế hoạt tính ATPase	Diện tích bề mặt surface area (-)
Österberg và Norinder [128]	2000	- 22 chất đa dạng giống thuốc - Bình phương tối thiểu từng phần đa biến (loại tamoxifen)	Ức chế hoạt tính ATPase	- R ² = 0,718; Q ² = 0,695; SD = 0,475; F = 48,37; RMSE ^{tr} = 0,452; p < 0,001 - Kích thước bề mặt phân tử, khả năng phân cực và liên kết hydro
Dearden và cộng sự [40]	2003	- 22 thuốc đa dạng - 22 thuốc đa dạng khác - 157 phenothiazin và các thuốc liên quan - Hồi quy bậc thang	Ức chế hoạt tính ATPase; bơm ngược thuốc	- Kích thước phân tử và tính phân cực: Quan trọng - Liên kết hydro hoặc tính kỵ nước: Không quan trọng
Lien và cộng sự [194]	2003	- Các chất nền/chất điều hòa, các tác nhân đảo ngược MDR triển vọng trên lâm sàng - Hồi quy bậc thang		Tính thân dầu (ClogP), trọng lượng phân tử (logMw), chuỗi dài nhất (Nlc) của phân tử, năng lượng orbital được chiếm giữ cao nhất (Ehomo) và tối thiểu một nguyên tử nitơ bậc ba base
Breier và cộng sự [87]	2004	- 25 methylxanthin - Hồi quy đa tuyến tính	Làm giảm sự đề kháng đa thuốc của tế bào	Khúc xạ phân tử, mật độ tinh thể (-) và hệ số phân chia (-)

Nghiên cứu	Năm	Cơ sở dữ liệu và phương pháp	Cơ chế tác động mục tiêu	Kết quả và thông số mô tả phân tử/tính chất liên quan đến sự điều hòa P-gp
Yang và cộng sự [196]	2005	- 57 flavonoid - Mạng nơron theo quy tắc Bayesian (BRNN); mạng nơron lan truyền phía sau và bình phương tối thiểu từng phần	Gắn vào domain gắn kết nucleotid C tận	- Mô hình tốt nhất (BRNN): $R^2_{\text{tập huấn luyện}} = 0,75$; $R^2_{\text{tập đánh giá}} = 0,73$ - 53 thông số mô tả Molconn-Z, 4 vector thành phần chính
Zhao và cộng sự [205]	2009	- 70 chất - Hồi quy đa tuyến tính đơn giản; máy vector hỗ trợ và phương pháp lai	Ức chế bơm ngược	- Mô hình lai tốt nhất: $R^2_{\text{tập huấn luyện}} = 0,85$; $R^2_{\text{tập đánh giá}} = 0,81$; $R^2_{\text{toàn tập}} = 0,84$ - 4 thông số mô tả Codessa được chọn
Fernandes và cộng sự [171]	2013	- 51 chất diterpen có hoạt tính sinh học - Hồi quy đa tuyến tính	Ức chế bơm ngược	- Bốn mô hình (mô hình tốt nhất: $R^2 = 0,806$; $R^2_{\text{dự đoán}} = 0,765$; $q^2 = 0,758$) - Tính thân dầu của các phân tử, số lượng nguyên tử tích điện dương và số lượng hệ thống liên hợp (-)
Ecker và cộng sự [72]	2013	- 35 benzopyran và benzopyrano [3,4b][1,4]oxazin - Hồi quy đa tuyến tính	Ức chế bơm ngược	- $R^2 = 0,67$; $q^2(\text{LOO}) = 0,63$; RMSE = 0,48 - Diện tích bề mặt vdW (\AA^2) của các nguyên tử kỵ nước
Ecker và cộng sự [139]	2014	- 22 dẫn xuất chalcon tổng hợp - Hồi quy đa tuyến tính	Ức chế bơm ngược	- $R^2 = 0,79$; $q^2(\text{LOO}) = 0,71$; RMSE = 0,51 - Số lượng liên kết có thể xoay (-) và diện tích bề mặt kỵ nước
Luco và cộng sự [191]	2014	- 62 flavonoid - Hồi quy đa biến	Ái lực gắn kết: Hằng số phân ly (K_d)	Tính kỵ nước, các yếu tố hình học và các tương tác phân tán liên phân tử ở vị trí gắn kết (chalcon)
Li và cộng sự [167]	2014	- 88 flavonoid - Mạng nơron nhân tạo lan truyền phía sau và hồi quy đa tuyến tính	Ái lực gắn kết: Hằng số phân ly (K_d)	- Bốn mô hình (mô hình tốt nhất: $R^2 \approx 0,92$; $q^2 \approx 0,77$; $R^2_{\text{ngoại}} \approx 0,87$) - 14 và 8 thông số mô tả PaDEL được chọn
Shayanfar và cộng sự [59]	2014	- 51 dẫn xuất dựa trên imidazol - Hồi quy đa tuyến tính; máy vector hỗ trợ và mạng nơron nhân tạo	Ức chế bơm ngược	- Mô hình tốt nhất (mạng nơron nhân tạo): $R^2_{\text{tập huấn luyện}} = 0,83$; $R^2_{\text{tập đánh giá}} = 0,81$ - 4 thông số mô tả Dragon được chọn bởi phương pháp bình phương tối thiểu-thuật toán di truyền (GA-PLS) và hồi quy bậc thang
Ghafourian và cộng sự [165]	2016	- 219 cặp chất ức chế/chất nền - Cây hồi quy; cây tương tác sử dụng CART; CHAID; cây tăng cường; rừng ngẫu nhiên và hồi quy đa biến tự thích nghi	Hằng số ức chế K_i	- Hai mô hình tốt nhất là CHAID và cây tăng cường với MAE trên tập huấn luyện lần lượt là 0,399 và 0,322; MAE trên tập đánh giá lần lượt là 0,511 và 0,554 - Các chất ức chế P-gp mạnh có tính thân dầu cao hơn và trọng lượng phân tử lớn hơn các phân tử giống như thuốc

Nghiên cứu	Năm	Cơ sở dữ liệu và phương pháp	Cơ chế tác động mục tiêu	Kết quả và thông số mô tả phân tử/tính chất liên quan đến sự điều hòa P-gp
Ghaemian và Shayanfar [57]	2017	- 52 dẫn xuất polyphenol methyl hóa (epigallocatechin; gallocatechin; dihydromyricetin và epigallocatechin gallat) - Hồi quy đa tuyến tính; mạng nơron nhân tạo và máy vector hỗ trợ	Số lần đảo ngược RF (paclitaxel)	- Giá trị R^2 cho các kiểu phân chia dữ liệu khác nhau: Hồi quy đa tuyến tính: 0,86; 0,87; 0,81 Mạng nơron nhân tạo: 0,88; 0,90; 0,86 Máy vector hỗ trợ: 0,95; 0,93; 0,81 - Các thông số mô tả Dragon (các thông số chung và quan trọng là 3D-MoRSE)
Ghaemian và Shayanfar [58]	2019	- 36 dẫn xuất polyphenol (epigallocatechin và gallocatechin) - Hồi quy thành phần chính; mạng nơron nhân tạo và máy vector hỗ trợ	Số lần đảo ngược RF (paclitaxel)	- Mô hình tốt nhất (mạng nơron nhân tạo): $R^2_{\text{tập huấn luyện}} = 0,99$; $R^2_{\text{tập đánh giá}} = 0,80$; CCC = 0,83 - Các thông số mô tả phân tích hình ảnh thay cho các thông số mô tả cấu trúc
Xu và cộng sự [207]	2019	- 31 flavonoid - Bình phương tối thiểu từng phần	Số lần đảo ngược RF (daunorubicin)	- $R^2 = 0,892$; $Q^2 = 0,829$; $R^2_{\text{dự đoán}} = 0,905$ - Thông số mô tả Sybyl và MOE (hoạt tính ức chế P-gp cao: vsurf_DW23 và vsurf_G nhỏ; E_sol và dipole lớn) - Sự vắng mặt của liên kết đôi 2,3; 3'-OH và 4'-OH và tăng số nhóm thế methoxy là có lợi cho sự ức chế P-gp
Nghiên cứu này	2016	- 499 chất đa dạng - Mạng nơron; C&R Tree; CHAID; máy vector hỗ trợ; hồi quy; tuyến tính suy rộng và kết hợp	Ức chế bơm ngược	- Mô hình kết hợp (Ensemble): Tập huấn luyện đa dạng: $R^2 = 0,84$; $Q^2_{\text{Loo}} = 0,70$ Tập đánh giá nội: $Q^2_{\text{Fi}} \geq 0,81$; CCC = 0,90 Tập đánh giá ngoại: $Q^2_{\text{Fi}} \geq 0,82$; CCC = 0,90 - 7 thông số mô tả MOE và 27 thông số mô tả PaDEL

*(-): Đóng góp ngược chiều.

Trong nghiên cứu này, thông tin cấu trúc và hoạt tính được tập hợp từ nhiều nguồn khác nhau với cùng loại thử nghiệm, dòng tế bào và tác nhân gây độc tế bào để vừa đảm bảo các thông tin này có thể so sánh với nhau một cách tương đối, vừa đảm bảo tập dữ liệu đủ lớn và đa dạng để sử dụng cho việc xây dựng mô hình. Hơn nữa, các điều kiện đánh giá chặt chẽ cũng giúp khẳng định chất lượng của các mô hình QSAR được tạo ra. Trong đó, mô hình kết hợp được kỳ vọng là một công cụ *in silico* mạnh để sử dụng cho các quá trình sàng lọc đầu vào cao, giúp dự đoán nhanh

và chính xác khả năng ức chế P-gp của các ứng viên thuốc. Mặt khác, nghiên cứu này cũng nhấn mạnh các tính chất lý hóa cần thiết cho sự ức chế P-gp là diện tích bề mặt (surface area), khả năng phân cực (polarizability), liên kết hydro (hydrogen bonding), tính thân dầu (lipophilicity), trọng lượng phân tử (molecular weight), chuỗi béo dài nhất (longest aliphatic chain) và khả năng tích điện dương (positive charge) đã được đề cập trong các nghiên cứu trước đó (**Bảng 4.2**), thông qua các thông số mô tả được chọn là PEOE_VSA_FHYD, PEOE_VSA_FPNEG, PEOE_VSA_FPPOS; AATSC6p; minHBint6, MLFER_BH; LipoaffinityIndex; SpMax2_Bhm, SpMin6_Bhm, SpMin7_Bhm; nAtomLAC và PEOE_RPC+ (**Phụ lục 4**). Những thông tin này có thể giúp định hướng cho sự tổng hợp, với các biến đổi cấu trúc phù hợp giúp tăng cường hoạt tính.

Một trong những vấn đề thách thức đối với các mô hình QSAR dự đoán là khả năng ngoại suy [96]. Để khắc phục khó khăn trên, việc sử dụng chiến lược như đã trình bày trong nghiên cứu này là đáng được cân nhắc, cụ thể là phát triển các mô hình hồi quy từ một tập dữ liệu lớn và có cấu trúc đa dạng (số lượng chất được sử dụng cho nghiên cứu là lớn nhất so với các công trình đã công bố trước đó, với gần 500 phân tử khác nhau), đồng thời kết hợp thực hiện xác định phạm vi khả năng ứng dụng, kiểm tra các điều kiện dựa vào MAE, phát hiện chất bất thường và xem xét đồ thị thành phần chính khi xây dựng và đánh giá mô hình, cũng như khi tiến hành dự đoán cho các chất từ các tập dữ liệu có không gian hóa học khác. Ngoài ra, các giá trị thống kê thu được từ mô hình kết hợp của chúng tôi là có thể so sánh với các mô hình tốt nhất trong **Bảng 4.2**, bao gồm cả các nghiên cứu được báo cáo từ năm 2016 trở về sau.

Trong một nghiên cứu gần nhất vào năm 2019 trên 31 hợp chất flavonoid (chứa các vòng A và B tương tự chalcon và thêm vòng C), Xu và cộng sự [207] đã thiết lập một phương trình hồi quy đa tuyến tính và tiết lộ hai thông số mô tả vsurf_DW23 và vsurf_G có mối tương quan tỷ lệ nghịch với hoạt tính ức chế P-gp, trong khi hai thông số mô tả khác là E_sol và dipole cho thấy mối tương quan ngược lại. Trong đó, vsurf_DW23 là thông số vùng kỵ nước, đại diện cho khoảng cách tiếp xúc của

vsurf_EWmin (năng lượng thân nước thấp nhất) khi một đầu dò nước tương tác với một phân tử mục tiêu; vsurf_G là thông số về hình dạng, mô tả dạng cầu bề mặt của phân tử; E_sol là thông số mô tả năng lượng tiềm năng, đại diện cho năng lượng solvat hóa; dipole là thông số mô tả điện tích phụ thuộc hình thể, đại diện cho momen lưỡng cực được tính từ điện tích từng phần của phân tử, thể hiện sự phân phối điện tích và mức độ phân chia điện tích âm, dương [207]. Hai thông số vsurf_DW23 và dipole lần lượt tương ứng với các thông số PEOE_VSA_FHYD; PEOE_VSA_FPNEG và PEOE_VSA_FPPOS trong nghiên cứu này. Bên cạnh đó, nhóm tác giả này cũng phát hiện là việc loại bỏ các nhóm thế hydroxyl ở hai vị trí 3', 4' thuộc vòng B và thêm các nhóm methoxy giúp làm tăng hoạt tính ức chế P-gp của các flavonoid (giống chalcon, ví dụ F88 và F90); trong khi sự hiện diện của liên kết đôi ở vị trí 2,3 lại không cần thiết (khác chalcon). Những ghi nhận trên góp phần chứng minh tính khách quan của các kết quả thu được từ đề tài nghiên cứu này, vốn đã được thực hiện và báo cáo trước đó.

4.1.3. Bản đồ nhận thức về sự ức chế bơm ngược qua trung gian P-gp và NorA

Từ hai bản đồ nhận thức (**Hình 3.2** và **Hình 3.3**), giá trị trung bình của mỗi thông số mô tả (**Phụ lục 7**) và tần số của mỗi dấu vân tay (**Phụ lục 8**) được tính cho mỗi lớp, bảy quy tắc được rút ra nhằm phân biệt giữa các lớp hoạt tính trên hai bơm ngược P-gp và NorA; và định hướng cho quá trình thiết kế thuốc với sự trợ giúp của máy tính (CADD) như sau:

Quy tắc 1 để phân biệt giữa (P, D) và (A, N)

Các chất P và D trội hơn về các thuộc tính nHCsat_u, minHCsat_u, ETA_BetaP_ns_d, MDEO-22, MACCSFP144 và Pubchem 2, trong khi các chất A và N trội hơn về các thuộc tính BCUT_PEOE_2 và GCUT_PEOE_2.

Quy tắc 2 để phân biệt giữa (P, A) và (D, N)

Các chất P và A trội hơn về thuộc tính ATSC2m, trong khi các chất D và N trội hơn về thuộc tính AATSC6v.

Quy tắc 3 để phân biệt giữa P, D và (A, N)

Đường kính của các chất D, P và A/N giảm theo thứ tự D (kích thước lớn) > A/N (kích thước trung bình) > P (kích thước nhỏ). Kết quả này phù hợp với phát hiện của Carosati và cộng sự từ số liệu thống kê đa biến của họ, sử dụng phân tích thành phần chính (PCA) [17].

Quy tắc 4 để phân biệt giữa P và D

Các chất P trội hơn về thuộc tính ATSC4m, trong khi các chất D trội hơn về các thuộc tính SpMAD_DzZ, ASP-3 và AVP-6.

Quy tắc 5 để phân biệt P với (A, D, N)

Các chất P trội hơn về các thuộc tính balabanJ, AATSC4s và MATS4s so với các chất thuộc các lớp khác. Trong một nghiên cứu gần đây, các mô hình hồi quy được báo cáo đã chứng minh balabanJ là thông số mô tả có ảnh hưởng đến hoạt tính ức chế NorA của các chất (pIC₅₀) [181].

Quy tắc 6 để phân biệt A với (P, D, N)

Các chất A trội hơn về thuộc tính Q_VSA_FNEG so với các chất thuộc các lớp khác. Thông số này phản ánh tính chất điện tích từng phần, có thể tương tự với yếu tố “khả năng phân cực” được đề cập trong công trình của Carosati và cộng sự [17].

Quy tắc 7 để phân biệt D với (P, A, N)

Các chất D trội hơn về các thuộc tính ATSC1s và MACCSFP128 so với các chất thuộc các lớp khác.

4.1.4. Các mô hình pharmacophore cho hoạt tính ức chế P-gp mạnh và cho sự ức chế chọn lọc NorA

Cả hai giả thuyết pharmacophore của chất ức chế P-gp mạnh và chất ức chế chọn lọc NorA được xem xét giúp bổ sung cho các kết luận của Carosati và cộng sự [17] như sau:

- Tính kỵ nước là cần thiết cho các phân tử nhằm đến sự ức chế cả hai bơm ngược, đặc biệt là P-gp (kỵ nước hơn);
- Các vòng thơm là các yếu tố quan trọng cho sự ức chế NorA nhưng không ức chế P-gp;

- Một nhóm nhận liên kết hydro giúp ức chế cả hai protein trong khi một nhóm cho liên kết hydro giúp ức chế chỉ NorA.

Trong khi sự ức chế P-gp được kỳ vọng giúp đảo ngược MDR trong điều trị kháng ung thư [93], tác dụng dược lý này lại được xem là một hiệu quả không mong muốn trong các lĩnh vực trị liệu khác bởi vì các chất ức chế P-gp có thể làm biến đổi hồ sơ dược động học của các thuốc dùng chung và gây ra các tương tác thuốc [47], [55], [99]. Trong khi đó, nghiên cứu này giúp có được nhận thức sâu hơn về việc thiết kế các chất ức chế bơm ngược NorA của *S. aureus* từ một tập dữ liệu hạn chế bao gồm 54 chất, sử dụng hai kỹ thuật máy tính là lập bản đồ nhận thức và mô hình hóa pharmacophore. Do không ức chế P-gp, các chất ức chế chọn lọc NorA có thể tránh được các tương tác thuốc nguy hiểm qua trung gian protein không mục tiêu này khi được sử dụng phối hợp trong điều trị nhiễm trùng.

4.2. Mô hình tương đồng của P-gp

Trước đó, nhiều nỗ lực nghiên cứu đã được thực hiện nhằm mô hình hóa tương đồng bơm ngược P-gp của người từ các cấu trúc đã được phân giải của bơm ngược vi khuẩn như MsbA [197], Sav1866 [38], [39] và BtuCD [70], [103] để sử dụng cho chiến lược thiết kế thuốc dựa vào cấu trúc [137]. Gần đây, một mô hình tương đồng P-gp của người cũng được tạo ra dựa trên cấu trúc P-gp của giun tròn *Caenorhabditis elegans* (giống nhau 46 %) ở độ phân giải 3,4 Å [75]. Tuy nhiên, để thu được mô hình tương đồng có chất lượng cao đòi hỏi protein đĩa vừa phải có độ phân giải cao, vừa phải có sự giống nhau ở mức độ cao về trình tự với protein mục tiêu [137]. Điều này giải thích tại sao việc tinh thể hóa và xác định cấu trúc của P-gp ở động vật có vú nhận được nhiều sự kỳ vọng. Cấu trúc tia X của P-gp chuột được sử dụng trong nghiên cứu này là 3g61 có độ phân giải 4,35 Å; 87 % trình tự giống với P-gp ở người và gần 100 % acid amin giống nhau trong khoang gắn kết, ngoại trừ mSer725/hAla729 [4], [5]. Mặc dù có sự chuyển dịch của một acid amin đăng ký (Tyr303/Ile302), cấu trúc này dường như vẫn là protein đĩa phù hợp để phát triển các mô hình tương đồng của P-gp ở người cho mục đích docking trong nghiên cứu này, thay vì các cấu trúc tinh thể không gắn kết với phối tử tại khoang trung tâm, mới được điều chỉnh của P-gp

chuột với mã số định danh trong ngân hàng dữ liệu protein là 4ksb, 4ksc và 4ksd [198]; bởi vì sự hiện diện của phối tử tham khảo (trong trường hợp này là QZ59-RRR) là cần thiết để dự đoán vị trí gắn kết bởi I-TASSER và để xác định chính xác vùng không gian này cho mục đích sàng lọc ảo bằng docking trong LeadIT.

4.3. Sàng lọc *in silico*

4.3.1. Khả năng ức chế P-gp

Kết quả sàng lọc *in silico* của 95 chalcon nội bộ (**Bảng 3.14**) cho thấy sự đồng thuận của tất cả các phương pháp máy tính (9 mô hình phân loại chất ức chế P-gp, mô hình kết hợp dự đoán IC₅₀ trên P-gp, các pharmacophore chất ức chế P-gp mạnh và chất ức chế chọn lọc NorA và các điểm số docking đều < -20,00 kJ/mol) trong việc dự đoán F89, F90 và F91 là ba chalcon nội bộ hứa hẹn nhất cho hoạt tính ức chế P-gp và có thể cả NorA. Tuy chỉ nhận được sự đồng thuận của 8 mô hình phân loại (trừ QUEST) nhưng với IC₅₀ dự đoán bởi mô hình kết hợp là 2,67 μM, điểm số docking xấp xỉ -20,00 kJ/mol và chỉ thỏa một mình pharmacophore chất ức chế NorA nhưng không ức chế P-gp, chalcon F88 có thể là một ứng viên tiềm năng để ức chế NorA.

Tương tự, sự đồng thuận của các phương pháp máy tính về khả năng ức chế P-gp tốt, trừ pharmacophore (ít nhất 9 mô hình phân loại chất ức chế P-gp, mô hình kết hợp dự đoán IC₅₀ trên P-gp và các điểm số docking đều âm) cũng được ghi nhận ở 17 hợp chất Ngân hàng Thuốc là DB01118 (amiodaron); DB04869 (olcegepant); DB02872; DB00877 (sirolimus); DB00224 (indinavir); DB01721 (đồng đẳng của indinavir); DB00503 (ritonavir); DB04748 (oximinoarylsulfonamid); DB02253; DB08889 (carfilzomib); DB02785; DB02888 (FKB-001); DB02009 (L-756,423); DB01180 (rescinnamin); DB04378; DB04318 và DB01232 (saquinavir) (**Bảng 3.15**).

4.3.2. Khả năng ức chế NorA

Bốn chalcon nội bộ F29, F88, F90 và F91 đáp ứng được hai tiêu chí bao gồm có kết quả sàng lọc *in silico* trên P-gp tốt (**Bảng 3.14**) và mẫu thử có sẵn cho thử nghiệm, được sử dụng để dự đoán hoạt tính ức chế NorA thông qua giá trị IC₅₀ cũng như khả năng gắn kết với bơm ngược này thông qua nghiên cứu docking. Với cấu

trúc tổng thể nhỏ gọn, linh hoạt cùng các nhóm thể có khả năng tạo tương tác với mục tiêu tác động, cả bốn chalcon thử nghiệm đều được gắn kết tốt vào mô hình tương đồng của NorA tại những vị trí có thể là khoang trung tâm và Walker B. Tuy nhiên, chỉ ba dẫn xuất F88, F90 và F91 được mô hình D dự đoán là có khả năng ức chế NorA hiệu quả do chúng có tính tan trong nước kém hơn ($\log S < -6,0$) (**Bảng 3.16**).

Các quy tắc về thuộc tính quan trọng (thông số mô tả và dấu vân tay) được đề nghị trong **Mục 4.1.3** để phân loại một chất bất kỳ theo khả năng ức chế của nó trên hai bơm ngược P-gp và NorA khó có thể được áp dụng trực tiếp cho mục đích sàng lọc ảo đầu vào cao (virtual high-throughput screening - vHTS) trên các tập dữ liệu phân tử mà không có sự tự động hóa. Mặt khác, khi một chất bị thiếu các thông số mô tả và/hoặc dấu vân tay có liên quan trong tổng số 21 thuộc tính (**Phụ lục 6**), do không tính toán được (như trường hợp F90 và F91) hoặc nằm ngoài không gian hóa học của tập dữ liệu 54 chất dùng để xây dựng các bản đồ nhận thức (như trường hợp F88) cũng không thể áp dụng được các quy tắc này. Ngoài các khó khăn đã nêu thì chúng có thể được sử dụng để xác nhận lại kết quả (các chất “hit”) của các quá trình sàng lọc *in silico* khác. Cho ví dụ, chalcon F29 với 21 thuộc tính liên quan được tính toán và sau đó các thông số mô tả của nó được chia tỷ lệ trong khoảng [0, 1], đã cho thấy sự trội hơn ở các thông số ETA_BetaP_ns_d = 0,94 và MDEO-22 = 0,30 (Quy tắc 1); ASP-3 = 0,63 (Quy tắc 4); balabanJ = 0,63, AATSC4s = 0,61 và MATS4s = 0,71 (Quy tắc 5) khi so với các giá trị trung bình được tính trong **Phụ lục 7**; cùng với sự hiện diện của các dấu vân tay MACCSFP144 và Pubchem 2 (Quy tắc 1) và kích thước phân tử nhỏ (Quy tắc 3) nên dẫn xuất này thiên về sự ức chế P-gp hơn là NorA.

4.4. Thử nghiệm *in vitro*

Thử nghiệm *in vitro* sử dụng đầu vào là mẫu thử sẵn có của bốn chalcon nội bộ F29, F88, F90 và F91 được chọn sau khi trải qua quá trình sàng lọc *in silico* kép về tác dụng ức chế trên cả P-gp và NorA. Dựa trên kết quả sàng lọc (**Bảng 3.14** và **Bảng 3.16**) và phân tích (**Mục 4.3.1** và **Mục 4.3.2**), các chalcon F90 và F91 được dự đoán là có khả năng ức chế tốt cả P-gp lẫn NorA, trong khi dẫn xuất F88 có thể ức chế NorA hiệu quả hơn P-gp và ngược lại với dẫn xuất F29.

Tác dụng ức chế bơm ngược NorA trên *S. aureus* của các chalcon nói trên được đánh giá thông qua việc kiểm tra khả năng làm tăng tính nhạy cảm hay làm giảm MIC của kháng sinh chất nền ciprofloxacin trên các chủng vi khuẩn đề kháng do sản xuất quá mức bơm ngược. Cụ thể trong các thử nghiệm xác định MIC của ciprofloxacin trên *S. aureus* SA-1199, SA-1199B và trên các chủng SA phân lập từ lâm sàng có đề kháng bằng bơm ngược (**Bảng 3.17** và **Bảng 3.19**):

- Chalcon F88 giúp tăng tính nhạy cảm với ciprofloxacin khi làm giảm giá trị MIC của kháng sinh 2 lần trên chủng đề kháng chuẩn và từ 2 - 4 lần trên một số chủng lâm sàng ở các nồng độ thử nghiệm; chalcon F90 cũng cho thấy tác dụng điều hòa MDR tương tự F88 nhưng trên ít chủng lâm sàng hơn.

- Chalcon F29 giúp cải thiện sự nhạy cảm của kháng sinh giống như F88 và F90 nhưng ở nồng độ thử nghiệm cao hơn; trong khi sự có mặt của chalcon F91 ở các nồng độ thử nghiệm lại không cho thấy bất kỳ sự khác biệt nào về hiệu quả so với khi vắng mặt nó, cùng trên chủng chuẩn. Cả F29 và F91 cùng làm giảm MIC của kháng sinh trên ít chủng lâm sàng hơn F88 và F90.

Các kết quả thu được lần lượt trên *in silico* và *in vitro* đã chứng minh hai chalcon F88 và F90 trong tập nội bộ (tham khảo các thông số đặc trưng của chúng [43] trong **Phụ lục 9**) là các ứng viên thuốc tiềm năng giúp điều hòa sự bài xuất thuốc ra khỏi tế bào vi khuẩn qua trung gian bơm NorA, và qua đó làm giảm hiện tượng kháng thuốc trong các bệnh nhiễm trùng. Các kết quả này cũng đồng thời cho thấy mối liên hệ hữu cơ giữa thử nghiệm *in vitro* và sàng lọc *in silico*.

Bằng nghiên cứu thực nghiệm, khả năng áp dụng của các công cụ *in silico* được tạo ra trong nghiên cứu này hướng đến sự ức chế P-gp và NorA được khẳng định lại một lần nữa, bên cạnh các đánh giá thống kê trên máy tính. Trong đó, các kết quả dự đoán trên P-gp đã được sử dụng cho NorA một cách hiệu quả và có thể được giải thích là do sự trùng lặp phối tử của hai loại bơm ngược này [17]. Qua sự phân tích như vậy, các chalcon “lead” này cũng được kỳ vọng sẽ cho hiệu quả ức chế tương tự trên P-gp ở tế bào ung thư, giúp khôi phục hiệu quả của liệu pháp hóa trị ở các bệnh nhân ung thư.

Dựa trên các nghiên cứu *in silico* và *in vitro* đã thực hiện, một số nhận xét về mối quan hệ cấu trúc - tác dụng (SAR) được rút ra từ 04 chalcon F29, F88, F90 và F91 như sau:

- Vòng A với nhóm thế hydroxy ở vị trí ortho (nhóm cho liên kết hydro, kết hợp với nhóm nhận liên kết hydro là carbonyl lân cận) hoặc thay vòng A bằng nhóm phenothiazin (nhóm cho liên kết hydro (NH) và tương tác vòng thơm-vòng thơm) giúp tăng sự tương tác giữa phối tử với khoang gắn kết của protein và có thể làm tăng khả năng ức chế bơm.
- Vòng B với các nhóm thế methoxy (nhóm nhận liên kết hydro, tương tác kỵ nước) và đặc biệt là nhóm thế halogen clor (tương tác kỵ nước) là có lợi cho hoạt tính ức chế so với nhóm thế phân cực hơn là dimethyl amino (tạo tương tác kém và giảm tính thấm).
- Nhóm thế ở vị trí para của vòng B cần cho sự ức chế P-gp hơn là sự ức chế NorA (theo 02 giả thuyết pharmacophore được xây dựng, cũng như kết quả thử hoạt tính của 03 chalcon F88, F90 và F91).

Trong nghiên cứu hoạt tính gây độc tế bào, F88 và F90 cũng là hai trong số các chalcon biểu lộ độc tính hứa hẹn trên các tế bào rhabdomyosarcoma với $IC_{50} < 20 \mu M$, nhưng hầu như không cho thấy độc tính trên dòng tế bào lành tính LLC-PK1 [43]. Sự hiện diện của dị vòng phenothiazin ở vòng A và nhóm thế tích điện âm (clor) ở vị trí số 2 (F88) hoặc hai nhóm methoxy ở các vị trí số 2 và số 4 (F90) của vòng B được xác định là giúp làm gia tăng đáng kể độc tính tế bào. Như vậy, các yếu tố cấu trúc này là cơ sở chung cho tác dụng của các chalcon dị vòng trên các tác nhân gây bệnh, bao gồm độc tính mạnh và chọn lọc trên tế bào ung thư và khả năng đảo ngược MDR qua trung gian bơm ngược trên vi khuẩn.

KẾT LUẬN VÀ KIẾN NGHỊ

Nghiên cứu này đã thực hiện được mục tiêu đặt ra ban đầu với bốn nội dung, bao gồm thiết lập các mô hình máy tính dựa trên phối tử (nội dung 1) và dựa trên cấu trúc (nội dung 2), sàng lọc *in silico* sử dụng các công cụ máy tính thu được (nội dung 3) và sàng lọc *in vitro* dựa trên các kết quả máy tính thu được (nội dung 4). Cụ thể như sau:

1. Trong phương pháp dựa vào phối tử:

- Các mô hình học máy đơn lẻ và kết hợp có khả năng phân loại tốt chất ức chế và chất không ức chế P-gp (biến nhị phân 1 và 0) được xây dựng và đánh giá từ một tập dữ liệu hơn 2000 chất đa dạng về cấu trúc. Tương tự, các mô hình đơn lẻ và kết hợp mạnh cũng được tạo ra từ xấp xỉ 500 chất có cấu trúc khác nhau cho mục đích dự đoán hoạt tính ức chế P-gp (biến liên tục IC₅₀).
- Hai bản đồ nhận thức cho bốn lớp hoạt tính liên quan đến sự ức chế P-gp của người và NorA của *S. aureus* đã được tạo ra bằng các kỹ thuật đo lường đa hướng (MDS) và phân tích tương hợp (CA), từ một tập dữ liệu nhỏ 54 chất. Dựa trên các bản đồ này, một tập hợp bảy quy tắc cho các thông số mô tả và dấu vân tay được rút ra để giải quyết vấn đề về sự hỗn tạp phối tử giữa hai loại bơm ngược.
- Các yếu tố pharmacophore của chất ức chế P-gp mạnh và chất ức chế NorA nhưng không ức chế P-gp được xác định bằng chiến lược dựa trên phối tử để bổ sung cho các điều kiện sàng lọc. Trong đó, tính kỵ nước (cao hơn với P-gp) và một nhóm nhận liên kết hydro là cần thiết cho sự ức chế cả P-gp và NorA, còn các vòng thơm và một nhóm cho liên kết hydro lại giúp cho sự ức chế chỉ NorA.

2. Trong phương pháp dựa vào cấu trúc của protein, các mô hình tương đồng của P-gp được tạo ra để thay cho cấu trúc tinh thể tia X vẫn chưa được phân giải của nó. Qua đánh giá, mô hình tương đồng tốt nhất với khoang gắn kết thuốc có thể tích lớn ở vị trí trung tâm được xác định cho nghiên cứu docking.

3. Quá trình sàng lọc *in silico* bằng các công cụ máy tính được áp dụng cho hai thư viện nội bộ và Ngân hàng Thuốc để tìm kiếm các hợp chất mới hoặc đã biết, có khả năng ức chế P-gp và/hoặc NorA một cách hiệu quả. Kết quả thu được từ hai tập dữ liệu này là các chất “hit” mà chúng có thể giúp làm giảm hiện tượng đề kháng đa thuốc qua trung gian bơm ngược.
4. Nghiên cứu thực nghiệm tác dụng ức chế P-gp gặp nhiều khó khăn và chưa thể triển khai với các điều kiện hiện có tại Việt Nam. Tuy nhiên với sự tồn tại của các chất đồng ức chế P-gp của người và NorA của *S. aureus* như được đề cập, tác dụng ức chế bơm ngược có thể được kiểm tra thông qua các thử nghiệm trên vi khuẩn. Kết quả sàng lọc *in vitro* trên *S. aureus* đã xác nhận F88 và F90 là các chalcon “lead” có tác dụng ức chế bơm NorA, bởi vì hai chất này giúp làm tăng tính nhạy cảm với ciprofloxacin của các chủng SA kháng thuốc ở các nồng độ xác định mà ở đó chúng không có tác dụng kháng khuẩn.

Qua đó, một số kiến nghị được rút ra từ đề tài nghiên cứu này, gợi ý cho các hướng nghiên cứu tiếp theo trên bơm ngược và các chất ức chế bơm:

1. Về phương diện máy tính: (i) có thể ứng dụng các phần mềm và thuật toán ưu việt hơn, chẳng hạn như giúp khắc phục những hạn chế của các phương pháp hiện có, giải quyết được tính linh động về hình thể túi gắn kết của protein, ...; (ii) sử dụng cấu trúc tinh thể được phân giải của các protein màng (P-gp, NorA) nếu chúng sẵn có trong thời gian tới để thay cho mô hình tương đồng khi nghiên cứu docking; (iii) phát triển các tập dữ liệu sinh học lớn, đa dạng và đồng nhất, trong đó các giá trị hoạt tính của các chất (chẳng hạn như IC₅₀) có thể so sánh được với nhau, giúp làm tăng khả năng ngoại suy của các mô hình được xây dựng từ các tập dữ liệu này.
2. Có thể mở rộng nghiên cứu riêng lẻ hoặc kết hợp trên một số bơm ngược khác ở người như MRP1/ABCC1, MRP2/ABCC2, BCRP/ABCG2, ... và ở vi khuẩn như NorB, NorC, MepA, ... của *S. aureus*; MexAB-OprM, MexCD-OprJ, MexEF-OprN của *P. aeruginosa*; AcrAB-TolC của *E. coli*; ... Ngoài ra, cần

xem xét khả năng tương tác thuốc có thể xảy ra do hiệu quả ức chế bơm (tính chất antitarget), cũng như độ an toàn và độc tính của các chất nghiên cứu.

3. Đối với các chất “hit”, có thể tiến hành tổng hợp và thử nghiệm với nhiều mô hình *in vitro* và *in vivo* khác nhau để kiểm tra khả năng ức chế bơm. Còn đối với các chất “lead”, có thể nghiên cứu tối ưu hóa cấu trúc nhằm thu được các ứng viên thuốc có tiềm lực mạnh hơn và an toàn hơn trước khi tiến hành các bước thực nghiệm sâu hơn.
4. Có thể cân nhắc một số chiến lược Hóa Dược khác như phát triển các chất hướng đặc hiệu đến các tế bào đề kháng, các chất làm giảm sự biểu lộ của bơm ngược bằng cách biến đổi con đường tín hiệu trong tế bào, ... hoặc ngoài phạm vi Hóa Dược như sử dụng các kỹ thuật nano (vi hạt, polymer) giúp chuyên giao thuốc liều cao đến đúng mục tiêu điều trị.

DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ CÓ LIÊN QUAN

1. Thai K. M., Ngo T. D., Tran T. D. (2013), *Molecular modelling on Staphylococcus aureus Nor-a efflux pump inhibitors*, Proceeding of The Eighth Indochina Conference on Pharmaceutical Sciences, Ho Chi Minh City, Vietnam, ISBN 604660159-2, PO-PC-06, pp. 134-139.
2. Thai K. M., Ngo T. D., Phan T. V., Tran T. D., Nguyen N. V., Nguyen T. H., Le M. T. (2015), “Virtual Screening for Novel *Staphylococcus Aureus* NorA Efflux Pump Inhibitors From Natural Products”, *Medicinal Chemistry*, 11(2), pp. 135-155.
3. Ngo T. D., Tran T. D., Le M. T., Thai K. M. (2016), “Computational predictive models for P-glycoprotein inhibition of in-house chalcone derivatives and drug-bank compounds”, *Molecular Diversity*, 20(4), pp. 945-961.
4. Ngo T. D., Tran T. D., Le M. T., Thai K. M. (2016), “Machine learning-, rule- and pharmacophore-based classification on the inhibition of P-glycoprotein and NorA”, *SAR and QSAR in Environmental Research*, 27(9), pp. 747-780.
5. Ngô Triều Dũ, Trần Thành Đạo, Nguyễn Thị Thu Hà, Lê Minh Trí, Thái Khắc Minh (2016), “Dự đoán hoạt tính ức chế p-glycoprotein bằng các phương pháp máy học và docking”, *Y học Thành Phố Hồ Chí Minh*, 20(2), tr. 131-138.

TÀI LIỆU THAM KHẢO

TIẾNG VIỆT

1. Phạm Hùng Vân, Phạm Thái Bình (2013), *Kháng sinh - Đề kháng kháng sinh - Kỹ thuật kháng sinh đồ: Các vấn đề cơ bản thường gặp*, Nhà xuất bản y học.

TIẾNG ANH

2. Abu-Qare A. W., Elmasry E., Abou-Donia M. B. (2003), “A Role for P-Glycoprotein in Environmental Toxicology”, *Journal of Toxicology and Environmental Health, Part B, Critical Reviews*, 6(3), pp. 279-288.
3. Aeschlimann J. R., Dresser L. D., Kaatz G. W., Rybak M. J. (1999), “Effects of NorA Inhibitors on *In Vitro* Antibacterial Activities and Postantibiotic Effects of Levofloxacin, Ciprofloxacin, and Norfloxacin In Genetically Related Strains of *Staphylococcus aureus*”, *Antimicrobial Agents and Chemotherapy*, 43(2), pp. 335-340.
4. Aller S. G., Yu J., Ward A., Weng Y., Chittaboina S., Zhuo R., Harrell P. M., Trinh Y. T., Zhang Q., Urbatsch I. L., Chang G. (2009), “Structure of P-Glycoprotein Reveals a Molecular Basis for Poly-specific Drug Binding”, *Science*, 323(5922), pp. 1718-1722.
5. Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W., Lipman D. J. (1997), “Gapped BLAST and PSI-BLAST: a new generation of protein database search programs”, *Nucleic Acids Research*, 25(17), pp. 3389-3402.
6. Amyes S. G., Smith J. T. (1978), “R-factor mediated dihydrofolate reductases which confer trimethoprim resistance”, *Journal of General Microbiology*, 107(2), pp. 263-271.
7. Astolfi A., Felicetti T., Iraci N., Manfroni G., Massari S., Pietrella D., Tabarrini O., Kaatz G. W., Barreca M. L., Sabatini S., Cecchetti V. (2017), “Pharmacophore-Based Repositioning of Approved Drugs as Novel *Staphylococcus aureus* NorA Efflux Pump Inhibitors”, *Journal of Medicinal Chemistry*, 60(4), pp. 1598-1604.
8. Avendano C., Menendez J. C. (2008), “Drugs that modulate resistance to antitumor agents”, *Medicinal Chemistry of Anticancer Drugs*, Elsevier Science, Amsterdam, chapter 12, pp. 387-416.
9. Baguley B. C. (2010), “Multiple Drug Resistance Mechanisms in Cancer”, *Molecular Biotechnology*, 46(3), pp. 308-316.
10. Balouiri M., Sadiki M., Ibsouda S. K. (2016), “Methods for *in vitro* evaluating antimicrobial activity: A review”, *Journal of Pharmaceutical Analysis*, 6(2), pp. 71-79.

11. Bansal T., Jaggi M., Khar R. K., Talegaonkar S. (2009), "Emerging Significance of Flavonoids as P-Glycoprotein Inhibitors in Cancer Chemotherapy", *Journal of Pharmacy & Pharmaceutical Sciences*, 12(1), pp. 46-78.
12. Bengio Y., Grandvalet Y. (2004), "No Unbiased Estimator of the Variance of K-Fold Cross-Validation", *Journal of Machine Learning Research*, 5, pp. 1089-1105.
13. Bento A. P., Gaulton A., Hersey A., Bellis L. J., Chambers J., Davies M., Kruger F. A., Light Y., Mak L., McGlinchey S., Nowotka M., Papadatos G., Santos R., Overington J. P. (2014), "The ChEMBL bioactivity database: an update", *Nucleic Acids Research*, 42(Database issue), pp. D1083-1090.
14. Bhardwaj A. K., Mohanty P. (2012), "Bacterial Efflux Pumps Involved in Multidrug Resistance and their Inhibitors: Rejuvenating the Antimicrobial Chemotherapy", *Recent Patents on Anti-Infective Drug Discovery*, 7(1), pp. 73-89.
15. Binkhathlan Z., Lavasanifar A. (2013), "P-glycoprotein Inhibition as a Therapeutic Approach for Overcoming Multidrug Resistance in Cancer: Current Status and Future Perspectives", *Current Cancer Drug Targets*, 13(3), pp. 326-346.
16. Breiman L., Friedman J. H., Olshen R. A., Stone C. J. (1984), *Classification and regression trees*, Taylor & Francis.
17. Brincat J. P., Broccatelli F., Sabatini S., Frosini M., Neri A., Kaatz G. W., Cruciani G., Carosati E. (2012), "Ligand Promiscuity between the Efflux Pumps Human P-Glycoprotein and *S. aureus* NorA", *ACS Medicinal Chemistry Letters*, 3(3), pp. 248-251.
18. Broccatelli F., Carosati E., Neri A., Frosini M., Goracci L., Oprea T. I., Cruciani G. (2011), "A Novel Approach for Predicting P-Glycoprotein (ABCB1) Inhibition Using Molecular Interaction Fields", *Journal of Medicinal Chemistry*, 54(6), pp. 1740-1751.
19. Callaghan R., Luk F., Bebawy M. (2014), "Inhibition of the Multidrug Resistance P-Glycoprotein: Time for a Change of Strategy?", *Drug Metabolism and Disposition*, 42(4), pp. 623-631.
20. Chang C., Bahadduri P. M., Polli J. E., Swaan P. W., Ekins S. (2006), "Rapid Identification of P-glycoprotein Substrates and Inhibitors", *Drug Metabolism and Disposition*, 34(12), pp. 1976-1984.
21. ChemBioDrawUltra 12.0, CambridgeSoft, 100 Cambridge Park Drive, Cambridge, MA 02140.
22. Chen L., Li Y., Yu H., Zhang L., Hou T. (2012), "Computational models for predicting substrates or inhibitors of P-glycoprotein", *Drug Discovery Today*, 17(7-8), pp. 343-351.

23. Chen L., Li Y., Zhao Q., Peng H., Hou T. (2011), "ADME Evaluation in Drug Discovery. 10. Predictions of P-Glycoprotein Inhibitors Using Recursive Partitioning and Naive Bayesian Classification Techniques", *Molecular Pharmaceutics*, 8(3), pp. 889-900.
24. Cheng F., Li W., Zhou Y., Shen J., Wu Z., Liu G., Lee P. W., Tang Y. (2012), "admetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties", *Journal of Chemical Information and Modeling*, 52(11), pp. 3099-3105.
25. Chiba P., Mihalek I., Ecker G. F., Kopp S., Lichtarge O. (2006), "Role of Transmembrane Domain/Transmembrane Domain Interfaces of P-Glycoprotein (ABCB1) in Solute Transport. Convergent Information from Photoaffinity Labeling, Site Directed Mutagenesis and *in Silico* Importance Prediction", *Current Medicinal Chemistry*, 13(7), pp. 793-805.
26. Chirico N., Gramatica P. (2011), "Real External Predictivity of QSAR Models: How To Evaluate It? Comparison of Different Validation Criteria and Proposal of Using the Concordance Correlation Coefficient", *Journal of Chemical Information and Modeling*, 51(9), pp. 2320-2335.
27. Chirico N., Gramatica P. (2012), "Real External Predictivity of QSAR Models. Part 2. New Intercomparable Thresholds for Different Validation Criteria and the Need for Scatter Plot Inspection", *Journal of Chemical Information and Modeling*, 52(8), pp. 2044-2058.
28. Chufan E. E., Kapoor K., Sim H. M., Singh S., Talele T. T., Durell S. R., Ambudkar S. V. (2013), "Multiple Transport-Active Binding Sites Are Available for a Single Substrate on Human P-Glycoprotein (ABCB1)", *PLoS ONE*, 8(12), e82463.
29. Clementine 12.0, SPSS Inc., 233 South Wacker Drive, 11th Floor, Chicago, USA.
30. Consonni V., Ballabio D., Todeschini R. (2009), "Comments on the Definition of the Q^2 Parameter for QSAR Validation", *Journal of Chemical Information and Modeling*, 49(7), pp. 1669-1678.
31. Consonni V., Ballabio D., Todeschini R. (2010), "Evaluation of model predictive ability by external validation techniques", *Journal of Chemometrics*, 24(3-4), pp. 194-201.
32. Cortes-Selva F., Campillo M., Reyes C. P., Jimenez I. A., Castanys S., Bazzocchi I. L., Pardo L., Gamarro F., Ravelo A. G. (2004), "SAR Studies of Dihydro- β -agarofuran Sesquiterpenes as Inhibitors of the Multidrug-Resistance Phenotype in a *Leishmania tropica* Line Overexpressing a P-Glycoprotein-Like Transporter", *Journal of Medicinal Chemistry*, 47(3), pp. 576-587.
33. Costa S. S., Viveiros M., Amaral L., Couto I. (2013), "Multidrug Efflux Pumps in *Staphylococcus aureus*: an Update", *The Open Microbiology Journal*, 7, pp. 59-71.

34. Crivori P., Reinach B., Pezzetta D., Poggesi I. (2006), "Computational Models for Identifying Potential P-Glycoprotein Substrates and Inhibitors", *Molecular Pharmaceutics*, 3(1), pp. 33-44.
35. Dano K. (1973), "Active outward transport of daunomycin in resistant Ehrlich ascites tumor cells", *Biochimica et Biophysica Acta*, 323(3), pp. 466-483.
36. Dantzig A. H., de Alwis D. P., Burgess M. (2003), "Considerations in the design and development of transport inhibitors as adjuncts to drug therapy", *Advanced Drug Delivery Reviews*, 55(1), pp.133-150.
37. Darby R. A., Callaghan R., McMahon R. M. (2011), "P-glycoprotein Inhibition: The Past, the Present and the Future", *Current Drug Metabolism*, 12(8), pp. 722-731.
38. Dawson R. J., Locher K. P. (2006), "Structure of a bacterial multidrug ABC transporter", *Nature*, 443(7108), pp. 180-185.
39. Dawson R. J., Locher K. P. (2007) "Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP", *FEBS Letters*, 581(5), pp. 935-938.
40. Dearden J. C., Al-Noobi A., Scott A. C., Thomson S. A. (2003), "QSAR studies on P-glycoprotein-regulated multidrug resistance and on its reversal by phenothiazines", *SAR and QSAR in Environmental Research*, 14(5-6), pp. 447-454.
41. Demel M. A., Janecek A. G. K., Thai K. M., Ecker G. F., Gansterer W. N. (2008), "Predictive QSAR Models for Polyspecific Drug Targets: The Importance of Feature Selection", *Current Computer-Aided Drug Design*, 4(2), pp. 91-110.
42. Dickens D., Owen A., Alfirovic A., Pirmohamed M. (2009), *Determination of the inhibitory potencies of p-glycoprotein inhibitors by transcellular permeability of Caco-2 cells*, British Pharmacological Society Winter Meeting.
43. Do T. H., Nguyen D. M., Truong V. D., Do T. H. T., Le M. T., Pham T. Q., Thai K. M., Tran T. D. (2016), "Synthesis and Selective Cytotoxic Activities on Rhabdomyosarcoma and Noncancerous Cells of Some Heterocyclic Chalcones", *Molecules*, 21(3), 329.
44. Dobchev D. A., Pillai G. G., Karelson M. (2014), "In Silico Machine Learning Methods in Drug Development", *Current Topics in Medicinal Chemistry*, 14(16), pp. 1913-1922.
45. Dolgih E., Bryant C., Renslo A. R., Jacobson M. P. (2011), "Predicting Binding to P-Glycoprotein by Flexible Receptor Docking", *PLoS Computational Biology*, 7(6), e1002083.
46. Dupree T. (2005), *Pharmacophore development and validation for inhibitors of the bacterial NorA efflux pump*, Master of Science (Research) thesis, University of Wollongong.

47. Eberl S., Renner B., Neubert A., Reisig M., Bachmakov I., König J., Dorje F., Murdter T. E., Ackermann A., Dormann H., Gassmann K. G., Hahn E. G., Zierhut S., Brune K., Fromm M. F. (2007), "Role of P-Glycoprotein Inhibition for Drug Interactions: Evidence from *In Vitro* and Pharmacoepidemiological Studies", *Clinical Pharmacokinetics*, 46(12), pp. 1039-1049.
48. Ekins S., Kim R. B., Leake B. F., Dantzig A. H., Schuetz E. G., Lan L. B., Yasuda K., Shepard R. L., Winter M. A., Schuetz J. D., Wikel J. H., Wrighton S. A. (2002), "Three-Dimensional Quantitative Structure-Activity Relationships of Inhibitors of P-Glycoprotein", *Molecular Pharmacology*, 61(5), pp. 964-973.
49. Ekins S., Kim R. B., Leake B. F., Dantzig A. H., Schuetz E. G., Lan L. B., Yasuda K., Shepard R. L., Winter M. A., Schuetz J. D., Wikel J. H., Wrighton S. A. (2002), "Application of Three-Dimensional Quantitative Structure-Activity Relationships of P-Glycoprotein Inhibitors and Substrates", *Molecular Pharmacology*, 61(5), pp. 974-981.
50. Éric S., Kalinic M., Ilic K., Zloh M. (2014), "Computational classification models for predicting the interaction of drugs with P-glycoprotein and breast cancer resistance protein", *SAR and QSAR in Environmental Research*, 25(12), pp. 939-966.
51. Everitt B. S. (2005), "Multidimensional scaling and correspondence analysis", *An R and S-PLUS® Companion to Multivariate Analysis*, Springer Texts in Statistics, Springer, London, pp. 91-114.
52. Ferreira A., Pousinho S., Fortuna A., Falcão A., Alves G. (2015), "Flavonoid compounds as reversal agents of the P-glycoprotein-mediated multidrug resistance: biology, chemistry and pharmacology", *Phytochemistry Reviews*, 14(2), pp. 233-272.
53. Ferreira L. G., dos Santos R. N., Oliva G., Andricopulo A. D. (2015), "Molecular Docking and Structure-Based Drug Design Strategies", *Molecules*, 20(7), pp. 13384-13421.
54. Ferreira R. J., dos Santos D. J., Ferreira M. J., Guedes R. C. (2011), "Toward a Better Pharmacophore Description of P-Glycoprotein Modulators, Based on Macrocyclic Diterpenes from *Euphorbia* Species", *Journal of Chemical Information and Modeling*, 51(6), pp. 1315-1324.
55. Fromm M. F. (2003), "Importance of P-glycoprotein for drug disposition in humans", *European Journal of Clinical Investigation*, 33(S2), pp. 6-9.
56. Garrigues A., Loiseau N., Delaforge M., Ferte J., Garrigos M., Andre F., Orłowski S. (2002), "Characterization of Two Pharmacophores on the Multidrug Transporter P-Glycoprotein", *Molecular Pharmacology*, 62(6), pp. 1288-1298.

57. Ghaemian P., Shayanfar A. (2017), "Quantitative Structure Activity Relationship (QSAR) of Methylated Polyphenol Derivatives as Permeability Glycoprotein (P-gp) Inhibitors: A Comparison of Different Training and Test Set Selection Methods", *Letters in Drug Design & Discovery*, 14(9), pp. 999-1007.
58. Ghaemian P., Shayanfar A. (2019), "Image-based QSAR Model for the Prediction of P-gp Inhibitory Activity of Epigallocatechin and Gallic acid Derivatives", *Current Computer-Aided Drug Design*, 15(3), pp. 212-224.
59. Ghandadi M., Shayanfar A., Hamzeh-Mivehroud M., Jouyban A. (2014), "Quantitative structure activity relationship and docking studies of imidazole-based derivatives as P-glycoprotein inhibitors", *Medicinal Chemistry Research*, 23, pp. 4700-4712.
60. Gibbons S., Oluwatuyi M., Kaatz G. W. (2003), "A novel inhibitor of multidrug efflux pumps in *Staphylococcus aureus*", *Journal of Antimicrobial Chemotherapy*, 51(1), pp. 13-17.
61. Gillespie S. H., McHugh T. D. (2010), *Antibiotic Resistance Protocols - Second Edition*, Humana Press, pp. 143-157.
62. Glaeser H., Fromm M. F., König J. (2008), "Transporters and Drugs - An Overview", *Antitargets: Prediction and Prevention of Drug Side Effects*, Volume 38, Eds.: Vaz R. J., Klabunde T., Wiley-VCH Verlag GmbH & Co. KGaA, pp. 341-366.
63. Golbraikh A., Tropsha A. (2002), "Beware of q²!", *Journal of Molecular Graphics and Modelling*, 20(4), pp. 269-276.
64. Gramatica P. (2007), "Principles of QSAR models validation: Internal and external", *QSAR & Combinatorial Science*, 26(5), pp. 694-701.
65. Guner O. F., Henry D. R. (2000), "Metric for Analyzing Hit Lists and Pharmacophores", *Pharmacophore Perception, Development, and Use in Drug Design*, International University Line, California, pp. 195-211.
66. Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I. H. (2009), "The WEKA data mining software: An update", *ACM SIGKDD Explorations Newsletter*, 11(1), pp. 10-18.
67. Handzlik J., Matys A., Kieć-Kononowicz K. (2013), "Recent Advances in Multi-Drug Resistance (MDR) Efflux Pump Inhibitors of Gram-Positive Bacteria *S. aureus*", *Antibiotics*, 2(1), pp. 28-45.
68. Hengzhuang W., Ciofu O., Yang L., Wu H., Song Z., Oliver A., Høiby N. (2013), "High β -Lactamase Levels Change the Pharmacodynamics of β -Lactam Antibiotics in *Pseudomonas aeruginosa* Biofilms", *Antimicrobial Agents and Chemotherapy*, 57(1), pp. 196-204.
69. Holler J. G., Slotved H. C., Molgaard P., Olsen C. E., Christensen S. B. (2012), "Chalcone inhibitors of the NorA efflux pump in *Staphylococcus aureus* whole cells and enriched everted membrane vesicles", *Bioorganic & Medicinal Chemistry*, 20(14), pp. 4514-4521.

70. Hvorup R. N., Goetz B. A., Niederer M., Hollenstein K., Perozo E., Locher K. P. (2007), "Asymmetry in the Structure of the ABC Transporter-Binding Protein Complex BtuCD-BtuF", *Science*, 317(5843), pp. 1387-1390.
71. Ivanciuc O. (2007), "Applications of Support Vector Machines in Chemistry", *Reviews in Computational Chemistry*, Volume 23, Eds.: Lipkowitz K. B., Cundari T. R., Wiley-VCH, Weinheim, pp. 291-400.
72. Jabeen I., Wetwitayaklung P., Chiba P., Pastor M., Ecker G. F. (2013), "2D- and 3D-QSAR studies of a series of benzopyranes and benzopyrano[3,4b][1,4]-oxazines as inhibitors of the multidrug transporter P-glycoprotein", *Journal of Computer-Aided Molecular Design*, 27(2), pp. 161-171.
73. Jacoby G. A., Medeiros A. A. (1991), "More extended-spectrum beta-lactamases", *Antimicrobial Agents and Chemotherapy*, 35(9), pp. 1697-1704.
74. Jaworska J., Nikolova-Jeliazkova N., Aldenberg T. (2005), "QSAR Applicability Domain Estimation by Projection of the Training Set in Descriptor Space: A Review", *Alternatives to Laboratory Animals*, 33(5), pp. 445-459.
75. Jin M. S., Oldham M. L., Zhang Q., Chen J. (2012), "Crystal structure of the multidrug transporter P-glycoprotein from *Caenorhabditis elegans*", *Nature*, 490(7421), pp. 566-569.
76. Jorgensen J., Turnidge J. (2015), "Susceptibility Test Methods: Dilution and Disk Diffusion Methods*", *Manual of Clinical Microbiology, Eleventh Edition*, Eds.: Jorgensen J., Pfaller M., Carroll K., Funke G., Landry M., Richter S., Warnock D., ASM Press, Washington, DC, pp. 1253-1273.
77. Joshi P., Singh S., Wani A., Sharma S., Jain S. K., Singh B., Gupta B. D., Satti N. K., Koul S., Khan I. A., Kumar A., Bharate S. B., Vishwakarma R. A. (2014), "Osthonol and curcumin as inhibitors of human Pgp and multidrug efflux pumps of *Staphylococcus aureus*: Reversing the resistance against frontline antibacterial drugs", *Medicinal Chemistry Communication*, 5(10), pp. 1540-1547.
78. Kaatz G. W. (2002), "Inhibition of bacterial efflux pumps: A new strategy to combat increasing antimicrobial agent resistance", *Expert Opinion on Emerging Drugs*, 7(2), pp. 223-233.
79. Kalia N. P., Mahajan P., Mehra R., Nargotra A., Sharma J. P., Koul S., Khan I. A. (2012), "Capsaicin, a novel inhibitor of the NorA efflux pump, reduces the intracellular invasion of *Staphylococcus aureus*", *Journal of Antimicrobial Chemotherapy*, 67(10), pp. 2401-2408.
80. Kass G. (1980), "An exploratory technique for investigating large quantities of categorical data", *Applied Statistics*, 29(2), pp. 119-127.

81. Kelly R. J., Robey R. W., Chen C. C., Draper D., Luchenko V., Barnett D., Oldham R. K., Caluag Z., Frye A. R., Steinberg S. M., Fojo T., Bates S. E. (2012), "A Pharmacodynamic Study of the P-glycoprotein Antagonist CBT-1® in Combination With Paclitaxel in Solid Tumors", *The Oncologist*, 17(4), 512.
82. Khan I. A., Mirza Z. M., Kumar A., Verma V., Qazi G. N. (2006), "Piperine, a Phytochemical Potentiator of Ciprofloxacin against *Staphylococcus aureus*", *Antimicrobial Agents and Chemotherapy*, 50(2), pp. 810-812.
83. Klepsch F., Vasanathanathan P., Ecker G. F. (2014), "Ligand and Structure-Based Classification Models for Prediction of P-Glycoprotein Inhibitors", *Journal of Chemical Information and Modeling*, 54(1), pp. 218-229.
84. Knox C., Law V., Jewison T., Liu P., Ly S., Frolkis A., Pon A., Banco K., Mak C., Neveu V., Djoumbou Y., Eisner R., Guo A. C., Wishart D. S. (2011), "DrugBank 3.0: a comprehensive resource for "Omics" research on drugs", *Nucleic Acids Research*, 39(Database issue), pp. D1035-D1041.
85. Kourtesi C., Ball A. R., Huang Y. Y., Jachak S. M., Vera D. M., Khondkar P., Gibbons S., Hamblin M. R., Tegos G. P. (2013), "Microbial Efflux Systems and Inhibitors: Approaches to Drug Discovery and the Challenge of Clinical Implementation", *The Open Microbiology Journal*, 7, pp. 34-52.
86. Kumar A., Khan I. A., Koul S., Koul J. L., Taneja S. C., Ali I., Ali F., Sharma S., Mirza Z. M., Kumar M., Sangwan P. L., Gupta P., Thota N., Qazi G. N. (2008), "Novel structural analogues of piperine as inhibitors of the NorA efflux pump of *Staphylococcus aureus*", *Journal of Antimicrobial Chemotherapy*, 61(6), pp. 1270-1276.
87. Kupsáková I., Rybár A., Dočolomanský P., Drobná Z., Stein U., Walther W., Barančík M., Breier A. (2004), "Reversal of P-glycoprotein mediated vincristine resistance of L1210/VCR cells by analogues of pentoxifylline: A QSAR study", *European Journal of Pharmaceutical Sciences*, 21(2-3), pp. 283-293.
88. Langer T. (2010), "Pharmacophores in Drug Research", *Molecular Informatics*, 29(6-7), pp. 470-475.
89. Langer T., Eder M., Hoffmann R. D., Chiba P., Ecker G. F. (2004), "Lead Identification for Modulators of Multidrug Resistance based on *in silico* Screening with a Pharmacophoric Feature Model", *Archiv der Pharmazoe*, 337(6), pp. 317-327.
90. Laskowski R. A., MacArthur M. W., Moss D. S., Thornton J. M. (1993), "PROCHECK: a program to check the stereochemical quality of protein structures", *Journal of Applied Crystallography*, 26(2), pp. 283-291.
91. Law V., Knox C., Djoumbou Y., Jewison T., Guo A. C., Liu Y., Maciejewski A., Arndt D., Wilson M., Neveu V., Tang A., Gabriel G., Ly C., Adamjee S., Dame Z. T., Han B., Zhou Y., Wishart D. S. (2014), "DrugBank 4.0: shedding new light on drug metabolism", *Nucleic Acids Research*, 42(Database issue), pp. D1091-D1097.

92. LeadIT 2.0.2, BioSolveIT GmbH, An der Ziegelei 79, 53757 St. Augustin, Germany.
93. Lehne G. (2000), "P-glycoprotein as a Drug Target in the Treatment of Multidrug Resistant Cancer", *Current Drug Targets*, 1(1), pp. 85-99.
94. Leitner I., Nemeth J., Feurstein T., Abraham A., Matzneller P., Lagler H., Erker T., Langer O., Zeitlinger M. (2011), "The third-generation P-glycoprotein inhibitor tariquidar may overcome bacterial multidrug resistance by increasing intracellular drug concentration", *Journal of Antimicrobial Chemotherapy*, 66(4), pp. 834-839.
95. Leong M. K., Chen H. B., Shih Y. H. (2012), "Prediction of Promiscuous P-Glycoprotein Inhibition Using a Novel Machine Learning Scheme", *PLoS One*, 7(3), e33829.
96. Lewis R. A. (2005), "A General Method for Exploiting QSAR Models in Lead Optimization", *Journal of Medicinal Chemistry*, 48(5), pp. 1638-1648.
97. Li J., Jaimes K. F., Aller S. G. (2014), "Refined structures of mouse P-glycoprotein", *Protein Science: A Publication of the Protein Society*, 23(1), pp. 34-46.
98. Li X. Z., Nikaido H. (2004), "Efflux-mediated drug resistance in bacteria", *Drugs*, 64(2), pp. 159-204.
99. Lin J. H., Yamazaki M. (2003), "Role of P-glycoprotein in Pharmacokinetics: Clinical Implications", *Clinical Pharmacokinetics*, 42(1), pp. 59-98.
100. Lin L. I. (1989), "A Concordance Correlation Coefficient to Evaluate Reproducibility", *Biometrics*, 45(1), pp. 255-268.
101. Litman T., Zeuthen T., Skovsgaard T., Stein W. D. (1997), "Structure-activity relationships of P-glycoprotein interacting drugs: kinetic characterization of their effects on ATPase activity", *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1361(2), pp. 159-168.
102. Liu H., Ma Z., Wu B. (2013), "Structure-activity relationships and *in silico* models of P-glycoprotein (ABCB1) inhibitors", *Xenobiotica*, 43(11), pp. 1018-1026.
103. Locher K. P., Lee A. T., Rees D. C. (2002), "The *E. coli* BtuCD Structure: A Framework for ABC Transporter Architecture and Mechanism", *Science*, 296(5570), pp. 1091-1098.
104. Loh W. Y., Shih Y. S. (1997), "Split selection methods for classification trees", *Statistica Sinica*, 7, pp. 815-840.
105. Lower M., Proschak E. (2011), "Structure-Based Pharmacophores for Virtual Screening", *Molecular Informatics*, 30(5), pp. 398-404.
106. MacGowan A. P., Wise R. (2001), "Establishing MIC breakpoints and interpretation of *in vitro* susceptibility tests", *Journal of Antimicrobial Chemotherapy*, 48(Suppl 1), pp. 17-38.
107. Marquez B. (2005), "Bacterial efflux systems and efflux pumps inhibitors", *Biochimie*, 87(12), pp. 1137-1147.

108. McCullagh P., Nelder J. A. (1989), *Generalized Linear Models - 2nd Edition*, Chapman & Hall, London.
109. Mei H., Zhou Y., Liang G., Li Z. (2005), "Support vector machine applied in QSAR modelling", *Chinese Science Bulletin*, 50(20), pp. 2291-2296.
110. Meng X. Y., Zhang H. X., Mezei M., Cui M. (2011), "Molecular Docking: A powerful approach for structure-based drug discovery", *Current Computer-Aided Drug Design*, 7(2), pp. 146-157.
111. MOE 2008.10, Chemical Computing Group Inc., 1010 Sherbrooke St. W, Suite 910, Montreal, Quebec, Canada.
112. Montanari F., Ecker G. F. (2015), "Prediction of drug-ABC-transporter interaction - Recent advances and future challenges", *Advanced Drug Delivery Reviews*, 86, pp. 17-26.
113. Morris G. M., Lim-Wilby M. (2008), "Molecular Docking", *Molecular Modeling of Proteins - Methods in Molecular Biology*, Volume 443, Ed.: Kukol A., Humana Press, chapter 19.
114. Mullin S., Mani N., Grossman T. H. (2004), "Inhibition of Antibiotic Efflux in Bacteria by the Novel Multidrug Resistance Inhibitors Biricodar (VX-710) and Timcodar (VX-853)", *Antimicrobial Agents and Chemotherapy*, 48(11), pp. 4171-4176.
115. NCBI, Multidrug resistance protein 1 isoform 2 [Homo sapiens] - NCBI Reference Sequence: NP_000918.2, <http://www.ncbi.nlm.nih.gov/protein/42741659?report=fasta>. Accessed on 22/09/2015.
116. Nelder J. A., Wedderburn R. W. M. (1972), "Generalized Linear Models", *Journal of the Royal Statistical Society. Series A (General)*, 135(3), pp. 370-384.
117. Neter J., Wasserman W., Kutner M. H. (1990), *Applied Linear Statistical Models - 3rd Edition*, CRC Press.
118. Netzeva T. I., Worth A., Aldenberg T., Benigni R., Cronin M. T., Gramatica P., Jaworska J. S., Kahn S., Klopman G., Marchant C. A., Myatt G., Nikolova-Jeliazkova N., Patlewicz G. Y., Perkins R., Roberts D., Schultz T., Stanton D. W., van de Sandt J. J., Tong W., Veith G., Yang C. (2005), "Current status of Methods for Defining the Applicability Domain of (Quantitative) Structure-Activity Relationships. The report and recommendations of ECVAM Workshop 52", *Alternatives to Laboratory Animals*, 33(2), pp. 155-173.
119. Neyfakh A. A. (1992), "The multidrug efflux transporter of *Bacillus subtilis* is a structural and functional homolog of the *Staphylococcus* NorA protein", *Antimicrobial Agents and Chemotherapy*, 36(2), pp. 484-485.
120. Ng E. Y., Trucksis M., Hooper D. C. (1994), "Quinolone resistance mediated by norA: physiologic characterization and relationship to flqB, a quinolone resistance locus on the *Staphylococcus aureus* chromosome", *Antimicrobial Agents and Chemotherapy*, 38(6), pp. 1345-1355.

121. Nikaido H. (2009), "Multidrug Resistance in Bacteria", *Annual Review of Biochemistry*, 78, pp. 119-146.
122. Nikaido H., Pagès J. M. (2012), "Broad Specificity Efflux pumps and Their Role in Multidrug Resistance of Gram Negative Bacteria", *FEMS Microbiology Reviews*, 36(2), pp. 340-363.
123. Ojha P. K., Mitra I., Das R. N., Roy K. (2011), "Further exploring r_m^2 metrics for validation of QSPR models", *Chemometrics and Intelligent Laboratory Systems*, 107(1), pp. 194-205.
124. Ojha P. K., Roy K. (2011), "Comparative QSARs for antimalarial endochins: Importance of descriptor-thinning and noise reduction prior to feature selection", *Chemometrics and Intelligent Laboratory Systems*, 109(2), pp. 146-161.
125. Oldham R. K., Reid W. K., Barnett D. (2000), "Phase I Study of CBT-1 and Taxol in Patients with Taxol Resistant Cancers", *Cancer Biotherapy & Radiopharmaceuticals*, 15(2), pp. 153-159.
126. Oldham R. K., Reid W. K., Preisler H. D., Barnett D. (1998), "A Phase I and Pharmacokinetic Study of CBT-1 as a Multidrug Resistance Modulator in the Treatment of Patients with Advanced Cancer", *Cancer Biotherapy and Radiopharmaceuticals*, 13(2), pp. 71-80.
127. Organisation for Economic Co-operation and Development (OECD) (2014), "Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models", *OECD Series on Testing and Assessment*, Volume 69, OECD Publishing, Paris.
128. Österberg T., Norinder U. (2000) "Theoretical calculation and prediction of P-glycoprotein-interacting drugs using MolSurf parametrization and PLS statistics", *European Journal of Pharmaceutical Sciences*, 10(4), pp. 295-303.
129. Ounpraseuth S., Lensing S. Y., Spencer H. J., Kodell R. L. (2012), "Estimating misclassification error: a closer look at cross-validation based methods", *BMC Research Notes*, 5, 656.
130. Ozben T. (2006), "Mechanisms and strategies to overcome multiple drug resistance in cancer", *FEBS Letters*, 580(12), pp. 2903-2909.
131. Pajeva I., Globisch C., Fleischer R., Tsakovska I., Wiese M. (2005), "Molecular Modeling of P-Glycoprotein and Related Drugs", *Medicinal Chemistry Research*, 14(2), pp. 106-117.
132. Pajeva I. K., Globisch C., Wiese M. (2009), "Combined Pharmacophore Modeling, Docking, and 3D QSAR Studies of ABCB1 and ABCC1 Transporter Inhibitors", *ChemMedChem*, 4(11), pp. 1883-1896.
133. Pajeva I. K., Wiese M. (2002), "Pharmacophore Model of Drugs Involved in P-Glycoprotein Multidrug Resistance: Explanation of Structural Variety (Hypothesis)", *Journal of Medicinal Chemistry*, 45(26), pp. 5671-5686.

134. Palmeira A., Rodrigues F., Sousa E., Pinto M., Vasconcelos M. H., Fernandes M. (2010), "Pharmacophore-Based Screening as a Clue for the Discovery of New P-Glycoprotein Inhibitors", *Advances in Bioinformatics - Advances in Intelligent and Soft Computing*, Volume 74, Eds.: Rocha M. P., Riverola F. F., Shatkay H., Corchado J. M., Springer, Berlin, pp. 175-180.
135. Palmeira A., Rodrigues F., Sousa E., Pinto M., Vasconcelos M. H., Fernandes M. X. (2011), "New Uses for Old Drugs: Pharmacophore-Based Screening for the Discovery of P-Glycoprotein Inhibitors", *Chemical Biology & Drug Design*, 78(1), pp. 57-72.
136. Palmeira A., Sousa E., Vasconcelos M. H., Pinto M. M. (2012), "Three Decades of P-gp Inhibitors: Skimming Through Several Generations and Scaffolds", *Current Medicinal Chemistry*, 19(13), pp. 1946-2025.
137. Palmeira A., Sousa E., Vasconcelos M. H., Pinto M., Fernandes M. X. (2012), "Structure and Ligand-based Design of P-glycoprotein Inhibitors: A Historical Perspective", *Current Pharmaceutical Design*, 18(27), pp. 4197-4214.
138. Pandya R., Pandya J. (2015), "C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning", *International Journal of Computer Applications*, 117(16), pp. 18-21.
139. Parveen Z., Brunhofer G., Jabeen I., Erker T., Chiba P., Ecker G. F. (2014), "Synthesis, biological evaluation and 3D-QSAR studies of new chalcone derivatives as inhibitors of human P-glycoprotein", *Bioorganic & Medicinal Chemistry*, 22(7), pp. 2311-2319.
140. Patel D., Kosmidis C., Seo S. M., Kaatz G. W. (2010), "Ethidium Bromide MIC Screening for Enhanced Efflux Pump Gene Expression or Efflux Activity in *Staphylococcus aureus*", *Antimicrobial Agents and Chemotherapy*, 54(12), pp. 5070-5073.
141. Pesic M., Bankovic J., Tanic N. (2014), "Targeted Anti-Cancer Therapy, Acquiring and Overcoming Multi-Drug Resistance", *Frontiers in Anti-Cancer Drug Discovery*, Eds.: Atta-ur-Rahman, Choudhary M. I., Sharjah: Bentham Science Publishers, pp. 109-150.
142. Poongavanam V., Haider N., Ecker G. F. (2012), "Fingerprint-based *in silico* models for the prediction of P-glycoprotein substrates and inhibitors", *Bioorganic & Medicinal Chemistry*, 20(18), pp. 5388-5395.
143. Prachayasittikul V., Worachartcheewan A., Shoombuatong W., Prachayasittikul V., Nantasenamat C. (2015), "Classification of P-glycoprotein-interacting compounds using machine learning methods", *EXCLI journal*, 14, pp. 958-970.
144. Prachayasittikul V., Worachartcheewan A., Toropova A. P., Toropov A. A., Schaduangrat N., Prachayasittikul V., Nantasenamat C. (2017), "Large-scale classification of P-glycoprotein inhibitors using SMILES-based descriptors", *SAR and QSAR in Environmental Research*, 28(1), pp. 1-16.

145. RapidMiner 5.3.008, Rapid-I and contributors, Stockumer Str. 475, 44227 Dortmund, Germany.
146. Rapposelli S., Coi A., Imbriani M., Bianucci A. M. (2012), "Development of Classification Models for Identifying "True" P-glycoprotein (P-gp) Inhibitors Through Inhibition, ATPase Activation and Monolayer Efflux Assays", *International Journal of Molecular Sciences*, 13(6), pp. 6924-6943.
147. Rautio J., Humphreys J. E., Webster L. O., Balakrishnan A., Keogh J. P., Kunta J. R., Serabjit-Singh C. J., Polli J. W. (2006), "In vitro P-glycoprotein inhibition assays for assessment of clinical drug interaction potential of new drug candidates: A recommendation for probe substrates", *Drug Metabolism and Disposition*, 34(5), pp. 786-792.
148. Ravna A.W., Sylte I. (2012), "Homology Modeling of Transporter Proteins (Carriers and Ion Channels)", *Homology Modeling. Methods in Molecular Biology (Methods and Protocols)*, Volume 857, Eds.: Orry A., Abagyan R., Humana Press, pp. 281-299.
149. Reyes C. P., Munoz-Martinez F., Torrecillas I. R., Mendoza C. R., Gamarro F., Bazzocchi I. L., Nunez M. J., Pardo L., Castanys S., Campillo M., Jimenez I. A. (2007), "Biological Evaluation, Structure-Activity Relationships, and Three-Dimensional Quantitative Structure-Activity Relationship Studies of Dihydro- β -agarofuran Sesquiterpenes as Modulators of P-Glycoprotein-Dependent Multidrug Resistance", *Journal of Medicinal Chemistry*, 50(20), pp. 4808-4817.
150. Ronchi E., Sanfilippo O., Fronzo G. D., Bani M. R., Torre G. D., Catania S., Silvestrini R. (1989), "Detection of the 170 kDa P-Glycoprotein in Neoplastic and Normal Tissues", *Tumori Journal*, 75(6), pp. 542-546.
151. Roy A., Kucukural A., Zhang Y. (2010), "I-TASSER: a unified platform for automated protein structure and function prediction", *Nature Protocols*, 5(4), pp. 725-738.
152. Roy K., Das R. N., Ambure P., Aher R. B. (2016), "Be aware of error measures. Further studies on validation of predictive QSAR models", *Chemometrics and Intelligent Laboratory Systems*, 152, pp. 18-33.
153. Roy K., Kar S. (2015), "Importance of applicability domain of QSAR models", *Quantitative Structure-Activity Relationships in Drug Design, Predictive Toxicology, and Risk Assessment*, Ed.: Roy K., IGI Global, USA, chapter 5, pp. 180-211.
154. Roy K., Kar S., Ambure P. (2015), "On a simple approach for determining applicability domain of QSAR models", *Chemometrics and Intelligent Laboratory Systems*, 145, pp. 22-29.
155. Roy K., Mitra I., Kar S., Ojha P. K., Das R. N., Kabir H. (2012), "Comparative Studies on Some Metrics for External Validation of QSPR Models", *Journal of Chemical Information and Modeling*, 52(2), pp. 396-408.

156. Roy P. P., Paul S., Mitra I., Roy K. (2009), "On Two Novel Parameters for Validation of Predictive QSAR Models", *Molecules*, 14(5), pp. 1660-1701.
157. Roy P. P., Roy K. (2008), "On Some Aspects of Variable Selection for Partial Least Squares Regression Models", *QSAR & Combinatorial Science*, 27(3), pp. 302-313.
158. Rumelhart D. E., Hinton G. E., McClelland J. L. (1986), "A general framework for parallel distributed processing", *Parallel distributed processing: explorations in the microstructure of cognition*, Volume 1, MIT Press Cambridge, USA, pp. 45-76.
159. Sak K. (2014), "Cytotoxicity of dietary flavonoids on different human cancer types", *Pharmacognosy Reviews*, 8(16), pp. 122-146.
160. Saneja A., Dubey R. D., Alam N., Khare V., Gupta P. N. (2014), "Co-formulation of P-glycoprotein Substrate and Inhibitor in Nanocarriers: An Emerging Strategy for Cancer Chemotherapy", *Current Cancer Drug Targets*, 14(5), pp. 419-433.
161. Schindler B. D., Kaatz G. W. (2016), "Multidrug efflux pumps of Gram-positive bacteria", *Drug Resistance Updates*, 27, pp. 1-13.
162. Schmitz F. J., Fluit A. C., Luckefahr M., Engler B., Hofmann B., Verhoef J., Heinz H. P., Hadding U., Jones M. E. (1998), "The effect of reserpine, an inhibitor of multidrug efflux pumps, on the in-vitro activities of ciprofloxacin, sparfloxacin and moxifloxacin against clinical isolates of *Staphylococcus aureus*", *Journal of Antimicrobial Chemotherapy*, 42, pp. 807-810.
163. Schuurmann G., Ebert R. U., Chen J., Wang B., Kuhne R. (2008), "External Validation and Prediction Employing the Predictive Squared Correlation Coefficient - Test Set Activity Mean vs Training Set Activity Mean", *Journal of Chemical Information and Modeling*, 48(11), pp. 2140-2145.
164. Schyman P., Liu R., Wallqvist A. (2016), "Using the Variable-Nearest Neighbor Method To Identify P-Glycoprotein Substrates and Inhibitors". *ACS Omega*, 1, pp. 923-929.
165. Sharifi M., Rayevsky A. V., Ghafourian T. (2016), "Effect of Molecular Structure, Substrate and Docking Scores on the Prediction of the Inhibition Constants of P-glycoprotein Inhibitors", *Journal of Drug Metabolism & Toxicology*, 7(4), pp. 1-12.
166. Sharom F. J. (2006), "Multidrug Resistance Protein: P-glycoprotein", *Drug transporters: Molecular Characterization and Role in Drug Disposition*, Eds.: You G., Morris M. E., John Wiley & Sons, Inc., chapter 10, pp. 223-262.
167. Shen J., Cui Y., Gu J., Li Y., Li L. (2014), "A Genetic Algorithm- Back Propagation Artificial Neural Network Model to Quantify the Affinity of Flavonoids Toward P-Glycoprotein", *Combinatorial Chemistry & High Throughput Screening*, 17(2), pp. 162-172.

168. Shi L. M., Fang H., Tong W., Wu J., Perkins R., Blair R. M., Branham W. S., Dial S. L., Moland C. L., Sheehan D. M. (2001), "QSAR Models Using a Large Diverse Set of Estrogens", *Journal of Chemical Information and Computer Sciences*, 41(1), pp. 186-195.
169. Shintre C. A., Pike A. C., Li Q., Kim J. I., Barr A. J., Goubin S., Shrestha L., Yang J., Berridge G., Ross J., Stansfeld P. J., Sansom M. S., Edwards A. M., Bountra C., Marsden B. D., von Delft F., Bullock A. N., Gileadi O., Burgess-Brown N. A., Carpenter E. P. (2013), "Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states", *Proceedings of the National Academy of Sciences of the United States of America*, 110(24), pp. 9710-9715.
170. Shityakov S., Forster C. (2014), "In silico structure-based screening of versatile P-glycoprotein inhibitors using polynomial empirical scoring functions", *Advances and Applications in Bioinformatics and Chemistry*, 7, pp. 1-9.
171. Sousa I. J., Ferreira M. J. U., Molnár J., Fernandes M. X. (2013), "QSAR studies of macrocyclic diterpenes with P-glycoprotein inhibitory activity", *European Journal of Pharmaceutical Sciences*, 48(3), pp. 542-553.
172. SPSS 20.0, IBM Corp., 1 New Orchard Road, Armonk, New York, United States.
173. Srivalli K. M. R., Lakshmi P. K. (2012), "Overview of P-glycoprotein inhibitors: A rational outlook", *Brazilian Journal of Pharmaceutical Science*, 48(3), pp. 353-367.
174. Sun H. (2005), "A Naive Bayes Classifier for Prediction of Multidrug Resistance Reversal Activity on the Basis of Atom Typing", *Journal of Medicinal Chemistry*, 48(12), pp. 4031-4039.
175. Sun J., Deng Z., Yan A. (2014), "Bacterial multidrug efflux pumps: Mechanisms, physiology and pharmacological exploitations", *Biochemical and Biophysical Research Communications*, 453(2), pp. 254-267.
176. Szakács G., Paterson J. K., Ludwig J. A., Booth-Genthe C., Gottesman M. M. (2006), "Targeting multidrug resistance in cancer", *Nature Reviews Drug Discovery*, 5(3), pp. 219-234.
177. Teodori E., Dei S., Martelli C., Scapecchi S., Gualtieri F. (2006), "The functions and structure of ABC transporters: implications for the design of new inhibitors of Pgp and MRP1 to control multidrug resistance (MDR)", *Current Drug Targets*, 7(7), pp. 893-909.
178. Thai K. M., Do T. N., Nguyen T. V. P., Nguyen D. K. T., Tran T. D. (2015), "QSAR studies on bacterial efflux pump inhibitors", *Quantitative Structure-Activity Relationships in Drug Design, Predictive Toxicology, and Risk Assessment*, Ed.: Roy K., IGI Global, USA, chapter 7, pp. 238-268.
179. Thai K. M., Ecker G. F. (2008), "Classification models for HERG inhibitors by counter-propagation neural networks", *Chemical Biology & Drug Design*, 72(4), pp. 279-289.

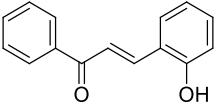
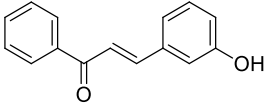
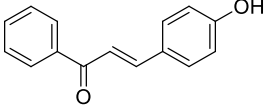
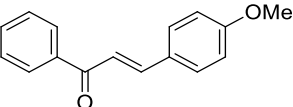
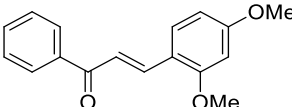
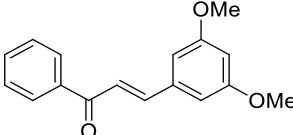
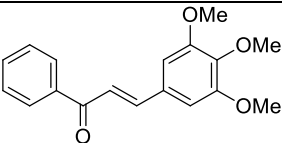
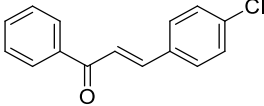
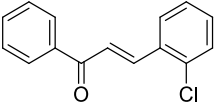
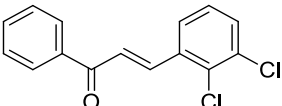
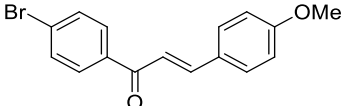
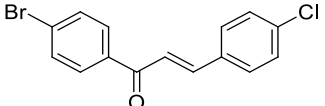
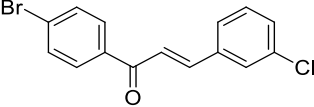
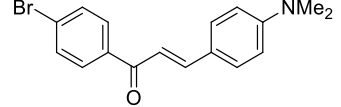
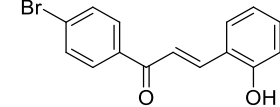
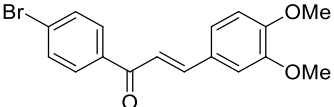
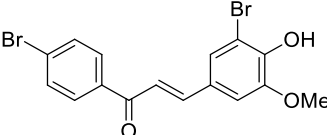
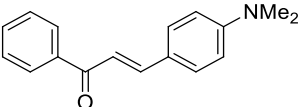
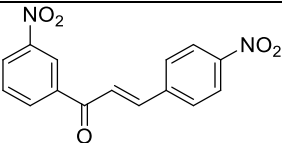
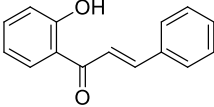
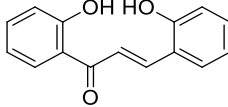
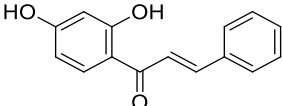
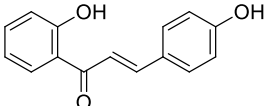
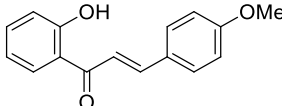
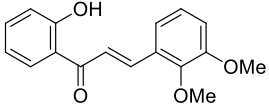
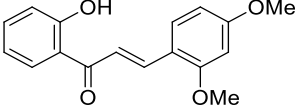
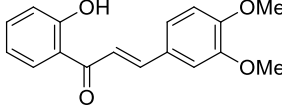
180. Thai K. M., Huynh N. T., Ngo T. D., Mai T. T., Nguyen T. H., Tran T. D. (2015), "Three- and four-class classification models for P-glycoprotein inhibitors using counter-propagation neural networks", *SAR and QSAR in Environmental Research*, 26(2), pp. 139-163.
181. Thai K. M., Ngo T. D., Phan T. V., Tran T. D., Nguyen N. V., Nguyen T. H., Le M. T. (2015), "Virtual Screening for Novel *Staphylococcus Aureus* NorA Efflux Pump Inhibitors From Natural Products", *Medicinal Chemistry*, 11(2), pp. 135-155.
182. Thai K. M., Ngo T. D., Tran T. D., Le M. T. (2013), "Pharmacophore modeling for antitargets", *Current Topics in Medicinal Chemistry*, 13(9), pp. 1002-1014.
183. Thomas H., Coley H. M. (2003), "Overcoming Multidrug Resistance in Cancer: An Update on the Clinical Strategy of Inhibiting P-Glycoprotein", *Cancer Control*, 10(2), pp. 159-165.
184. Tropsha A. (2010), "Best Practices for QSAR Model Development, Validation, and Exploitation", *Molecular Informatics*, 29(6-7), pp. 476-488.
185. Tropsha A., Gramatica P., Gombar V. K. (2003), "The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models", *QSAR & Combinatorial Science*, 22(1), pp. 69-77.
186. Ughachukwu P., Unekwe P. (2012), "Efflux Pump-Mediated Resistance in Chemotherapy", *Annals of Medical and Health Sciences Research*, 2(2), pp. 191-198.
187. U.S. Food and Drug Administration, Drug Development and Drug Interactions: Table of Substrates, Inhibitors and Inducers. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/DrugInteractionsLabeling/ucm093664.htm>. Accessed on 22/09/2015.
188. Van Bambeke F., Balzi E., Tulkens P. M. (2000), "Antibiotic efflux pumps", *Biochemical Pharmacology*, 60(4), pp. 457-470.
189. Varma M. V., Ashokraj Y., Dey C. S., Panchagnula R. (2003), "P-glycoprotein inhibitors and their screening: a perspective from bioavailability enhancement", *Pharmacological Research*, 48(4), pp. 347-359.
190. Vaz R. J., Klabunde T. (2008), "A personal foreword", *Antitargets: Prediction and Prevention of Drug Side Effects*, Volume 38, Eds.: Vaz R. J., Klabunde T., Wiley-VCH Verlag GmbH & Co. KGaA, pp. XXI-XXIV.
191. Vazquez R. N., Camargo A. B., Marchevsky E. J., Luco J. M. (2014), "Molecular Factors Influencing the Affinity of Flavonoid Compounds on P-glycoprotein Efflux Transporter", *Current Computer-Aided Drug Design*, 10(3), pp. 250-258.
192. Veerasamy R., Rajak H., Jain A., Sivadasan S., Varghese C. P., Agrawal R. K. (2011), "Validation of QSAR Models - Strategies and Importance", *International Journal of Drug Design and Discovery*, 2(3), pp. 511-519.

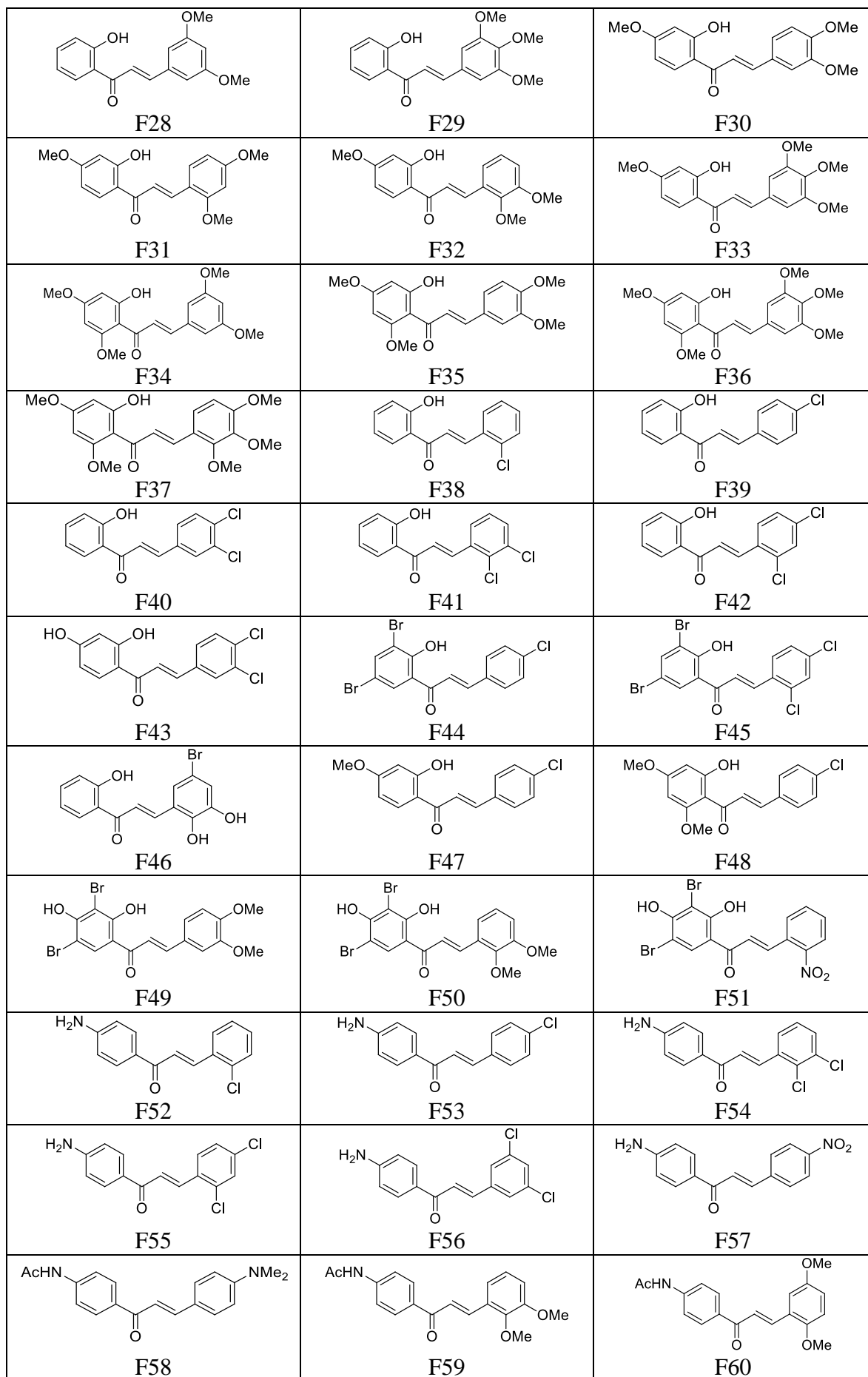
204. World Health Organization (2017), *Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics*.
205. Wu J., Li X., Cheng W., Xie Q., Liu Y., Zhao C. (2009), "Quantitative Structure Activity Relationship (QSAR) Approach to Multiple Drug Resistance (MDR) Modulators Based on Combined Hybrid System", *QSAR & Combinatorial Science*, 28(9), pp. 969-978.
206. Wu S., Skolnick J., Zhang Y. (2007), "Ab initio modeling of small proteins by iterative TASSER simulations", *BMC Biology*, 5, 17.
207. Xia M., Fang Y., Cao W., Liang F., Pan S., Xu X. (2019), "Quantitative Structure-Activity Relationships for the Flavonoid-Mediated Inhibition of P-Glycoprotein in KB/MDR1 Cells", *Molecules*, 24, 1661.
208. Xu J., Zhang Y. (2010), "How significant is a protein structure similarity with TM-score = 0.5?", *Bioinformatics*, 26(7), pp. 889-895.
209. Yap C. W. (2011), "PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints", *Journal of Computational Chemistry*, 32(7), pp. 1466-1474.
210. Yin Y., He X., Szewczyk P., Nguyen T., Chang G. (2006), "Structure of the Multidrug Transporter EmrD from *Escherichia coli*", *Science*, 312(5774), pp. 741-744.
211. Yoshida H., Bogaki M., Nakamura S., Ubukata K., Konno M. (1990), "Nucleotide sequence and characterization of the *Staphylococcus aureus* norA gene, which confers resistance to quinolones", *Journal of Bacteriology*, 172(12), pp. 6942-6949.
212. Youden W. J. (1950), "Index for rating diagnostic tests", *Cancer*, 3(1), pp. 32-35.
213. Zdrazil B., Pinto M., Vasanthanathan P., Williams A. J., Balderud L. Z., Engkvist O., Chichester C., Hersey A., Overington J. P., Ecker G. F. (2012), "Annotating Human P-glycoprotein Bioassay Data", *Molecular Informatics*, 31(8), pp. 599-609.
214. Zhang J., Huan J. (2010), "Comparison of chemical descriptors for protein-chemical interaction prediction", *International Journal of Computational Bioscience*, 1(1), pp. 13-21.
215. Zhang L., Ma S. (2010), "Efflux Pump Inhibitors: A Strategy to Combat P-Glycoprotein and the NorA Multidrug Resistance Pump", *ChemMedChem*, 5(6), pp. 811-822.
216. Zhang S., Morris M. E. (2003), "Effects of the Flavonoids Biochanin A, Morin, Phloretin, and Silymarin on P-Glycoprotein-Mediated Transport", *Journal of Pharmacology and Experimental Therapeutics*, 304(3), pp. 1258-1267.
217. Zhang Y. (2008), "I-TASSER server for protein 3D structure prediction", *BMC Bioinformatics*, 9, 40.
218. Zhang Y. (2009), "I-TASSER: fully automated protein structure prediction in CASP8", *Proteins*, 77(Suppl 9), pp. 100-113.

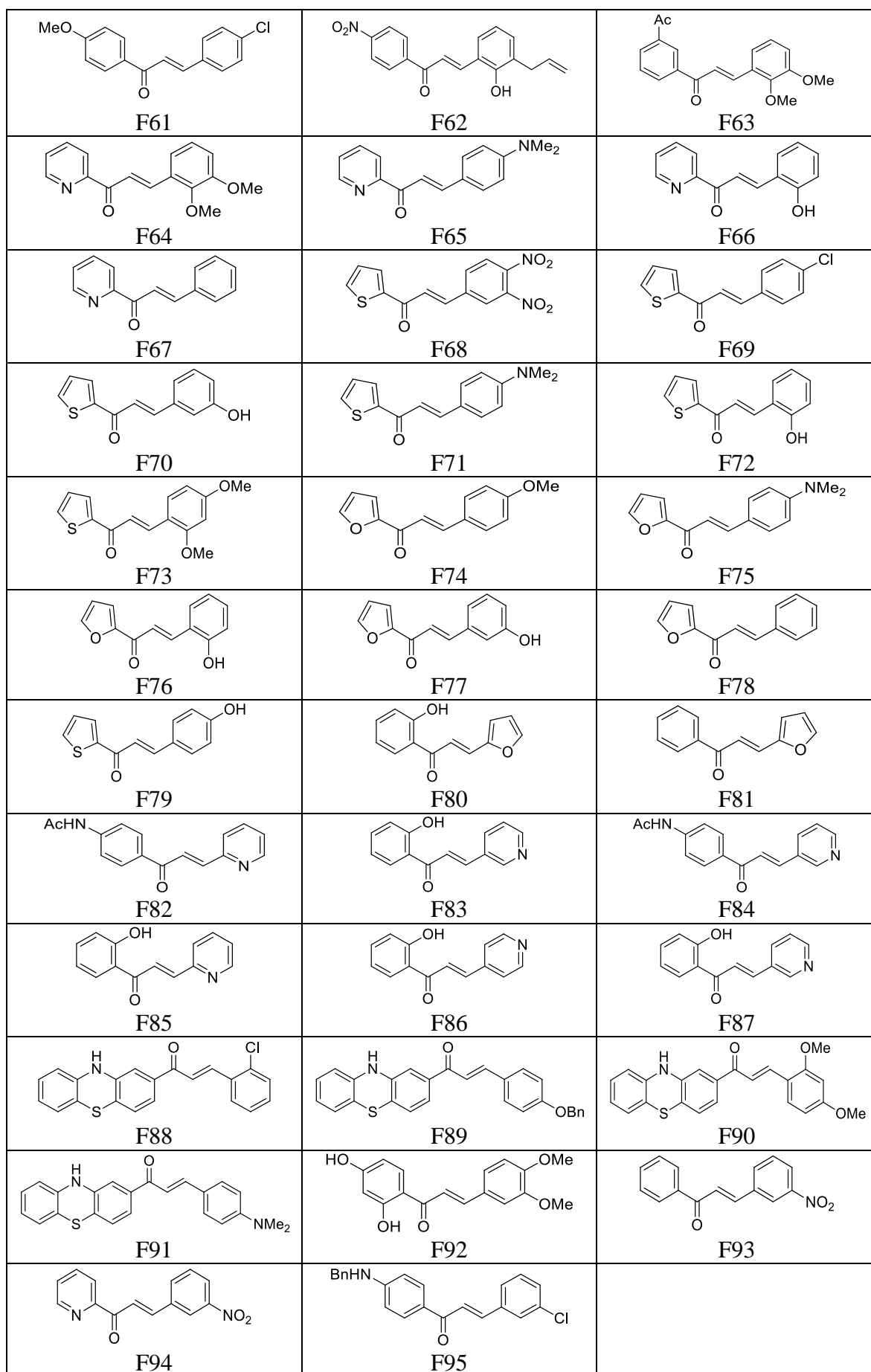
219. Zhou H., Wu S., Zhai S., Liu A., Sun Y., Li R., Zhang Y., Ekins S., Swaan P. W., Fang B., Zhang B., Yan B. (2008), "Design, Synthesis, Cytoselective Toxicity, Structure–Activity Relationships, and Pharmacophore of Thiazolidinone Derivatives Targeting Drug-Resistant Lung Cancer Cells", *Journal of Medicinal Chemistry*, 51(5), pp. 1242-1251.
220. Zhou S. F. (2008), "Structure, function and regulation of P-glycoprotein and its clinical relevance in drug disposition", *Xenobiotica*, 38(7-8), pp. 802-832.

PHỤ LỤC

PHỤ LỤC 1. Cấu trúc 95 chalcon nội bộ.

 F1	 F2	 F3
 F4	 F5	 F6
 F7	 F8	 F9
 F10	 F11	 F12
 F13	 F14	 F15
 F16	 F17	 F18
 F19	 F20	 F21
 F22	 F23	 F24
 F25	 F26	 F27





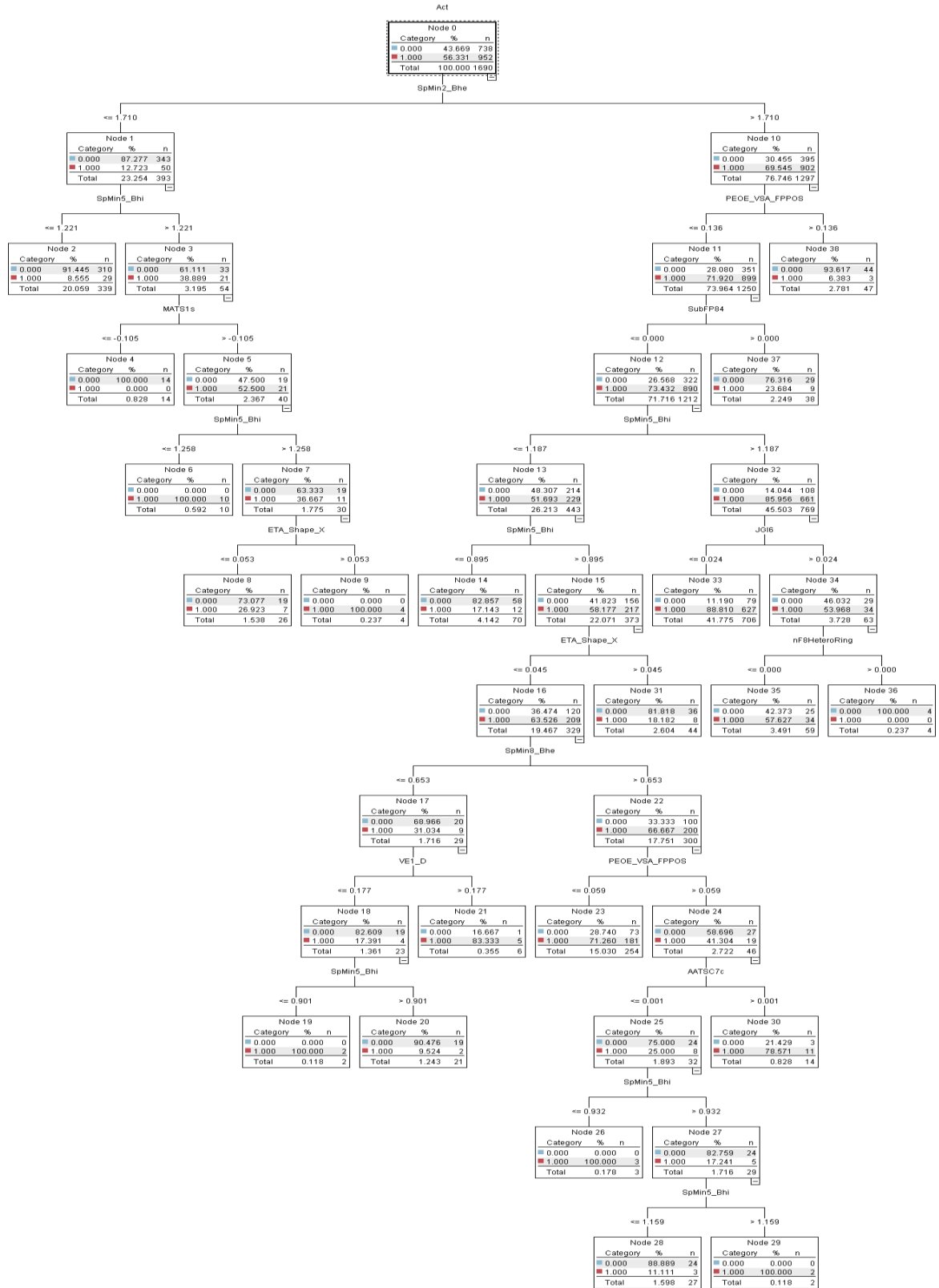
PHỤ LỤC 2. Tập hợp 24 thuộc tính đóng vai trò là biến độc lập trong các mô hình phân loại chất ức chế và chất không ức chế P-gp.

Mã thuộc tính	Loại	Mô tả
PEOE_VSA_FPPOS	Thông số mô tả MOE 2D	Diện tích bề mặt vdW dương, có cực, theo phần
AATSC7c		Tự tương quan Broto-Moreau trung tâm trung bình - lag 7/được trọng số hóa bởi điện tích
MATS1s		Tự tương quan Moran - lag 1/được trọng số hóa bởi trạng thái I
SpMax3_Bhv		Giá trị eigen tuyệt đối lớn nhất của ma trận biến đổi Burden - n 3/được trọng số hóa bởi thể tích van der Waals tương đối
SpMin2_Bhe		Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 2/được trọng số hóa bởi điện âm Sanderson tương đối
SpMin8_Bhe	Thông số mô tả PaDEL 2D	Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 8/được trọng số hóa bởi điện âm Sanderson tương đối
SpMin5_Bhi		Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 5/được trọng số hóa bởi khả năng ion hóa đầu tiên tương đối
ETA_Shape_Xn8HeteroRing		Chỉ số hình dạng X
nF8HeteroRing		Số vòng 8 chứa dị tố (N, O, P, S, hoặc halogen)
JGI6		Số vòng 8 hợp nhất chứa dị tố (N, O, P, S, hoặc halogen)
JGI7		Chỉ số tích điện topo trung bình của thứ tự 6
VE1_D		Chỉ số tích điện topo trung bình của thứ tự 7
		Tổng hệ số của vector eigen sau cùng từ ma trận khoảng cách topo
MACCSFP30	Chuỗi MACCS	CQ(C)(C)A* (Q4 không phải C, được gắn kết >= 3 C)
PubchemFP122		>= 2 vòng cỡ 3 bất kỳ
PubchemFP534	Dấu vân tay	S-C:C-O
PubchemFP741	Pubchem	Oc1cc(S)ccc1
PubchemFP749		Nc1cc(N)ccc1
PubchemFP772		Nc1c(Br)cccc1
SubFP41		1,2-Diol
SubFP48	Dấu vân tay	Aldehyd
SubFP62	dưới cấu trúc	Giống halogen acetal
SubFP84	(substructure)	Acid carboxylic
SubFP157		Semicarbazon

*A: Bất kỳ ký hiệu nguyên tố bảng tuần hoàn hợp lệ nào; Q: Dị tố; bất kỳ nguyên tử không phải C hoặc H.

PHỤ LỤC 3. Các mô hình phân loại được tạo ra từ tập huấn luyện đa dạng.
 (Mô hình mạng nơron và SVM: Tham khảo các thông số trong **Mục 2.2.4**)

Mô hình C5.0



Mô hình hồi quy logistic

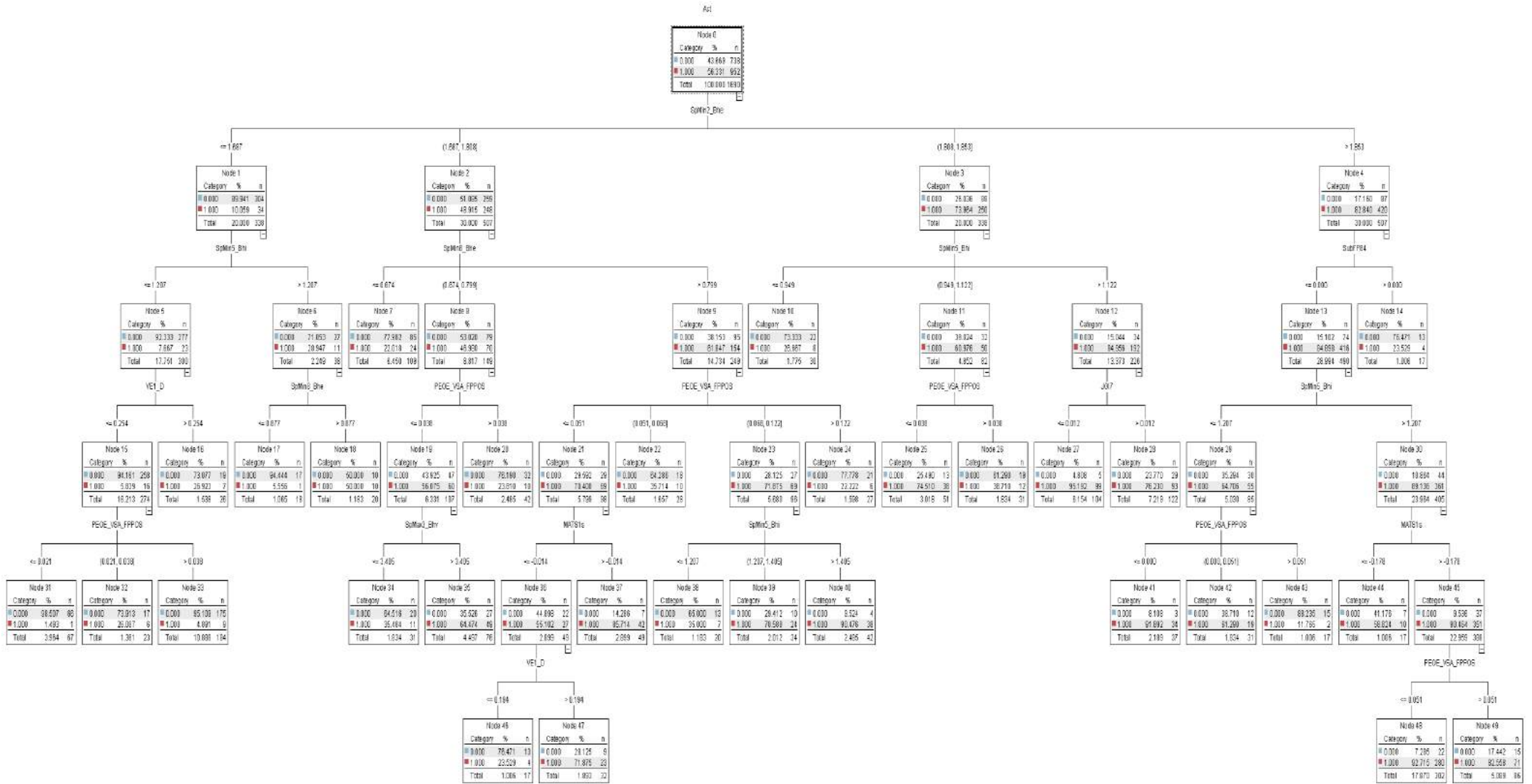
Phương trình cho “0”

+ 0,00000000000000000000

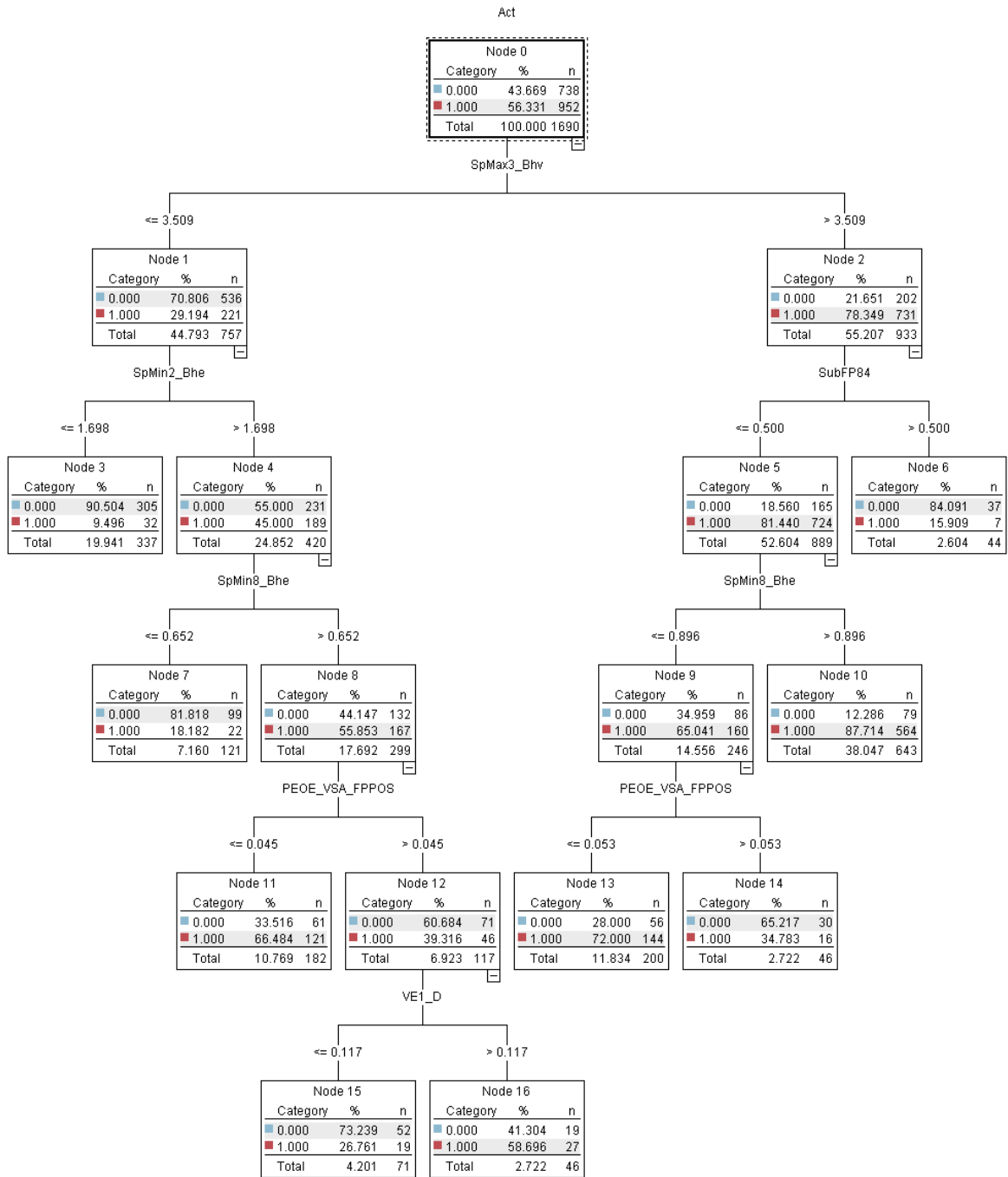
Phương trình cho “1”

115,7*AATSC7c - 5,207*ETA_Shape_X + 6,81*JGI6 + 32,75*JGI7 -
4,006*MACCSFP30 + 1,62*MATS1s - 6,913*PEOE_VSA_FPPOS -
16,33*PubchemFP122 - 17,24*PubchemFP534 - 4,213*PubchemFP741 -
2,503*PubchemFP749 - 18,07*PubchemFP772 + 0,9461*SpMax3_Bhv +
5,928*SpMin2_Bhe + 2,32*SpMin5_Bhi + 1,423*SpMin8_Bhe + 18,64*SubFP157
- 2,177*SubFP41 - 17,88*SubFP48 - 19,4*SubFP62 - 2,539*SubFP84 +
2,614*VE1_D - 19,3*n8HeteroRing - 1,337*nF8HeteroRing - 17,69

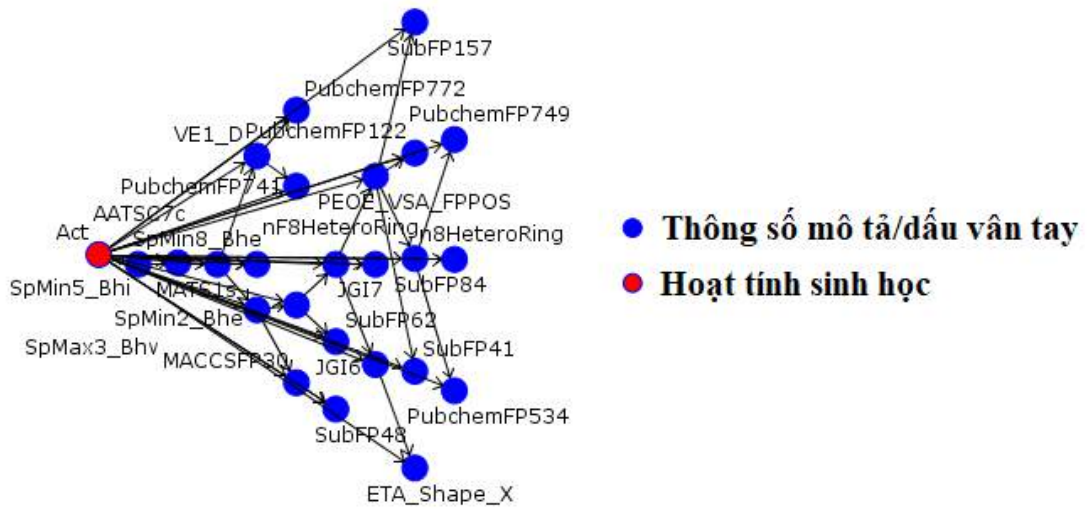
Mô hình CHAID



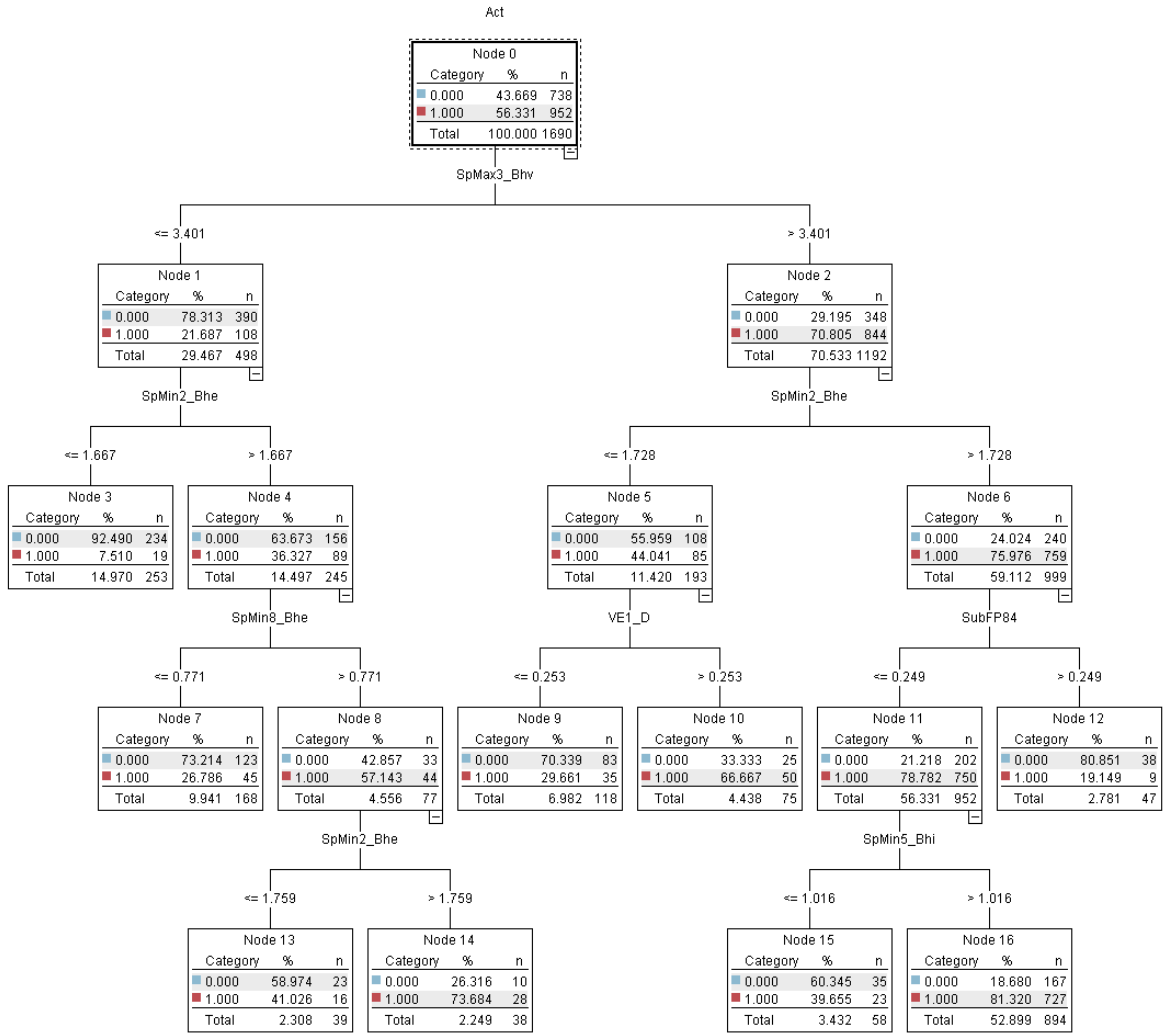
Mô hình C&R Tree



Mô hình mạng Bayesian



Mô hình QUEST



Mô hình mặt nghiêng quyết định

STT	Phân đoạn	Điểm	Độ phủ (n)	Tần số	Xác suất
	Tất cả phân đoạn bao gồm phần còn lại		1690	952	56,33 %
1	SpMin2_Bhe, SpMin5_Bhi SpMin2_Bhe > 1,89 và SpMin5_Bhi > 1,41 và SpMin5_Bhi ≤ 1,48	1	59	59	100 %
2	PEOE_VSA_FPPOS, SpMax3_Bhv PEOE_VSA_FPPOS > 0,02 và PEOE_VSA_FPPOS ≤ 0,03 và SpMax3_Bhv > 3,7	1	57	56	98,25 %
3	SpMin5_Bhi SpMin5_Bhi > 1,41 và SpMin5_Bhi ≤ 1,48	1	99	84	84,85 %
4	SpMax3_Bhv SpMax3_Bhv > 3,78	1	86	71	82,56 %
5	SpMin2_Bhe, AATSC7c SpMin2_Bhe > 1,85 và AATSC7c > 0 và AATSC7c ≤ 0	1	61	56	91,80 %
	Phần còn lại		1328	626	47,14 %

PHỤ LỤC 4. Tập hợp 34 thuộc tính đóng vai trò là biến độc lập trong các mô hình dự đoán hoạt tính ức chế P-gp.

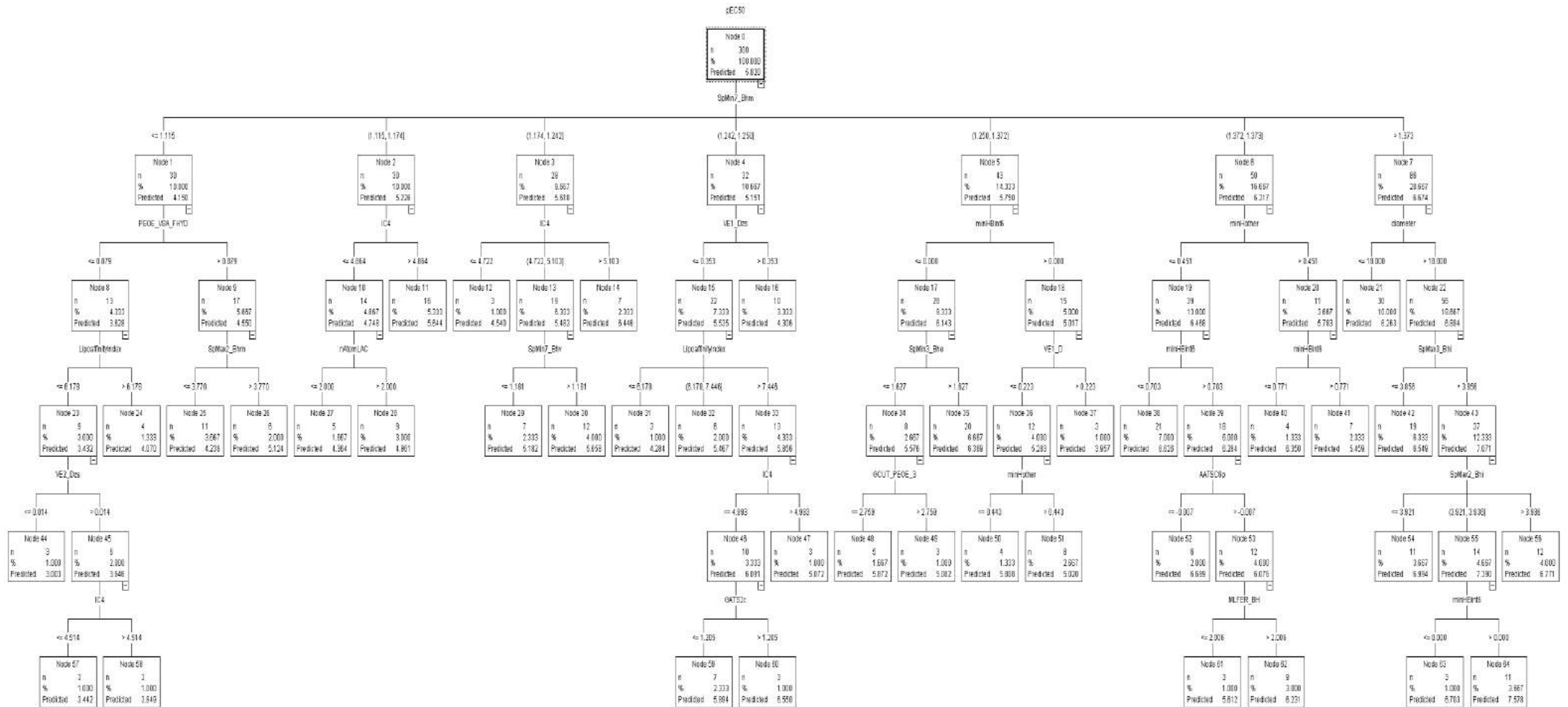
Mã thuộc tính	Loại	Mô tả
diameter		Độ lệch tâm đỉnh lớn nhất trong biểu đồ
GCUT_PEOE_3		PEOE tích điện GCUT (3/3)
a_nCl	Thông số	Số nguyên tử clor
PEOE_RPC+	mô tả	Điện tích từng phần dương tương đối
PEOE_VSA_FHYD	MOE 2D	Diện tích bề mặt vdW kỵ nước, theo phần
PEOE_VSA_FPNEG		Diện tích bề mặt vdW âm, có cực, theo phần
PEOE_VSA_FPPOS		Diện tích bề mặt vdW dương, có cực, theo phần
AATSC6p		Tự tương quan Broto-Moreau trung tâm trung bình - lag 6/được trọng số hóa bởi khả năng phân cực
MATS8e		Tự tương quan Moran - lag 8/được trọng số hóa bởi điện âm Sanderson
MATS8s		Tự tương quan Moran - lag 8/được trọng số hóa bởi trạng thái I
GATS2c		Tự tương quan Geary - lag 2/được trọng số hóa bởi điện tích
VE2_DzZ		Tổng hệ số trung bình của vector eigen sau cùng từ ma trận Barysz/được trọng số hóa bởi số nguyên tử
VE1_Dzi		Tổng hệ số của vector eigen sau cùng từ ma trận Barysz/được trọng số hóa bởi khả năng ion hóa đầu tiên
VE1_Dzs		Tổng hệ số của vector eigen sau cùng từ ma trận Barysz/được trọng số hóa bởi trạng thái I
VE2_Dzs		Tổng hệ số trung bình của vector eigen sau cùng từ ma trận Barysz/được trọng số hóa bởi trạng thái I
SpMax2_Bhm		Giá trị eigen tuyệt đối lớn nhất của ma trận biến đổi Burden - n 2/được trọng số hóa bởi số khối tương đối
SpMin6_Bhm		Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 6/được trọng số hóa bởi số khối tương đối
SpMin7_Bhm	Thông số mô tả	Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 7/được trọng số hóa bởi số khối tương đối
SpMax2_Bhv	PaDEL 2D	Giá trị eigen tuyệt đối lớn nhất của ma trận biến đổi Burden - n 2/được trọng số hóa bởi thể tích van der Waals tương đối
SpMin7_Bhv		Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 7/được trọng số hóa bởi thể tích van der Waals tương đối
SpMin2_Bhe		Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 2/được trọng số hóa bởi điện âm Sanderson tương đối
SpMin3_Bhe		Giá trị eigen tuyệt đối nhỏ nhất của ma trận biến đổi Burden - n 3/được trọng số hóa bởi điện âm Sanderson tương đối
SpMax2_Bhi		Giá trị eigen tuyệt đối lớn nhất của ma trận biến đổi Burden - n 2/được trọng số hóa bởi khả năng ion hóa đầu tiên tương đối
SpMax3_Bhi		Giá trị eigen tuyệt đối lớn nhất của ma trận biến đổi Burden - n 3/được trọng số hóa bởi khả năng ion hóa đầu tiên tương đối
minHBint6		Các thông số mô tả trạng thái E tối thiểu của độ mạnh cho các liên kết hydro tiềm năng của độ dài đường 6
minHother		Trạng thái E loại nguyên tử H tối thiểu: H trên aaCH, dCH2 hoặc dsCH
maxaaCH		Trạng thái E loại nguyên tử tối đa: :CH:
LipoaffinityIndex		Chỉ số ái lực thân dầu
IC4		Chỉ số nội dung thông tin (đôi xứng lân cận của thứ tự 4)

Mã thuộc tính	Loại	Mô tả
nAtomLAC		Số nguyên tử trong chuỗi béo dài nhất
MLFER_BH	Thông số	Tính kiểm liên kết hydro hòa tan tổng thể hoặc tổng cộng
n5HeteroRing	mô tả	Số vòng 5 chứa dị tố (N, O, P, S, hoặc halogen)
GGI9	PaDEL	Chỉ số tích điện topo của thứ tự 9
VE1_D	2D	Tổng hệ số của vector eigen sau cùng từ ma trận khoảng cách topo

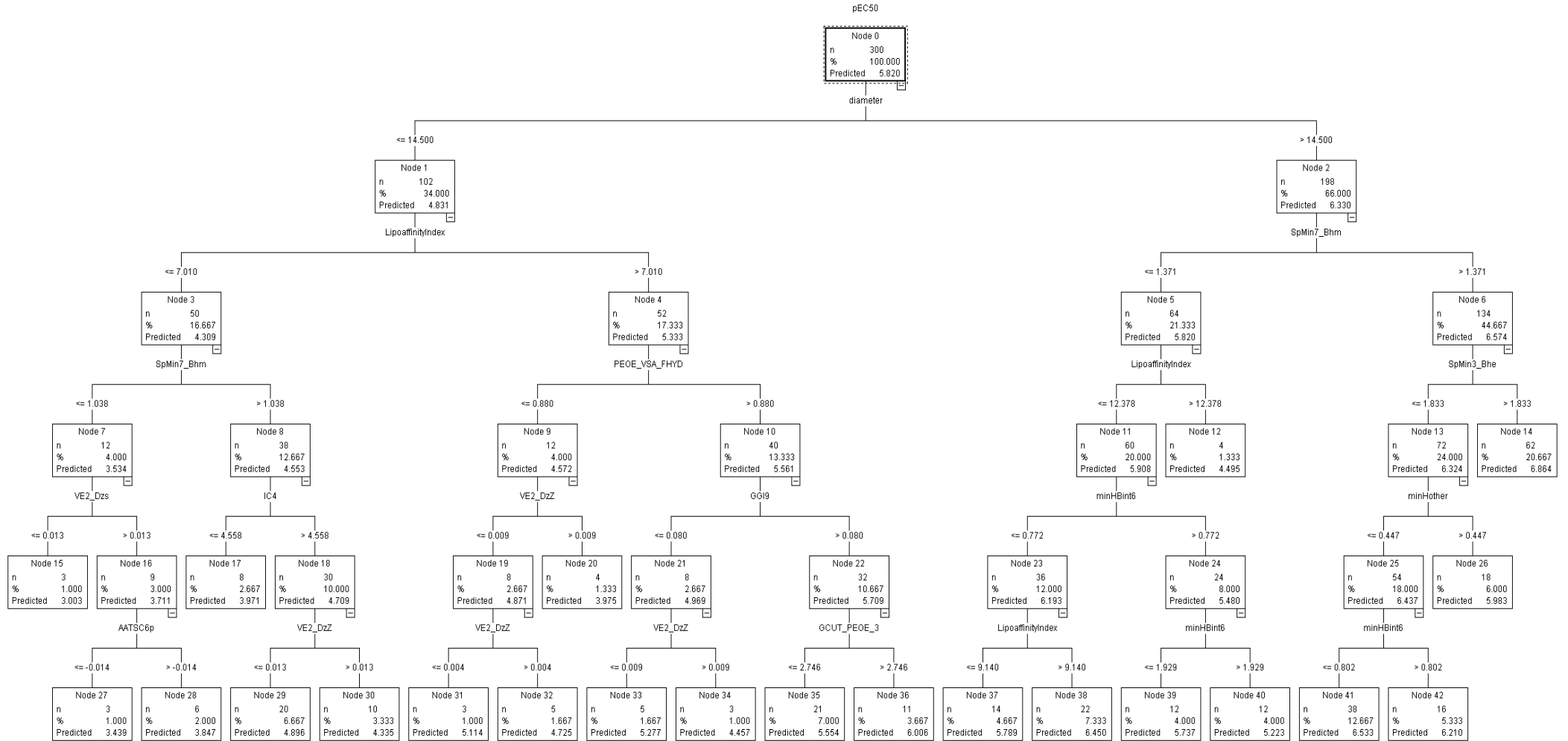
PHỤ LỤC 5. Các mô hình dự đoán được tạo ra từ tập huấn luyện đa dạng.

(Mô hình mạng nơ-ron và SVM: Tham khảo các thông số trong Mục 2.2.4)

Mô hình CHAID



Mô hình C&R Tree



Mô hình hồi quy

$$\begin{aligned} pIC_{50} = & 0,06205 * diameter + 0,4806 * GCUT_PEOE_3 + 0,1267 * a_nCl + \\ & 1,69 * PEOE_RPC + 8,605 * PEOE_VSA_FHYD + 8,454 * PEOE_VSA_FPPOS + \\ & 1,108 * AATSC6p + 0,5715 * MATS8e + 0,1708 * MATS8s + 0,8089 * GATS2c + \\ & 21,67 * VE2_DzZ - 1,046 * VE1_Dzi - 1,139 * VE1_Dzs + 28,5 * VE2_Dzs + \\ & 1,91 * SpMax2_Bhm + 0,9249 * SpMin6_Bhm + 1,739 * SpMin7_Bhm - \\ & 0,4505 * SpMax2_Bhv - 1,427 * SpMin7_Bhv + 0,7218 * SpMin2_Bhe - \\ & 0,01277 * SpMin3_Bhe - 0,622 * SpMax2_Bhi + 2,18 * SpMax3_Bhi - \\ & 0,1878 * minHBint6 - 0,8197 * minHother + 0,08064 * maxaaCH - \\ & 0,113 * LipoaffinityIndex + 0,3266 * IC4 + 0,1595 * nAtomLAC - \\ & 0,1851 * MLFER_BH - 0,284 * n5HeteroRing + 2,085 * GGI9 + 0,1197 * VE1_D - \\ & 20,78 \end{aligned}$$

PHỤ LỤC 6. Tập hợp 21 thuộc tính đóng vai trò là biến trong bản đồ nhận thức.

Mã thuộc tính	Loại	Mô tả	
diameter (dia)	Thông số mô tả MOE 2D	Độ lệch tâm đỉnh lớn nhất trong biểu đồ	
BCUT_PEOE_2 (BP2)		PEOE tích điện BCUT (2/3)	
GCUT_PEOE_2 (GP2)		PEOE tích điện GCUT (2/3)	
balabanJ (bJ)		Khả năng kết nối tổng khoảng cách trung bình Balaban	
Q_VSA_FNEG (QVF)		Diện tích bề mặt vdW âm, theo phần	
ATSC2m (A2m)	Thông số mô tả PaDEL 2D	Tự tương quan Broto-Moreau trung tâm - lag 2/được trọng số hóa bởi số khối	
ATSC4m (A4m)		Tự tương quan Broto-Moreau trung tâm - lag 4/được trọng số hóa bởi số khối	
ATSC1s (A1s)		Tự tương quan Broto-Moreau trung tâm - lag 1/được trọng số hóa bởi trạng thái I	
AATSC6v (AA6v)		Tự tương quan Broto-Moreau trung tâm trung bình - lag 6/được trọng số hóa bởi thể tích van der Waals	
AATSC4s (AA4s)		Tự tương quan Broto-Moreau trung tâm trung bình - lag 4/được trọng số hóa bởi trạng thái I	
MATS4s (M4s)		Tự tương quan Moran - lag 4/được trọng số hóa bởi trạng thái I	
SpMAD_DzZ (SpM)		Độ lệch phổ tuyệt đối trung bình từ ma trận Barysz/được trọng số hóa bởi số nguyên tử	
ASP-3 (ASP3)		Đường dẫn đơn giản trung bình, thứ tự 3	
AVP-6 (AVP6)		Đường dẫn hóa trị trung bình, thứ tự 6	
nHCsat (nHCs)		Tính trạng thái E loại nguyên tử H: H trên C sp ³ gắn với C chưa bão hòa	
minHCsat (minHCs)		Trạng thái E loại nguyên tử H tối thiểu: H trên C sp ³ gắn với C chưa bão hòa	
ETA_BetaP_ns_d (EBPnsd)		Đo các electron đơn lẻ đi vào cộng hưởng liên quan đến kích thước phân tử	
MDEO-22 (MDEO22)		Ngưỡng khoảng cách phân tử giữa tất cả nguyên tử oxy thứ yếu	
MACCSFP128 (MFP128)		Chuỗi MACCS	ACH2AAACH2A* (các nhóm CH ₂ được phân cách bởi 4 liên kết)
MACCSFP144 (MFP144)		MACCS	Anot%A% Anot%A* (các nguyên tử được phân cách bởi (!):(!:))
PubchemFP2 (PFP2)	Dấu vân tay Pubchem	>= 16 H	

*A: Bất kỳ ký hiệu nguyên tố bảng tuần hoàn hợp lệ nào.

PHỤ LỤC 7. Giá trị trung bình của 18 thông số mô tả được tính cho 4 lớp hoạt tính khi xây dựng bản đồ MDS.

	P	A	D	N
diameter	0,21	0,38	0,50	0,35
BCUT_PEOE_2	0,48	0,78	0,51	0,77
GCUT_PEOE_2	0,44	0,68	0,48	0,71
balabanJ	0,51	0,35	0,32	0,39
Q_VSA_FNEG	0,49	0,77	0,45	0,68
ATSC2m	0,57	0,51	0,38	0,30
ATSC4m	0,58	0,51	0,39	0,54
ATSC1s	0,65	0,54	0,81	0,59
AATSC6v	0,27	0,32	0,47	0,51
AATSC4s	0,52	0,39	0,37	0,37
MATS4s	0,57	0,38	0,35	0,40
SpMAD_DzZ	0,18	0,29	0,56	0,28
ASP-3	0,41	0,52	0,58	0,57
AVP-6	0,39	0,50	0,59	0,39
nHCsatu	0,37	0,00	0,21	0,00
minHCsatu	0,38	0,00	0,40	0,00
ETA_BetaP_ns_d	0,34	0,20	0,36	0,23
MDEO-22	0,18	0,00	0,14	0,00

PHỤ LỤC 8. Tần số của 3 dấu vân tay được tính cho 4 lớp hoạt tính khi xây dựng bản đồ CA.

	P	A	D	N
MACCSFP128	3	1	16	1
MACCSFP144	16	1	17	4
PubchemFP2	13	7	19	7

PHỤ LỤC 9. Các thông số đặc trưng của 02 chalcon “lead” F88 và F90.

Các thông số đặc trưng của F88

Tên IUPAC: (E)-3-(2-Chlorophenyl)-1-(10H-phenothiazin-2-yl)prop-2-en-1-on.

Hiệu suất: 57 %.

Nhiệt độ nóng chảy: 202 - 203 °C.

EIMS m/z: 386,0331 [M + Na]⁺.

UV (λ_{max} nm, MeOH): 204, 248, 309, 449.

IR (KBr) cm^{-1} : 3354, 1654, 1590, 754.

¹H-NMR (DMSO) δ ppm: 8,79 (s, 1H, NH); 8,16 (d, $J_{3''-4''} = 7,5$ Hz, 1H, H3''); 8,00 (d, $J_{3-2} = 15,5$ Hz, 1H, H3); 7,84 (d, $J_{2-3} = 15,5$ Hz, 1H, H2); 7,64 (d, $J_{3'-4'} = 8$ Hz, 1H, H3'); 7,57 (d, $J_{6''-5''} = 7,5$ Hz, 1H, H6''); 7,46 (m, 2H, H4'', H5''); 7,30 (s, 1H, H1'); 7,09 (d, $J_{4'-3'} = 8$ Hz, 1H, H4'); 7,01 (m, 1H, H8'); 6,92 (d, $J_{6'-7'} = 7,5$ Hz, 1H, H6'); 6,66 (m, 1H, H7'); 6,66 (d, $J_{9'-8'} = 8$ Hz, 1H, H9').

¹³C-NMR (DMSO) δ ppm: 187,7 (C1=O); 142,1 (C3); 141,0 (C13'); 138,2 (C11'); 136,5 (C2'); 134,2 (C2''); 132,2 (C1''); 131,9 (C3''); 130 (C4''); 128,4 (C8'); 127,9 (C6''); 127,6 (C5'); 126,2 (C5''); 126,1 (C4'); 124,5 (C7'); 124,4 (C2); 122,7 (C3'); 115,1 (C12'); 114,5 (C14'); 112,8 (C9').

Các thông số đặc trưng của F90

Tên IUPAC: (E)-3-(2,4-Dimethoxyphenyl)-1-(10H-phenothiazin-2-yl)prop-2-en-1-on.

Hiệu suất: 49 %.

Nhiệt độ nóng chảy: 178 - 179 °C.

EIMS m/z: 388,1154 [M - H]⁻.

UV (λ_{max} nm, MeOH): 204, 248, 304, 365.

IR (KBr) cm^{-1} : 3310, 1643, 1570, 1272, 1150.

¹H-NMR (DMSO) δ ppm: 8,76 (s, 1H, NH); 7,95 (d, $J_{3-2} = 15,5$ Hz, 1H, H3); 7,86 (d, $J_{6''-5''} = 8,5$ Hz, 1H, H6''); 7,61 (d, $J_{2-3} = 15,5$ Hz, 1H, H2); 7,52 (dd, $J_{3'-4'} = 8$ Hz, $J_{3'-1'} = 1,5$ Hz, 1H, H3'); 7,28 (d, $J_{1'-3'} = 1,5$ Hz, 1H, H1'); 7,06 (d, $J_{4'-3'} = 8$ Hz, 1H, H4'); 7,01 (m, 1H, H8'); 6,92 (d, $J_{6'-7'} = 7$ Hz, 1H, H6'); 6,80 (m, 1H, H7'); 6,64 (m, 3H, H9', H3'', H5''); 3,90 (s, 3H, 2''-OMe); 3,84 (s, 3H, 4'-OMe).

¹³C-NMR (DMSO) δ ppm: 187,9 (C1=O); 163,1 (C4''); 160 (C2''); 142,1 (C13'); 141,2 (C11'); 138,6 (C3); 137,3 (C2'); 130,1 (C6''); 127,9 (C8'); 126,2 (C6'); 126,1 (C4'); 123,0 (C7'); 122,1 (C2); 122,0 (C3'); 118,9 (C1''); 115,9 (C1'); 115,2 (C12'); 114,5 (C14'); 112,9 (C9'); 106,3 (C5''); 98,3 (C3''); 55,8 (2'-MeO); 55,5 (4'-CH₃O).