

ĐẠI HỌC QUỐC GIA TP.HCM
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA VẬT LÝ-VẬT LÝ KỸ THUẬT

KHÓA LUẬN TỐT NGHIỆP
Chuyên ngành: Vật lý Lý thuyết

Nguyễn Hữu Quý Ngân

KHẢO SÁT TƯƠNG TÁC GIỮA PHÂN TỬ PROTEIN VP35 VÀ dsRNA CỦA VIRUS EBOLA

Người hướng dẫn:
TS. Nguyễn Hà Hùng Chương

TP.HCM 07/2015

Mục lục

Mục lục	i
Lời cảm ơn	iii
Từ điển chú giải	v
Danh sách hình vẽ	vi
Lời mở đầu	viii
1 Tổng quan	1
1.1 Bệnh do Virus Ebola	1
1.1.1 Các protein của EBOV	1
1.2 Cơ chế hoạt động của EBOV	3
1.2.1 Quá trình xâm nhập và sao chép của EBOV	3
1.2.2 Quá trình hệ thống miễn dịch hoạt động	4
1.2.3 VP35 vô hiệu hệ thống miễn dịch	4
1.2.4 Các phương án ngăn chặn EBOV	5
1.3 Các nghiên cứu trên thế giới về VP35	5
2 Đối tượng và phương pháp nghiên cứu	7
2.1 Cấu hình của VP35	7
2.1.1 Cấu hình 3L25	7
2.1.2 B-conf	9
2.2 Phương pháp Mô phỏng động học phân tử	9
2.2.1 Ý tưởng chính	9

2.2.2	Trường lực	10
2.2.3	Thuật toán Verlet sử dụng trong mô phỏng MD	12
2.2.4	Thuật toán kiểm soát nhiệt độ	13
2.3	Phương pháp mô phỏng Steered Molecular Dynamics (SMD)	17
2.3.1	SMD sử dụng lực kéo không đổi	18
2.3.2	SMD sử dụng vận tốc kéo không đổi	18
2.4	Tham số sử dụng trong các quá trình mô phỏng	19
2.5	Các công cụ phân tích kết quả	20
2.5.1	Lực kéo tối đa	20
2.5.2	Thế năng tương tác theo thời gian	22
2.5.3	Liên kết Hydro	22
3	Kết quả và thảo luận	23
3.1	Thiết lập hệ mô phỏng	23
3.1.1	Cực tiểu hoá năng lượng	23
3.1.2	Cân bằng chính tắc (NVT)	24
3.1.3	Cân bằng đẳng nhiệt - đẳng tích (NPT)	25
3.2	Khoảng cách giữa dsRNA và VP35 theo thời gian	26
3.3	Thế năng tương tác giữa amino acid của VP35 với dsRNA	26
3.4	Kết quả phân tích liên kết Hydro giữa các amino acid của VP35 với dsRNA	32
3.5	So sánh lực kéo	33
3.6	Đột biến Lys282Ala	34
4	Kết luận và hướng phát triển tiếp theo của đề tài	37
	Tài liệu tham khảo	44
A	Các amino acid tương tác với dsRNA	45
B	Khảo sát liên kết Hydro	47
C	Các amino acid và các ký hiệu tương ứng	49

Lời cảm ơn

Đề tài hoàn thành được là nhờ sự hỗ trợ rất lớn từ quý thầy, các anh chị và các bạn trong nhóm nghiên cứu.

Em xin cảm ơn TS. Nguyễn Hà Hùng Chương rất nhiều vì thầy đã luôn dành thời gian và kinh nghiệm để hỗ trợ em hết mình trong việc giải quyết các khó khăn. Em xin cảm ơn TS. Võ Thành Văn đã nỗ lực truyền đạt kiến thức cho tập thể lớp, đồng thời tạo cho em nếp suy nghĩ logic và kỹ năng tính toán cần có. Em xin cảm ơn TS. Vũ Quang Tuyên đã luôn lắng nghe và hỗ trợ em cùng các bạn trong lớp. Em xin cảm ơn TS. Võ Quốc Phong vì đã chia sẻ những ý kiến thiết thực về nhân cách sống. Em xin cảm ơn TS. Nguyễn Hữu Nhã đã tận tình hỗ trợ và tìm mọi cách để tạo điều kiện học tập tốt nhất cho sinh viên bộ môn.

Em xin cảm ơn anh Nguyễn Quý Trường và anh Trần Phước Duy đã luôn dành thời gian giúp đỡ em về mặt kỹ thuật. Em xin cảm ơn anh Minh Trung, anh Minh Nhật, anh Anh Tuấn đã cung cấp cho em bản khoá luận, giúp em thừa kế kiến thức các anh đã bỏ công nghiên cứu.

Xin cảm ơn các bạn lớp Vật lý lý thuyết khoá 2011. Các bạn đã luôn hỗ trợ lẫn nhau hết mình.

Trên hết, con xin cảm ơn ba mẹ đã luôn tin tưởng con có thể học và theo đuổi nghiên cứu khoa học. Em xin cảm ơn chị gái đã luôn tin tưởng và giúp đỡ em. Xin cảm ơn em gái đã chăm sóc anh trong suốt thời gian làm việc căng thẳng.

Từ điển chú giải

amber99sb ff trường lực AMBER99SB. 10, 22

Arg312Ala Đột biến amino acid Arginine ở vị trí thứ 312 thành Alanine.. 19, 34, 35

dsRNA chuỗi kép RNA - double-strain RNA -. vi-ix, 2, 4-9, 21, 22, 27, 28, 32-35, 38, 48, 49

EBOV virus Ebola. vi, viii, 1-5

EHF Sốt xuất huyết Ebola - Ebola Haemorrhagic Fever -. 1

EVD Bệnh Ebola - Ebola Virus Disease -. 1, 5

GP Glycoprotein. 2, 5

IFN Interferon (IFN) là một nhóm các protein tự nhiên được sản xuất bởi các tế bào của hệ miễn dịch ở hầu hết các động vật nhằm chống lại các tác nhân ngoại lai như virus, vi khuẩn, kí sinh trùng và tế bào ung thư.[9]. 2, 4

Lys282Ala Đột biến amino acid Lysine ở vị trí thứ 282 thành Alanine.. vii, 19, 27, 36

Lys339Ala Đột biến amino acid Lysine ở vị trí thứ 339 thành Alanine.. 19, 32, 34, 35

MD phương pháp động học phân tử - Molecular dynamics -. viii, 9, 17

NP Nucleoprotein. 2

phối tử ligand là một (hoặc một nhóm) phân tử có khả năng gắn vào một nhóm phân tử khác để tạo nên phức hợp với chức năng riêng biệt.. 17, 18

RIG-I retinoic acid-inducible gene 1 là một loại enzym có chức năng phát hiện chuỗi vật chất di truyền của các virus cúm A, virus Sendai, họ flavivirus (virus sốt xuất huyết, viêm não Nhật Bản,...). Tuy nhiên RIG-I không phát hiện được retrovirus hay virus có vật chất di truyền là DNA.[29]. vi, 4

SMD Steered Molecular Dynamics. viii, 7, 17–19, 27, 28, 34, 35, 38

thụ thể receptor là phân tử nằm ở màng tế bào có chức năng thu nhận tín hiệu hóa học cho tế bào đó.. 17, 18

vị trí amino acid mỗi amino acid tại vị trí cụ thể sẽ có một vai trò cụ thể, vì vậy cần xác định tên amino acid và số thứ tự của nó trong chuỗi protein.. vii, 5, 21, 22, 28–34, 48, 49

ZEBOV Zaire Ebola. 1, 5

Danh sách hình vẽ

1.2	Vòng đời của virus Ebola (EBOV). Nguồn: http://www.expasy.org/ .	3
1.3	Quá trình VP35 vô hiệu RIG-I thông qua việc bám vào chuỗi kép RNA - double-strain RNA - (dsRNA). Nguồn [35].	4
1.4	Vị trí tương đối của ligand 1DK so với VP35 (màu xanh da trời) và dsRNA (màu vàng) đã được công bố trên [39].	6
2.1	Cấu trúc VP35 được khảo sát.	8
2.2	Tương tác giữa các nguyên tử có liên kết	11
2.3	Bằng cách di chuyển nhóm liên kết (constraint group) với vận tốc không đổi, lực tác động lên nhóm phân tử cần kéo ra sẽ lớn dần đến một giá trị đủ để tách nhóm đó ra.	18
2.4	(A) kéo ligand ra khỏi protein bằng cách đặt một thế năng điều hoà thay đổi theo thời gian, tương đương với việc đặt một lực kéo của một lò xo với hệ số đàn hồi k lên ligand. (B) Đường biểu diễn lực kéo theo toạ độ. Lực kéo tối đa F_{max} đặc trưng cho ái lực liên kết giữa ligand và protein. Nguồn [37]	21
2.5	Đồ thị tương quan giữa lực kéo tối đa (Rupture Force) và năng lượng tương tác giữa phân tử R-125489 với thụ thể NA của virus cúm và giữa phân tử CS-8959 với thụ thể NA của virus cúm [37]. Hệ số tương quan (correlation level) $R = 0.97$	21
2.6	Khoảng cách và góc giữa phân tử Donor, Acceptor và nguyên tử Hydro	22
3.1	Đồ thị thế năng của hệ theo thời gian trong bước cực tiểu hoá năng lượng.	23

3.2	Nhiệt độ trung bình của hệ theo thời gian trong bước cân bằng chính tắc (NVT)	24
3.4	Đường biểu diễn tọa độ khối tâm trung bình theo thời gian giữa dsRNA và VP35 với phân tử không đột biến (3L25) và đột biến (3L27, 3L28, Lys282Ala)	26
3.5	Đồ thị thế năng theo thời gian của các amino acid	28
3.6	Đồ thị thế năng theo thời gian của các amino acid	29
3.7	Đồ thị thế năng theo thời gian của các amino acid	30
3.8	vị trí amino acid 275–279, tiêu biểu Cys275, Ile278, Glu279. Glu279 tương tác Coulomb là chủ yếu. Cys275, Ile278 tương tác VdW là chủ yếu với thế năng cùng scale với thế năng của Lys339.	31
3.9	Liên kết Hydro theo thời gian giữa các vị trí amino acid và dsRNA. Lần kéo thứ 2 cấu hình B-conf của phân tử có PDBID: 3L25.	32
3.10	Mạch chính (Backbone) của VP35 được tô màu xanh da trời. Amino acid Ile340 nằm tại C-terminal nên là amino acid cuối cùng, có gốc OH^- có khả năng tạo liên kết Hydro với dsRNA. Nguyên tử Oxy được tô màu đỏ.	33
B.1	Liên kết Hydro theo thời gian giữa các vị trí amino acid và dsRNA. Lần kéo thứ 1 cấu hình B-conf phân tử có PDBID: 3L25	47
B.2	Liên kết Hydro theo thời gian giữa các vị trí amino acid và dsRNA. Lần kéo thứ 3 cấu hình B-conf phân tử có PDBID: 3L25	48

Lời mở đầu

Bệnh Ebola - Ebola Virus Disease - (EVD) hay Sốt xuất huyết Ebola - Ebola Haemorrhagic Fever - (EHF) có tỷ lệ tử vong cao và chưa có thuốc đặc trị cũng như vaccine để phòng ngừa. Để có thể tìm được thuốc, chúng ta cần tìm hiểu các protein tham gia vào quá trình virus Ebola xâm nhập tế bào và sao chép. Một trong những protein có nhiều vai trò trong quá trình sao chép của virus Ebola là VP35. Một trong các chức năng của VP35 là gắn vào vật chất di truyền của virus Ebola (chuỗi kép RNA - double-strain RNA - (dsRNA)) và ngăn cản hệ thống miễn dịch phát hiện ra sự có mặt của dsRNA trong tế bào.

Đề tài tập trung vào việc phân tích tương tác của phân tử VP35 và dsRNA của virus Ebola (EBOV) nhằm đánh giá các amino acid có vai trò quan trọng trong tương tác giữa hai phân tử trên. Đề tài sử dụng phương pháp mô phỏng Steered Molecular Dynamics (SMD) nhằm khảo sát sự chuyển trạng thái từ VP35-dsRNA tương tác với nhau sang trạng thái VP35-dsRNA không tương tác. Kết quả của khóa luận cho thấy các amino acid Arg312 và Lys339 của VP35 đóng vai trò quan trọng trong tương tác VP35-dsRNA. Kết quả này phù hợp với thực nghiệm nghiên cứu cấu trúc của VP35 [34]. Đồng thời đề tài cũng cho thấy đột biến amino acid Lys282 từ Lysine thành Alanine (Lys282Ala) làm giảm khả năng liên kết giữa VP35 và dsRNA. Do vậy Lys282 được dự đoán đóng vai trò quan trọng trong tương tác giữa VP35 và dsRNA.

Đề tài gồm 4 chương:

- Chương 1: Bệnh Ebola, EBOV và sự cần thiết của việc nghiên cứu EBOV, quá trình EBOV sao chép và vai trò của protein VP35, một số phương án ngăn chặn EBOV và một số nghiên cứu gần đây.
- Chương 2: Các cấu hình của VP35 được khảo sát, phương pháp động học phân

tử - Molecular dynamics - (MD) và phần mềm, công cụ sử dụng để khảo sát tương tác giữa VP35 và dsRNA.

- Chương 3: Trình bày các kết quả thu được và thảo luận tương ứng.
- Chương 4: Các kết luận và hướng phát triển tiếp theo của đề tài.

Xin chân thành cảm ơn!

Chương 1

Tổng quan

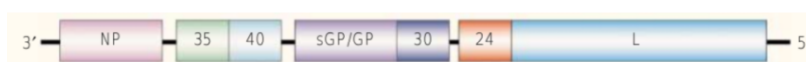
1.1 Bệnh do Virus Ebola

Bệnh Ebola - Ebola Virus Disease - (EVD) hay Sốt xuất huyết Ebola - Ebola Haemorrhagic Fever - (EHF) là bệnh truyền nhiễm trên động vật có vú gây bởi virus Ebola (EBOV). Bệnh được phát hiện lần đầu năm 1976 tại Trung Phi[15]. Tỷ lệ tử vong nếu không được chữa trị lên đến 90% ở chủng Zaire Ebola (ZEBOV). Đây là bệnh có tốc độ bùng phát cao và khó kiểm soát do Ebola có thể lây truyền qua nhiều động vật là vật chủ trung gian, đồng thời lây truyền qua tiếp xúc với dịch tiết từ cơ thể người bệnh[6].

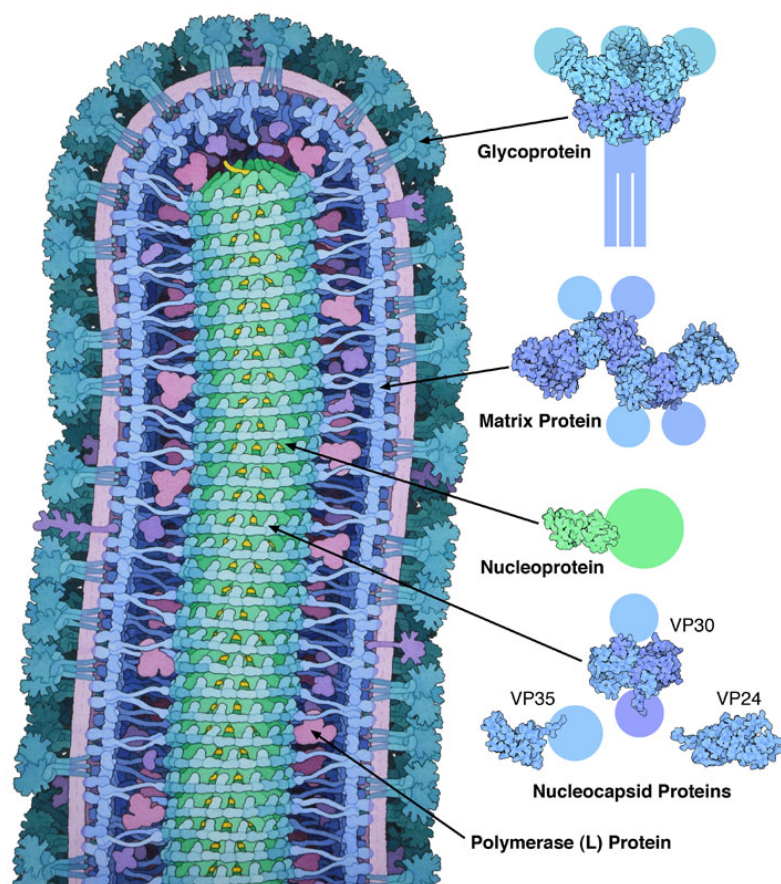
EVD đặc biệt nguy hiểm và có khả năng xuất hiện ở các nước có khí hậu nhiệt đới, cận nhiệt[38]. Đồng thời, EVD lây truyền qua dơi ăn trái cây và các động vật như tinh tinh, khỉ, heo[33], nên khả năng xuất hiện ở Việt Nam là rất cao nếu không được kiểm soát và theo dõi chặt chẽ. Trong nhiều thập kỷ, việc tìm hiểu về EBOV được quan tâm nhiều khi vũ khí sinh học trở nên phát triển[15, 30].

1.1.1 Các protein của EBOV

Việc tìm hiểu các protein có trong EBOV có ý nghĩa quan trọng để tìm ra các phương án ngăn chặn EBOV. EBOV thuộc họ Filoviridae, bộ Mononegavirales, tức là EBOV thuộc họ virus có vật chất di truyền RNA gram âm và có hình sợi[33]. EBOV được cấu tạo bởi 7 loại protein với các chức năng khác nhau[10]. Xem hình 1.1a. Các



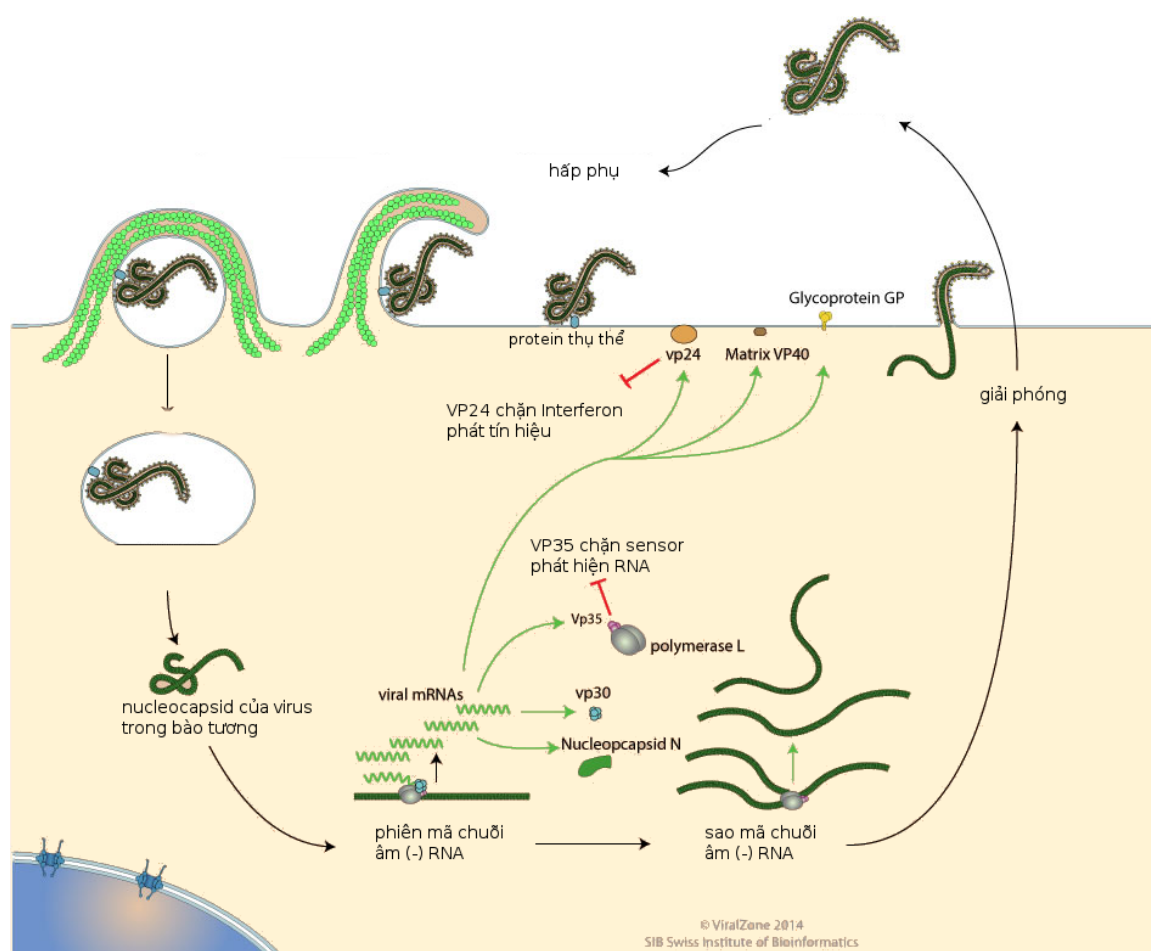
(a) Bảy loại protein được mã hoá trên RNA của EBOV



(b) Vị trí tương đối của các protein trong EBOV

protein có của EBOV gồm:

- Glycoprotein (GP) là protein nằm ở vỏ của virus, mang các thụ thể như các chìa khóa giúp EBOV xâm nhập vào tế bào vật chủ [25, 30].
- VP40 là protein ma trận, giúp định hình cấu trúc cho EBOV[7, 11].
- VP24 là protein có vai trò ngăn chặn các IFN, một loại protein có chức năng kích hoạt các gene mã hóa các protein mang tính báo động của cơ thể[26].
- VP30 là protein có vai trò hỗ trợ phiên mã dsRNA[12].
- L protein và Nucleoprotein (NP)[15, 40] có vai trò quan trọng trong quá trình phiên mã dsRNA, đồng thời NP giữ cho chuỗi đơn RNA của EBOV không bị



Hình 1.2: Vòng đời của EBOV. Nguồn: <http://www.expasy.org/>

tổn hại[40]

- VP35 là phân tử có nhiều cấu trúc thực nghiệm và nắm giữ nhiều vai trò quan trọng trong quá trình phiên mã và vô hiệu cơ chế phát hiện RNA lạ của cơ thể người[4, 5, 13, 23, 28, 31].

1.2 Cơ chế hoạt động của EBOV

Biết được vòng đời của EBOV, ta có thể đề ra các hướng ngăn chặn EBOV

1.2.1 Quá trình xâm nhập và sao chép của EBOV

EBOV xâm nhập tế bào thành mạch máu và thực hiện quá trình phiên mã, dịch mã, tổng hợp vật chất. Trong quá trình sao chép, protein VP35 sẽ bất hoạt phân tử

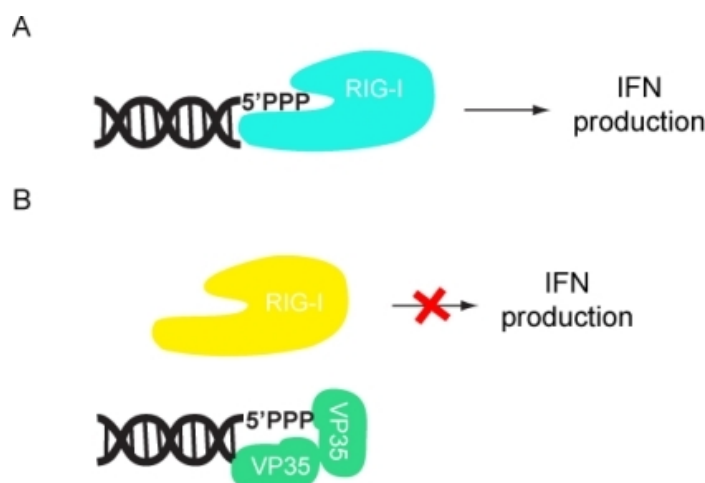
RIG-I (là một loại enzyme có chức năng phát hiện chuỗi vật chất di truyền của các virus cúm A, virus Sendai, họ flavivirus) và vô hiệu hóa hệ thống miễn dịch. Sau khi nhân lên với số lượng lớn, EBOV sẽ giải phóng khỏi tế bào vật chủ và xâm nhập các tế bào khác[6] (xem hình 1.2).

1.2.2 Quá trình hệ thống miễn dịch hoạt động

Phân tử RIG-I có vai trò phát hiện các dsRNA lạ và kích hoạt tế bào sản xuất $\text{IFN}\alpha/\beta$ [22]. Sau khi tế bào bệnh bị phá vỡ do virus nhân lên với số lượng lớn và giải phóng ra ngoài, $\text{IFN}\alpha/\beta$ (protein tín hiệu của tế bào bị nhiễm bệnh) cũng được giải phóng theo và bám vào thụ thể của các tế bào lân cận. Sự bám này là tín hiệu cho hệ thống miễn dịch hoạt động và phá hủy tất cả các tế bào có chứa các tác nhân lạ[9].

1.2.3 VP35 vô hiệu hệ thống miễn dịch

VP35 bám vào dsRNA của EBOV và không cho RIG-I phát hiện dsRNA, do RIG-I phát hiện dsRNA bằng cách liên kết với dsRNA. Theo đó, không có tín hiệu kích hoạt tế bào sản xuất $\text{IFN}\alpha/\beta$ và hệ thống miễn dịch không được báo động khi EBOV phá vỡ thành tế bào và chui ra ngoài[17] (Xem hình 1.3)



Hình 1.3: Quá trình VP35 vô hiệu RIG-I thông qua việc bám vào dsRNA. Nguồn [35].

1.2.4 Các phương án ngăn chặn EBOV

Bằng chứng cho thấy những trường hợp được chữa khỏi EVD có nồng độ Immunoglobulin G (IgG - kháng thể G) và Immunoglobulin T (IgT) cao. Thử nghiệm sử dụng kháng thể trên cơ thể ngựa và chuột đã thành công[15].

Vaccine tiềm năng nhất được phát triển và đưa vào thử nghiệm là VSV-ZEBOV. Bằng cách cấy đoạn RNA mã hoá GP của EBOV vào vesicular stomatitis virus (một loại virus cùng họ với virus dại), GP của EBOV sẽ được tổng hợp khi vesicular stomatitis virus được đưa vào cơ thể và hệ miễn dịch sẽ dễ dàng nhận diện phân tử GP do các phân tử khác của EBOV không được tổng hợp để vô hiệu hệ thống miễn dịch. Vaccine này đã được thử nghiệm thành công đối với động vật có vú[43].

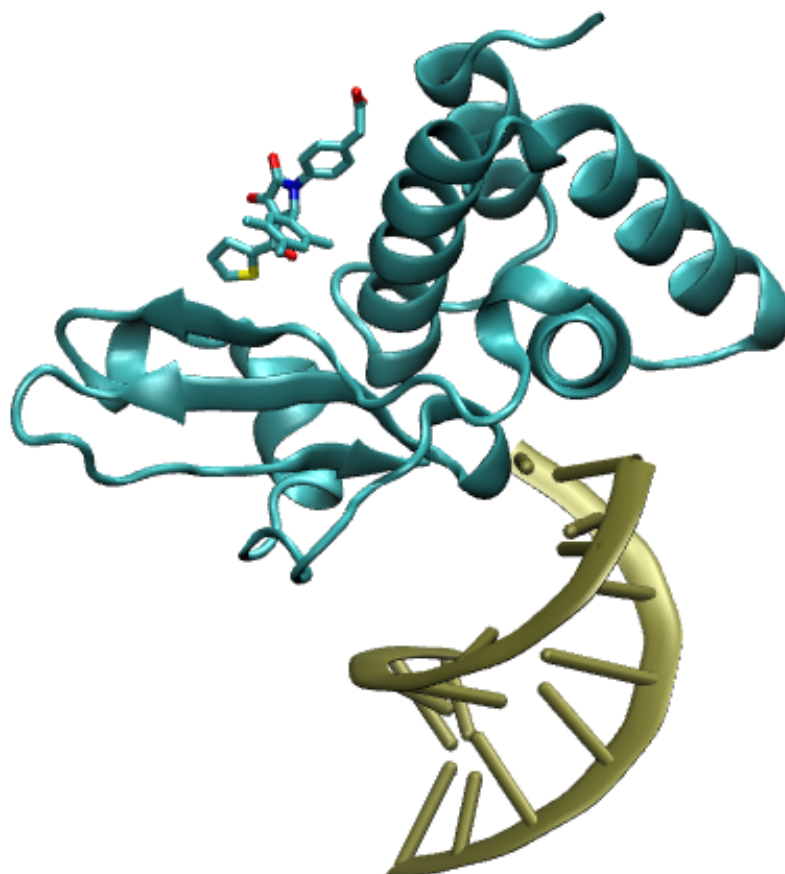
Ngoài ra, việc điều chế thuốc có tác dụng đặc hiệu lên các protein của EBOV cũng là một phương pháp tiềm năng. Các GP của EBOV đôi khi chịu đột biến và các vaccine hoạt động dựa trên việc tổng hợp GP sẽ không còn hiệu quả. Phân tử VP35 là một trong những phân tử có tính bảo tồn cao của họ Filoviridae nên nếu có thể tìm được phân tử thuốc bất hoạt được VP35 thì phân tử thuốc này sẽ có khả năng bất hoạt nhiều chủng virus thuộc họ Filoviridae.

1.3 Các nghiên cứu trên thế giới về VP35

Ngân hàng cấu trúc protein thế giới (Protein Data Bank) lưu trữ các cấu trúc thực nghiệm của VP35. Trong các cấu trúc thực nghiệm thu được, có một cấu trúc của phân tử VP35 chủng Zaire Ebola (ZEBOV) nguyên thủy gắn với dsRNA của chính virus này[34]. Đồng thời, việc đột biến các vị trí amino acid thứ 312, 322, 339 đã được xác nhận là làm mất hoàn toàn chức năng gắn vào dsRNA của VP35[22, 17].

Có 9 phân tử nhỏ (ligand) đã được tiên đoán là có khả năng tương tác tốt với VP35 và có tiềm năng trở thành thuốc đặc trị[39, 42] (xem hình 1.4). 9 phân tử này có PDBID lần lượt là: 1DK, 12G, 1DL, 11Y, 1D5, 1D6, 1D8, 1D9, 1DE. Các phân tử này can thiệp vào vùng tương tác với nucleoprotein (NP) của EBOV và gây cản trở quá trình nhân lên của virus trong điều kiện phòng thí nghiệm (*in vitro*).

Cho đến nay, chưa có nghiên cứu mô phỏng nào đã được tiến hành để khảo sát



Hình 1.4: Vị trí tương đối của ligand 1DK so với VP35 (màu xanh da trời) và dsRNA (màu vàng) đã được công bố trên [39].

tính chất động học của tương tác dsRNA–VP35 của virus Ebola. Việc khảo sát động học tương tác dsRNA–VP35 sẽ giúp tìm ra những amino acid nào thực sự quan trọng cho sự gắn kết giữa hai phân tử này. Ngoài ra, có một nghiên cứu mô phỏng để khảo sát các tính chất động học của dsRNA–VP35 của virus Marburg, một trong số rất ít các virus thuộc họ Filoviridae[41].

Chương 2

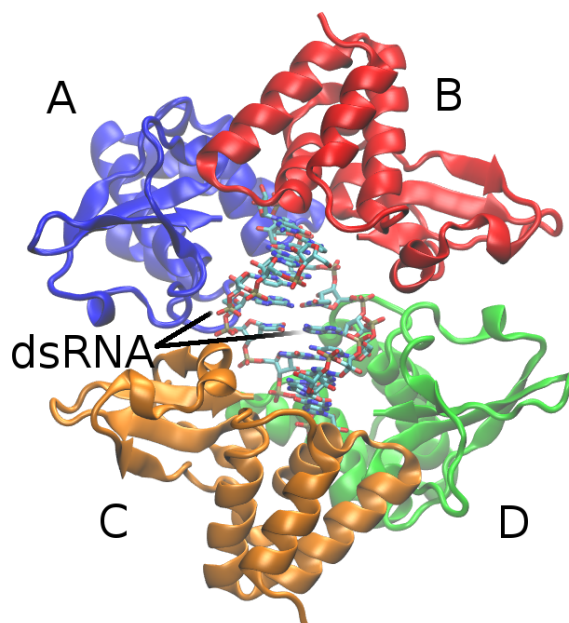
Đối tượng và phương pháp nghiên cứu

Khóa luận tập trung nghiên cứu cấu hình có PDBID là 3L25 của phân tử VP35, sử dụng phương pháp SMD để đánh giá tương tác giữa VP35 và dsRNA. Đây là cấu trúc thực nghiệm duy nhất của VP35 không đột biến gắn với dsRNA.

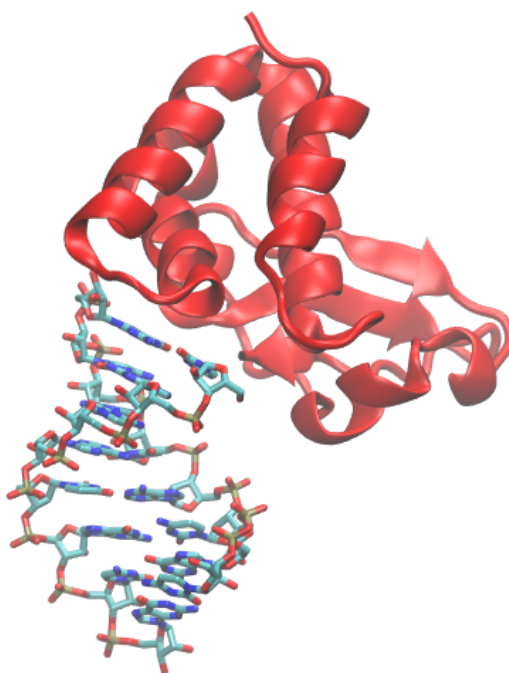
2.1 Cấu hình của VP35

2.1.1 Cấu hình 3L25

Gồm chuỗi dsRNA dài 8 đơn vị (8bps) liên kết với 4 phân tử VP35 (chuỗi A, B, C và D). Chuỗi A và B tạo thành một nhóm, với chuỗi B gắn với đầu chuỗi dsRNA và chuỗi A gắn với dsRNA dọc theo chiều dài chuỗi. Tương tự với chuỗi C và D. Chuỗi A và C hình thành tương tác Van der Waals với dsRNA. Chuỗi B và D xuất hiện liên kết nhánh phụ trực tiếp với nhóm phosphodiester của dsRNA[34] (Xem 2.1a). Tuy nhiên, để khảo sát cấu trúc 3L25, chúng tôi chọn khảo sát cấu hình chỉ gồm chuỗi B và dsRNA (B-conf) vì VP35 tại vị trí này đóng vai trò chính trong việc gắn vào dsRNA[32]. Chuỗi B này còn được gọi là chuỗi "end-cap" vì nó chặn đầu của dsRNA[34].



(a) 4 đơn vị phân tử VP35 bám vào dsRNA được tô màu xanh dương, đỏ, cam, xanh lá cây.



(b) Chuỗi VP35 B bám vào dsRNA dài 8bps được tô màu đỏ.

Hình 2.1: Cấu trúc VP35 được khảo sát.

2.1.2 B-conf

Gồm chuỗi B và dsRNA. B-conf giúp phân tích tương tác tại đầu chuỗi dsRNA của VP35. Đây là phân tử VP35 tương tác chính với dsRNA và có vai trò quan trọng trong tương tác giữa hệ các phân tử VP35 và dsRNA. Xem hình 2.1b

2.2 Phương pháp Mô phỏng động học phân tử

Mô phỏng động lực học phân tử (Molecular Dynamics - MD) là một phương pháp hữu ích giúp tìm hiểu các tính chất vật lý của hệ phân tử sinh học.

2.2.1 Ý tưởng chính

Khi khảo sát một hệ nhiều hạt, người ta sử dụng các quy luật thống kê do việc gần như không thể tìm được lời giải giải tích cho mọi tính chất của hệ. Với điều kiện hệ nhiều hạt này tuân theo giả thiết ergodic, việc khảo sát hệ theo thời gian có thể cho ta bức tranh sơ bộ về không gian pha của hệ.

MD cho biết quỹ đạo của từng nguyên tử trong hệ một cách cổ điển, tức là xem mỗi nguyên tử là mô hình một "hạt" tuân theo phương trình chuyển động Newton. Sau khi biết được quỹ đạo theo thời gian, các thông số như vận tốc, động năng, nhiệt độ,... sẽ được tính toán gián tiếp.

Cụ thể, các bước thuật toán dùng để mô phỏng động học phân tử gồm:

1. Đọc các tham số ban đầu: nhiệt độ, bước tính δt , nhóm cố định (restrained groups),...
2. Chuẩn bị hệ: sử dụng tham số vận tốc, tọa độ đã có từ mô phỏng MD trước đó hoặc ngẫu nhiên chọn tọa độ và vận tốc đầu thỏa phân bố Boltzmann.
3. Tính toán lực tác dụng lên tất cả các nguyên tử, sử dụng trường lực (sẽ trình bày ở phần 2.2.2).
4. Tích phân các phương trình chuyển động Newton. Thuật toán sử dụng thường là tích phân Verlet được trình bày ở phần 2.2.3. Trong bước này, các hiệu chỉnh

cần thiết để kiểm soát nhiệt độ, áp suất, liên kết hoá học,... cũng được sử dụng. Khoá luận sẽ trình bày thuật toán kiểm soát nhiệt độ ở phần 2.2.4. Các phương trình chuyển động cần tính:

$$ma_i = \sum_i^N F_i \quad (2.1)$$

$$v_i = a_i \times dt \quad (2.2)$$

$$\vec{x}_i = v_i \times dt \quad (2.3)$$

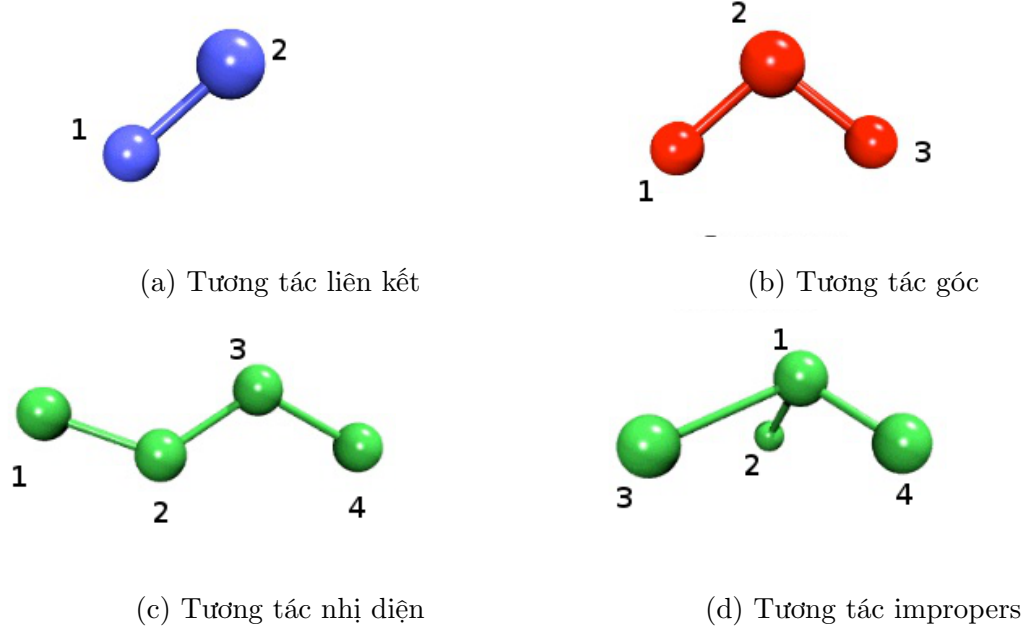
5. Sau khi tính tích phân phương trình chuyển động, tính các đại lượng ta quan tâm như năng lượng, áp suất, nhiệt độ,... và ghi kết quả.

2.2.2 Trường lực

Trường lực AMBER99SB (amber99sb ff) là tên gọi của trường lực đặc trưng cho tương tác giữa các nguyên tử trong hệ. Tuy có tên gọi là "trường lực" nhưng nó không phải là đại lượng có thứ nguyên của lực. Trường lực (force field - ff) chính là hàm thế năng tương tác giữa các phân tử trong hóa học và sinh học tính toán. Sử dụng gần đúng cổ điển, người ta thường bỏ qua các ảnh hưởng gây ra bởi chuyển động của electron và xét năng lượng của hệ nguyên tử như hàm của các "hạt". Có rất nhiều ff được xây dựng. Các trường lực thường xuyên được sử dụng nhất được mang tên AMBER[14, 2, 24, 16, 19], GROMOS[21, 20, 36, 8, 2], CHARMM[19, 2], OPLS[19, 2]. Với mỗi trường lực được phát triển còn có nhiều phiên bản khác nhau với mỗi ưu điểm riêng.

Khoá luận đã sử dụng phiên bản amber99sb ff để mô phỏng tương tác giữa các phân tử vì đây là trường lực có rất nhiều ưu điểm trong việc mô tả hệ protein-RNA[24]. Amber99sb ff có đầy đủ tham số tương tác cho các nguyên tử trong phân tử RNA, protein, kể cả hydro không phân cực. Trong khi đó các trường lực united-atom xem các nguyên tử Hydro và Carbon trong mỗi methyl terminal và mỗi methylene bridge như một điểm tương tác duy nhất. Để mô phỏng thời gian dài, các trường lực thô (coarse-grained) được sử dụng.

Amber99sb ff được mô tả bởi một hàm thế như sau:



Hình 2.2: Tương tác giữa các nguyên tử có liên kết

$$\begin{aligned}
 V(r^N) = & \sum_{ij}^{bond} k_{ij} (r_{ij} - r_{ij}^0)^2 + \sum_{ijkl}^{dihedrals} k_{ijkl} [1 + \cos(n_{ijkl}\phi_{ijkl} - \phi_{ijkl}^0)] + \\
 & + \sum_{ijkl}^{improper} \frac{1}{2} k_{\xi} [\xi_{ijkl} - \xi_{ijkl}^0]^2 + \sum_{ijk}^{angles} k_a (\theta_{ijk} - \theta_{ijk}^0)^2 + \sum_{ij} \left[\frac{q_i q_j}{\epsilon r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right]
 \end{aligned} \tag{2.4}$$

Bốn tổng đầu (bốn tổng liên kết - bonded), lần lượt được lấy tổng theo:

- tất cả các liên kết trực tiếp (tương tác 1-2, xem hình 2.2a)
- tất cả các góc (tương tác 1-3, xem hình 2.2b)
- tất cả các góc nhị diện (tương tác 1-4, xem hình 2.2c)
- tất cả các impropers (tương tác 1-4, xem hình 2.2d)

Trong đó, tổng liên kết trực tiếp và các số hạng góc được biểu diễn như các dao động tử điều hòa đơn giản. Hai nguyên tử xem như dao động quanh một vị trí cân bằng.

Các đại lượng k_{ij}, r_{ij}^0, r_{ij} lần lượt là hệ số đàn hồi, độ dài liên kết ở vị trí cân bằng và độ dài liên kết thực tế giữa hai nguyên tử có liên kết hóa trị với nhau i và j .

Trong tổng theo các góc, k_{ijk} , θ_{ijk}^0 , θ_{ijk} , lần lượt là hệ số đàn hồi, góc cân bằng và góc thực tế tạo bởi các nguyên tử i , j , k .

Trong tổng theo góc nhị diện, k_{ijkl} , ϕ_{ijkl}^0 , ϕ_{ijkl} lần lượt là hằng số lực góc nhị diện, góc nhị diện cân bằng và góc nhị diện thực tế tạo bởi các nguyên tử i , j , k , l . Hệ số n_{ijkl} là hằng số đối xứng nhận giá trị 1, 2 hoặc 3 ứng với số vị trí cực tiểu khi cho góc nhị diện quay một góc 360° .

Tương tự, trong tổng theo các impropers, k_ξ , ξ_{ijkl}^0 , ξ_{ijkl} lần lượt là hằng số lực improper, góc improper cân bằng và góc improper thực tế tạo bởi các nguyên tử i , j , k , l .

Các tổng cuối trong phương trình (2.4) là các tổng không liên kết (non-bonded). Các tổng này bao gồm tương tác Van der Waals và tương tác tĩnh điện. Tương tác Van der Waals được mô tả bởi thế Lennard-Jones (6-12) và tương tác tĩnh điện được mô tả bởi thế Coulomb với giả thiết là điện tích điểm nằm tại tâm của nguyên tử [16]. Hằng số điện môi ϵ đặc trưng cho môi trường giữa hai nguyên tử, tham số B_{ij} và A_{ij} là các tham số 6-12 đặc trưng cho từng nguyên tố.

Tất cả các hằng số lực, các tham số ứng với tọa độ cân bằng, các tham số 6-12 và hệ số đối xứng n_{ijkl} được dẫn ra từ các tính toán lượng tử với độ chính xác cao và từ thực nghiệm[16].

Khi biết được thông tin chi tiết về trường lực, lực tác dụng lên mỗi hạt sẽ được tính theo:

$$\vec{F}_i = -\vec{\nabla}V_i \quad (2.5)$$

Giá trị lực này được dùng để giải phương trình chuyển động 2.1.

2.2.3 Thuật toán Verlet sử dụng trong mô phỏng MD

Trong quá trình giải phương trình chuyển động Newton bằng phương pháp số, có một thủ thuật nhằm tính đại lượng tọa độ của hệ hạt với bậc chặt cụt cao hơn so với những phương pháp thông thường.

Tích phân Verlet[1] là một phương pháp tính toán số được sử dụng trong phương trình chuyển động Newton. Nó thường dùng để tính quỹ đạo của hạt trong mô phỏng động học và đồ họa máy tính. Thuật toán được phát triển bởi Delambre (1791) và

được Verlet sử dụng nhiều trong động học phân tử những năm 60. Nó cũng được dùng để tính toán quỹ đạo sao chổi Halley.

Tích phân Verlet có tính hội tụ cao, có tính chất suy ngược (time-reversible), tính bảo toàn của hệ ngẫu đối của không gian pha và không tốn quá nhiều bước tính so với thuật toán Euler.

Thay vì tính trực tiếp phương trình (2.2) để tìm tọa độ \vec{x}_i , sử dụng khai triển Taylor đến bậc 3 của tọa độ hạt theo thời gian ta được hai phương trình sau:

$$\vec{x}(t + \Delta t) = \vec{x}(t) + \vec{v}(t) \Delta t + \frac{\vec{a}(t) \Delta t^2}{2} + \frac{\vec{b}(t) \Delta t^3}{6} + \mathcal{O}(\Delta t^4) \quad (2.6)$$

$$\vec{x}(t - \Delta t) = \vec{x}(t) - \vec{v}(t) \Delta t + \frac{\vec{a}(t) \Delta t^2}{2} - \frac{\vec{b}(t) \Delta t^3}{6} + \mathcal{O}(\Delta t^4) \quad (2.7)$$

Cộng hai vế hai phương trình trên, rút ra được:

$$\vec{x}(t + \Delta t) = 2\vec{x}(t) - \vec{x}(t - \Delta t) + \vec{a}(t) \Delta t^2 + \mathcal{O}(\Delta t^4) \quad (2.8)$$

Từ phương trình (2.8), có thể thấy rằng tích phân Verlet đem lại bậc sai số cao hơn nhiều so với phương pháp Euler.

2.2.4 Thuật toán kiểm soát nhiệt độ

Để có thể mô phỏng (i) hệ phân tử lớn trong (ii) thời gian dài sao cho kết quả mô phỏng phù hợp với thực nghiệm, người ta tìm các thuật toán để có thể kiểm soát các đại lượng nhiệt động như mong muốn. Cụ thể, người ta áp dụng mô hình Vi chính tắc, cùng các điều kiện ràng buộc tương ứng để mô phỏng hệ (số hạt, thể tích, năng lượng tổng của hệ không đổi theo thời gian). Nhưng mô hình Vi chính tắc này vẫn chưa đem lại những quan sát có được trong thực nghiệm, người ta sử dụng mô hình Chính tắc để mô tả hệ với mong muốn nhận được những kết quả phù hợp với thực nghiệm hơn. Thuật toán đơn giản có thể kiểm soát nhiệt độ của hệ thông qua kiểm soát vận tốc của hạt trong hệ được gọi chung là V-rescale.

Có nhiều phương pháp khác nhau để kiểm soát nhiệt độ của hệ. Khóa luận chỉ sử dụng phương pháp Berendsen và Nosé-Hoover nên sẽ chỉ trình bày sơ lược hai phương pháp này[27].

Một cách tổng quát, kiểm soát nhiệt độ thông qua kiểm soát vận tốc hạt trong hệ dựa trên mối liên hệ giữa nhiệt độ và vận tốc hạt được trình bày trong phương trình

(2.9) sau:

$$\sum_{i=1}^N \frac{\vec{p}_i}{2m_i} = \frac{k_b T}{2} (3N - N_c) \quad (2.9)$$

với $3N - N_c = N_{df}$ là số bậc tự do. Để thực hiện được việc điều chỉnh nhiệt độ về một giá trị cho trước thông qua điều chỉnh ΔT , ta sử dụng tham số λ như sau:

$$\Delta T = \frac{1}{2} \sum 2 \frac{m_i (\lambda v_i)^2}{N_{df} k_B} - 2 \frac{m_i v_i^2}{N_{df} k_B} \quad (2.10)$$

$$\Delta T = (\lambda^2 - 1) T(t) \quad (2.11)$$

$$\lambda = \sqrt{\frac{T_0}{T(t)}} \quad (2.12)$$

Chỉ việc nhân vận tốc hạt trong mỗi bước lặp với λ vừa tính được ở bước liền trước thì vận tốc của các hạt sẽ dần dần phân bố một cách tương ứng với nhiệt độ T_0 .

Việc kiểm soát nhiệt độ một cách chặt chẽ bởi tham số λ khiến cho thăng giáng nhiệt bị bóp hẹp và cần thiết sử dụng các phương pháp hiệu chỉnh ở cấp độ cao hơn.

Phương pháp Berendsen

Ta có thể hình dung sơ lược về phương pháp này như sau:

- Phương pháp Berendsen chủ yếu được dùng để đưa hệ về một nhiệt độ ổn định, nhưng hệ không có tính chất thăng giáng nhiệt của hệ Chính tắc.
- Sau mỗi bước lặp, vận tốc của hạt trong hệ thay đổi và làm thay đổi nhiệt độ của hệ một lượng ΔT . Thuật toán sẽ giúp điều chỉnh độ thay đổi này sao cho nhiệt độ ổn định quanh giá trị T_0 .
- Bằng cách hiệu chỉnh như trên, có thể thực hiện quá trình mô phỏng trong thời gian dài mà vẫn tránh được dịch chuyển mức năng lượng sinh ra bởi sai số của phép tính toán số.
- Hệ mô phỏng có thể được chọn một cách tối giản để giảm thiểu tài nguyên tính toán, do sai số đã được giảm thiểu.

Thay vì tính tham số λ một cách trực tiếp như trước, người ta tìm cách để nhiệt độ T có thể thăng giáng quanh một giá trị với biên độ thăng giáng có thể đánh giá được

qua tham số τ như sau:

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_0 - T(t)) \quad (2.13)$$

Ở đây, τ được gọi là tham số cặp nhiệt độ (T-coupling). Gọi như vậy vì nó cho thấy tương tác giữa hệ và bể nhiệt trao đổi nhiệt với nhau nhanh hay không. Với cách kiểm soát nhiệt độ như vậy, nhiệt độ của hệ hội tụ theo hàm số mũ và thu được hệ số λ cho vận tốc hạt như sau:

$$\lambda^2 = 1 + \frac{\delta t}{\tau} \left[\frac{T_0}{T(t - \delta t)} - 1 \right] \quad (2.14)$$

Hiệu chỉnh τ một cách kỹ lưỡng sẽ cho phép nhiệt độ T tiến đến giá trị T_0 cho trước với thang giáng nhiệt thu được từ phương trình sau:

$$\Delta T = \frac{\delta t}{\tau} (T_0 - T(t)) \quad (2.15)$$

Với $t \rightarrow \infty$ thì hệ hạt có các tính chất của hệ Vi chính tắc, không có hiệu chỉnh Berendsen tham gia vào quá trình mô phỏng. Nhưng với giá trị $\tau \sim \delta t$ thì hệ số λ trở về với giá trị như đã trình bày ở phần tổng quát của V-rescale.

Trong chương trình mô phỏng mà đề tài sử dụng, τ được hiệu chỉnh về giá trị $0.4\delta t$.

Từ phương trình (2.13) có thể thấy rằng nhiệt độ T của hệ tiến về giá trị T_0 theo hàm mũ. Như vậy, hệ sẽ tiến về mô hình Vi chính tắc, thay vì tiến về mô hình Chính tắc mà ta mong muốn.

Phương pháp Nosé-Hoover

Để khắc phục các nhược điểm của phương pháp Berendsen, Nosé phát triển một phương pháp khác nhằm mở rộng hệ bằng cách xem bể nhiệt như một phần bổ sung của hệ hạt đang xét, nhờ thêm vào biến phụ \bar{s} , cũng như giả khối lượng Q ($\gg 0$) và giả vận tốc $\dot{\bar{s}}$. Cụ thể, biến thời gian được hiệu chỉnh lại như sau:

$$d\bar{t} = \bar{s} dt \quad (2.16)$$

Các biến tọa độ không gian của hệ hạt vẫn không bị thay đổi, theo đó:

$$\vec{r} = \vec{r} \quad (2.17)$$

$$\vec{\dot{r}} = \bar{s}^{-1} \dot{\vec{r}} \quad (2.18)$$

$$\bar{s} = s^{-1} \quad (2.19)$$

$$\dot{\bar{s}} = \bar{s}^{-1} \dot{s} \quad (2.20)$$

Sử dụng bốn hệ thức biến đổi tọa độ trên, ta có thể xây dựng được Lagrangian:

$$\mathcal{L} = \sum_i \frac{m_i}{2} \bar{s}^2 \dot{\vec{r}}_i^2 - U(\vec{r}) + \frac{1}{2} Q \dot{\bar{s}}^2 - g k_B T_0 \ln \bar{s} \quad (2.21)$$

với $g = N_{df}$. Quan sát phương trình (2.21) dễ dàng thấy được số hạng động năng bây giờ gồm tổng động năng có được do chuyển động của hạt và động năng sinh bởi biến phụ \bar{s} và thế năng gồm thế năng thực của hệ hạt cộng với thế năng phát sinh bởi biến phụ \bar{s} . Theo đó, năng lượng thực của hệ hạt mà ta đang quan sát (gồm động năng và thế năng của hệ hạt) không còn là hằng số mà thay đổi theo \bar{s} . Điều đó có nghĩa là sử dụng Lagrangian vừa xây dựng được, ta có thể mô tả hệ hạt trao đổi nhiệt với bể nhiệt.

Phương trình chuyển động có dạng:

$$\ddot{\vec{r}}_i = \frac{\vec{F}_i}{m_i \bar{s}^2} - \frac{2 \dot{\bar{s}} \dot{\vec{r}}_i}{\bar{s}} \quad (2.22)$$

$$\ddot{\bar{s}} = \frac{1}{Q \bar{s}} \left(\sum_i m_i \bar{s}^2 \dot{\vec{r}}_i^2 - g k_B T_0 \right) \quad (2.23)$$

Đây là phương trình chuyển động của hệ hạt trong hệ tọa độ suy rộng $(\vec{r}, \bar{p}, \bar{t})$. Tuy nhiên, (2.23) là phương trình đạo hàm bậc 2 theo thời gian và \bar{s} có xu hướng dao động tuần hoàn, nên thăng giáng nhiệt cũng mang tính chu kỳ.

Nosé và Hoover đã tìm cách biến đổi lại phương trình (2.22) trở về phương trình sử dụng hệ tọa độ thực, và phương trình (2.23) trở về phương trình đạo hàm bậc nhất theo thời gian, sử dụng các hệ thức biến đổi sau:

$$s = \bar{s}, \quad \dot{s} = \bar{s} \dot{\bar{s}}, \quad \ddot{s} = \bar{s}^2 \ddot{\bar{s}} + \bar{s} \dot{\bar{s}}^2, \quad (2.24)$$

$$\vec{r} = \vec{r}, \quad \dot{\vec{r}} = \bar{s} \dot{\vec{r}}, \quad \ddot{\vec{r}} = \bar{s}^2 \ddot{\vec{r}} + \bar{s} \dot{\vec{r}}^2 \quad (2.25)$$

và thay

$$\gamma = \frac{\dot{s}}{s} \quad (2.26)$$

phương trình chuyển động (2.23) được viết lại thành:

$$\ddot{\vec{r}}_i = \frac{\vec{F}_i}{m_i} - \gamma \vec{r}_i, \quad (2.27)$$

$$\dot{\gamma} = \frac{-k_B N_{df}}{Q} T(t) \left(\frac{T_0}{T(t)} - 1 \right) \quad (2.28)$$

Đặt $\tau_{relax} = \frac{Q}{N_{df} k_B T_0}$ là thời gian phục hồi đặc trưng cho chu kỳ dao động nhỏ của biến \bar{s} thì phương trình (2.28) có thể viết lại:

$$\dot{\gamma} = -\frac{1}{\tau_{relax}} \left(\frac{T_0}{T(t)} - 1 \right) \quad (2.29)$$

Trong phương trình (2.28), khối lượng ảo Q cần được chọn một cách phù hợp. Hệ trở thành hệ Vi chính tắc nếu cho $Q \rightarrow \infty$. Nếu chọn Q quá lớn, sẽ phải mô phỏng trong thời gian dài mới đạt được phân bố Chính tắc. Ngược lại, chọn Q quá nhỏ sẽ dẫn đến tần số nhiễu loạn của hệ hạt trong không gian thực là rất lớn, và ta không quan sát được sự trao đổi động năng giữa bể nhiệt và hệ phân tử đang quan sát.

Trong thực tế, hệ phân tử mô phỏng thường có khối lượng khác nhau tùy theo nguyên tố. Để chọn được Q , cần thiết phải chọn giá trị Q với mỗi nhóm phân tử khác nhau để thời gian mô phỏng là tối thiểu mà hệ vẫn đạt được phân bố chính tắc. Cách chọn Q theo nhóm được trình bày cụ thể hơn trong phần 2.4

2.3 Phương pháp mô phỏng Steered Molecular Dynamics (SMD)

Chính là MD nhưng đặt thế năng phụ thuộc vào thời gian lên các nhóm phân tử xác định để tách các nhóm này ra. Việc tách hai nhóm phân tử ra có tác dụng hỗ trợ kiểm tra giả thiết về chuyển trạng thái của hệ phân tử. Một cách cụ thể, SMD áp dụng đối với phối tử và thụ thể sẽ giúp quan sát quá trình hai nhóm phân tử này tách ra khỏi nhau. Đây là một quá trình xảy ra trong thực tế, tuy nhiên cần khảo sát hệ trong thời gian rất dài mới có thể ngẫu nhiên quan sát được quá trình này. Do đó, SMD được sử dụng với mục đích giảm thiểu thời gian mô phỏng mà vẫn giúp quan sát hệ chuyển trạng thái ra sao.

Có hai cách xác định tham số của thế năng, tương ứng với hai cách "kéo" các nhóm phân tử ra: kéo với lực kéo không đổi hoặc kéo nhóm tách ra với vận tốc không đổi. Đề tài sử dụng phương pháp kéo với vận tốc không đổi[3]. Trong suốt quá trình kéo, dễ dàng để tính toán lực kéo các nhóm. Từ đó phác thảo được tương tác giữa các nhóm phân tử. Hơn nữa, có thể xác định được độ thay đổi năng lượng khi hệ phân tử chuyển từ trạng thái trước khi kéo sang trạng thái sau khi kéo.

2.3.1 SMD sử dụng lực kéo không đổi

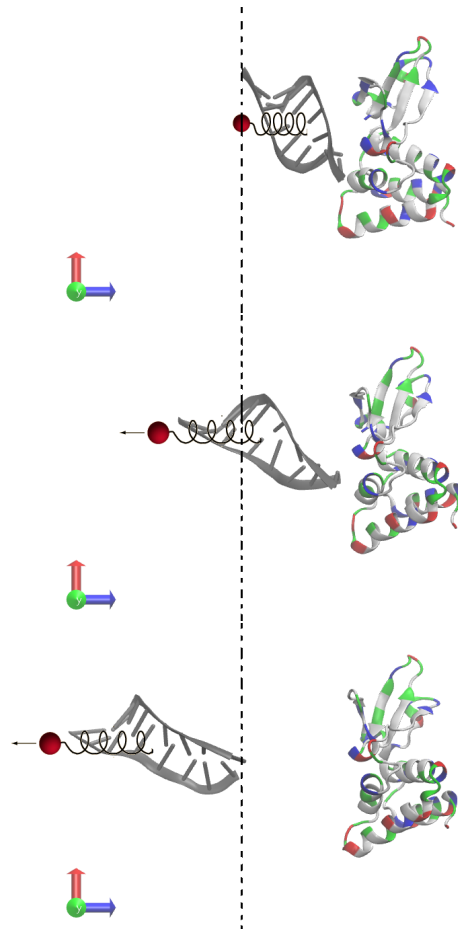
Bằng cách cộng thêm lực kéo với độ lớn không đổi vào nhóm các phân tử cần kéo và thực hiện quá trình mô phỏng. Đối với nhóm phân tử không cần kéo, cần cố định một cách tương đối vị trí của chúng, nhằm tránh việc bị kéo theo do tương tác với nhóm phân tử được kéo ra. Thực hiện SMD nhiều lần với giá trị lực kéo khác nhau giúp xác định lực tương tác giữa hai nhóm phân tử nằm trong khoảng giá trị nào. Qua đó, có thể đánh giá sơ bộ tương tác giữa hai nhóm phân tử (ví dụ phối tử và thụ thể được nêu ở trên).

2.3.2 SMD sử dụng vận tốc kéo không đổi

Có thể hình dung về cách kéo nhóm phân tử như hình 2.3 được trình bày dưới đây: Thế năng đàn hồi đặt lên nhóm phân tử cần kéo có dạng như sau:

$$U(\vec{r}) = \frac{1}{2}k[vt - (\vec{r} - \vec{r}_0) \cdot \vec{n}]^2 \quad (2.30)$$

với k là hệ số đàn hồi, $v = constant$ là tốc độ di chuyển của nhóm liên kết, \vec{r} là vị trí của hạt được



Hình 2.3: Bằng cách di chuyển nhóm liên kết (constraint group) với vận tốc không đổi, lực tác động lên nhóm phân tử cần kéo ra sẽ lớn dần đến một giá trị đủ để tách nhóm đó ra.

kéo và \vec{r}_0 là vị trí ban đầu của hạt, \vec{n} là hướng kéo. Dưới tác dụng của thế đàn hồi, hạt sẽ chịu một lực kéo là $\vec{F} = -\vec{\nabla}U = k(\vec{r} - \vec{r}_0 - \vec{v} \times t)$

Cách kéo nhóm phân tử này có ưu điểm so với áp dụng lực kéo không đổi là không phải thực hiện rà soát quá nhiều giá trị lực kéo để ước lượng lực tương tác giữa hai nhóm phân tử. Hơn nữa, công trung bình tác dụng lên hệ sẽ tiến đến một giá trị xác định, do nhóm liên kết di chuyển với vận tốc không đổi và công tác dụng lên hệ (một cách tương đối) sẽ bằng năng lượng liên kết giữa hai nhóm và thế năng đàn hồi $U(\vec{r})$ trung bình và động năng trung bình của nhóm phân tử kéo. Nói cách khác, hệ sẽ ít bị can thiệp một cách "thô bạo" hơn so với cách kéo sử dụng lực không đổi được trình bày ở trên.

2.4 Tham số sử dụng trong các quá trình mô phỏng

Mô phỏng đã được thực hiện 3 lần đối với mỗi phân tử 3L25 (VP35 tự nhiên), 3L27 (đột biến Arg312Ala), 3L28 (đột biến Lys339Ala) và phân tử VP35 đột biến Lys282Ala.

Hệ gồm 200000 nguyên tử. Mô hình phân tử nước được sử dụng là TIP-4P [18]. Hệ phân tử được phân làm hai nhóm khi đưa nhiệt độ vào hệ thông qua khối lượng ảo Q tương ứng: Protein và không phải Protein (Protein group & non-Protein group). Các bước chạy mô phỏng được thực hiện:

1. Cực tiểu hóa năng lượng hệ phân tử
2. Mô phỏng trạng thái (NVT) của hệ không đổi
3. Mô phỏng trạng thái (NPT) của hệ không đổi
4. Mô phỏng SMD. Các bước trước đó được xem như các bước thiết lập hệ mô phỏng thể thực hiện mô phỏng SMD

Cực tiểu hóa năng lượng

Nhằm mục đích để hệ hạt có vị trí ổn định nhất và không có những "va đập" quá mạnh. Cụ thể, hệ phân tử phân bố trong không gian sao cho thế năng tổng

của hệ là thấp nhất. Điều kiện dừng chương trình: lực tối đa tác dụng lên hạt bằng 1000 kJ/mol/nm . Cách tính tương tác Coulomb ở khoảng cách xa: Particle Mesh Ewald (PME). Bán kính chặt cắt đối với tương tác Coulomb và Van der Waals là $r_{Coulomb} = 1 \text{ nm}$ và $r_{VDW} = 1 \text{ nm}$.

Cân bằng chính tắc (NVT)

Các thông số bán kính chặt cắt và phương pháp tính tương tác Coulomb ở khoảng cách xa vẫn như trên. Áp dụng phương pháp Berendsen để ổn định nhiệt độ của hệ tại $300K$ với $\tau_t = 0.1 \text{ (ps)}$. Vận tốc của hạt được đưa vào và ổn định một cách ngẫu nhiên sao cho thoả phân bố Maxwell. Áp dụng thuật toán Verlet để giải phương trình chuyển động. Thời gian mô phỏng bước cân bằng NVT là 100 (ps) .

Cân bằng đẳng nhiệt - đẳng tích (NPT)

Các thông số vẫn tương tự như trên. Để ổn định áp suất của hệ tại 1 (bar) , sử dụng phương pháp Parrinello-Rahman với $\tau_p = 2 \text{ (ps)}$, hằng số nén $\beta = -\frac{1}{V} \frac{\partial V}{\partial p} = 4.5 \times 10^{-5} \text{ (bar}^{-1}\text{)}$. Thời gian mô phỏng bước cân bằng NPT là 100 (ps) .

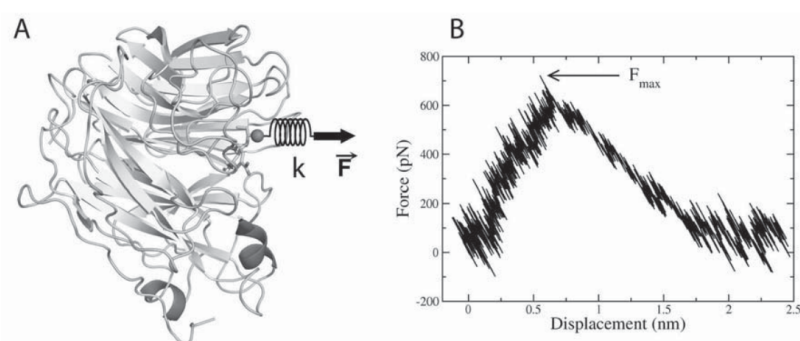
SMD

Bán kính chặt cắt được sử dụng trong bước này là 1.4 (nm) . Tốc độ kéo của nhóm liên kết (constraint group) bằng 5 (nm/ns) . Hệ số đàn hồi (độ cứng) được sử dụng để kéo là $k = 1000 \text{ (kJ.mol}^{-1}\text{nm}^{-2}\text{)}$.

2.5 Các công cụ phân tích kết quả

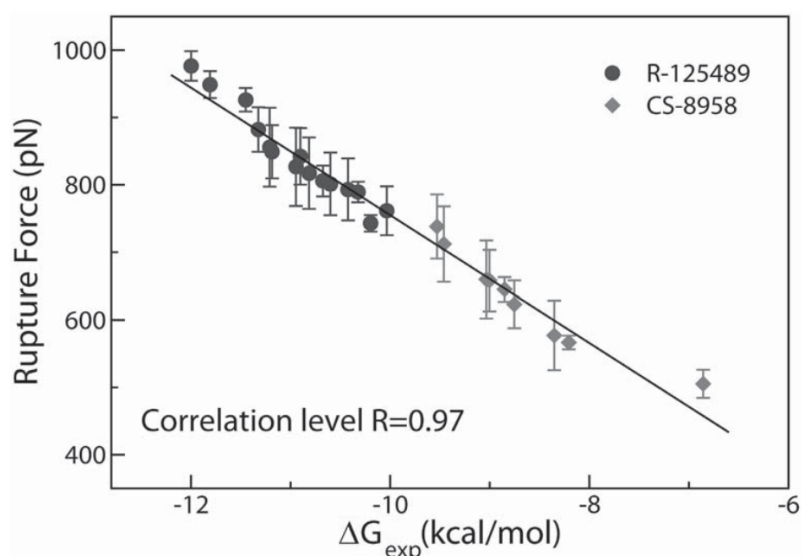
2.5.1 Lực kéo tối đa

Là lực lớn nhất dùng để kéo hai phân tử ra trong suốt quá trình mô phỏng. Do hệ phân tử đã ở vị trí ứng với đáy của giếng thế địa phương, nên để tách hai nhóm phân tử protein và dsRNA ra thì lực dùng để tách hai nhóm ra sẽ lớn dần đến một giá trị mà tại đó hai nhóm không còn tương tác đáng kể với nhau nữa (Xem hình 2.4). Việc so sánh lực kéo tối đa giữa các nhóm phân tử giúp đánh giá tương đối tương tác giữa



Hình 2.4: (A) kéo ligand ra khỏi protein bằng cách đặt một thế năng điều hoà thay đổi theo thời gian, tương đương với việc đặt một lực kéo của một lò xo với hệ số đàn hồi k lên ligand. (B) Đường biểu diễn lực kéo theo tọa độ. Lực kéo tối đa F_{max} đặc trưng cho ái lực liên kết giữa ligand và protein. Nguồn [37]

các nhóm phân tử này. Điều này đặc biệt có ý nghĩa khi so sánh lực kéo các cấu hình đột biến một vị trí amino acid để tìm hiểu vai trò của nó trong tương tác giữa các nhóm. Quan sát đồ thị 2.5 có thể thấy mối tương quan giữa lực kéo tối đa để tách



Hình 2.5: Đồ thị tương quan giữa lực kéo tối đa (Rupture Force) và năng lượng tương tác giữa phân tử R-125489 với thụ thể NA của virus cúm và giữa phân tử CS-8959 với thụ thể NA của virus cúm [37]. Hệ số tương quan (correlation level) $R = 0.97$

hai nhóm phân tử và năng lượng tương tác giữa hai nhóm phân tử. Đường biểu diễn gần đúng cho thấy năng lượng tương tác giữa hai nhóm phân tử tỷ lệ tuyến tính với lực kéo tối đa để tách hai nhóm phân tử này. Như vậy, việc khảo sát lực kéo tối đa

giữa các phân tử VP35 không đột biến và VP35 đột biến ở các vị trí amino acid xác định với dsRNA giúp so sánh tốt năng lượng liên kết của VP35 không đột biến và VP35 đột biến với dsRNA.

2.5.2 Thế năng tương tác theo thời gian

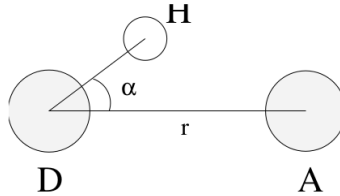
Từ các tham số tọa độ và từ amber99sb ff đã được định nghĩa, có thể dễ dàng tính lại được đóng góp của tương tác Coulomb hoặc tương tác Van der Waals vào tương tác của một vị trí amino acid. Điều này giúp đánh giá một cách tương đối vai trò của các vị trí amino acid, đồng thời giúp dự đoán các amino acid có khả năng đóng góp cho tương tác giữa dsRNA và VP35. Một cách cụ thể, thành phần tương tác Coulomb có dạng:

$$V_i(t) = \sum_j \left[\frac{q_i q_j}{\epsilon r_{ij}} \right] \quad (2.31)$$

và thành phần tương tác theo thể Lennard-Jones có dạng:

$$V_i(t) = \sum_j \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right] \quad (2.32)$$

2.5.3 Liên kết Hydro



Hình 2.6: Khoảng cách và góc giữa phân tử Donor, Acceptor và nguyên tử Hydro

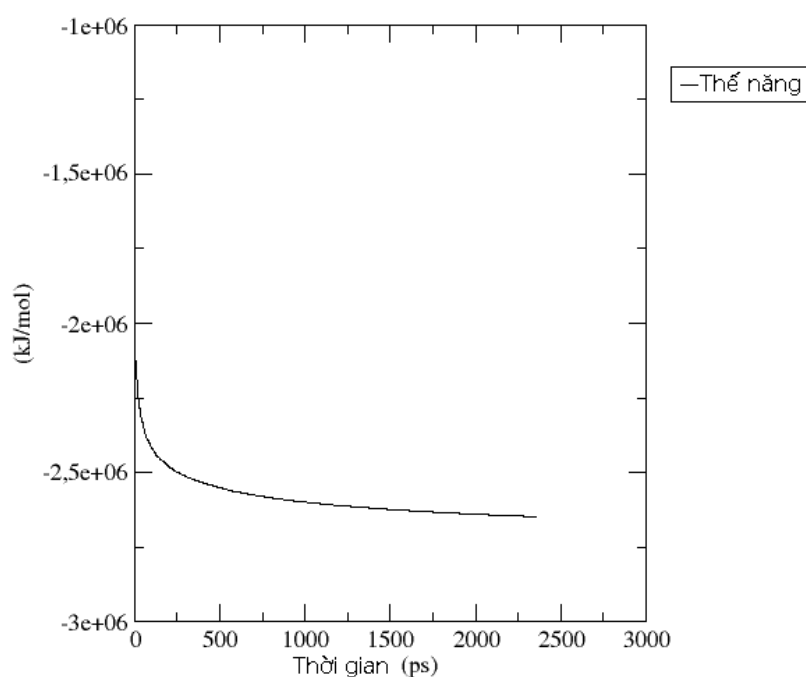
Liên kết Hydro (H-bond) là liên kết đóng góp rất nhiều cho tương tác giữa hai nguyên tử. Đặc biệt trong môi trường nước, liên kết Hydro là liên kết chiếm ưu thế. Liên kết Hydro là liên kết giữa nguyên tử phân cực dương (donor) và nguyên tử phân cực âm (acceptor) thỏa điều kiện (i) cách nhau một khoảng 3.5Å và (ii) góc nhị diện tạo bởi donor và acceptor là lớn hơn hoặc bằng 30°. Xem hình 2.6.

Chương 3

Kết quả và thảo luận

3.1 Thiết lập hệ mô phỏng

3.1.1 Cực tiểu hoá năng lượng

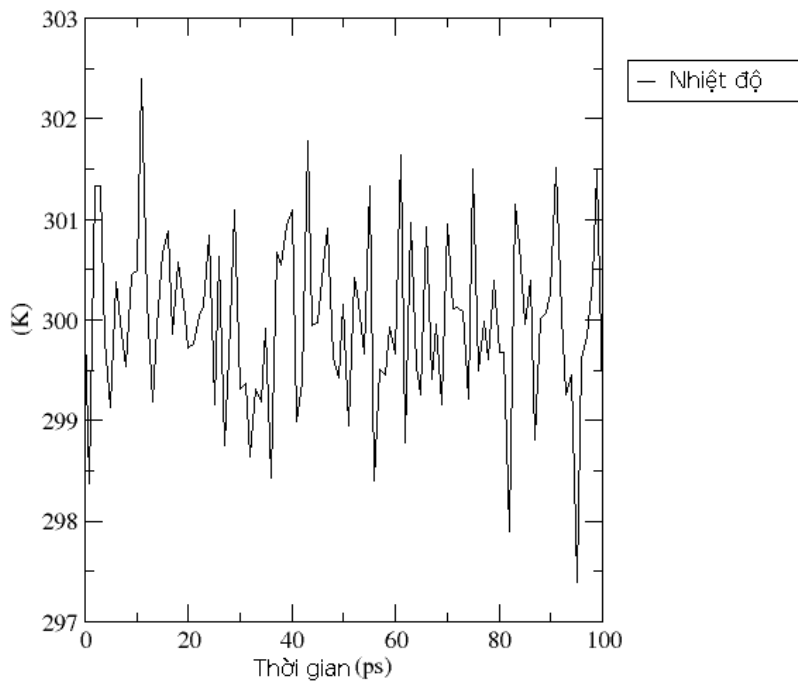


Hình 3.1: Đồ thị thế năng của hệ theo thời gian trong bước cực tiểu hoá năng lượng.

Như đã trình bày trong chương trước, đây là bước chạy mô phỏng giúp hệ hạt phân bố một cách ổn định trong không gian. Quan sát đồ thị 3.1 có thể thấy thế năng của hệ giảm theo thời gian. Điều này có nghĩa các nguyên tử trong hệ dần dần

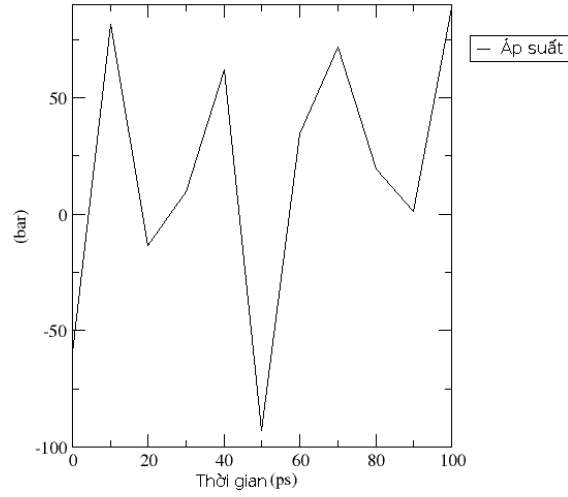
"chuyển" về các vị trí sao cho tương tác lực lên lẫn nhau là tối thiểu. Cụ thể, thế năng của hệ sau quá trình cực tiểu hoá năng lượng là dưới $-2.5 \times 10^{-6} kJ/mol$. Đây là một con số thường thấy đối với các hệ mô phỏng. Điều kiện để dừng phép mô phỏng chính là *lực lớn nhất tác dụng lên bất kỳ nguyên tử nào trong hệ phải nhỏ hơn $1000 kJ.mol^{-1}.nm^{-1}$* . Đây là điều kiện cần, bảo đảm các nguyên tử tương tác với nhau với giá trị lực không quá lớn. Xem xét các dữ liệu xuất ra từ quá trình mô phỏng (cụ thể, dữ liệu từ file **em.log**) thu được lực có giá trị lớn nhất là lực đặt lên nguyên tử thứ 210: $999 kJ.mol^{-1}.nm^{-1}$.

3.1.2 Cân bằng chính tắc (NVT)

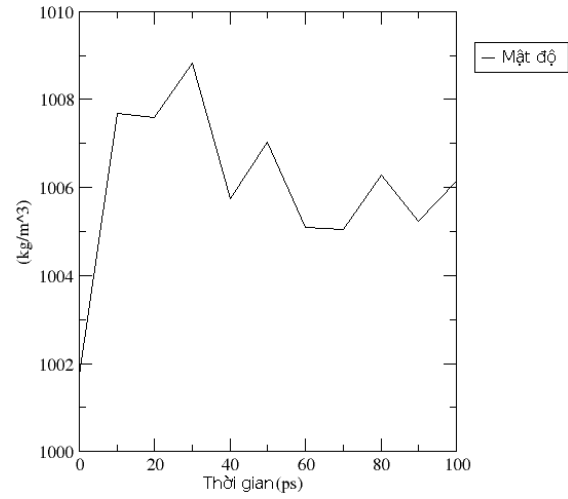


Hình 3.2: Nhiệt độ trung bình của hệ theo thời gian trong bước cân bằng chính tắc (NVT)

Ở bước này, các nguyên tử ở chuỗi chính ($C_{\alpha}-C_{\beta}$) được giữ cố định và nhiệt độ được đưa dần vào hệ thông qua điều khiển động năng của hệ như đã trình bày ở phần 2.2.4. Từ đồ thị rút ra được thăng giáng nhiệt trung bình của hệ khoảng 1%.



(a) Áp suất của hệ theo thời gian trong bước cân bằng đẳng nhiệt - đẳng tích (NPT)

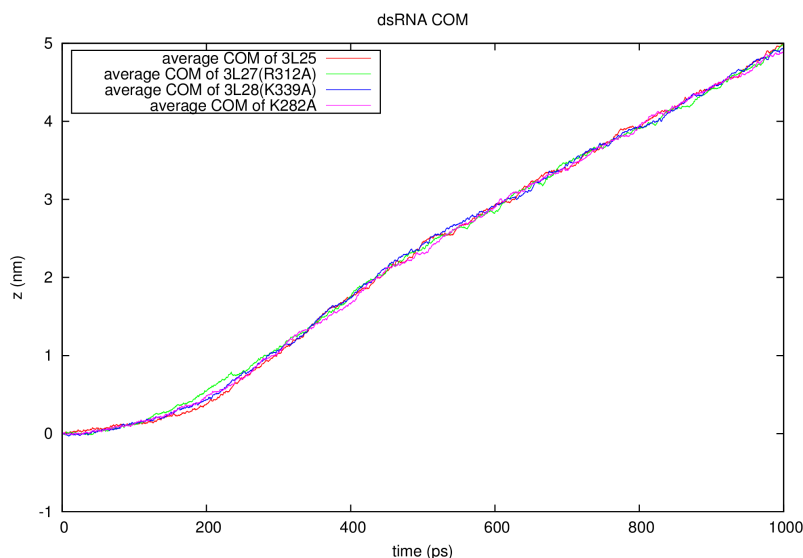


(b) Mật độ của hệ theo thời gian trong bước cân bằng đẳng nhiệt - đẳng tích (NPT)

3.1.3 Cân bằng đẳng nhiệt - đẳng tích (NPT)

Đồ thị 3.3a cho thấy áp suất của hệ mô phỏng thay đổi theo thời gian với biên độ lớn, nhưng nếu lấy trung bình áp suất theo thời gian dài thì giá trị trung bình này lại không lớn. Đồ thị 3.3b cho thấy mật độ khối lượng của hệ vào khoảng 1006 kg/m^3 . Giá trị này tương đối phù hợp với mô hình nước TIP-4P[18] đã được sử dụng.

3.2 Khoảng cách giữa dsRNA và VP35 theo thời gian



Hình 3.4: Đường biểu diễn tọa độ khối tâm trung bình theo thời gian giữa dsRNA và VP35 với phân tử không đột biến (3L25) và đột biến (3L27, 3L28, Lys282Ala)

Từ hình 3.4 có thể nhận thấy thời gian VP35 tương tác trực tiếp với dsRNA vào khoảng 0 đến khoảng 300 ps. Độ dốc của đồ thị từ đoạn 0 đến 200ps thấp, chứng tỏ đây là giai đoạn dsRNA tương tác với VP35 để chống lại lực kéo tách hai nhóm ra. Sau đó là giai đoạn dsRNA di chuyển với tốc độ ổn định. Điều này là hợp lý vì quá trình mô phỏng SMD sử dụng chế độ kéo với nhóm liên kết có vận tốc không đổi được trình bày ở chương trước. dsRNA được kéo ra và di chuyển đến khi không còn tương tác với VP35 nữa thì di chuyển gần như tự do.

3.3 Thế năng tương tác giữa amino acid của VP35 với dsRNA

Cấu hình B-conf của VP35 có 126 amino acid từ amino acid thứ 215 đến 340. Cần chọn các amino acid đáng quan tâm để thực hiện tính toán thế năng tương tác theo thời gian. Việc chọn các amino acid có khả năng tham gia liên kết với dsRNA được

thực hiện như sau:

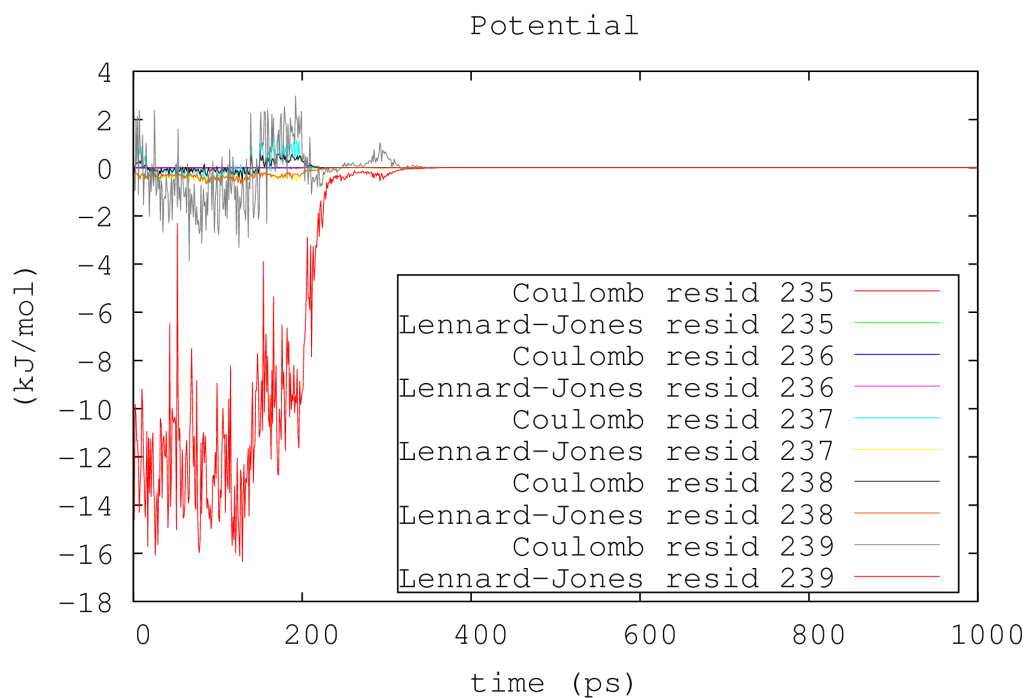
1. chọn phân tử bất kỳ thuộc dsRNA làm tâm
2. tất cả các amino acid của VP35 có bất kỳ phân tử nào nằm trong bán kính $r_{cutoff} = r_{Coulomb} = r_{VDW} = 1.4nm$ (đã trình bày trong 2.4) được xem là có khả năng tương tác với dsRNA

Kết quả của phép chọn trên được liệt kê trong phụ lục A. Quan sát 81 vị trí amino acid này, có thể chọn ra các vị trí amino acid tiêu biểu có khả năng đóng góp cho tương tác giữa hai nhóm phân tử này dựa trên các tiêu chí sau:

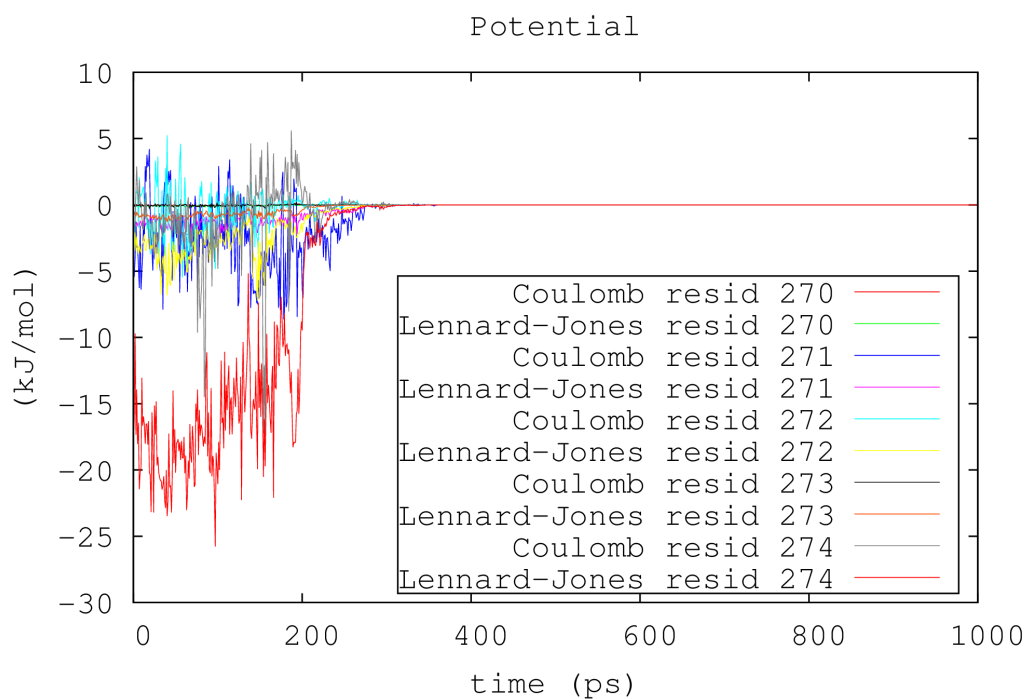
- Có thể năng biến đổi theo thời gian trong quá trình mô phỏng SMD. Điều này là hợp lý vì nếu vị trí amino acid đó nếu thực sự đóng góp vào tương tác dsRNA–VP35 thì khi tách hai nhóm phân tử này ra, tổng thể năng tương tác của vị trí amino acid này với các phân tử của dsRNA phải giảm một cách rõ rệt.
- Độ giảm của thế năng được đánh giá dựa vào hai vị trí amino acid đã được thực nghiệm xác minh tính quan trọng của nó trong tương tác dsRNA–VP35 là Arg312 và Lys339[34]. Nếu độ biến thiên thế năng này có scale nhỏ hơn hẳn so với cả hai vị trí amino acid trên thì đó không phải là vị trí amino acid ta cần quan tâm.
- Từ kết quả thu được, độ thay đổi thế năng của Lys339 là $10kJ/mol$ được chọn làm giới hạn chọn lọc các amino acid cần quan tâm.

Từ các điều kiện trên, có thể rút ra được từ kết quả chạy mô phỏng các vị trí amino acid đáng quan tâm gồm: Phe239, Glu274, Cys275, IleLE278, GLu279, Lys282, Arg312, Asp321, Arg322, Lys339, Ile340. Quan sát các đồ thị 3.5, 3.6, 3.7, 3.8 có thể thấy thời gian các amino acid tương tác với dsRNA là vào khoảng 200ps đầu tiên, kể từ 400ps trở về sau thế năng tương tác giữa các amino acid và dsRNA giảm về 0. Điều này phù hợp với kết quả quan sát 3.2.

Các đồ thị trong hình 3.5, 3.6, 3.7, 3.8 lấy dữ liệu từ lần mô phỏng SMD thứ 2 đối với cấu hình 3L25 (tức cấu hình không đột biến). Dữ liệu từ lần mô phỏng SMD thứ 1 và thứ 3 cũng cho kết quả tương tự.

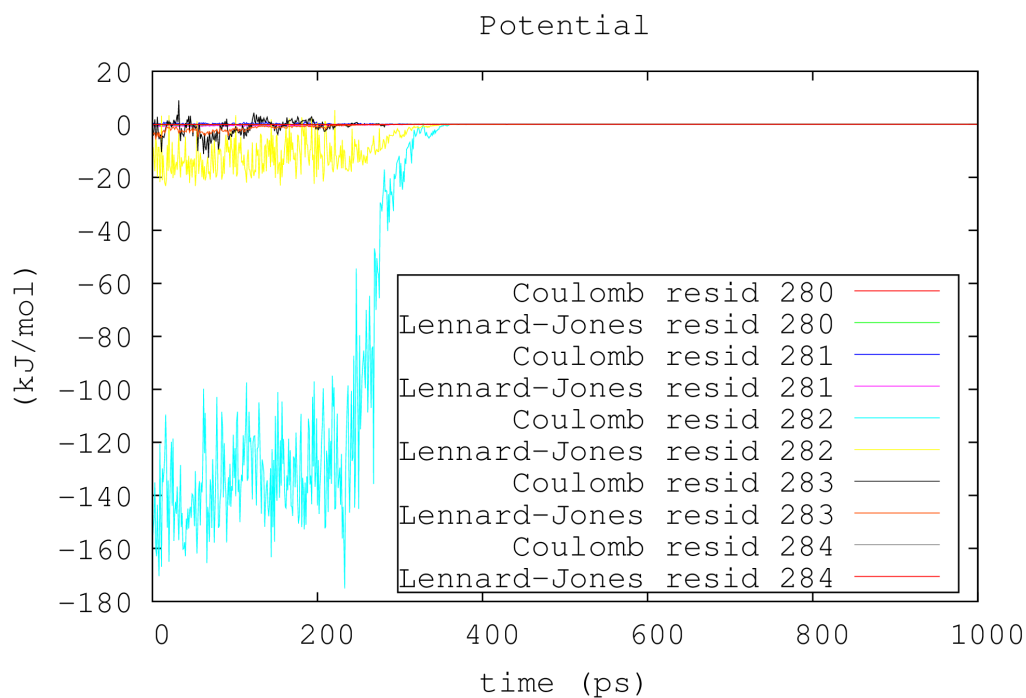


(a) vị trí amino acid 235–239, tiêu biểu Phe239

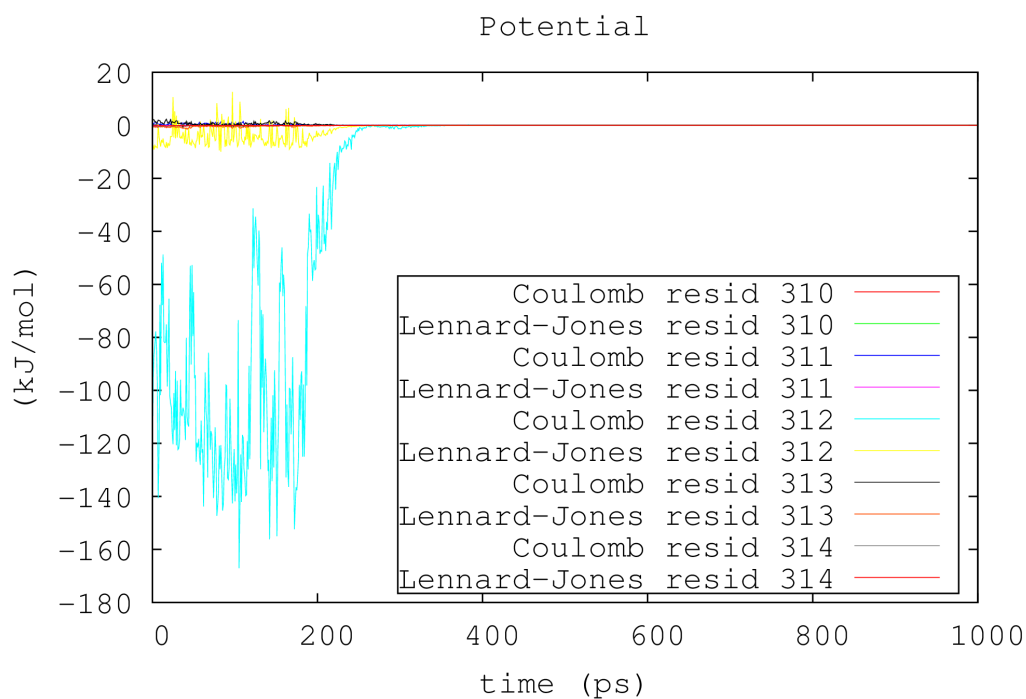


(b) vị trí amino acid 270–274, tiêu biểu Q274

Hình 3.5: Đồ thị thế năng theo thời gian của các amino acid

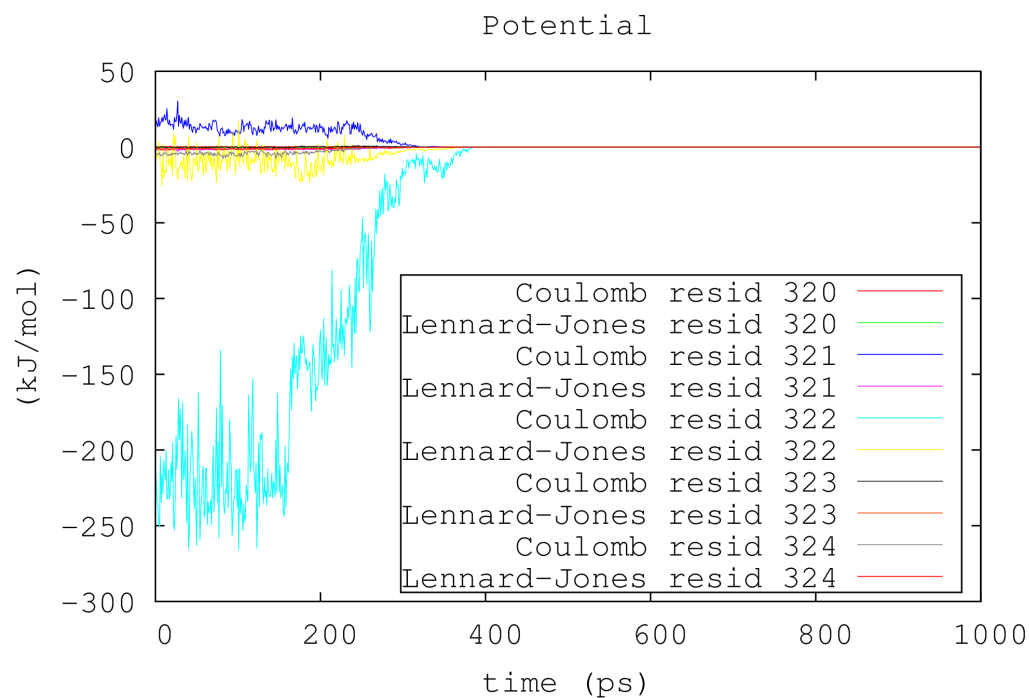


(a) vị trí amino acid 280–284, tiêu biểu Lys282

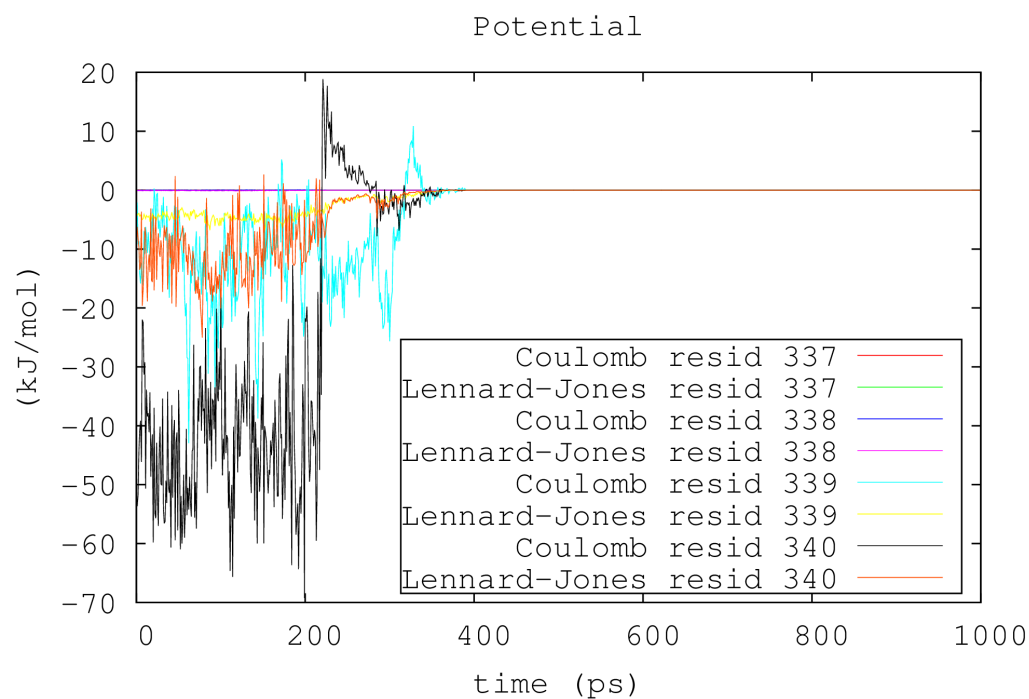


(b) vị trí amino acid 310–314, tiêu biểu Arg312 (khoảng 120 kJ/mol)

Hình 3.6: Đồ thị thế năng theo thời gian của các amino acid

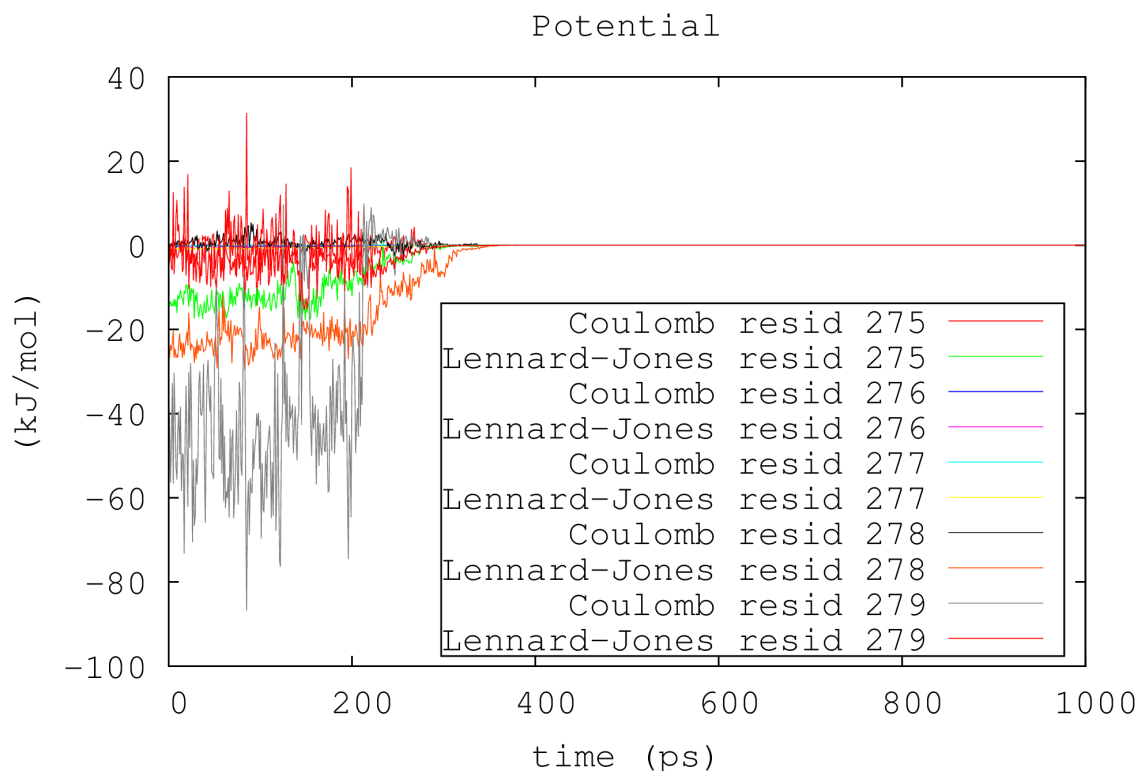


(a) vị trí amino acid 320–324, tiêu biểu Asp321, Arg322



(b) vị trí amino acid 337–340, tiêu biểu ILE340, Lys339 (khoảng 10–20 kJ/mol)

Hình 3.7: Đồ thị thế năng theo thời gian của các amino acid

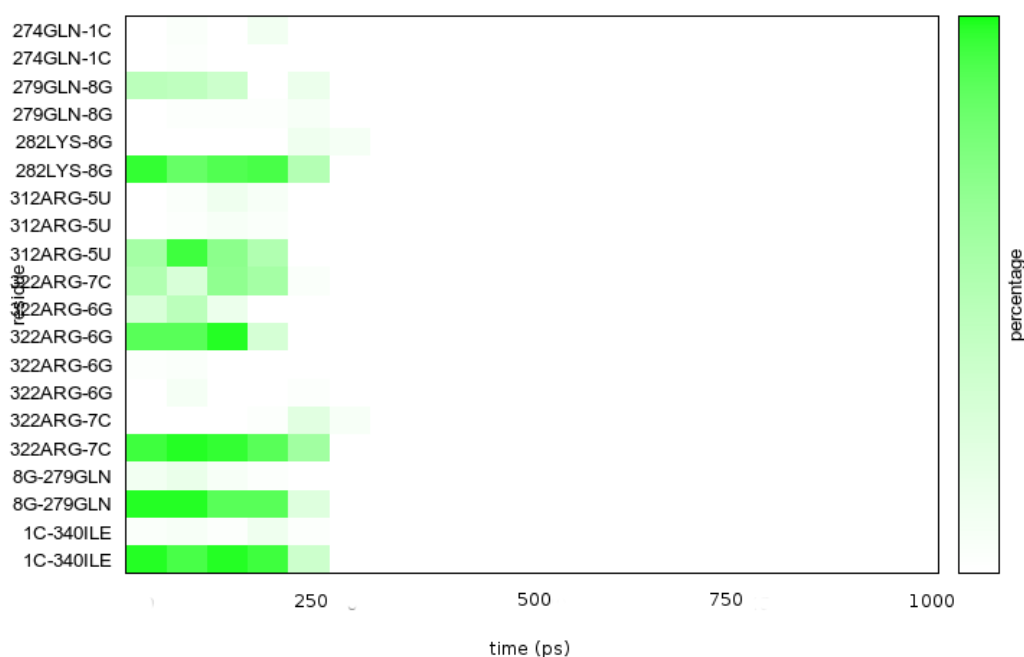


Hình 3.8: vị trí amino acid 275–279, tiêu biểu Cys275, Ile278, Glu279. Glu279 tương tác Coulomb là chủ yếu. Cys275, Ile278 tương tác VdW là chủ yếu với thế năng cùng scale với thế năng của Lys339.

Dựa vào đồ thị 3.5,3.6,3.7, có thể thấy rằng đóng góp của tương tác Coulomb là rất đáng kể. Tuy nhiên vẫn có những vị trí amino acid có đóng góp của tương tác Van der Waals là đáng kể, đó là Cys275 và Ile278 (xem đồ thị 3.8).

Dựa vào đồ thị 3.7b và 3.6b có thể thấy Arg312 đóng vai trò tương tác trực tiếp với dsRNA. Mặc dù đột biến Lys339Ala cũng gây mất liên kết dsRNA–VP35 trong thực nghiệm, nhưng độ thay đổi thế năng tương tác do Lys339 gây ra không nhiều, chỉ khoảng 10% khi so với Arg312 gây ra. Ngoài ra, Lys282 và Arg322 cũng tương tác nhiều với dsRNA do thế năng tương tác giữa các vị trí amino acid này với dsRNA lớn (Lys282 có thế năng tương tác lên hơn 140kJ/mol , Arg322 có thế năng tương tác lên đến 250kJ/mol , Arg312 có thế năng tương tác chỉ ở mức 140kJ/mol). Xem hình 3.6a, 3.6b, 3.7a. Các amino acid này có thể đóng vai trò quan trọng trong tương tác giữa dsRNA–VP35.

3.4 Kết quả phân tích liên kết Hydro giữa các amino acid của VP35 với dsRNA



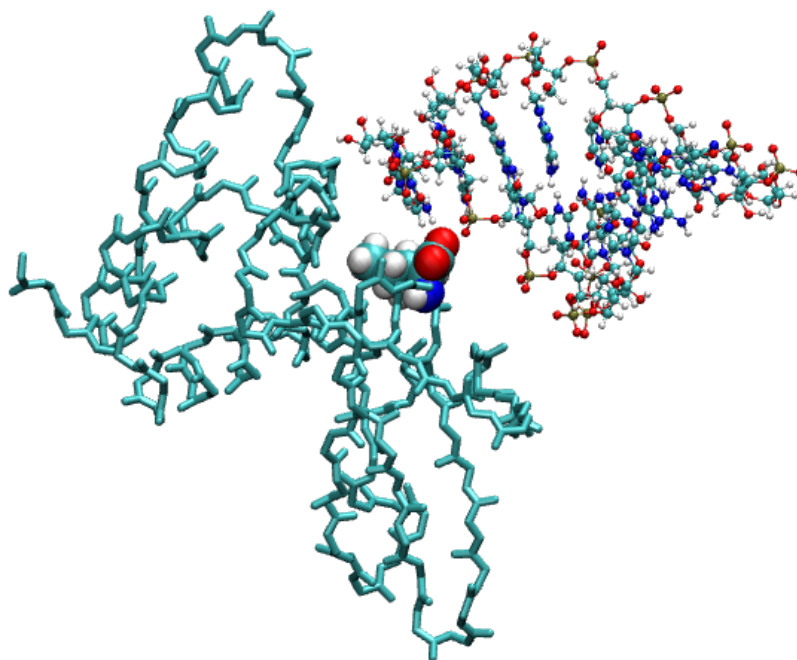
Hình 3.9: Liên kết Hydro theo thời gian giữa các vị trí amino acid và dsRNA. Lần kéo thứ 2 cấu hình B-conf của phân tử có PDBID: 3L25.

Hình 3.9 cho thấy kết quả phân tích liên kết Hydro theo thời gian. Số vị trí amino acid tham gia tạo liên kết Hydro không nhiều nên từ kết quả này có thể sàng lọc ra các vị trí amino acid hiệu quả hơn. Đây là kết quả từ lần kéo thứ 2 cấu hình B-conf của 3L25. Các lần kéo khác cũng có kết quả tương tự. Xem phụ lục B.

Phân tích kết quả thu được trong cả ba lần kéo, dễ dàng nhận thấy các vị trí amino acid có khả năng đóng vai trò chủ yếu trong tương tác trực tiếp giữa dsRNA–VP35 gồm Q274, Glu279, Lys282, Arg312, Arg322, I340.

Cụ thể, các vị trí amino acid nổi bật với tỉ lệ tạo liên kết Hydro cao (màu xanh đậm trên hình 3.9) gồm Lys282, Arg312, Arg322, Glu279 và Ile340. Ile340 là amino

acid ở C-terminal nên dù có nhánh bên kỵ nước nhưng gốc OH^- của mạch chính vẫn có khả năng tạo liên kết Hydro. Xem hình 3.10. Các amino acid này có thể đóng vai trò quan trọng trong tương tác giữa dsRNA–VP35.



Hình 3.10: Mạch chính (Backbone) của VP35 được tô màu xanh da trời. Amino acid Ile340 nằm tại C-terminal nên là amino acid cuối cùng, có gốc OH^- có khả năng tạo liên kết Hydro với dsRNA. Nguyên tử Oxy được tô màu đỏ.

3.5 So sánh lực kéo

Để khảo sát vai trò của các vị trí amino acid có thể năng tương tác lớn và tạo liên kết Hydro với dsRNA, chúng tôi tiến hành mô phỏng SMD cho các hệ đột biến và so sánh lực kéo tối đa so với trường hợp không đột biến. Vì lý do thời gian thực hiện khoá luận ngắn nên chúng tôi tiến hành mô phỏng cho 3 hệ đột biến các vị trí amino acid:

- Arg312Ala (PDBID: 3L27).
- Lys339Ala (PDBID: 3L28).
- Lys282Ala.

Quan sát hình 3.11a dễ dàng nhận thấy rằng lực kéo tối đa dùng để tách dsRNA với phân tử không đột biến VP35 là lớn nhất. Điều này phù hợp với kết quả thực nghiệm[34]. Cũng từ kết quả thu được, lực kéo tối đa dùng để kéo phân tử VP35 đột biến tại Arg312 (Arg312Ala - cấu hình 3L27) thấp hơn hẳn so với lực kéo phân tử không đột biến, điều này cho thấy tầm quan trọng của amino acid tại vị trí này trong tương tác dsRNA-VP35.

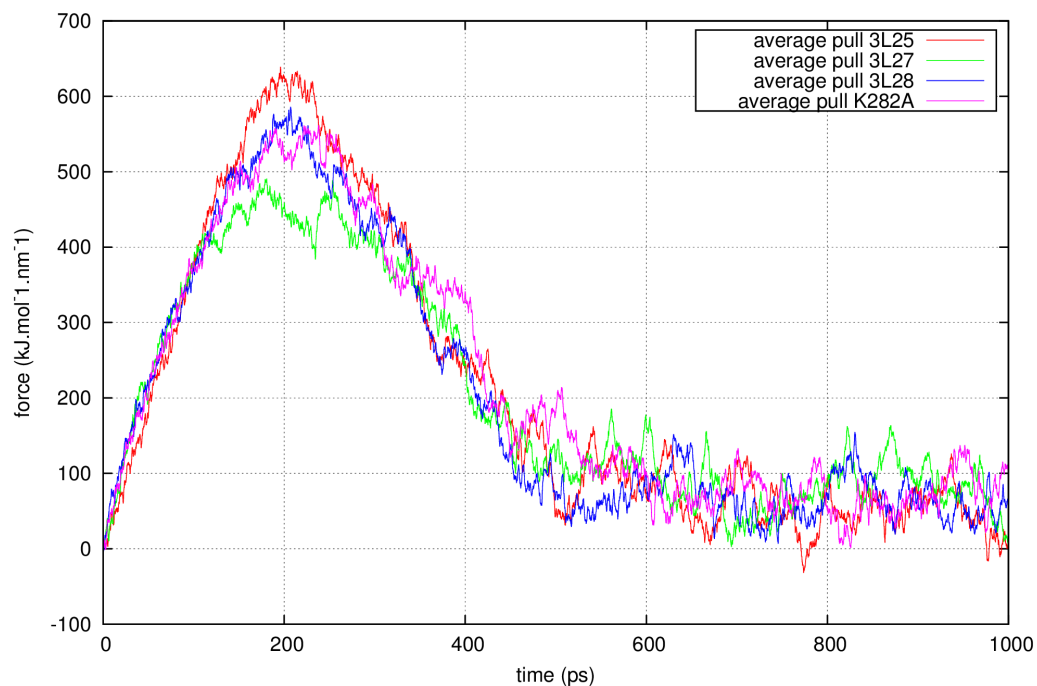
Tương tự với cấu hình 3L28 (đột biến amino acid 339 từ Lysine sang Alanine - Lys339Ala), lực kéo tối đa dùng để tách hai nhóm phân tử thấp hơn so với lực kéo trong quá trình mô phỏng phân tử gốc 3L25, điều này phù hợp với kết quả thực nghiệm[34].

3.6 Đột biến Lys282Ala

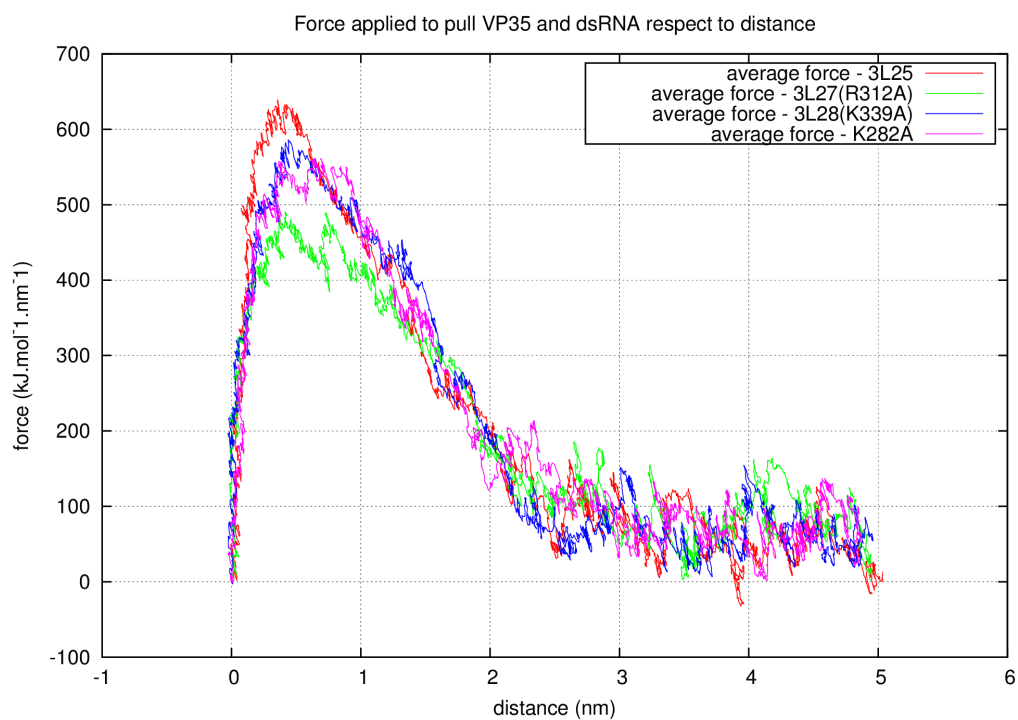
Từ dữ liệu thu được trong quá trình mô phỏng các cấu hình không đột biến và đột biến, đề tài phát triển thêm một bước ở việc dự đoán amino acid Arginine thứ 282 cũng đóng vai trò quan trọng trong tương tác dsRNA-VP35 dựa trên các quan sát sau:

- Lys282 là phân tử Lysine mang điện tích, vì vậy nó có thể đóng góp cho tương tác giữa hai nhóm phân tử trong tương tác tĩnh điện. Hơn nữa, kết quả phân tích thể năng cho thấy Lys282 đóng góp nhiều vào tương tác dsRNA-VP35 (xem hình 3.6a). Độ thay đổi thể năng tương tác giữa Lys282 với dsRNA (trên $140kJ/mol$) lớn hơn so với độ thay đổi thể năng tương tác giữa Arg312 với dsRNA (khoảng $140kJ/mol$). Thực nghiệm cho thấy đột biến Arg312Ala làm mất hoàn toàn khả năng gắn vào dsRNA của VP35.
- Lys282 là một trong những amino acid tạo liên kết Hydro mạnh trong cả 3 lần mô phỏng SMD (xem hình B.1,3.9,B.2). Cụ thể, Lys282 luôn tạo liên kết Hydro ở mức 100% trong suốt thời gian VP35 vẫn còn tương tác với dsRNA.

Để đánh giá kết quả tiên đoán, đường biểu diễn lực kéo trung bình cấu hình đột biến Lys282 cho thấy lực kéo tối đa của phân tử này thấp hơn so với phân tử không đột biến và tương đối xấp xỉ với lực kéo tối đa cho cấu hình đột biến 3L28 (đột biến



(a) Đường biểu diễn lực kéo trung bình theo thời gian giữa dsRNA và VP35 với phân tử không đột biến (3L25) và đột biến (3L27, 3L28, Lys282Ala)



(b) Đường biểu diễn lực kéo trung bình theo tọa độ khối tâm của dsRNA của phân tử không đột biến (3L25) và đột biến (3L27, 3L28, Lys282Ala)

Lys339Ala) có được từ thực nghiệm. Như vậy, Lys282 có thể là một trong những amino acid chính đóng góp vào tương tác dsRNA-VP35. Việc đột biến phân tử này gây ảnh hưởng lên tương tác dsRNA-VP35 một cách rõ rệt.

Chương 4

Kết luận và hướng phát triển tiếp theo của đề tài

Bằng phương pháp mô phỏng SMD cùng với các phân tích về thể năng tương tác, liên kết Hydro và lực kéo tối đa, khoá luận đã chỉ ra được vai trò của các amino acid Arg312 và Lys339. Kết quả này phù hợp với kết quả trong thực nghiệm, đồng thời dự đoán amino acid đóng góp vào tương tác giữa dsRNA-VP35.

Kết quả thu được từ khóa luận đem lại gợi ý cho việc phát triển các phân tử có thể tác động vào các amino acid đóng góp vào tương tác dsRNA-VP35. Đồng thời gợi ý amino acid Lys282 đóng vai trò quan trọng trong tương tác dsRNA-VP35.

Tuy nhiên, đề tài cần được thực hiện thêm quá trình mô phỏng động học thời gian dài để quan sát các trạng thái khác nhau của hệ phân tử, qua đó đánh giá chính xác hơn các amino acid đóng góp chủ yếu cho tương tác dsRNA-VP35.

Hướng phát triển tiếp theo của đề tài:

- Khảo sát phân tử VP35 đột biến Arg322Ala để kiểm tra giả thiết ở phần 3.3 và 3.4.
- Tiên đoán các phân tử có khả năng gắn vào khu vực tương tác dsRNA-VP35, theo đó gợi ý phân tử thuốc có khả năng bất hoạt VP35.
- Đánh giá và tiên đoán dựa trên mô phỏng động học phân tử thời gian dài nhằm tìm ra các amino acid đóng góp cho tương tác dsRNA-VP35.

Tài liệu tham khảo

- [1] Loup Verlet. “Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules”. In: *Physical Review* 159.1 (1967), pp. 98–103. ISSN: 0031899X. DOI: 10.1103/PhysRev.159.98.
- [2] Wendy D. Cornell. “A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules”. In: 117.19 (1995), pp. 5179–5197. DOI: 10.1021/ja00124a002.
- [3] H Lu et al. “Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation.” In: *Biophysical journal* 75.2 (Aug. 1998), pp. 662–71. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(98)77556-3. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1299741%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [4] E Mühlberger et al. “Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems.” In: *Journal of virology* 73.3 (Mar. 1999), pp. 2333–42. ISSN: 0022-538X. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=104478%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [5] C F Basler et al. “The Ebola virus VP35 protein functions as a type I IFN antagonist.” In: *Proceedings of the National Academy of Sciences of the United States of America* 97.22 (Oct. 2000), pp. 12289–94. ISSN: 0027-8424. DOI: 10.1073/pnas.220398297. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=17334%5C&tool=pmcentrez%5C&rendertype=abstract>.

- [6] R Colebunders and M Borchert. “Ebola haemorrhagic fever—a review.” In: *The Journal of infection* 40.1 (Jan. 2000), pp. 16–20. ISSN: 0163-4453. DOI: 10.1053/jinf.1999.0603. URL: <http://www.ncbi.nlm.nih.gov/pubmed/10762106>.
- [7] a Dessen et al. “Crystal structure of the matrix protein VP40 from Ebola virus.” In: *The EMBO journal* 19.16 (2000), pp. 4228–4236. ISSN: 0261-4189. DOI: 10.1093/emboj/19.16.4228.
- [8] Lukas D Schuler, Xavier Daura, and Wilfred F V a N Gunsteren. “An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase”. In: 22.11 (2001), pp. 1205–1218.
- [9] Marco De Andrea et al. “The interferon system: an overview.” In: *European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society* 6 Suppl A (Jan. 2002), A41–6, discussion A55–8. ISSN: 1090-3798. URL: <http://www.ncbi.nlm.nih.gov/pubmed/12365360>.
- [10] Yue Huang et al. “The Assembly of Ebola Virus Nucleocapsid Requires Virion-Associated Proteins 35 and 24 and Posttranslational Modification of Nucleoprotein”. English. In: *Molecular Cell* 10.2 (Aug. 2002), pp. 307–316. ISSN: 10972765. DOI: 10.1016/S1097-2765(02)00588-9. URL: <http://www.cell.com/article/S1097276502005889/fulltext>.
- [11] Takeshi Noda et al. “Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP.” In: *Journal of virology* 76.10 (2002), pp. 4855–4865. ISSN: 0022-538X. DOI: 10.1128/JVI.76.10.4855-4865.2002.
- [12] M. Weik et al. “Ebola Virus VP30-Mediated Transcription Is Regulated by RNA Secondary Structure Formation”. In: *Journal of Virology* 76.17 (Sept. 2002), pp. 8532–8539. ISSN: 0022-538X. DOI: 10.1128/JVI.76.17.8532-8539.2002. URL: <http://jvi.asm.org/cgi/content/long/76/17/8532>.
- [13] Christopher F Basler et al. “The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3.” In: *Journal of virology* 77.14 (July 2003), pp. 7945–56. ISSN: 0022-538X. URL: <http://www.pubmedcentral.nih.gov/>

articlerender.fcgi?artid=161945%5C&tool=pmcentrez%5C&rendertype=abstract.

- [14] Yong Duan et al. “A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations”. In: *Journal of Computational Chemistry* 24.16 (Dec. 2003), pp. 1999–2012. ISSN: 01928651. DOI: 10.1002/jcc.10349. URL: <http://www.ncbi.nlm.nih.gov/pubmed/14531054>.
- [15] Heinz Feldmann et al. “Ebola virus: from discovery to vaccine.” In: *Nature reviews. Immunology* 3.8 (Aug. 2003), pp. 677–85. ISSN: 1474-1733. DOI: 10.1038/nri1154. URL: <http://www.ncbi.nlm.nih.gov/pubmed/12974482>.
- [16] Jay W. Ponder and David a. Case. “Force fields for protein simulations”. In: *Advances in Protein Chemistry* 66 (2003), pp. 27–85. ISSN: 00653233. DOI: 10.1016/S0065-3233(03)66002-X.
- [17] Amy L Hartman, Jonathan S Towner, and Stuart T Nichol. “A C-terminal basic amino acid motif of Zaire ebolavirus VP35 is essential for type I interferon antagonism and displays high identity with the RNA-binding domain of another interferon antagonist, the NS1 protein of influenza A virus.” In: *Virology* 328.2 (Oct. 2004), pp. 177–84. ISSN: 0042-6822. DOI: 10.1016/j.virol.2004.07.006. URL: <http://www.ncbi.nlm.nih.gov/pubmed/15464838>.
- [18] Hans W Horn et al. “Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew.” In: *The Journal of chemical physics* 120.20 (May 2004), pp. 9665–78. ISSN: 0021-9606. DOI: 10.1063/1.1683075. URL: <http://www.ncbi.nlm.nih.gov/pubmed/15267980>.
- [19] Alexander D. Mackerell. *Empirical force fields for biological macromolecules: Overview and issues*. 2004. DOI: 10.1002/jcc.20082.
- [20] Chris Oostenbrink et al. “A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6”. In: *Journal of Computational Chemistry* 25.13 (2004), pp. 1656–1676. ISSN: 01928651. DOI: 10.1002/jcc.20090.

- [21] Thereza a. Soares et al. “An improved nucleic acid parameter set for the GRO-MOS force field”. In: *Journal of Computational Chemistry* 26.7 (2005), pp. 725–737. ISSN: 01928651. DOI: 10.1002/jcc.20193.
- [22] Washington B Cárdenas et al. “Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling.” In: *Journal of virology* 80.11 (June 2006), pp. 5168–5178. ISSN: 0022-538X. DOI: 10.1128/JVI.02199-05. URL: <http://jvi.asm.org/cgi/content/long/80/11/5168>.
- [23] Sven Enterlein et al. “VP35 knockdown inhibits Ebola virus amplification and protects against lethal infection in mice.” In: *Antimicrobial agents and chemotherapy* 50.3 (Mar. 2006), pp. 984–93. ISSN: 0066-4804. DOI: 10.1128/AAC.50.3.984-993.2006. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1426423%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [24] Viktor Hornak et al. “Comparison of multiple Amber force fields and development of improved protein backbone parameters.” In: *Proteins* 65.3 (Nov. 2006), pp. 712–25. ISSN: 1097-0134. DOI: 10.1002/prot.21123. URL: <http://www.ncbi.nlm.nih.gov/pubmed/16981200>.
- [25] Reed F Johnson, Peter Bell, and Ronald N Harty. “Effect of Ebola virus proteins GP, NP and VP35 on VP40 VLP morphology.” In: *Virology journal* 3.1 (Jan. 2006), p. 31. ISSN: 1743-422X. DOI: 10.1186/1743-422X-3-31. URL: <http://www.virologyj.com/content/3/1/31>.
- [26] St Patrick Reid et al. “Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation.” In: *Journal of virology* 80.11 (June 2006), pp. 5156–67. ISSN: 0022-538X. DOI: 10.1128/JVI.02349-05. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1472181%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [27] Giovanni Bussi, Davide Donadio, and Michele Parrinello. “Canonical sampling through velocity rescaling”. In: *Journal of Chemical Physics* 126.1 (2007), pp. 1–7. ISSN: 00219606. DOI: 10.1063/1.2408420. arXiv: arXiv:0803.4060v1.

- [28] Zongdi Feng et al. “The VP35 protein of Ebola virus inhibits the antiviral effect mediated by double-stranded RNA-dependent protein kinase PKR.” In: *Journal of virology* 81.1 (Jan. 2007), pp. 182–92. ISSN: 0022-538X. DOI: 10.1128/JVI.01006-06. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1797262%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [29] Hiroki Kato et al. “Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5.” In: *The Journal of experimental medicine* 205.7 (July 2008), pp. 1601–10. ISSN: 1540-9538. DOI: 10.1084/jem.20080091. URL: <http://jem.rupress.org/content/205/7/1601>.
- [30] Jeffrey E Lee and Erica Ollmann Saphire. “Ebola virus glycoprotein structure and mechanism of entry.” In: *Future virology* 4.6 (Jan. 2009), pp. 621–635. ISSN: 1746-0794. DOI: 10.2217/fvl.09.56. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2829775%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [31] Kathleen C Prins, Washington B Cárdenas, and Christopher F Basler. “Ebola virus protein VP35 impairs the function of interferon regulatory factor-activating kinases IKKepsilon and TBK-1.” In: *Journal of virology* 83.7 (Apr. 2009), pp. 3069–77. ISSN: 1098-5514. DOI: 10.1128/JVI.01875-08. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2655579%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [32] Christopher R Kimberlin et al. “Ebola virus VP35 uses a bimodal strategy to bind dsRNA for innate immune suppression.” In: *Proceedings of the National Academy of Sciences of the United States of America* 107 (2010), pp. 314–319. ISSN: 0027-8424. DOI: 10.1073/pnas.0910547107.
- [33] Jens H Kuhn et al. “Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations.” In: *Archives of virology* 155.12 (Dec. 2010), pp. 2083–103. ISSN: 1432-8798. DOI: 10.1007/s00705-010-0814-x. URL: <http://www.pubmedcentral.nih.gov/>

- articlerender.fcgi?artid=3074192%5C&tool=pmcentrez%5C&rendertype=abstract.
- [34] Daisy W Leung et al. “Structural basis for dsRNA recognition and interferon antagonism by Ebola VP35.” In: *Nature structural & molecular biology* 17.2 (Feb. 2010), pp. 165–172. ISSN: 1545-9993. DOI: 10.1038/nsmb.1765. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2872155%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [35] Parameshwaran Ramanan et al. “Filoviral Immune Evasion Mechanisms”. In: *Viruses* 3.9 (Sept. 2011), pp. 1634–1649. ISSN: 1999-4915. DOI: 10.3390/v3091634. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3187693/%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3187693%5C&tool=pmcentrez%5C&rendertype=abstract>.
- [36] Nathan Schmid et al. “Definition and testing of the GROMOS force-field versions 54A7 and 54B7”. In: *European Biophysics Journal* 40.7 (2011), pp. 843–856. ISSN: 01757571. DOI: 10.1007/s00249-011-0700-9.
- [37] Mai Suan Li and Binh Khanh Mai. “Steered Molecular Dynamics-A Promising Tool for Drug Design”. In: *Current Bioinformatics* 7.4 (2012), pp. 342–351. ISSN: 15748936. DOI: 10.2174/157489312803901009. URL: <http://www.eurekaselect.com/openurl/content.php?genre=article%5C&issn=1574-8936%5C&volume=7%5C&issue=4%5C&spage=342>.
- [38] Aftab A. Ansari. “Clinical features and pathobiology of Ebolavirus infection”. In: *Journal of Autoimmunity* 55 (Sept. 2014), pp. 1–9. ISSN: 08968411. DOI: 10.1016/j.jaut.2014.09.001. URL: <http://www.sciencedirect.com/science/article/pii/S0896841114001309>.
- [39] Craig S Brown et al. “In silico derived small molecules bind the filovirus VP35 protein and inhibit its polymerase cofactor activity.” In: *Journal of molecular biology* 426.10 (May 2014), pp. 2045–58. ISSN: 1089-8638. DOI: 10.1016/j.jmb.2014.01.010. URL: <http://www.ncbi.nlm.nih.gov/pubmed/24495995%20http://www.sciencedirect.com/science/article/pii/S0022283614000485>.

- [40] Paulina J Dziubańska et al. “The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein.” In: *Acta crystallographica. Section D, Biological crystallography* 70.Pt 9 (Sept. 2014), pp. 2420–9. ISSN: 1399-0047. DOI: 10.1107/S1399004714014710. URL: <http://dx.doi.org/10.1107/S1399004714014710> 20<http://www.ncbi.nlm.nih.gov/pubmed/25195755>.
- [41] Qiao Xue et al. “Exploring the mechanism how Marburg virus VP35 recognizes and binds dsRNA by molecular dynamics simulations and free energy calculations”. In: *Biopolymers* 101.8 (2014), pp. 849–860. ISSN: 10970282. DOI: 10.1002/bip.22463.
- [42] F. Dapiaggi, S. Pieraccini, and M. Sironi. “In silico study of VP35 inhibitors: from computational alanine scanning to essential dynamics”. In: *Mol. BioSyst.* (2015). ISSN: 1742-206X. DOI: 10.1039/C5MB00348B. URL: <http://xlink.rsc.org/?DOI=C5MB00348B>.
- [43] Chad E Mire et al. “Single-dose attenuated Vesiculovax vaccines protect primates against Ebola Makona virus.” In: *Nature* 520.7549 (Apr. 2015), pp. 688–691. ISSN: 1476-4687. DOI: 10.1038/nature14428. URL: <http://dx.doi.org/10.1038/nature14428>.

Phụ lục A

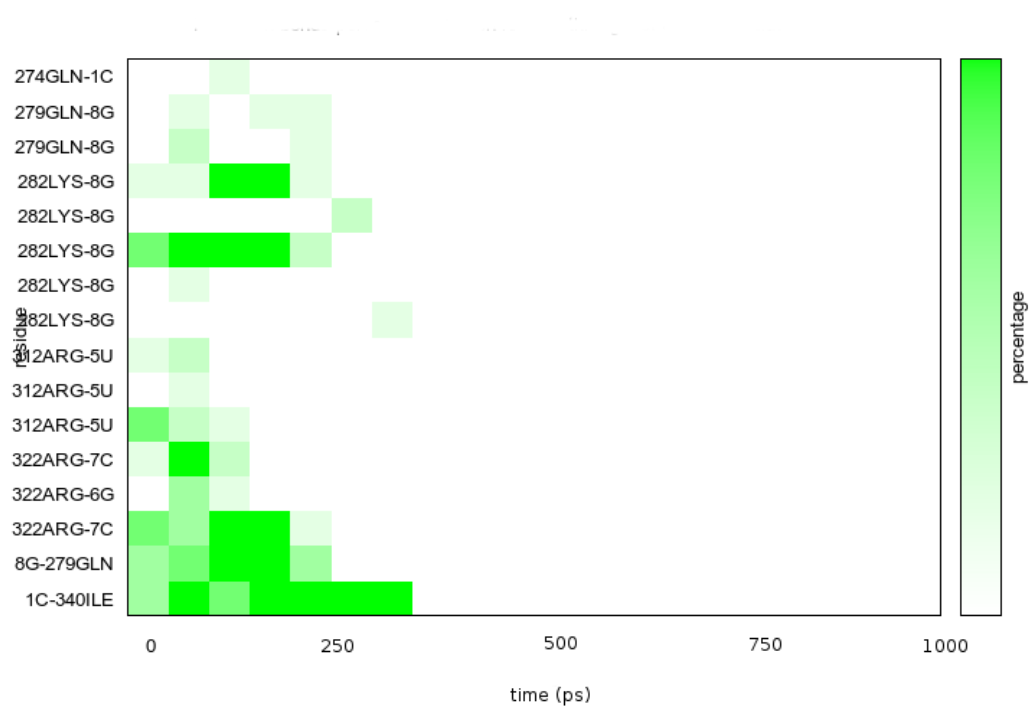
Các amino acid tương tác với dsRNA

STT	Vị trí	Tên	STT	Vị trí	Tên	STT	Vị trí	Tên
1	228	Met	17	246	Ile	33	275	Cys
2	229	Tyr	18	259	Ile	34	276	Ala
3	230	Asp	19	261	Ala	35	277	Leu
4	231	His	20	262	Glu	36	278	Ile
5	232	Leu	21	263	Phe	37	279	Gln
6	233	Pro	22	264	Gln	38	280	Ile
7	234	Gly	23	265	Ala	39	281	Thr
8	235	Phe	24	266	Ser	40	282	Lys
9	236	Gly	25	267	Leu	41	283	Arg
10	237	Thr	26	268	Ala	42	284	Val
11	238	Ala	27	269	Glu	43	285	Pro
12	239	Phe	28	270	Gly	44	287	Phe
13	240	His	29	271	Asp	45	288	Gln
14	241	Gln	30	272	Ser	46	289	Asp
15	242	Leu	31	273	Pro	47	290	Ala
16	243	Val	32	274	Gln	48	292	Pro

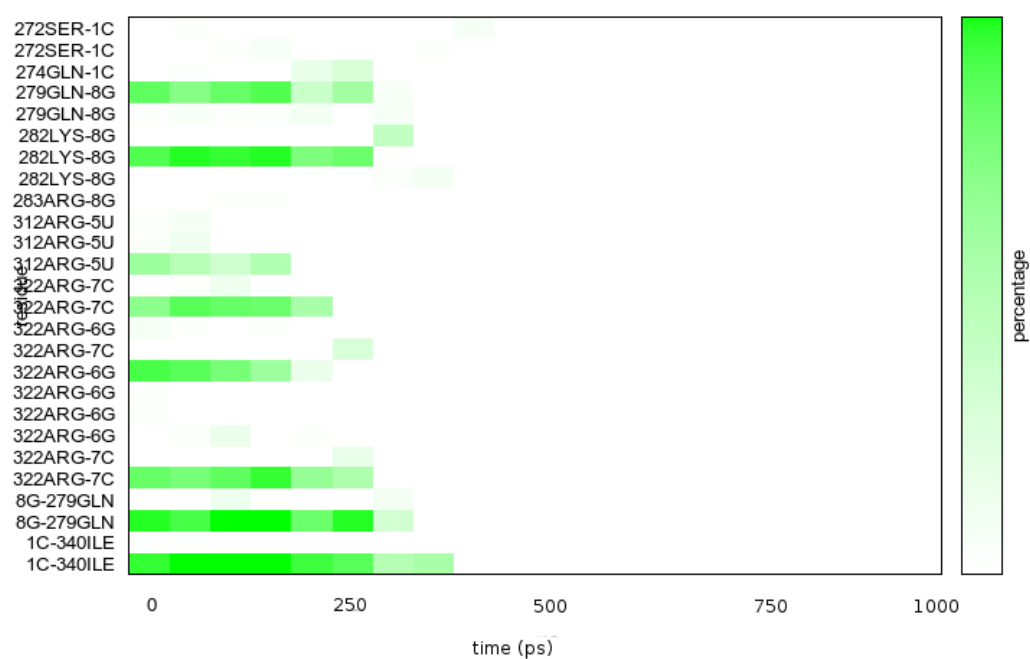
STT	Vị trí	Tên	STT	Vị trí	Tên	STT	Vị trí	Tên
49	300	Arg	65	317	Ser	81	340	Ile
50	301	Gly	66	318	Pro			
51	303	Ile	67	319	Lys			
52	304	Pro	68	320	Ile			
53	305	Arg	69	321	Asp			
54	306	Ala	70	322	Arg			
55	307	Cys	71	323	Gly			
56	308	Gln	72	324	Trp			
57	309	Lys	73	325	Val			
58	310	Ser	74	326	Cys			
59	311	Leu	75	328	Phe			
60	312	Arg	76	335	Thr			
61	313	Pro	77	336	Leu			
62	314	Val	78	337	Gly			
63	315	Pro	79	338	Leu			
64	316	Pro	80	339	Lys			

Phụ lục B

Khảo sát liên kết Hydro



Hình B.1: Liên kết Hydro theo thời gian giữa các vị trí amino acid và dsRNA. Lần kéo thứ 1 cấu hình B-conf phân tử có PDBID: 3L25



Hình B.2: Liên kết Hydro theo thời gian giữa các vị trí amino acid và dsRNA. Lần kéo thứ 3 cấu hình B-conf phân tử có PDBID: 3L25

Phụ lục C

Các amino acid và các ký hiệu tương ứng

Tên đầy đủ	3 ký tự	1 ký tự	Tên đầy đủ	3 ký tự	1 ký tự
Alanine	ALA	A	Methionine	MET	M
Cysteine	CYS	C	Asparagine	ASN	N
Aspartic acid	ASP	D	Proline	PRO	P
Glutamic acid	GLU	E	Glutamine	GLN	Q
Phenylalanine	PHE	F	Arginine	ARG	R
Glycine	GLY	G	Serine	SER	S
Histidine	HIS	H	Threonine	THR	T
Isoleucine	ILE	I	Valine	VAL	V
Lysine	LYS	K	Tryptophan	TRP	W
Leucine	LEU	L	Tyrosine	TYR	Y