

**ĐẠI HỌC QUỐC GIA TP. HCM
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

PHẠM ĐĂNG LÂN

**MÔ PHỎNG QUÁ TRÌNH CUỘN, PHÂN RÃ SỢI
VÀ ẢNH HƯỞNG TỐC ĐỘ DỊCH MÃ LÊN
DIMER HOÁ CỦA PROTEIN**

Ngành: Vật lý lý thuyết và vật lý toán

Mã số ngành: 62440103

TÓM TẮT LUẬN ÁN TIẾN SĨ VẬT LÝ

Công trình được hoàn thành tại:.....

.....

Người hướng dẫn khoa học: GS. TSKH Mai Xuân Lý

Phản biện 1: PGS.TS. Phạm Cẩm Nam

Phản biện 2: PGS.TS. Đỗ Ngọc Sơn

Phản biện 3: TS. Nguyễn Minh Tâm

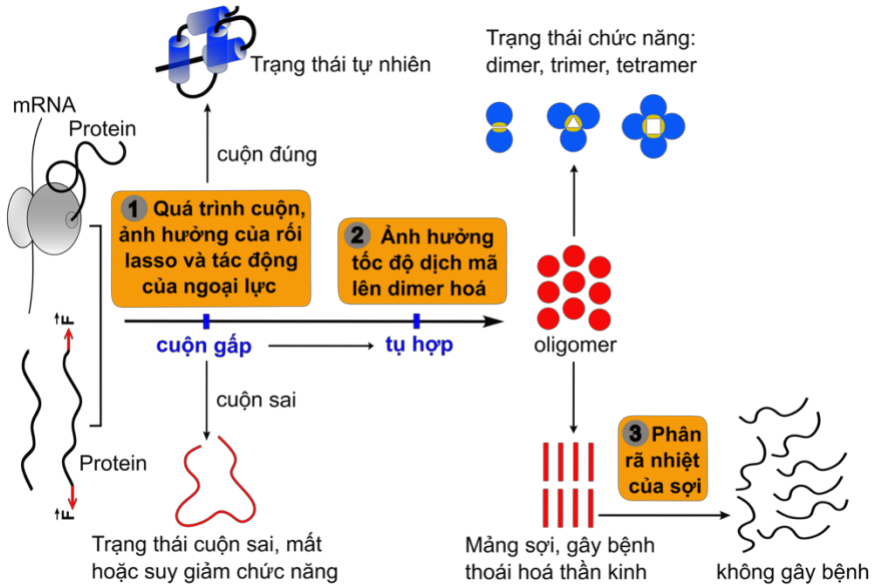
Phản biện độc lập 1: PGS.TS. Trịnh Xuân Hoàng

Phản biện độc lập 2: PGS.TS. Trần Thị Thu Hạnh

L luận án sẽ được bảo vệ trước Hội đồng chấm luận án cấp Cơ sở đào tạo họp tại Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, vào hồi giờ, ngày tháng năm

MỞ ĐẦU

Để thực hiện chức năng sinh học, protein cần phải cuộn gấp chính xác về trạng thái tự nhiên sau khi được tổng hợp từ chuỗi mRNA trong ribosome của tế bào. Khả năng cuộn gấp hiệu quả và biểu hiện chức năng của protein được cho là đã tối ưu trong quá trình chọn lọc tiến hóa. Tuy nhiên, ngày càng có nhiều bằng chứng cho thấy protein cuộn sai là nguyên nhân gây ra nhiều bệnh lý khác nhau. Một số nghiên cứu thực nghiệm và lý thuyết đã chỉ ra rằng các biến động trong quá trình dịch mã hoặc tương tác với các phân tử khác trong tế bào có thể tác động lên hành vi cuộn gấp và tụ hợp của protein, khiến chúng cuộn kém hiệu quả, bị suy giảm chức năng ở dạng đơn phân tử (monomer) hoặc khả năng tụ hợp thành các cấu trúc đa phân tử (oligomer) bậc cao. Protein cuộn sai, bị biến tính nếu không bị phân giải có thể tích tụ lại thành các đám rối làm thoái hóa tế bào. Do đó, việc làm sáng tỏ cơ chế cuộn gấp và quá trình tụ hợp của protein là điều cần thiết. Hiểu biết sâu sắc về các quá trình này, đánh giá ảnh hưởng của các yếu tố nội tại và môi trường bên ngoài đến các biểu hiện chức năng của protein có ý nghĩa quan trọng đối với các lĩnh vực hóa sinh và lý sinh nhằm xác định nguyên nhân protein cuộn sai và dự đoán khả năng tích tụ. Từ đó đề ra các giải pháp hỗ trợ để ngăn ngừa và điều trị các bệnh lý liên quan. Mặc dù vậy, các kỹ thuật thực nghiệm hiện nay còn nhiều hạn chế trong khả năng quan sát chi tiết các biến đổi cấu trúc và các quá trình động học của protein ở cấp độ phân tử. Trong luận án này, tác giả sử dụng các công cụ mô phỏng động lực học phân tử và Monte Carlo để tìm hiểu về cơ chế cuộn gấp và tụ hợp của protein nhằm làm rõ một số giả thuyết và giải thích kết quả thu được từ các quan sát thực nghiệm gần đây.



Hình 1.1: Sơ đồ mô tả lộ trình cuộn và tụ hợp của protein. Protein có thể cuộn chính xác về trạng thái tự nhiên hoặc bị cuộn sai. Protein cuộn sai nếu không bị phân giải bởi protease có thể bị suy giảm chức năng dạng đơn phân tử (monomer) hoặc đa phân tử (oligomer), hoặc tích tụ thành các mảng sợi bền gây tổn hại tế bào. Các ô được tô màu cam để làm nổi bật ba vấn đề được nghiên cứu trong luận án.

Mục tiêu tổng quát: Luận án khảo sát một số vấn đề liên quan đến động học cuộn và tụ hợp protein bao gồm: mô tả quá trình cuộn và tìm hiểu tác động của sự hình thành rối lasso và ngoại lực lên lộ trình, thời gian cuộn và khả năng cuộn gấp chính xác của protein. Phân tích ảnh hưởng của động học dịch mã lên hành vi cuộn và tụ hợp dimer của protein. Khảo sát cơ chế phân rã nhiệt của các cụm sợi, hình thành do sự tích tụ của các protein cuộn sai. Các nội dung nghiên cứu nhằm mục đích cung cấp thêm kiến thức về cơ chế cuộn, nguyên nhân protein cuộn sai, mối liên hệ giữa việc cuộn sai và sự suy giảm

chức năng, khả năng tự hợp của protein. Hình 1.1 cho thấy mối liên kết giữa các nội dung được nghiên cứu trong luận án.

Các mục tiêu cụ thể:

1. Mô phỏng và so sánh quá trình tái cuộn từ trạng thái duỗi của miền bám vào thụ thể (Receptor binding domain - RBD) của protein gai của vi-rút SARS-CoV-1 và SARS-CoV-2 trong hai trường hợp có và không có ngoại lực không đối tác dụng lên hai đầu của protein, đánh giá tương quan giữa các tác động của ngoại lực và sự hình thành rổ lasso lên lộ trình, thời gian cuộn và khả năng cuộn gấp chính xác của protein. Rổ lasso là hình thái cấu trúc của một lớp protein, hình thành trong quá trình cuộn khi một phần của chuỗi protein đóng thành một vòng kín do liên kết không cộng hoá trị giữa hai axit amin (aa) và được luồn qua bởi một hoặc hai đầu của phần chuỗi còn lại.

2. Khảo sát cơ chế ảnh hưởng của tốc độ dịch mã lên cấu trúc và khả năng kết tụ dimer chức năng của hai protein oligoribonuclease và ribonuclease T trong tế bào vi khuẩn *E. Coli*: mô phỏng quá trình các protein được dịch mã ở các tốc độ khác nhau từ chuỗi mRNA tự nhiên và các biến thể mRNA giả định có tốc độ dịch mã nhanh nhất hoặc chậm nhất so với tốc độ tự nhiên; phân tích ảnh hưởng của việc thay đổi tốc độ dịch mã lên quá trình cuộn và tự hợp của protein sau khi quá trình dịch mã kết thúc.

3. Tìm hiểu quá trình phân rã nhiệt sợi Amyloid beta ($A\beta$) ở các nhiệt độ khác nhau: Phân tích động học phân rã các đơn phân tử (monomer) từ cụm sợi mẹ và sự suy giảm tương ứng của số liên kết sợi cho hai trường hợp các monomer phân rã được phép hoặc không được phép tái hợp vào cụm sợi mẹ.

PHƯƠNG PHÁP NGHIÊN CỨU

Mô phỏng trên máy tính là một công cụ hữu ích để hỗ trợ các nghiên cứu trong lĩnh vực lý sinh. Mô phỏng cho phép ghi lại các trạng thái thể hiện sự biến đổi động học của protein trong quá trình cuộn và tụ hợp. Một số đại lượng đặc trưng tính được từ dữ liệu mô phỏng có thể dùng để đối chiếu với kết quả thu được từ thực nghiệm. Các nội dung nghiên cứu của luận án được tiến hành bằng các chương trình mô phỏng động lực học phân tử (MD) và Monte Carlo (MC) kết hợp với mô hình đầy đủ nguyên tử hoặc thu gọn - mô hình thô (một nguyên tử $\text{C}\alpha$ đại diện cho một aa) của protein. Ngoài ra, dữ liệu thí nghiệm tái cuộn của protein oligoribonuclease và ribonuclease T trong vi khuẩn *E. Coli* được sàng lọc, phân tích diện tích tiếp xúc bề mặt dung môi để đánh giá hiệu quả cuộn gấp và so sánh với kết quả mô phỏng.

CÁC NỘI DUNG VÀ KẾT QUẢ NGHIÊN CỨU

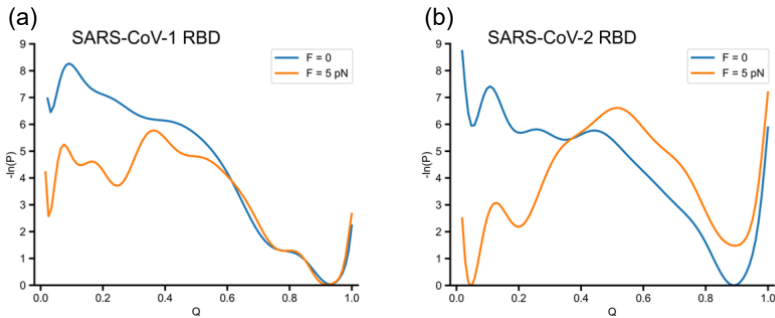
Nội dung 1: Khảo sát quá trình tái cuộn của protein, ảnh hưởng của rối lasso và tác động của ngoại lực

I. Nội dung nghiên cứu: Một nghiên cứu gần đây nhận định rằng sự hình thành rối lasso trong quá trình cuộn của các protein có mô-típ rối lasso chỉ phối thời gian cuộn toàn cục của protein. Quá trình này có thể dẫn đến các trạng thái cuộn sai, ảnh hưởng đến chức năng và hoạt động của protein. Mặt khác, tương tác giữa protein với các phân tử xung quanh trong môi trường tế bào có thể hỗ trợ hoặc cản trở quá trình cuộn gấp của chúng. Một câu hỏi thú vị được đặt ra là: các rối lasso ảnh hưởng như thế nào đến quá trình cuộn khi có và không có tác động từ bên ngoài? Để trả lời câu hỏi này, chúng tôi đã thực hiện mô phỏng động lực học phân tử kết hợp mô hình thô để so sánh lộ trình cuộn của SARS-CoV-1-RBD và SARS-CoV-2-RBD xuất phát từ các cấu hình duỗi khi có và không có một lực hãm không đổi ở hai đầu protein.

II. Các kết quả đạt được:

1. SARS-CoV-2-RBD cuộn hiệu quả hơn so với SARS-CoV-1-RBD

Do tương đồng về trình tự chuỗi và cấu trúc tự nhiên, quá trình cuộn từ trạng thái duỗi của SARS-CoV-1-RBD và SARS-CoV-2-RBD khá giống nhau. Hai trạng thái phổ biến trong bề mặt xác suất theo phân bố của tỷ lệ liên kết tự nhiên Q (các đường màu xanh dương, Hình 4.3a, 4.3b) tương ứng với các cấu hình ban đầu chưa có cấu trúc bậc hai ($Q \sim 0.05$) và các cấu hình ổn định hơn ($Q \sim 0.9$). Tuy nhiên phân tích tỷ lệ phần trăm các quỹ đạo đã cuộn thành công trên tổng số 100 quỹ đạo mô phỏng kéo dài 6 μ s cho thấy SARS-CoV-2-RBD cuộn hiệu quả hơn: 97% đối với SARS-CoV-2-RBD và 56% đối với SARS-CoV-1-RBD. Thời gian cuộn tương ứng được ước tính lần lượt là 257.4 ns và 369.0 ns.



Hình 4.3: Sự phụ thuộc của $-\ln(P)$ vào Q , trong đó P là xác suất tìm thấy một giá trị Q từ các cấu hình protein được lấy mẫu trong mô phỏng của (a) SARS-CoV-1-RBD và (b) SARS-CoV-2-RBD. Khi có lực kéo $F = 5$ pN, một cực tiểu mới được quan sát ở $Q \sim 0.2$. Giá trị của $-\ln(P)$ tính theo đơn vị năng lượng $k_B T$.

2. SARS-CoV-2-RBD nhạy cảm hơn với lực so với SARS-CoV-1-RBD

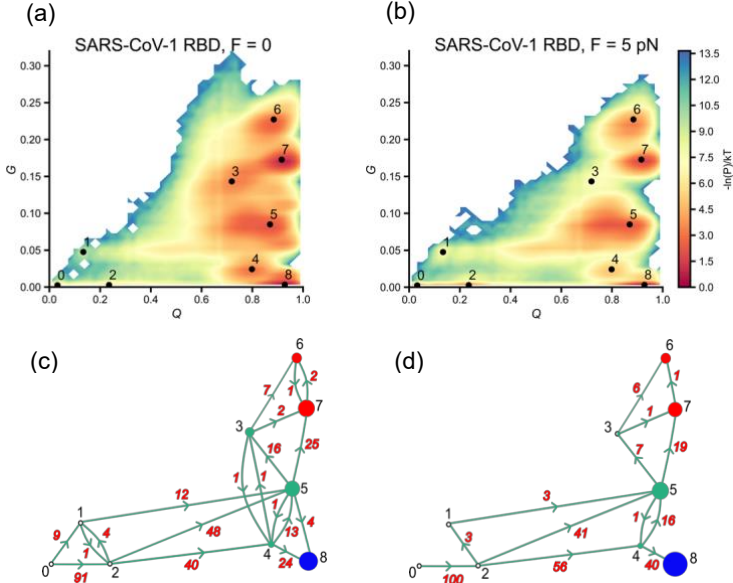
Bề mặt xác suất theo phân bố của tham số Q trong điều kiện có lực được minh họa bằng các đường màu cam trong Hình 4.3. So với trường hợp không có lực, sự xuất hiện thêm một cực tiểu tại $Q \sim 0.2$ cho thấy quá trình cuộn

của cả SARS-CoV-1-RBD và SARS-CoV-2-RBD bị chậm lại do tác động của lực. SARS-CoV-2-RBD tỏ ra nhạy cảm hơn với ngoại lực, với sự hiện diện của một rào thế cao trên đồ thị $-\ln(P)$ tại $Q \sim 0.5$. Trong điều kiện có lực, 58% các quỹ đạo mô phỏng của SARS-CoV-2-RBD không tái cuộn thành công vào cuối thời gian mô phỏng 6 μ s, so với chỉ 3% của trường hợp không có lực. Như dự đoán, thời gian cuộn của SARS-CoV-2-RBD chậm hơn đáng kể khi có lực (5646 ns) so với khi không có lực (257 ns). SARS-CoV-1-RBD ít nhạy cảm hơn với lực, biểu hiện bởi một rào thế thấp hơn trên đồ thị $-\ln(P)$ tại $Q \sim 0.4$. Sự khác biệt về tỷ lệ các quỹ đạo đã cuộn xong khi có và không có lực là không đáng kể, lần lượt là 58% và 56%. Thời gian cuộn ước tính của SARS-CoV-1-RBD cũng tương đương nhau trong cả hai trường hợp (369 ns khi $F = 0$ và 304.95 ns khi $F = 5$ pN).

3. Các trạng thái trung gian chứa rối không tự nhiên xuất hiện trong quá trình cuộn của SARS-CoV-1-RBD và SARS-CoV-2-RBD

Để mô tả chi tiết hơn quá trình cuộn, chúng tôi đã phân tích bề mặt xác suất kết hợp các tham số Q và G , trong đó G đặc trưng cho sự thay đổi về trạng thái rối so với cấu hình tự nhiên. Không gian cấu hình (Q, G) được phân cụm và phân chia thành các trạng thái metastable nhằm xác định vai trò của các trạng thái trung gian, nhận diện các trạng thái cuộn sai trong quá trình cuộn của các RBD và diễn đạt các lộ trình cuộn. Có thể thấy rõ từ các bề mặt xác suất (Q, G) của SARS-CoV-1-RBD (Hình 4.6a, 4.6b) và SARS-CoV-2-RBD (Hình 4.7a, 4.7b), một lượng lớn cấu hình được lấy mẫu có liên quan đến các trạng thái trung gian có thể dẫn đến cuộn sai do sự hình thành các liên kết tự nhiên có sự thay đổi về rối. Điều này giải thích vì sao nhiều quỹ đạo không cuộn thành công sau khi các mô phỏng kết thúc. Trong số các trạng thái trung gian, trạng thái 7 (đối với SARS-CoV-1-RBD) và trạng thái 5 (đối với SARS-CoV-2-RBD) hình thành cấu trúc gần giống trạng thái tự nhiên nhất với tỷ lệ liên kết tự nhiên cao ($Q \sim 0.9$) nhưng lại có tới 15%-20% các liên kết có sự

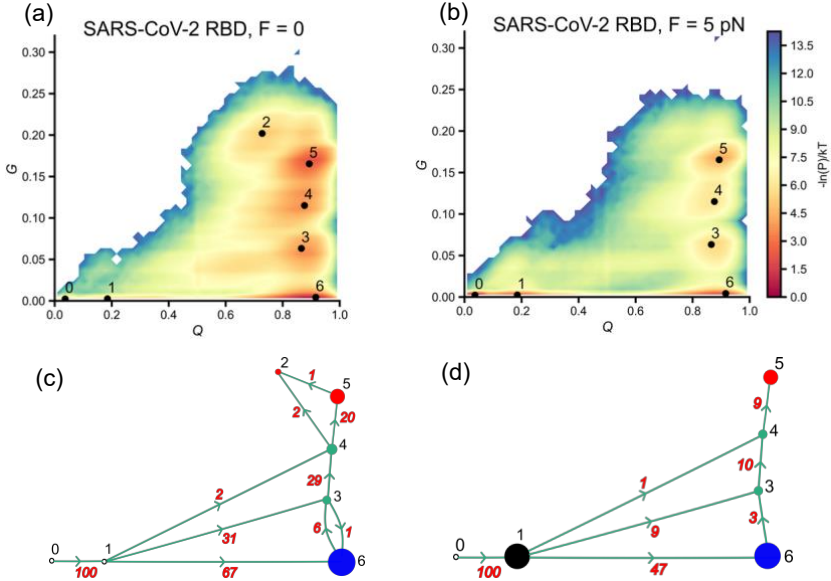
thay đổi về rồi ($G \sim 0.15-0.2$), tức là không thể phân biệt chúng với trạng thái tự nhiên nếu không phân tích chi tiết các khác biệt về cấu trúc.



Hình 4.6: (a) và (b) Bề mặt xác suất được đặc trưng bởi phân bố $-\ln(P)$, trong đó P là xác suất tìm thấy một cặp giá trị (Q, G) từ các cấu hình protein được lấy mẫu trong mô phỏng của SARS-CoV-1-RBD với $F = 0$ và $F = 5$ pN. (c) và (d) Mạng lưới chuyển tiếp của các quỹ đạo độc lập giữa các trạng thái metastable (được biểu diễn bằng các nút). Các nút màu xanh dương, xanh lá và đỏ tương ứng với các trạng thái cuộn đúng, trung gian và cuộn sai. Các số màu đỏ cho biết số lần chuyển đổi giữa các trạng thái.

4. Ngoại lực giúp SARS-CoV-1-RBD cuộn theo các lộ trình hiệu quả hơn, tránh được các trạng thái cuộn sai chứa các rối không tự nhiên

Hình 4.6c, 4.7c mô tả lộ trình cuộn với $F = 0$ của SARS-CoV-1-RBD và SARS-CoV-2-RBD. Nhìn chung, SARS-CoV-1-RBD đi theo hai lộ trình song song



Hình 4.7: (a) và (b) Bề mặt xác suất được đặc trưng bởi phân bố $-\ln(P)$, trong đó P là xác suất tìm thấy một cặp giá trị (Q, G) từ các cấu hình protein được lấy mẫu trong mô phỏng của SARS-CoV-2-RBD với $F = 0$ và $F = 5$ pN. (c) và (d) Mạng lưới chuyển tiếp của các quỹ đạo độc lập giữa các trạng thái metastable, mô tả bằng các nút. Các nút màu xanh dương, xanh lá và đỏ tương ứng với các trạng thái cuộn đúng, trung gian và cuộn sai. Nút màu đen (trạng thái 1) mô tả 43/100 quỹ đạo vẫn ở trong trạng thái chưa cuộn ở cuối mô phỏng. Các số màu đỏ cho biết số lần chuyển đổi giữa các trạng thái.

đến các trạng thái trung gian ($2 \rightarrow 4$ và $1 \mid 2 \rightarrow 5$). 40% các quỹ đạo đến trạng thái 4 đặc trưng bởi G thấp ($G < 0.05$), trong khi 60% đến trạng thái 5 có G cao hơn ($G \sim 0.05-0.1$). Các quỹ đạo đến trạng thái 5 phần lớn sẽ chuyển tiếp đến các trạng thái 6 và 7. Các trạng thái này có G cao nhất ($G > 0.15$) và được xem là các trạng thái cuộn sai vì khi các quỹ đạo thăm dò những trạng thái này có xu hướng bị mắc kẹt, không thể quay lại lộ trình cuộn đúng về trạng thái tự nhiên (Hình 4.6c). Quá trình cuộn của SARS-CoV-2-RBD đi theo

một lộ trình chủ đạo (67% tổng số quỹ đạo) hướng đến trạng thái cuộn đúng (trạng thái 6), trong khi một lộ trình ít phổ biến hơn dẫn đến trạng thái trung gian 3 với G trong khoảng 0.05-0.1 (Hình 4.7c). Tương tự như những gì quan sát được với SARS-CoV-1-RBD, các trạng thái 2 và 5 là các trạng thái cuộn sai vì các quỹ đạo bị mắc kẹt trong những trạng thái này.

Ngoại lực gây ra các ảnh hưởng không đồng nhất lên hành vi cuộn của hai protein. So sánh các sơ đồ cuộn của $F = 0$ và $F = 5$ pN của SARS-CoV-1-RBD (Hình 4.6c và 4.6d), có thể thấy rằng lực góp phần cải thiện số lượng quỹ đạo đi theo các lộ trình dẫn đến cuộn gấp chính xác. Cụ thể, khi có lực 56% các quỹ đạo theo lộ trình $2 \rightarrow 4$ (có xu hướng cuộn đúng) và 44% theo lộ trình $1 \mid 2 \rightarrow 5$ (có xu hướng cuộn sai). Những con số này tương ứng là 40% và 60% trong trường hợp không có lực. Do số lượng quỹ đạo đi đến trạng thái 4 (có G thấp) tăng lên, tỷ lệ các quỹ đạo đã cuộn thành công đến trạng thái tự nhiên 5 trong mô phỏng có lực (40%) cao hơn so với mô phỏng không có lực (28%). Lực cũng hạn chế sự hình thành các rối không tự nhiên ở giai đoạn đầu của quá trình cuộn, vì không có quỹ đạo nào theo lộ trình $0 \rightarrow 1$ khi có lực tác động (Hình 4.6d). Đối với SARS-CoV-2-RBD, việc đưa ngoại lực vào các mô phỏng dẫn đến 43% các quỹ đạo vẫn ở trong các cấu hình chưa cuộn ($Q \sim 0.2$) ở cuối mô phỏng (trạng thái 1, Hình 4.7b, 4.7d). Nhìn chung ngoại lực làm chậm quá trình cuộn của SARS-CoV-2-RBD như dự đoán.

Nội dung 2: Ảnh hưởng của tốc độ dịch mã lên dimer hóa của protein

I. Nội dung nghiên cứu: Quan sát thực nghiệm chỉ ra rằng việc thay thế các codon hiếm bằng các codon đồng nghĩa có tốc độ dịch mã nhanh hơn đã làm giảm 60% năng lượng liên kết của protein FRQ với WC-2, làm mất chức năng đồng hồ sinh học của phức dimer này trong nhiều ngày. Để làm rõ cơ chế ảnh hưởng của động học dịch mã lên quá trình dimer hóa, chúng tôi tiến hành các mô phỏng với hai protein có cấu trúc chức năng dạng homodimer

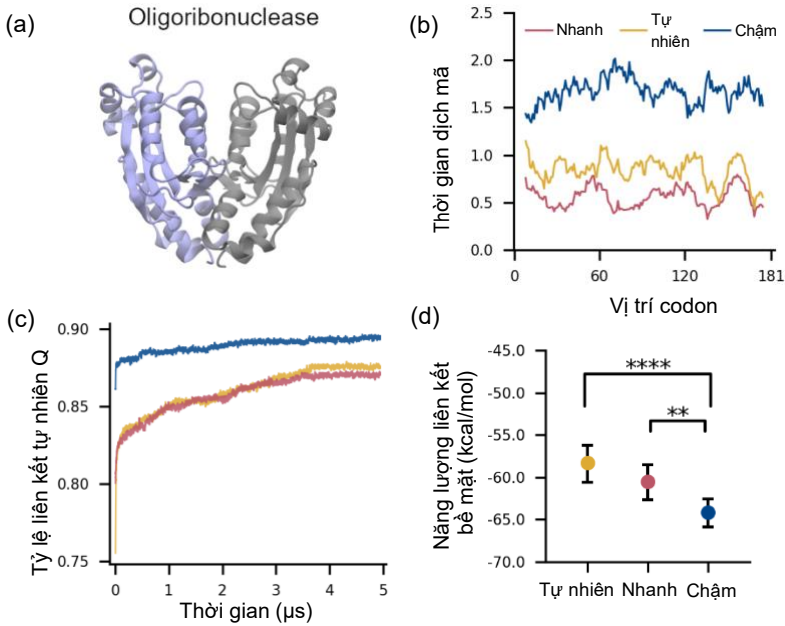
trong tế bào chất của vi khuẩn *E. Coli* là oligoribonuclease và ribonuclease T do protein FRQ về bản chất không có cấu trúc ổn định. Các bước mô phỏng bao gồm: quá trình tổng hợp protein trong ribosome, quá trình cuộn và tự hợp sau khi protein thoát ra khỏi ribosome. Để nghiên cứu ảnh hưởng của việc thay đổi tốc độ dịch mã lên quá trình cuộn và dimer hóa của protein sau khi quá trình dịch mã kết thúc, các mô phỏng tổng hợp protein được thực hiện ở các tốc độ dịch mã khác nhau: tốc độ bình thường tương ứng chuỗi mRNA tự nhiên và các tốc độ nhanh hoặc chậm với các biến thể mRNA giả định. Các biến thể này được tạo ra bằng cách tạo đột biến codon đồng nghĩa từ chuỗi mRNA tự nhiên để có tốc độ dịch mã nhanh nhất hoặc chậm nhất so với tốc độ tự nhiên.

II. Các kết quả đạt được:

1. Các đột biến đồng nghĩa làm thay đổi cấu trúc sau dịch mã và khả năng dimer hóa của oligoribonuclease trong khi không có ảnh hưởng đáng kể đối với ribonuclease T

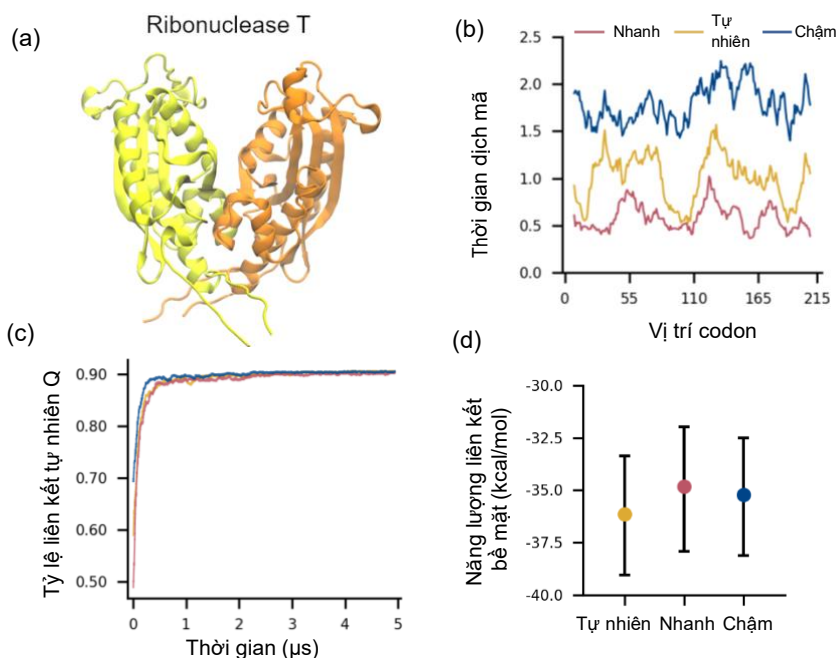
Phân tích về tỷ lệ liên kết tự nhiên, được tính trung bình trên 200 quỹ đạo mô phỏng (100 monomer A và 100 monomer B) chỉ ra rằng khả năng cuộn gấp chính xác của oligoribonuclease thay đổi tùy thuộc vào tốc độ dịch mã từ các chuỗi mRNA khác nhau (Hình 5.1c). Cụ thể, tốc độ dịch mã chậm làm tăng tỷ lệ các cấu trúc cuộn đúng, thể hiện ở tỷ lệ liên kết trung bình cao ($Q = 0.89$) với khoảng tin cậy (CI) 95% là $[0.87, 0.90]$, được tính toán từ 10^6 lần gieo mẫu ngẫu nhiên bằng phương pháp Giá trị này cao hơn so với protein được dịch mã từ chuỗi mRNA tự nhiên ($Q = 0.86$, 95% CI: $[0.85, 0.88]$). Phân bố năng lượng của các cấu trúc dimer thu được từ các mô phỏng ủ nhiệt cho thấy sự gia tăng tỷ lệ các cấu trúc cuộn đúng từ quá trình dịch mã chậm dẫn đến năng lượng liên kết bề mặt trung bình cao hơn: -64.1 kcal/mol (95% CI: $[-65.7, -62.4]$) so với các dimer được tạo ra từ chuỗi mRNA tự nhiên: -58.3 kcal/mol (95% CI: $[-60.3, -55.9]$). Kết quả kiểm chứng bằng phép thử

hoán vị 10^6 lần trên hai tập dữ liệu về năng lượng liên kết của các cấu hình dimer thu được từ dịch mã chậm và tự nhiên cho thấy sự khác biệt có ý nghĩa thống kê ($p = 2 \times 10^{-5}$). Năng lượng liên kết trung bình



Hình 5.1: (a) Cấu trúc tự nhiên ở dạng dimer của oligoribonuclease. (b) Thời gian dịch mã trung bình của các vị trí codon trong các chuỗi mRNA dịch mã nhanh, tự nhiên và chậm. (c) Sự phụ thuộc thời gian của tỷ lệ liên kết tự nhiên trung bình của các cấu hình monomer kể từ lúc thoát ra khỏi ribosome đối với từng chuỗi mRNA. (d) Năng lượng liên kết bề mặt trung bình giữa monomer A và B tính trung bình trên 200 cặp lấy ngẫu nhiên từ các mô phỏng ủ nhiệt. Các dấu móc vuông và dấu “*” để chỉ khác biệt có ý nghĩa thống kê của việc so sánh các giá trị trung bình bằng phép thử hoán vị với 10^6 lần gieo mẫu. Một, hai, hoặc ba dấu “*” lần lượt ứng với $p \leq 0.05$, $p \leq 0.01$, $p \leq 0.001$.

của các cấu trúc dimer từ protein được dịch mã nhanh là -60.5 kcal/mol (95% CI: $[-62.5, -58.4]$) cũng cho thấy sự khác biệt có ý nghĩa thống kê so với protein được dịch mã chậm (Hình 5.1d). Các kết quả thu được chứng tỏ rằng khả năng dimer hóa của oligoribonuclease bị ảnh hưởng bởi tốc độ dịch mã. Trái với những quan sát trên oligoribonuclease, các đặc điểm cấu trúc sau dịch mã và năng lượng liên kết cho thấy quá trình dimer hóa của ribonuclease T không bị ảnh hưởng đáng kể bởi tốc độ dịch mã mRNA (Hình 5.2). Các tính toán về năng lượng liên kết của các cấu trúc dimer được hình thành từ các mRNA có tốc độ dịch mã tự nhiên, nhanh và chậm không cho thấy sự khác biệt đáng kể nào có ý nghĩa thống kê.



Hình 5.2: Tương tự Hình 5.1 nhưng cho ribonuclease T.

2. Các đột biến đồng nghĩa làm thay đổi sự phân bố các cấu trúc oligoribonuclease bị rối

Phân tích tỷ lệ cấu trúc rối xuất hiện trong các cấu trúc monomer (từ 200 quỹ đạo độc lập sau mô phỏng dịch mã) và dimer (200 cấu trúc sau mô phỏng ủ nhiệt) cho thấy tỷ lệ cấu trúc rối xuất hiện rất ít ở ribonuclease T ($G < 0.15$) trong cả dạng monomer và dimer và không có sự khác biệt đáng kể nào do tốc độ dịch mã gây ra. Mặc dù có sự khác biệt có ý nghĩa thống kê nhưng mức chênh lệch này nhỏ hơn 10.5%. Điều này giải thích vì sao năng lượng liên kết dimer của ribonuclease T không nhạy cảm với các đột biến đồng nghĩa. Mức độ phổ biến của các trạng thái cuộn sai là thấp và do đó không thể thấy được ảnh hưởng của tốc độ dịch mã lên khả năng dimer hóa của protein ribonuclease T. Ngược lại, oligoribonuclease cho thấy một lượng đáng kể các cấu trúc bị rối và tỷ lệ này thay đổi phụ thuộc vào tốc độ dịch mã. Tỷ lệ phổ biến của các trạng thái cuộn sai có chứa rối trong cấu trúc dimer tương ứng với các tốc độ dịch mã tự nhiên, nhanh và chậm lần lượt là 0.62 (95% CI: [0.54, 0.68]), 0.51 (95% CI: [0.43, 0.57]) và 0.37 (95% CI: [0.29, 0.43]). Kết quả này phù hợp với các tính toán năng lượng liên kết trung bình (Hình 5.1d), các giá trị tương ứng lần lượt là -58.3, -60.5 và -64.1 kcal/mol và cho thấy sự thay đổi trong tốc độ dịch mã dẫn đến sự thay đổi trong mức độ phổ biến của các trạng thái rối. Các vùng rối xuất hiện trong cấu trúc của monomer có tác động lên tương tác giữa các monomer từ đó làm thay đổi ái lực liên kết của các monomer trong quá trình dimer hóa.

3. Các trạng thái rối cuộn sai thường nằm ở vùng bề mặt của dimer oligoribonuclease, làm giảm ái lực liên kết dimer

Từ các phân tích tham số G đặc trưng cho sự thay đổi về trạng thái rối so với cấu trúc tự nhiên, chúng tôi xác định được hai vùng rối xuất hiện thường xuyên trong các cấu trúc mô phỏng sau dịch mã của oligoribonuclease: Dạng 1 là các aa 96-102 luôn qua vòng kín được đóng bởi liên kết tự nhiên giữa

hai aa 31-41, Dạng 2 là các aa từ 125-129 luồn qua vòng kín giữa hai aa 9-103. Một số cấu trúc rối rất giống với cấu trúc tự nhiên với $RMSD \leq 3$ Å. Sự tương đồng về cấu trúc, khó phân biệt được giữa trạng thái rối với trạng thái tự nhiên được giả thuyết là nguyên nhân protein vẫn có thể hòa tan và tránh được các hệ thống kiểm soát chất lượng protein.

Mặc dù cấu trúc tổng thể của monomer có thể giống với trạng thái cuộn gấp chính xác, sự xuất hiện của các rối có thể gây xáo trộn cấu trúc tại vùng tương tác bề mặt của dimer. Rối xuất hiện ở đoạn aa từ 125 đến 129 trong monomer A đã ngăn cản sự hình thành một phiến beta quan trọng tại bề mặt dimer của oligoribonuclease. Điều này cho thấy rằng những thay đổi trong năng lượng liên kết dimer có thể là do các rối làm gián đoạn cấu trúc bề mặt tương tác dimer. Do đó, chúng tôi phân tích sự xuất hiện của các rối tại bề mặt dimer bằng cách tính tỷ lệ các cấu trúc monomer và dimer có các aa bị rối nằm trong vùng tiếp xúc bề mặt. Đối với ribonuclease T, do tỷ lệ rối thấp trong monomer, tỷ lệ rối ở bề mặt dimer của cũng thấp chiếm khoảng 15%. Trong khi đó, các cấu trúc cuộn sai có rối ở bề mặt tương tác xuất hiện thường xuyên hơn ở oligoribonuclease trong cả cấu trúc monomer và dimer được bắt cặp ngẫu nhiên. Đáng chú ý, khi xét trên tổng số các cấu trúc dimer có rối ít nhất ở một trong hai monomer thành phần, tỷ lệ vùng rối xuất hiện ở bề mặt tương tác khá cao, lần lượt là 0.49 (95% CI: [0.41, 0.55]), 0.38 (95% CI: [0.31, 0.45]), và 0.26 (95% CI: [0.21, 0.32]), tương ứng với tốc độ dịch mã tự nhiên, nhanh và chậm. Do đó, các rối không tự nhiên xuất hiện phổ biến ở bề mặt dimer của oligoribonuclease.

Để xem xét liệu rằng các rối có làm giảm năng lượng liên kết giữa các monomer hay không, chúng tôi tạo ra một bảng ngẫu nhiên (2 x 2) phân loại tất cả 200 cấu trúc dimer sau quá trình ủ nhiệt thành hai nhóm các cấu trúc có liên kết mạnh và yếu, cùng với đó phân tích sự hiện diện của rối trong mỗi nhóm cấu trúc. Đối với oligoribonuclease, kết quả cho thấy tương quan giữa

sự hiện diện của các rỗi và sự xuất hiện của các dimer có năng lượng liên kết yếu. Ngược lại, ribonuclease T không cho thấy bất kỳ khác biệt có ý nghĩa thống kê nào ở cả ba tốc độ dịch mã. Điều này chỉ ra rằng nhóm cấu trúc dimer có rỗi của ribonuclease T không quyết định việc hình thành dimer có năng lượng liên kết mạnh hay yếu.

4. Trạng thái rỗi cuộn sai là bất động học trong quá trình cuộn của protein và có thể tồn tại lâu dài

Trong khi tất cả các monomer của ribonuclease T cuộn xong trong khoảng 0.8 μ s sau khi thoát ra khỏi ribosome, một số quỹ đạo oligoribonuclease không thể cuộn sau 5 μ s mô phỏng sau dịch mã ở cả ba tốc độ dịch mã. Cụ thể, tỷ lệ monomer chưa cuộn xong là 13% (95% CI: [8%, 17%]), 14% (95% CI: [9%, 19%]) và 10% (95% CI: [6%, 14%]) tương ứng với dịch mã tự nhiên, nhanh và chậm. Từ việc khớp dữ liệu động học, chúng tôi ước lượng một số quỹ đạo của oligoribonuclease cần từ 6 ms đến 14 ms để cuộn xong, chúng tôi các trạng thái này bị mắc kẹt về mặt động học. Để kiểm chứng kết quả từ mô phỏng mô hình thô, chúng tôi tái hiện cấu trúc đầy đủ nguyên tử của 5 dimer và ba monomer khác nhau của oligoribonuclease có xuất hiện rỗi và có năng lượng thấp nhất. Sau đó với mỗi cấu trúc, chúng tôi tiến hành mô phỏng ba quỹ đạo độc lập MD, mỗi quỹ đạo kéo dài 500 ns. Trong mỗi trường hợp, trạng thái rỗi trong mô hình thô vẫn tồn tại ở độ phân giải đầy đủ nguyên tử trong suốt thời gian mô phỏng. Ngoài ra, chúng tôi chọn ngẫu nhiên một cấu trúc dimer và monomer để kéo dài thời gian mô phỏng lên 1 μ s và thấy rằng các trạng thái rỗi vẫn tồn tại.

5. Các trạng thái rỗi phù hợp với các quan sát thực nghiệm về các thay đổi cấu trúc trong quá trình tái cuộn của oligoribonuclease

Dữ liệu từ thí nghiệm phân tích proteolysis khối phổ giới hạn (LiP-MS) thu được từ các nghiên cứu tái cuộn trên toàn bộ proteome được thực hiện trên

các chiết xuất *E. Coli* chỉ ra rằng nhiều loại protein của *E. Coli* không thể tái cuộn hoàn toàn từ các trạng thái bị biến tính. Chúng tôi đề xuất nhóm thực nghiệm phân tích lại dữ liệu tái cuộn để tìm kiếm các đoạn peptide bị cuộn sai, tập trung vào các protein oligoribonuclease và ribonuclease T. Trong thực nghiệm, các đoạn peptide cuộn sai sẽ thay đổi khả năng tiếp cận protease dẫn đến thay đổi phân bố so với trạng thái tự nhiên và được phát hiện. Phân tích cho thấy oligoribonuclease có hai đoạn peptide có sự thay đổi phân bố có ý nghĩa thống kê, tương ứng các đoạn aa 84-94 và 166-175. Trong số 5 vị trí được xác định có thay đổi trên ribonuclease T, không có thay đổi nào có ý nghĩa thống kê về độ nhạy với protease sau khi protein cuộn xong. Chúng tôi tiến hành phân tích bề mặt xác suất được xác định bởi các tham số Q và G cho toàn bộ cấu trúc mô phỏng bao gồm dịch mã tự nhiên, nhanh và chậm và tính toán sự thay đổi tương đối của diện tích bề mặt tiếp xúc với dung môi của các trạng thái rối đặc trưng ở các đoạn aa 84-94 và 166-175 của oligoribonuclease và các đoạn [S2], [204-215] của ribonuclease T để đối chiếu với các dữ liệu LiP-MS. Cả hai protein đều chứa một số trạng thái với tỷ lệ cao các liên kết tự nhiên có sự thay đổi về rối mặc dù đã cuộn về trạng thái giống tự nhiên. Hai trạng thái metastable của oligoribonuclease cho thấy sự gia tăng diện tích bề mặt tiếp xúc với dung môi của các aa 84-94 so với trạng thái tự nhiên từ 10% đến 16% trên toàn bộ tập hợp cấu hình mô phỏng. Tương tự, đối với các aa 166-175, hai trạng thái metastable đặc trưng cho thấy sự gia tăng có ý nghĩa thống kê diện tích bề mặt tiếp xúc với dung môi và chiếm 65-80% số cấu hình trong toàn bộ tập hợp cấu hình mô phỏng. Do đó, các trạng thái metastable trong tập hợp trạng thái rối quan sát từ mô phỏng phù hợp với sự gia tăng khả năng tiếp cận protease được quan sát thực nghiệm ở hai đoạn peptide tương ứng. Trong khi đó, ribonuclease T chỉ hiển thị một trạng thái metastable phổ biến nhất. Các phân tích cho thấy không có sự thay đổi đáng kể diện tích bề mặt tiếp xúc với dung môi so với trạng thái tự nhiên ở hai trong 5 đoạn peptide được chọn ngẫu nhiên để đối chiếu dữ

liệu LiP-MS. Quan sát này cũng có quan điểm rằng ribonuclease T cuộn hiệu quả và ít bị tác động bởi động học dịch mã.

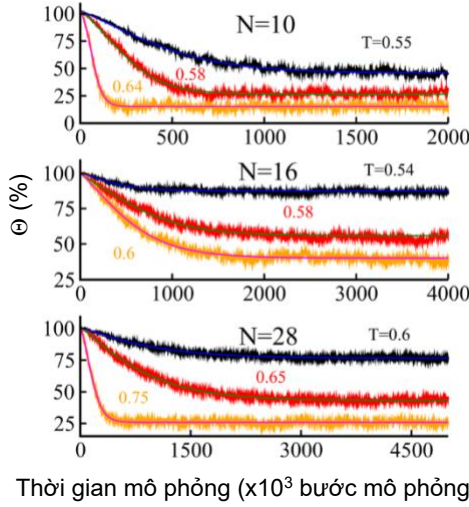
Nội dung 3: Sự phân rã nhiệt của sợi

I. Nội dung nghiên cứu: trong thí nghiệm của Kardos và cộng sự, động học phân rã nhiệt của sợi dựa trên sự suy giảm số lượng liên kết sợi đo được bằng tín hiệu huỳnh quang ThT có thể được mô tả bằng một hàm mũ đôi với hai thang thời gian khác biệt tương ứng với các pha suy giảm tỷ lệ sợi với tốc độ khác nhau. Gruning và đồng nghiệp cũng nghiên cứu quá trình phân giải của sợi A β bằng cách bố trí thí nghiệm để không cho các monomer tách ra có thể tái hợp lại vào cụm mẹ. Các phân tích về sự suy giảm số lượng sợi từ tín hiệu huỳnh quang tryptophan cho thấy quá trình phân rã của sợi tuân theo quy luật hàm mũ đơn. Sự khác biệt trong động học phân rã của số lượng monomer trong cụm sợi và số lượng liên kết sợi là một kết quả thú vị và cần được kiểm chứng. Vì vậy chúng tôi đã đề xuất lý thuyết giải tích, kết hợp với các phương pháp mô phỏng MD, MC và các mô hình protein có độ phân giải khác nhau để tìm hiểu cơ chế phân rã nhiệt của các sợi có kích thước khác nhau và so sánh với các quan sát thực nghiệm. Do các monomer phân rã có thể tái hợp lại cụm sợi nên các mô phỏng sẽ được tiến hành cho hai trường hợp các monomer được phép tái hợp và không được phép tái hợp (mô tả thí nghiệm của Gruning)

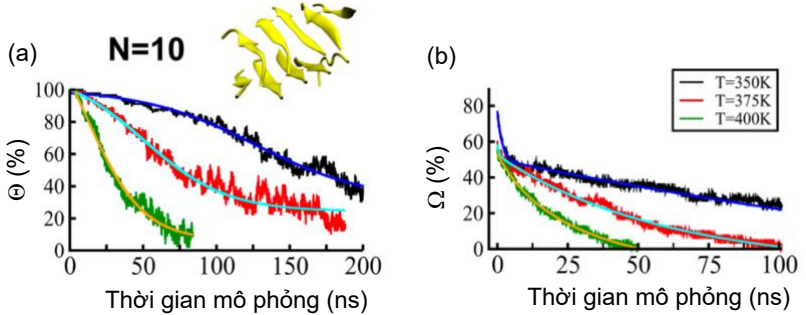
II. Các kết quả đạt được:

1. Sự suy giảm theo thời gian của số lượng monomer tự hợp tuân theo động học mô tả bởi hàm logistic trong khi sự suy giảm theo thời gian của số lượng liên kết sợi tuân theo động học mô tả bởi hàm mũ đôi

Trường hợp các monomer phân rã được phép tái hợp với cụm sợi:



Hình 6.1: Sự phụ thuộc thời gian của tỷ lệ các monomer tụ hợp $\Theta(t)$ cho trường hợp monomer được phép tái hợp ở các nhiệt độ khác nhau được khớp theo các đường cong logistic.

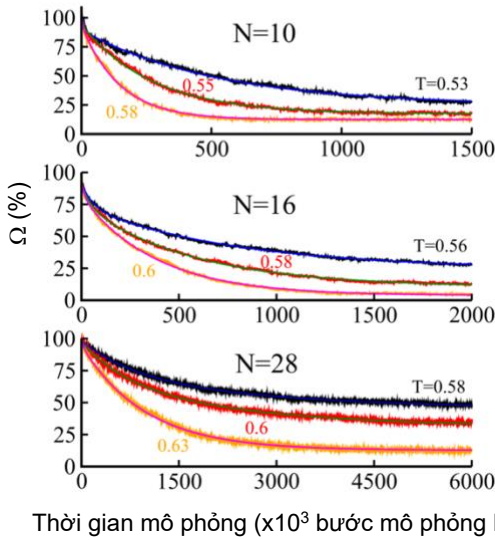


Hình 6.2: Cấu trúc giống sợi của 10 peptide $A\beta_{37-42}$. Cấu trúc này có tất cả 113 liên kết sợi. (a) Sự suy giảm của tỷ lệ các monomer tụ hợp $\Theta(t)$ theo thời gian mô phỏng được phép tái hợp. Đường khớp dữ liệu có dạng logistic được áp dụng cho các nhiệt độ khác nhau. (b) Sự suy giảm của tỷ lệ các liên kết sợi, $\Omega(t)$. Đường khớp dữ liệu có dạng hàm mũ đôi. Các kết quả thu được bằng mô phỏng đầy đủ nguyên tử.

Hình 6.1 thể hiện sự phụ thuộc thời gian của tỷ lệ phần trăm monomer trong cụm sợi với các hệ ban đầu gồm $N = 10, 16$ và 28 monomer ở các nhiệt độ khác nhau trên nhiệt độ phòng ($T > 0.39$). Phân tích kết quả mô phỏng phù hợp với lý thuyết về cơ chế phân rã của cụm sợi cho trường hợp monomer được phép tái hợp, tức là có thể được mô tả tốt bởi hàm logistic.

Để kiểm chứng kết quả từ mô phỏng bằng mô hình lưới, chúng tôi đã tiến hành mô phỏng MD sử dụng mô hình đầy đủ nguyên tử kết hợp dung môi nước. Mô phỏng bắt đầu với cấu trúc giống sợi gồm 10 đoạn peptide $A\beta_{37-42}$ ở các nhiệt độ $T = 350$ K, 375 K và 400 K (Hình 6.2). Kết quả thu được tương tự như kết quả trong mô hình lưới cho thấy tỷ lệ phần trăm số lượng monomer còn lại trong cụm sợi $\Theta(t)$ tuân theo quy luật logistic. Các hàm đa mũ cũng phù hợp tuy nhiên không mô tả được đầy đủ toàn quá trình phân rã. Cụ thể, các hàm mũ đôi và mũ ba đã được dùng để khớp dữ liệu mô phỏng nhưng cả hai loại hàm đều không mô tả được sự suy giảm chậm ban đầu của sợi ở $T = 350$ K và 375 K. Tuy nhiên với $T = 400$ K, cả hàm logistic và mũ đơn đều mô tả tốt. Như vậy khi nhiệt độ đủ cao, quá trình phân rã sẽ diễn ra đủ nhanh thì các hàm khớp đơn giản mới có thể mô tả đầy đủ quá trình rã trên toàn bộ thang thời gian quan sát.

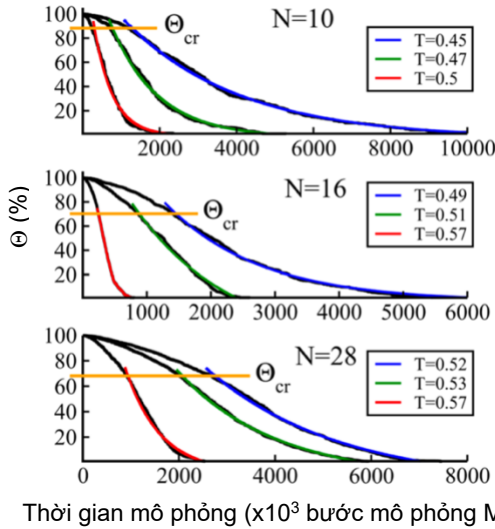
Thí nghiệm quan sát phân giải sợi bằng phổ huỳnh quang Thioflavin-T (ThT) cho thấy sự phụ thuộc thời gian của tỷ lệ phần trăm liên kết sợi, $\Omega(t)$ có thể được mô tả bởi một hàm mũ đôi. Từ dữ liệu mô phỏng như quan sát được trên Hình 6.3, tỷ lệ liên kết sợi còn lại khớp tốt với hàm mũ đôi, phù hợp tốt với thực nghiệm. Sự tồn tại của hai thang thời gian τ_1 và τ_2 ($\tau_1 < \tau_2$) cho thấy rằng quá trình phân rã của sợi trải qua một trạng thái trung gian. τ_1 và τ_2 lần lượt mô tả sự suy giảm nhanh của số lượng liên kết ở giai đoạn đầu và chậm hơn sau đó khi số lượng monomer trong cụm sợi giảm dần.



Hình 6.3: Sự phụ thuộc thời gian của tỷ lệ liên kết sợi $\Omega(t)$ cho trường hợp monomer được phép tái hợp ở các nhiệt độ khác nhau. Đường cong khớp dữ liệu tương ứng với dạng hàm mũ đôi.

Để kiểm chứng các kết quả thu được từ mô hình lưới và mô hình đầy đủ nguyên tử của 10 peptide $A\beta_{37-42}$, chúng tôi khảo sát thêm quá trình suy giảm của một cấu trúc sợi dài hơn bao gồm 5 đoạn peptide $A\beta_{17-42}$ (mã PDB: 2BEG). Cấu trúc 2BEG có tất cả 184 liên kết sợi. Kết quả cũng cho thấy tỷ lệ liên kết sợi ở 300 K giảm nhanh theo quy luật mũ. Mặc dù cả hàm mũ đơn và mũ đôi đều khớp tốt với dữ liệu mô phỏng nhưng chỉ hàm mũ đôi mới mô tả được giai đoạn ban đầu của quá trình phân rã. Tốc độ tương ứng giữa hai giai đoạn mô tả bởi hàm mũ đôi chênh nhau khoảng 22 lần. Như vậy, các kết quả mô phỏng này tương đồng về mặt định tính với quy luật hàm mũ được quan sát trong các thí nghiệm về sự phân rã của sợi amyloid $\beta 2m$ ở vùng nhiệt độ cao cũng như ở nhiệt độ phòng.

Trường hợp các monomer phân rã không được tái hợp với cụm sợi:



Hình 6.10: Sự phụ thuộc của tỷ lệ monomer tụ hợp Θ vào thời gian trong mô phỏng không tái hợp. Các đường cong được khớp bởi hàm mũ đơn sau khi $\Theta(t)$ đạt đến ngưỡng Θ_{cr} , được chú thích bằng đoạn thẳng màu cam.

Hình 6.10 mô tả sự phụ thuộc thời gian của tỷ lệ monomer tụ hợp trong mô hình lưới với $N = 10, 16$ và 28 trong các mô phỏng không tái hợp. Đối với tất cả các kích thước và nhiệt độ khác nhau, quá trình phân rã vẫn tuân theo quy luật logistic như trong trường hợp monomer được phép tái hợp.

Động học phân rã sợi không tái hợp $\Theta(t)$ giảm chậm ban đầu khi độ ổn định của sợi còn cao, nhưng sau khi đạt giá trị tới hạn Θ_{cr} , quá trình phân rã diễn ra nhanh và có thể được mô tả bằng hàm mũ đơn. Giá trị tới hạn Θ_{cr} phụ thuộc vào số lượng monomer ban đầu N và nhiệt độ T , nhưng theo quy tắc chung là N càng nhỏ và T càng cao thì Θ_{cr} càng lớn. Do đó khi số lượng monomer tăng hoặc nhiệt độ giảm, việc khớp dữ liệu bằng hàm mũ đơn càng kém chính xác. Đối với các nhiệt độ đã thiết lập mô phỏng, $\Theta_{cr} \approx 87\%, 70\%$, và 67% lần lượt cho $N = 10, 16$, và 28 . Khi giới hạn vùng khớp dữ liệu với

$\Theta \leq \Theta_{cr}$, quy luật mũ đơn phù hợp tốt với các mô phỏng. Kết quả này khớp với thực nghiệm của Gruning.

2. Liên kết sợi giảm nhanh hơn so với sự suy giảm số lượng monomer tự hợp

Từ mô phỏng lưới, có thể thấy rằng khi nhiệt độ tăng, $\Omega(t)$ giảm nhanh hơn $\Theta(t)$ (Hình 6.1 và 6.3). Với $N = 10$ và $T = 0.55$, và thời gian mô phỏng $t = 5 \times 10^5$ bước MC, $\Theta(t)$ khoảng 67%, cao hơn so với $\Omega(t)$ khoảng 29%. Với $N = 16$ và $T = 0.60$, sau $t = 1.5 \times 10^6$ bước MC, $\Theta(t)$ khoảng 49%, $\Omega(t)$ giảm xuống còn 8.5%. Điều này cũng đúng cho các mô phỏng với mô hình đầy đủ nguyên tử của $10A\beta_{37-42}$ (Hình 6.2): tại $t = 25$ ns, $\Theta(t)$ khoảng 96.7%, 87.5%, và 52.2%, trong khi $\Omega(t)$ khoảng 42.5%, 32.1%, và 15.7% lần lượt cho $T = 350$ K, 375 K, và 400 K. Tương tự với mô phỏng của 10 peptide $A\beta_{37-42}$, mô phỏng với 2BEG cho thấy ở thời điểm $t = 200$ ns, tỷ lệ liên kết sợi $\Omega(t)$ giảm dưới 50% nhưng $\Theta(t)$ vẫn ở mức 100% do không có monomer nào bị tách ra khỏi sợi. Điều này chứng tỏ rằng $\Omega(t)$ giảm nhanh hơn nhiều so với $\Theta(t)$ và không thể mô tả bởi hàm logistic.

KẾT LUẬN

Các kết quả chính của luận án

1. Ảnh hưởng của các rối lasso và ngoại lực lên lộ trình và thời gian cuộn của protein đã được khảo sát chi tiết bằng phương pháp mô phỏng MD kết hợp mô hình thô cho miền bám vào thụ thể (RBD) của SARS-CoV-1 và SARS-CoV-2. Bề mặt xác suất mô tả quá trình cuộn của các RBD cho thấy rằng các trạng thái trung gian với sự thay đổi trong mức độ rối (so với trạng thái tự nhiên) xuất hiện phổ biến trong các lộ trình cuộn gấp của cả hai protein và có xu hướng dẫn đến sự hình thành các trạng thái cuộn sai. Các trạng thái cuộn sai này chứa các rối lasso không tồn tại trong trạng thái tự nhiên của cả SARS-CoV-1-RBD và SARS-CoV-2-RBD. Nhìn chung, tác động của ngoại lực đối với quá trình hình thành rối lasso và lộ trình cuộn toàn bộ của các

protein là khác nhau. Do nhạy cảm với lực, thời gian cuộn của SARS-CoV-2-RBD tăng đáng kể. Hiện tượng này phù hợp với lý thuyết của Bell, tức là lực kéo tác động vào hai đầu protein sẽ làm chậm quá trình cuộn gấp. Tuy nhiên đối với SARS-CoV-2-RBD, ngoại lực giúp giảm thiểu việc protein cuộn sai thành các trạng thái bị mắc kẹt về mặt động học, do đó thúc đẩy quá trình cuộn hiệu quả và nhìn chung không làm chậm tốc độ cuộn của SARS-CoV-2-RBD so với trường hợp không có lực. Quan sát này đồng thuận với cơ chế hoạt động của các chaperone, giúp cho protein cuộn hiệu quả bằng cách làm chậm quá trình này để protein tránh bị cuộn sai và kết tụ [248, 257, 258]. Kết quả của luận án lần đầu tiên chỉ ra rằng rối lasso có thể ảnh hưởng đáng kể đến tác động của ngoại lực lên quá trình cuộn của protein.

2. Chúng tôi cũng đã quan sát được sự xuất hiện của các cấu trúc cuộn sai thông qua mô phỏng động học dịch mã protein ở các tốc độ khác nhau. Điều này củng cố quan điểm cho rằng sự tồn tại các trạng thái cuộn gấp sai là một hiện tượng phổ biến trong tế bào. Đáng chú ý, một số trạng thái cuộn sai có cấu trúc ổn định và rất giống với cấu trúc tự nhiên nên chúng có thể tránh khỏi sự kiểm soát của tế bào và ảnh hưởng sâu rộng và lâu dài đến các chức năng của protein.

Cách đây hơn 10 năm thực nghiệm đã cho thấy rằng quá trình dimer hóa của protein trong môi trường tế bào phụ thuộc vào tổ độ dịch mã của protein trong ribosome. Kết quả này rất bất ngờ vì theo quan niệm truyền thống, quá trình dimer hóa chỉ phụ thuộc vào trình tự aa của protein chứ không bị ảnh hưởng bởi động học tạo ra chúng. Trong luận án này, lần đầu tiên bằng mô phỏng chúng tôi đã thu được kết quả phù hợp với thực nghiệm. Như vậy, cả kết quả thực nghiệm và tính toán mô phỏng đã góp phần củng cố một giả thuyết (nguyên lý) hoàn toàn mới. Theo giả thuyết này, các quá trình xảy ra sau dịch mã trong môi trường tế bào chẳng hạn như trong trường hợp dimer hóa có thể bị ảnh hưởng bởi động học hình thành protein chứ không nhất

thiết chỉ phụ thuộc vào trình tự aa của chúng. Vấn đề này đòi hỏi nhiều nghiên cứu kỹ lưỡng hơn nữa trước khi có thể đưa ra kết luận cuối cùng.

3. Chúng tôi đã xây dựng một lý thuyết giải tích và hiện tượng luận để mô tả cơ chế logistic của sự phân rã monomer khỏi các cụm sợi trong cả hai trường hợp, có và không có khả năng tái hợp các monomer được giải phóng với cụm mẹ. Kết quả cũng cho thấy rằng, trong quá trình phân rã nhiệt, số lượng monomer liên kết trong cụm sợi, có thể được đo bằng phổ huỳnh quang tryptophan, giảm chậm hơn so với số liên kết sợi được đo bằng phổ huỳnh quang ThT. Ở một ngưỡng nồng độ sợi nhất định, quy luật phân rã logistic trở thành hàm mũ đơn, mô tả sự suy giảm của sợi theo thời gian phù hợp với thí nghiệm. Các mô phỏng lưới và toàn phần tử đã được thực hiện để xác nhận lý thuyết giải tích. Phân tích từ mô phỏng cũng cho thấy sự thay đổi theo thời gian của số lượng liên kết sợi được mô tả bằng hàm mũ đôi, trùng khớp với dữ liệu thực nghiệm quan sát quá trình phân rã sợi bằng tín hiệu huỳnh quang ThT. Các đặc điểm động học của quá trình phân rã nhiệt đã được giải thích từ quan điểm năng lượng tự do dưới dạng một bề mặt xác suất của các tọa độ phản ứng mô tả phân bố của số lượng monomer và số lượng liên kết trong cụm sợi. Cụ thể, động học phân rã của số lượng liên kết sợi được mô tả bởi hàm mũ đôi theo thời gian có liên quan đến hai cực tiểu rõ rệt trên bề mặt năng lượng xác suất, trong khi cơ chế logistic của số lượng monomer tương ứng với đặc điểm bề mặt xác suất chỉ chứa một cực tiểu toàn cục cùng với một số bẫy nông.

Tóm lại, sự phân rã nhiệt của các cấu trúc sợi đã được khảo sát chi tiết ở cấp độ phân tử, cung cấp thêm hiểu biết về sự hình thành và độ bền của các dạng tụ hợp liên quan đến các bệnh thoái hóa thần kinh như bệnh Alzheimer.

Các hướng nghiên cứu trong thời gian tới

1. Trong phần nghiên cứu ảnh hưởng của ngoại lực lên quá trình cuộn gấp protein, chúng tôi mới chỉ xem xét trường hợp ngoại lực $F = 5$ pN. Để làm

rõ hơn tác động của ngoại lực lên cấu trúc ròi lasso, chúng tôi dự định sẽ nghiên cứu thêm các trường hợp với nhiều giá trị F khác nhau. Chúng tôi nhận định rằng, khi ngoại lực tăng thì mức độ ròi giảm nhưng giả thiết này cần được chứng minh.

2. Mô hình thô mà chúng tôi đã sử dụng trong luận án có thể áp dụng rộng rãi để nghiên cứu các bài toán khác nhau, chẳng hạn như khảo sát tương tác protein-protein, protein-ligand. Tuy nhiên vì cấu trúc bị thu gọn, các kết luận dựa trên mô phỏng với mô hình thô có thể sai lệch so với thực tế. Do đó, các đối tượng nghiên cứu sẽ có ý nghĩa khoa học hơn nếu có thể kết hợp với các quan sát thực nghiệm để tiến hành kiểm chứng. Gần đây chúng tôi đã sử dụng mô hình thô kết hợp với phương pháp mô phỏng MD trao đổi bản sao (replica exchange) kết hợp kỹ thuật lấy mẫu ô (umbrella sampling) để tính toán năng lượng liên kết cho một số hệ tương tác liên quan đến nghiên cứu về COVID và kết quả thu được chỉ thấy sự phù hợp tốt với các dữ liệu thực nghiệm. Ngoài ra, khả năng mô phỏng quá trình tổng hợp protein trong ribosome mở ra nhiều hướng nghiên cứu tiềm năng khác như là: quá trình cuộn đồng dịch mã của protein ngay khi chúng đang được tổng hợp, tương tác giữa protein với ribosome và các phân tử chaperone, cơ chế ức chế quá trình dịch mã của vi khuẩn....

3. Trong luận án này, mô hình lưới đã được áp dụng để nghiên cứu quá trình phân rã sợi của các chuỗi peptide, từ đó đề xuất một bức tranh về động học của quá trình phân rã nhiệt của sợi chủ yếu dựa trên nguyên lý chung về sự cạnh tranh giữa các yếu tố năng lượng và entropy. Do đó, phân tích chi tiết hơn về quá trình này có thể được xem xét trong tương lai. Mô hình lưới cũng cho thấy tiềm năng ứng dụng trong việc tìm hiểu các thuộc tính của hình thái sợi cũng như có thể được phát triển cho lớp kép lipid và nghiên cứu tác động của nó lên quá trình kết tụ protein. Hơn nữa, mô hình này còn có thể được sử

dụng để nghiên cứu các ứng dụng trong lĩnh vực công nghệ vật liệu nano và vật liệu mới.

DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

1. Nguyen Truong Co[†], Pham Dang Lan[†], Pham Dinh Quoc Huy, and Mai Suan Li (2020). *Heat-induced degradation of fibrils: Exponential vs logistic kinetics*. Journal of Chemical Physics, **152**, 115101.
2. Pham Dang Lan, Daniel Allen Nissley, Ian Sitarik, Vu Van Quyen, Yang Jiang, Philip To, Yingzi Xia, Stephen David Fried, Mai Suan Li, Edward Patrick O'Brien (2024). *Synonymous mutations can alter protein dimerization through localized interface misfolding involving self-entanglements*. Journal of Molecular Biology, **436**(6), 168487-168490.
3. Pham Dang Lan, Edward Patrick O'Brien, and Li, M. S. (2024). *Pulling Forces Differentially Affect Refolding Pathways Due to Entangled Misfolded States in SARS-CoV-1 and SARS-CoV-2 Receptor Binding Domain*. Biomolecules, **14**, 1327-1340.

([†]) đồng tác giả chính