

# THIẾT KẾ THUỐC HỢP LÝ TRONG NGHIÊN CỨU TÁC DỤNG ỨC CHẾ TOPOISOMERASE-I CỦA CÁC CHẤT TƯƠNG ĐỒNG BENZO[C]PHENATHRIDIN

Thái Khắc Minh\*, Huỳnh Thị Ngọc Phương\*, Nguyễn Thúy Quyên\*, Đỗ Thị Ngọc Mai\*,  
Bùi Quang Huỳnh\*

## TÓM TẮT

**Mở đầu:** Trong những năm gần đây các enzym topoisomerase, enzym cần thiết cho hoạt động sao chép, phiên mã và tái tổ hợp ADN, trở thành một trong những mục tiêu đầy hứa hẹn trong việc nghiên cứu tổng hợp các thuốc kháng ung thư mới. Trong số những hợp chất thể hiện hoạt tính ức chế topoisomerase, nhiều chất tương đồng tổng hợp của họ alkaloid benzo[c]phenanthridin (BCP) được nghiên cứu. Trong đó, các chất ethoxidin, NK-109 và topoval (ARC 111) là những chất có tiềm năng sử dụng trong hóa trị liệu ung thư.

**Mục tiêu:** Nghiên cứu khả năng gắn kết của các chất tương đồng BCP với phức hợp TOP-I:ADN và xây dựng mô hình phân loại các chất tương đồng BCP có khả năng ức chế TOP-I bằng phương pháp máy vector hỗ trợ SVM.

**Phương pháp:** Phương pháp mô hình mô tả phân tử docking và phương pháp máy học SVM được sử dụng trên cơ sở dữ liệu các chất tương đồng BCP với hoạt tính ức chế TOP-I.

**Kết quả:** Sử dụng chương trình mô tả phân tử DOCK 6.2, khả năng gắn kết ở mức độ phân tử của các dẫn chất BCP với phức hợp DNA:TOP-I được phân tích. Hệ thống vòng của tất cả dẫn chất BCP có khả năng tạo liên kết  $\pi$ - $\pi$  với Guanin G11 của ADN và nằm ở vị trí song song và xen vào giữa các cặp base G11/C112 và A113/T10 của cấu trúc ADN và hạn chế khả năng tạo thành phức hợp giữa DNA và TOP-I. Kết quả docking cho thấy các chất tương đồng BCP có xu hướng gắn kết với DNA hơn là TOP-I. Phương pháp máy vector hỗ trợ SVM được sử dụng với mục đích xây dựng mô hình dự đoán và phân loại 73 chất tương đồng BCP trên hoạt tính kháng TOP-I. Mô hình SVM tốt nhất được xây dựng bằng gói SVM - e1071 trong R với thông số tối ưu của hàm kernel ( $C=4$ ,  $\gamma=0,25$ ) và thông số mô tả xác định bằng phương pháp rừng ngẫu nhiên RF. Mô hình SVM này có khả năng dự đoán đúng 93% các chất trong tập huấn luyện và 87% các chất trong tập hợp kiểm tra. Ngoài ra, mô hình còn được đánh giá lại trên tập hợp thử ngoại (không dùng để xây dựng mô hình) cho độ đúng 89% (dự đoán đúng 8/9 chất) và tập hợp ứng dụng cho độ đúng 80% (dự đoán đúng 8/10 chất). Mô hình SVM được xây dựng chứng tỏ khả năng phân loại chính xác chất có hoạt tính (80-100% tùy tập hợp).

**Kết luận:** Mô hình docking cho thấy các chất tương đồng BCP có xu hướng gắn kết mạnh với ADN hơn là TOP-I và cần thực hiện các thử nghiệm xác định mức độ gắn kết với ADN và enzym riêng rẽ để khẳng định lại kết quả này. Mô hình phân loại SVM được áp dụng trong sàng lọc, phân loại các chất tương đồng BCP dựa trên tác dụng sinh học của các chất tương đồng BCP đã và sắp tổng hợp. Từ nghiên cứu này, các cấu trúc BCP mới sẽ được thiết kế, tổng hợp và thử tác dụng với mục tiêu tìm kiếm chất có tác dụng ức chế TOP-I và kháng ung thư mạnh.

**Từ khóa:** Kháng ung thư, topoisomerase, benzo[c]phenanthridin, thiết kế thuốc, docking, máy vector hỗ trợ, SVM, phân loại.

---

\* Bộ môn Hóa Dược – Khoa Dược - Đại học Y Dược Thành phố Hồ Chí Minh

Địa chỉ liên hệ: TS.DS. Thái Khắc Minh

ĐT: 0909 680 385

Email: thaikhacminh@gmail.com

## ABSTRACT

### RATIONAL DRUG DESIGN OF BENZO[C]PHENANTHRIDINE DERIVATIVES WITH TOPOISOMERASE-I INHIBITORY ACTIVITY

Khac Minh Thai, Huynh Thi Ngoc Phuong, Nguyen Thuy Quyen, Do Thi Ngoc Mai,  
Bui Quang Huynh \* Y Hoc TP. Ho Chi Minh \* Vol. 14 - Supplement of No 1 - 2010: 6-14

**Background:** For recent years, the topoisomerases are promising antitumor drug targets for design new anticancer novels. Among the compounds expressing a topoisomerase-targeting activity, the alkaloid family of benzo[c]phenanthridine (BCP) is well known including nitidine, fagaronine, chelerythrine, sangunarine... Ethoxidine, NK-109 and topoival (ARC 111), BCP synthetic analogues, are expressed as potential agents using in cancer chemotherapy.

**Objective:** Aims of this research are the molecular modeling study of the interaction between BCP analogues and DNA-TOP-I complexes as well as design the support vector machine SVM model for classification BCP compounds with topoisomerase-I inhibitory property.

**Method:** Both ligand-based and structure-based approaches are applied on a series of BCP with antitopoisomerase activity. For the ligand-based design, the machine learning method namely support vector machine (SVM) is used to classify and to predict BCP compounds.

**Results:** Based on crystal structure of DNA:TOP-I complex, 53 BCP derivatives were successfully docked into the complex by DOCK 6.2 program. Docked BCP:DNA:TOP-I complexes indicated that BCPs tend to establish strong interaction with DNA more than TOP-I. The interaction between BCP analogues and DNA:TOP-I complex was also analyzed. The best SVM model based on 73 BCP analogues was built from the SVM-e1071 package in R with the optimal settings of the Kernel ( $C = 4$ ,  $\gamma = 0.25$ ). The final SVM model with total accuracy of 93% for training set of 58 compounds was archived using a set of 7 descriptors identified out of a large set via a random forest algorithm. Moreover, the power for SVM classifier was validated internally by a test set of 15 compounds. This SVM model gained the overall accuracy up to 87% and the Matthews correlation coefficient (MCC) of 0.71. For two external test sets, 89% and 80% BCP compounds, respectively, were correctly predicted. The SVM model has proved its ability to classify correctly BCP analogues that have a positive activity, with an accuracy from 80 to 100% overall.

**Conclusion:** These *in silico* models including molecular modeling and SVM classification could be applied to search and to design the new analogues of BCPs which express highly topoisomerase I inhibitory activity.

**Keywords:** anticancer, topoisomerase, benzo[c] phenanthridine, drug design, docking, support vector machine, SVM, classification

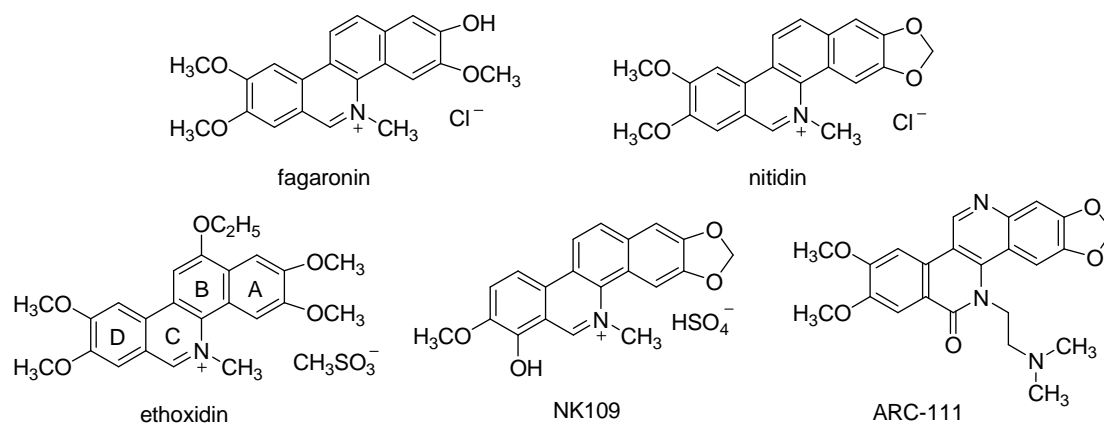
## ĐẶT VẤN ĐỀ

Sự xuất hiện các dòng tế bào ung thư kháng thuốc thúc đẩy các nhà khoa học tìm kiếm không ngừng những tác nhân kháng ung thư mới. Trong những năm gần đây, topoisomerase-ADN (TOP-ADN) - một enzym cần thiết cho các hoạt động sao chép, phiên mã và tái tổ hợp ADN bình thường của tế bào - trở thành một trong những mục tiêu đầy hứa hẹn cho việc tìm kiếm và phát triển các thuốc kháng ung thư (1,10,15). Nhiều thuốc nguồn

gốc tổng hợp kháng ung thư được sử dụng hiện nay có cơ chế kháng topoisomerase-2 như amsacrin, asulacrin, mitoxantron, loxoxantron, piroxantron..., hoặc kháng topoisomerase-1 (TOP-1) như topotecan, irrinotecan...(4,5,6) Trong số những hợp chất thể hiện tác động kháng TOP-ADN, nhóm alkaloid benzo[c]phenanthridin (BCP) được biết đến nhiều. Nhiều công trình nghiên cứu được thực hiện để tìm những con đường tổng hợp BCP thiên nhiên (fagaronin, nitidin... Hình 1) đáp

ứng cho việc nghiên cứu cơ chế tác động ở mức độ phân tử của chúng (6,14). Bên cạnh đó các chất tương đồng BCP cũng được tổng hợp và nghiên cứu độc tính tế bào cũng như khả năng ức chế TOP-ADN. Trong số những chất tương đồng tổng hợp này, ARC-111 và NK109 (Hình 1) hiện là hai ứng viên tiềm năng sử dụng trong hóa trị liệu ung thư (12). Bên cạnh đó ethoxidin, một dẫn chất 12-ethoxy BCP, thể hiện độc tính tế bào có ý nghĩa trong thử nghiệm *in vitro* trên những dòng tế bào ung thư máu (leukemia) ở người. Những thử nghiệm sinh hóa cũng cho thấy ethoxidin có ái lực gắn kết với ADN và khả năng ức chế TOP-1 đáng kể, trong đó vai trò của nhóm 12-ethoxy được cho là cần thiết trong tác động ức chế TOP-1 của phân tử này (6). Các nghiên cứu về mối quan hệ cấu trúc của các chất tương đồng BCP và độc tính tế bào cũng cho thấy chức amid lactam đóng vai trò quan trọng vì có thể làm tăng sinh khả dụng của các hợp chất này (1).

Với mục tiêu tìm hiểu thêm về cơ chế tác động ở mức độ phân tử cũng như thiết kế và tổng hợp ra các chất tương đồng BCP chống ung thư mạnh, các phương pháp của thiết kế thuốc hợp lý được ứng dụng nghiên cứu các chất tương đồng BCP và hoạt tính kháng topoisomerase-I (TOP-I). Nghiên cứu kết hợp 2 phương pháp (i) phương pháp dựa vào cấu trúc mục tiêu tác động là nghiên cứu mô hình mô tả phân tử docking và (ii) phương pháp dựa vào cấu tử ligand là phân loại các chất có hoạt tính ức chế TOP-I bằng máy vector hỗ trợ SVM. Kết quả từ nghiên cứu này có thể ứng dụng để dự đoán hoạt tính sinh học của các chất tương đồng BCP và định hướng tổng hợp các chất tương đồng BCP có tác dụng kháng ung thư mạnh. Từ đó, các chất tương đồng BCP được lựa chọn để tổng hợp và thử nghiệm độc tính tế bào cũng như khả năng ức chế enzym TOP-I.



**Hình 1.** Cấu trúc hóa học của các dẫn chất và chất tương đồng BCP có hoạt tính ức chế TOP-I

## ĐỐI TƯỢNG - PHƯƠNG PHÁP NGHIÊN CỨU

### Cơ sở dữ liệu

Tổng cộng 82 chất tương đồng BCP với hoạt tính trên TOP-I được thu thập từ các công trình nghiên cứu của La Voie và cộng sự (8, 10, 14, 18, 19). Hoạt tính trên TOP-I của các chất này được đánh giá dựa vào chỉ số thể hiện sự ức chế quá trình tách ADN gián tiếp thông qua TOP-I (TOP-I-mediated DNA cleavage). Chỉ số hoạt tính trên TOP-I được biểu hiện bằng nồng độ hiệu quả

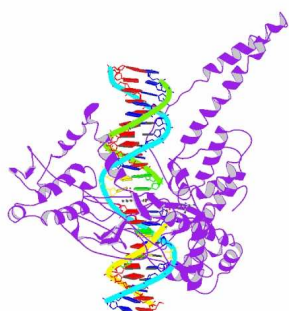
trương đổi (relative effective concentration - REC) giữa dẫn chất thử nghiệm và topotecan (một chất được chứng minh là có hiệu quả trên TOP-I). Chỉ số hoạt tính trên TOP-I của topotecan được ấn định là 1 (10, 14). Dựa trên hoạt tính TOP-I, các chất được phân làm 2 nhóm hoạt tính là mạnh hơn và yếu hơn topotecan. Các chất có hoạt tính mạnh hơn so với topotecan cho giá trị chỉ số hoạt tính trên TOP-I nhỏ hơn 1, và ngược lại, dẫn chất có hoạt tính yếu hơn cho giá trị lớn hơn 1. Đồng thời, các giá trị hoạt tính ức chế

TOP-I cũng như các giá trị độc tính tế bào ( $IC_{50}$ ) trên dòng tế bào ung thư RPMI8402 của 53 chất tương đồng BCP được thu thập (10,14).

### Phương pháp dựa vào cấu trúc mục tiêu tác động – Mô hình mô tả phân tử docking

Docking là phương pháp thiết kế thuốc dựa vào cấu trúc mục tiêu tác động, nghiên cứu khả năng gắn kết của một hay nhiều phân tử thuốc (hay cấu tử, ligand) lên trên mục tiêu tác động (thụ thể, enzym hay các protein, ADN) trong không gian 3 chiều. Chương trình DOCK 6.2 được sử dụng để tiến hành thí nghiệm (2).

Cấu trúc tinh thể tia X của ADN:TOP1:Indenoisoquinolein (pdb 1SC7) trình bày ở Hình 2 có độ phân giải là 3Å được tải về máy tính từ ngân hàng cơ sở dữ liệu protein (protein data bank <http://www.pdb.org>) (16, 17). Điểm tác động trên phức hợp ADN:TOP1 (pdb 1SC7) được xác định là tất cả các acid amin và acid nucleic trong phạm vi không gian 6Å tính từ trung tâm phân tử indenoisoquinolein(16).



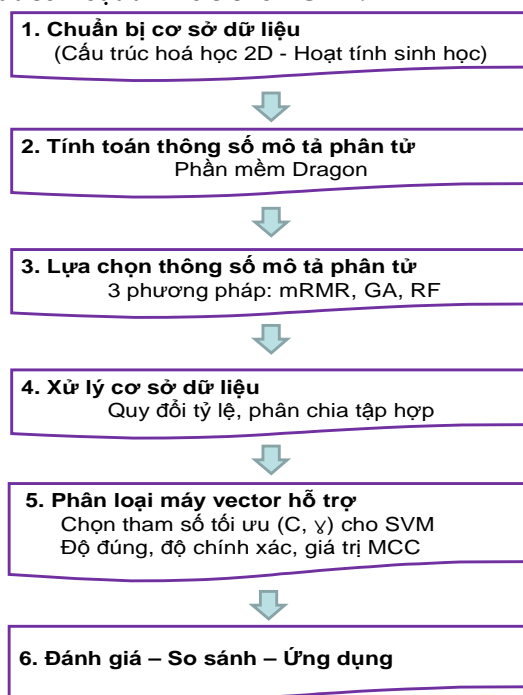
Hình 2. Cấu trúc tinh thể phức hợp ADN:TOP1:Indenoisoquinolein MJ-II-38 (pdb 1SC7)

### Phương pháp dựa vào cấu tử ligand – Phân loại các chất có hoạt tính ức chế TOP-I bằng phương pháp máy học SVM

Máy vector hỗ trợ SVM là phương pháp máy học có sự giám sát được sử dụng trong dự đoán phân loại và định lượng (13, 17). SVM giải quyết cả hai trường hợp, tuyến tính và không tuyến tính bằng một hàm kernel, cũng như các phương pháp thống kê khác, hàm này phụ thuộc vào một số tham số mà khi giải quyết được các tham số đó xem như đã cơ bản thực

hiện được sự phân loại. Hàm kernel hiện diện 4 loại cơ bản (9, 13, 21), mỗi loại tương ứng với những tham số khác nhau, trong nghiên cứu sử dụng hàm RBF kernel, với cặp tham số ( $C$  và  $\gamma$ ).

Trình tự tiến hành xây dựng mô hình phân loại SVM (7, 21) trình bày ở Hình 3 bao gồm các bước (i) Chuẩn bị cơ sở dữ liệu (cấu trúc và hoạt tính sinh học), (ii) Tính toán thông số mô tả phân tử (sử dụng phần mềm Dragon), (iii) Lựa chọn thông số mô tả phân tử, (iv) Xử lý cơ sở dữ liệu (tập hợp huấn luyện, kiểm tra), (v) Xây dựng mô hình phân loại SVM (xác định thông số hàm Kernel tối ưu bao gồm 2 giá trị  $[C, \gamma]$ , độ đúng, độ chính xác, đánh giá chéo Leave-one-out, hệ số tương quan Matthews MCC), (vi) Đánh giá lại mô hình trên tập hợp kiểm tra, tập hợp bên ngoài không dùng xây dựng mô hình, (vii) Ứng dụng mô hình phân loại các chất có hoạt tính ức chế TOP-I.



Hình 3. Tiến trình xây dựng mô hình phân loại máy vector hỗ trợ

## KẾT QUẢ VÀ BÀN LUẬN

### Mô hình mô tả phân tử docking

Khả năng gắn kết của BCP trên TOP-I ở mức độ phân tử

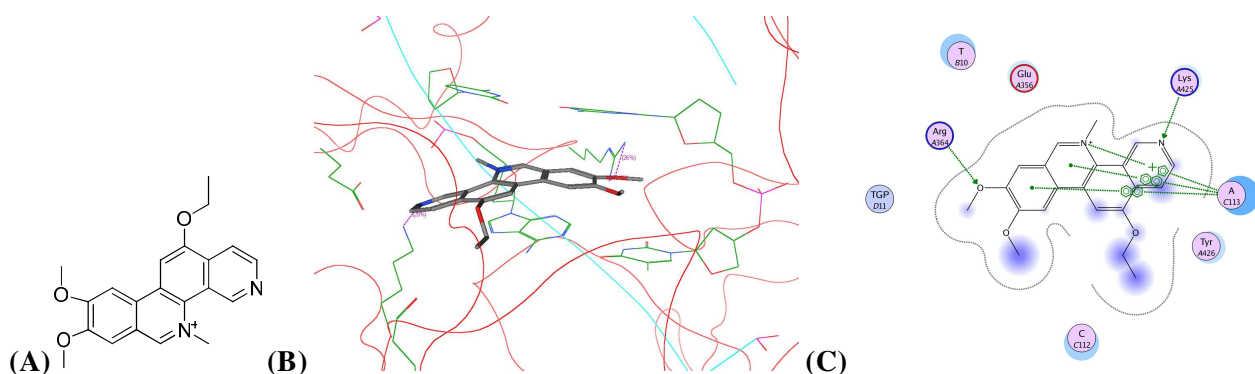
Các chất ức chế hoạt động của enzym TOP-I hiện nay được chia làm 2 nhóm (1, 15). Thứ nhất là nhóm chất kìm hãm (suppressor) hoạt tính enzym TOP-I. Nhóm này là những chất ức chế hoạt động enzym TOP-I nhưng không làm bền trạng thái trung gian của phức hợp DNA:TOP-I. Các chất nhóm kìm hãm hoạt động TOP-I thường gắn kết với ADN làm cho TOP-I không thể nhận diện ra điểm gắn kết chuyên biệt trên ADN và kết quả là ức chế sự thành lập phức hợp giữa TOP-I và ADN. Cơ chế kìm hãm này hoạt động theo cơ chế tương tranh (thuốc và TOP-I cạnh tranh gắn kết trên cơ chất là ADN) và enzym vẫn có hoạt tính nhưng không có cơ chất để hoạt động. Kết quả là làm ngừng chu trình tế bào. Thứ hai là nhóm gây độc TOP-I (poison). Nhóm này bao gồm những chất có tác động chống sự tách ra của phức ADN:TOP-I sau khi TOP-I đến gắn kết và cắt đoạn ADN. Các chất gây độc TOP-I hoạt động theo nguyên tắc là gắn kết với cả hai, TOP-I lẫn ADN, trong phức hợp ADN:TOP-I bằng liên kết đồng hóa trị, bền hóa phức và làm vỡ sợi kép ADN. Sự hiện diện của sợi ADN đơn được nhận diện và tín hiệu này làm ngừng chu trình tế bào. Kết quả cuối cùng là sự hủy bào.<sup>8</sup> Trong nghiên cứu này, sau khi tiến hành docking, phức hợp ADN:TOP-I:BCP được phân tích trên khả năng gắn kết với phức hợp hay ADN riêng lẻ.

Tổng số 53 chất tương đồng BCP docking thành công vào phức hợp. Điểm số gắn kết thay đổi từ -57.57 kcal/mol đến -36.27 kcal/mol. Tất cả các cấu dạng trong cấu trúc phức hợp được phân tích và so sánh nhằm giải thích mối liên quan giữa cấu trúc hóa học và tác dụng sinh học. Phân tích kết quả docking cho thấy tất cả các chất tương đồng BCP đều có khả năng tạo liên kết  $\pi$ - $\pi$  với Guanin G11. Hệ thống vòng của chất tương đồng BCP có xu hướng nằm ở vị trí song song và xen vào giữa các cặp base G11/C112 và A113/T10 của cấu trúc ADN và hạn chế khả năng gắn kết của ADN với TOP-I (với ATGC là tên của các

nucleotid). Kết quả docking cũng cho thấy nhóm chất tương đồng BCP có xu hướng cản trở không gian do tạo liên kết bề mặt với ADN nhiều hơn khả năng gắn kết với TOP-I. Điều này có thể giải thích do cấu trúc 'cứng' của hệ thống vòng và khả năng chèn vào cấu trúc ADN. Bên cạnh đó, một số chất tương đồng BCP cũng có khả năng tạo liên kết với TOP-I ở các acid amin Asn352, Arg364, Asn722, Glu356, Lys425. Tuy nhiên, kết quả docking cho thấy xu hướng gắn kết rõ rệt của những hợp chất này trên ADN nhiều hơn là trên TOP-I. Kết quả này cho thấy đa số các chất trong 53 chất tương đồng BCP được tiến hành docking hoạt động theo cơ chế kìm hãm hoạt tính TOP-I (suppressor).

### **Ứng dụng**

Nghiên cứu docking đồng thời được tiến hành với một số chất tương đồng BCP mới được tổng hợp (19). Chất K7511-NP22 với cấu trúc trình bày ở hình 4A có độc tính tế bào  $IC_{50} = 2.0 \mu M$  và  $5.97 \mu M$  lần lượt trên dòng tế bào ung thư KB-3-1 và L1210. Mô hình docking cho thấy chất này có khả năng gắn kết tốt với phức hợp TOP-I:DNA. K7511-NP22 có khả năng tạo 2 liên kết  $\pi$ - $\pi$  song song và 1 liên kết cation- $\pi$  với adenin. Khả năng gắn kết của trên TOP-I của chất này thể hiện bằng 2 liên kết hydro với Arg364 và Lys432, tuy nhiên hai liên kết này tương đối yếu. Khả năng gắn kết ở mức độ phân tử của K7511-NP22 có thể dùng giải thích cho khả năng ức chế hoạt động của TOP-I bằng cách cạnh tranh với TOP-I trong việc hình thành phức hợp với ADN, cản trở hoạt động của TOP-I trong việc tiếp cận và gắn kết với ADN và làm ngừng quá trình sao chép, phiên mã và tái tổ hợp ADN. Mô hình gắn kết ở mức độ phân tử trong không gian 3 chiều của K7511-NP22 với phức hợp ADN:TOP-I được trình bày ở Hình 4B và mô hình tương ứng thể hiện khả năng gắn kết ở không gian 2 chiều được trình bày ở Hình 4C (hình tạo bởi chương trình MOE (11)).



**Hình 4.** (A) Cấu trúc chất K7511-NP22 ; (B) Mô hình gắn kết ở mức độ phân tử trong không gian 3 chiều của K7511-NP22 với phức hợp ADN:TOP-I; (C) Mô hình biểu diễn khả năng gắn kết 2D của K7511-NP22 với phức hợp ADN:TOP-I.

### Mô hình phân loại các chất tương đồng BCP ức chế TOP-I

#### Xây dựng mô hình SVM

Cơ sở dữ liệu gồm 82 chất được phân chia thành 3 tập hợp (20) bao gồm (i) Tập hợp huấn luyện dùng để xây dựng mô hình có 58 chất, (ii) Tập hợp kiểm tra có 15 chất, (iii) Tập hợp đánh giá có 9 chất không sử dụng trong xây dựng mô hình. Để lựa chọn ra các thông số tương quan tốt và mô tả được cấu trúc liên quan đến hoạt tính sinh học, các bước chọn lọc được tiến hành. Từ các 2032 thông số mô tả phân tử tính toán từ phần mềm Dragon (3), các thông số chứa toàn bộ giá trị 0 được loại bỏ còn lại 533 thông số. Tiếp theo, 420 thông số được loại bỏ vì có: (i) hơn 80% dẫn chất trong tập hợp chứa giá trị 0, (ii) giá trị độ lệch chuẩn nhỏ hơn 0,5, (iii) hệ số tương quan

chéo giữa các thông số >90%. 103 thông số còn lại được áp dụng vào 3 thuật toán (i) mRMR: phù hợp tối đa, dư thừa tối thiểu - Max-Relevance, Min-Redundancy, (ii) thuật toán di truyền GA - genetic algorithm, (iii) “rừng ngẫu nhiên” - RF - Random Forest). Kết quả cho thấy 7 thông số lựa chọn từ phương pháp RF (Bảng 1) cho kết quả phân loại ổn định và tốt hơn 2 phương pháp lựa chọn còn lại. Đồng thời, một lợi điểm nữa khi sử dụng thông số mô tả từ RF là sử dụng 7 thông số mô tả phân tử nhưng khả năng phân loại tốt hơn khi sử dụng 10 thông số từ mRMR và 16 từ GA. Bên cạnh đó, mô hình SVM sử dụng thông số mô tả từ RF có số vector hỗ trợ sử dụng cũng thấp hơn các tập hợp thông số còn lại.

**Bảng 1.** Thông số mô tả phân tử được chọn lọc từ phương pháp RF và sử dụng để xây dựng mô hình SVM phân loại

Nhóm thông số	Tên thông số	Mô tả thông số
Thông số hình học Topo	D/Dr05	Chỉ số khoảng cách/vòng bậc 5
	D/Dr06	Chỉ số khoảng cách/vòng bậc 6
Thông số dựa vào đếm các nối và cấu trúc	MPC06	Số lượng chỉ điểm phân tử bậc 06
	MPC08	Số lượng chỉ điểm phân tử bậc 08
	MPC10	Số lượng chỉ điểm phân tử bậc 10
Chỉ điểm tần suất 2D	F05[C-C]	Tần suất của C-C ở khoảng cách Topo 05
	F08[N-O]	Tần suất của N-O ở khoảng cách Topo 08

Trong quá trình xây dựng mô hình SVM (13), việc xác định thông số tối ưu nhất cho hàm kernel cũng rất quan trọng nhằm hạn chế

lỗi phân loại. Thuật toán dò tìm Grid được sử dụng trong nghiên cứu với mục đích xác định thông số tối ưu của hàm kernel sao cho lỗi

đánh giá chéo khi phân loại là nhỏ nhất. Cặp thông số ( $C=4, \gamma=0,25$ ) cho kết quả lỗi đánh giá chéo thấp nhất (0,17) và độ đúng cao nhất được xác định và sử dụng để xây dựng mô hình. Từ kết quả trên, 7 thông số mô tả phân tử xác định từ phương pháp RF (Bảng 1) và thông số tối ưu của hàm kernel ( $C=4, \gamma=0,25$ ) được sử dụng để xây dựng mô hình SVM bằng cách sử dụng phần mềm SVM - e1071 trong R. Kết quả trình bày ở Bảng 2 cho thấy mô hình SVM có khả năng phân loại đúng 93% các chất trong tập hợp huấn luyện và 87% các chất trong tập hợp kiểm tra. Hệ số tương quan MCC của mô hình ổn định duy trì ở mức trên

0,7 đối với cả 3 tập hợp : huấn luyện, kiểm tra và đánh giá ngoại. Theo một số nghiên cứu, mô hình có hệ số MCC > 0,4 được xem là mô hình có khả năng dự đoán (9). Mô hình xây dựng được đánh giá trên tập hợp đánh giá (gồm những chất không sử dụng để xây dựng mô hình) cho kết quả phân loại đúng 8/9 chất, tương ứng với 89% các chất trong tập hợp. Phân tích kết quả cho thấy mô hình có khả năng phân loại chính xác các chất có hoạt tính ức chế TOP-I mạnh hơn hoặc yếu hơn topotecan và có thể ứng dụng trong thực tế cho các chất có cấu trúc tương đồng BCP khác.

**Bảng 2.** Mô hình phân loại SVM dựa vào 7 thông số mô tả phân tử xác định bằng phương pháp RF

Đánh giá	Tập huấn luyện	Tập kiểm tra	Tập đánh giá	Tập ứng dụng
TP	13	3	4	5
TN	41	10	4	3
FP	2	0	1	0
FN	2	2	0	2
Độ đúng của tập hợp	0,93	0,87	0,89	0,80
Độ đúng trên chất có hoạt tính	0,87	0,60	1,00	0,71
Độ đúng trên chất không hoạt tính	0,95	1,00	0,80	1,00
Độ chính xác của chất có hoạt tính	0,87	1,00	0,80	1,00
Độ chính xác của chất không hoạt tính	0,95	0,83	1,00	0,60
Hệ số tương quan Matthews MCC	0,82	0,71	0,80	0,65

Với TP và TN lần lượt là số lượng hợp chất có hoạt tính và không có hoạt tính được dự đoán đúng từ mô hình. FP và FN lần lượt là số lượng hợp chất không có hoạt tính và có hoạt tính được mô hình dự đoán sai.

### Ứng dụng mô hình phân loại SVM

Mô hình SVM được đánh giá mức độ chính xác và khả năng ứng dụng trên tập hợp 10 dẫn chất BCP mới được tổng hợp năm 2008-2009 bởi nhóm nghiên cứu La Voie và cộng sự (4, 5). Trong 10 dẫn chất này, 7 chất có hoạt tính mạnh hơn topotecan, 3 chất không hoạt tính. Kết quả được trình bày ở Bảng 2 cho thấy của mô hình phân loại đúng 8/10 chất của tập hợp và đạt độ đúng 80%. Bên cạnh đó, độ chính xác của chất có hoạt tính là 100%. Giá trị này cho thấy mô hình có khả năng phân loại chính xác các chất có hoạt tính hơn là các chất không có hoạt tính. Điều này hoàn toàn phù hợp với mục tiêu của nghiên cứu là tìm kiếm các chất có khả năng ức chế TOP-I để phát triển thành các thuốc kháng ung thư có khả năng sử dụng trong điều trị.

Từ chất khởi nguồn ethoxidin (hình 1), việc đưa thêm vào khung BCP các nguyên tử N có khả năng làm gia tăng điện tích âm và nhóm ethoxy C12, hy vọng là sẽ tìm ra những chất tương đồng BCP mới với tác động ức chế topoisomerase-1 mạnh hơn. Tổng cộng 12 cấu trúc tương đồng BCP được thiết kế và ứng dụng vào mô hình để dự đoán hoạt tính ức chế TOP-I. Kết quả cho thấy 6 trong 12 chất tương đồng BCP thiết kế được dự đoán là có khả năng ức chế TOP-I mạnh hơn topotecan. Trong số 6 hợp chất này, có 4 dẫn chất 12-ethoxy BCPs. Nhóm 12-ethoxy trong khung BCPs được cho là cần thiết trong việc định hướng vào rãnh nhỏ (small groove) của ADN và ngăn cản TOP-I tiến đến vị trí gắn kết của chúng trên ADN (6, 10). Các chất dự đoán có khả năng ức chế TOP-I sẽ được ưu tiên tổng hợp và thử

nghiệm *in vitro* về tác dụng sinh học bao gồm tác động ức chế TOP-1 và độc tính trên các dòng tế bào ung thư.

## KẾT LUẬN

Thiết kế thuốc hợp lý với sự hỗ trợ của máy tính và dược tính học là chuyên ngành mới của Hóa Dược và đã ứng dụng thành công trong nghiên cứu và phát triển thuốc mới. Trong nghiên cứu này, Sử dụng chương trình mô tả phân tử DOCK 6.2, khả năng gắn kết ở mức độ phân tử của các chất tương đồng BCP với phức hợp ADN:TOP1 được phân tích. Kết quả cho thấy các cấu trúc tương đồng BCP có xu hướng gắn kết với ADN hơn là TOP1 và cần thực hiện các thử nghiệm xác định mức độ gắn kết với ADN và enzym riêng rẽ để khẳng định lại kết quả này. Phương pháp máy vector hỗ trợ SVM được sử dụng với mục đích xây dựng mô hình dự đoán và phân loại 73 chất tương đồng benzo[c]phenanthridine trên hoạt tính kháng topoisomerase I. Mô hình SVM tốt nhất được xây dựng bằng gói SVM - e1071 trong R với thông số tối ưu của hàm kernel ( $C=4$ ,  $\gamma=0,25$ ). Ngoài ra, mô hình còn được đánh giá lại trên tập hợp thử ngoại (gồm các chất không dùng để xây dựng mô hình) cho độ đúng 89% (dự đoán đúng 8/9 chất) và tập hợp ứng dụng cho độ đúng 80% (dự đoán đúng 8/10 chất). Mô hình chứng tỏ khả năng phân loại chính xác chất có hoạt tính (80-100% tùy tập hợp). Kết quả này chứng tỏ mô hình có khả năng phân loại các chất có khả năng ức chế TOP-I trên dãy chất tương đồng BCP. Mô hình này có thể ứng dụng trong thực tế để dự đoán hoạt tính ức chế TOP-I của các cấu trúc tương đồng BCP mới trước khi tổng hợp, góp phần vào việc tìm kiếm các chất có hoạt tính kháng TOP-I tốt để phát triển thành các thuốc kháng ung thư có khả năng sử dụng trong điều trị. Kết quả nghiên cứu này cũng có thể ứng dụng để thiết kế ra các cấu trúc tương đồng BCP có hoạt tính sinh học cao hơn.

Ứng dụng phương pháp thiết kế thuốc hợp lý này giúp giảm chi phí trong nghiên cứu tìm ra thuốc mới vì loại bỏ ngay từ đầu các chất dự

đoán không có hoạt tính, tăng tỷ lệ thành công và đồng thời rút ngắn thời gian trong tìm kiếm các chất có hoạt tính kháng TOP-I. Bên cạnh ý nghĩa khoa học, đề tài còn xây dựng được protocol thực hiện quá trình docking bằng cách kết nối các chương trình miễn phí và ứng dụng triển khai ở qui mô phòng thí nghiệm.

## TÀI LIỆU THAM KHẢO

1. Choi I, Kim C, Choi S. (2007). Binding mode analysis of topoisomerase inhibitors, 6-arylamino-7-chloroquinazoline-5,8-diones, within the cleavable complex of human topoisomerase I and DNA. Arch. Pharm. Res., 30(12): 1526-35.
2. DOCK 6.2. UCSF, [http://dock.compbio.ucsf.edu/DOCK\\_6/index.htm](http://dock.compbio.ucsf.edu/DOCK_6/index.htm)
3. Dragon® (2009): The software for calculating molecular descriptors, [www.taletale.mi.it/](http://www.taletale.mi.it/)
4. Feng, W.; Satyanarayana, M.; Cheng, L.; Liu, A.; Tsai, Y.-C.; Liu, L. F.; LaVoie, E.J. (2008) Synthesis of N-substituted 5-[2-(N-alkylamino)ethyl]dibenzo[c,h][1,6]-naphthyridines as novel topoisomerase I-targeting anti-tumor agents. Bioorg. Med. Chem., 16: 9295-9301
5. Feng, W.; Satyanarayana, M.; Tsai, Y.-C.; Liu, A. A.; Liu, L. F.; LaVoie, E. J. 11-Substituted 2,3-dimethoxy-8,9-methylenedioxybenzo[i]phenanthridine derivatives as novel topoisomerase I-targeting agents. Bioorg. Med. Chem.. 2008, 16, 8598-8606
6. Fleury, F.; Sukhanova, A.; Ianoul, A.; Devy, J.; Kudelina, I.; Duval, O.; Alix, A. J.P.; Jardillier, J. C.; Nabiev, I. (2000). Molecular determinants of site-specific inhibition of human DNA topoisomerase I by fagaronine and ethoxidine. Relation to DNA binding. J. Biol. Chem., 275: 3501-9.
7. Hsu, C.-W.; Chang, C.-C.; Lin, C.-J. (2008). LIBSVM: A Practical Guide to Support Vector Classification. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
8. Huỳnh Thị Ngọc Phương, Thái Khắc Minh (2009). Ứng dụng QSAR trong định hướng tổng hợp các chất tương đồng benzo[c]phenanthridin có tác dụng kháng ung thư. Y học Thành Phố Hồ Chí Minh, 13 (phụ bản 1- chuyên đề Dược): 324-330.
9. Ji, L.; Wang, X.; Qin, L.; Luo, S.; Wang, L. (2009) Predicting the Androgenicity of Structurally Diverse Compounds from Molecular Structure Using Different Classifiers. QSAR Comb. Sci. 2009, 28(5), 542-50.
10. Lynch, M.A.; Duval, O.; Sukhanova, A.; Devy, J.; MacKay, S.P.; Waigh, R.D.; Nabiev, I. (2001). Synthesis, biological activity and comparative analysis of DNA binding affinities and human DNA topoisomerase I inhibitory activities of novel 12-alkoxy-benzo[c]phenanthridinium salts. Bioorg. Med. Chem. Lett., 11(19): 2643-2646.
11. MOE 2007.02. Chemical Computing Group Inc., Montreal, H3A 2R7 Canada, <http://www.chemcomp.com>
12. Nakanishi, T.; Suzuki, M.; Mashiba, A.; Ishikawa, K.; Yokotsuka, T. (1998). Synthesis of NK109, an Anticancer Benzo[c]phenanthridine Alkaloid. J. Org. Chem., 63: 4235-9.
13. R (2009): The R project for statistical computing, <http://r-project.org/>.

14. Ruchelman, A. L. et al. (2004). Dimethoxybenzo(i)phenanthridine-12-carboxylic acid derivatives and 6H-dibenzo(c,h) (2,6)naphthyridin-5-ones with potent topoisomerase I-targeting activity and cytotoxicity. *Bioorg. Med. Chem. Lett.*, 14: 5585-9.
15. Scaglioni, L.; Mazzini, S.; Mondelli, R.; Dallavalle, S.; Gattinoni, S.; Tinelli, S.; Beretta, G. L.; Zunino, F.; Ragg, E. (2009). Interaction between double helix DNA fragments and a new topopyrone acting as human topoisomerase I poison. *Bioorg. Med. Chem.*, 17: 484-491
16. Staker BL, Feese MD, Cushman M, Pommier Y, Zembower D, Stewart L, Burgin AB. (2005). Structures of three classes of anticancer agents bound to the human topoisomerase I-DNA covalent complex. *J. Med. Chem.*, 48: 2336-45.
17. Staker BL, Hjerrild K, Feese MD, Behnke CA, Burgin AB Jr, Stewart L. (2002). The mechanism of topoisomerase I poisoning by a camptothecin analog. *Proc. Natl. Acad. Sci. USA*, 99: 15387-92.
18. Thái Khắc Minh, Huỳnh Thị Ngọc Phương (2008). Ứng dụng Dược Tin Học trong nghiên cứu khả năng ức chế topoisomerase I và độc tính tế bào của dẫn chất benzo[c]phenanthridin. Hội nghị Khoa Học Kỹ Thuật và Công Nghệ Hóa Dược Toàn Quốc, Hà nội, 1: 81-86.
19. Thái Khắc Minh, Huỳnh Thị Ngọc Phương (2009). Dự đoán độc tính tế bào của các dẫn chất benzo[c]phenanthridin bằng mô hình QSAR. *Tạp chí Dược học*, 49 (2): 51-54.
20. Thai, K.-M.; Ecker, G. F. (2008). A Binary QSAR Model for Classification of hERG Potassium Channel Blockers. *Bioorg. Med. Chem.*, 16: 4107-19.
21. Yap, C. W.; Chen, Y. Z. (2005). Prediction of Cytochrome P450 3A4, 2D6, and 2C9 Inhibitors and Substrates by Using Support Vector Machines. *J. Chem. Inf. Model.*, 45: 982-92.



