

MỤC LỤC

ĐẶT VẤN ĐỀ.....	1
NỘI DUNG.....	3
PHẦN A: CƠ SỞ LÝ THUYẾT.....	3
I. PHIÊN MÃ.....	3
1. Các thành phần cơ bản tham gia phiên mã.....	3
1.1. Khuôn mẫu của quá trình phiên mã.....	3
1.2. Các ribonucleotide triphosphate (rNTP).....	4
1.3. Enzyme RNA polymerase.....	4
1.4. Các trình tự điều hòa phiên mã.....	5
1.5. Các yếu tố điều hòa phiên mã.....	6
2. Phiên mã ở prokaryote.....	6
2.1. Promoter.....	6
2.2. Các giai đoạn của phiên mã.....	7
3. Phiên mã ở eukaryote.....	15
3.1. Các RNA polymerase ở eukaryote.....	16
3.2. mRNA của eukaryote.....	22
3.3. Biến đổi RNA sau phiên mã.....	23
II. DỊCH MÃ.....	43
1. Bản chất của mã di truyền.....	43
1.1. Garrod đề xuất rằng một số gene mã hóa cho sự tạo thành các enzyme.....	43
1.2. Thí nghiệm của Beadle và Tatum với <i>Neurospora</i> dẫn đến đề xuất giả thuyết một gene - một enzyme.....	44
1.3. Mã di truyền.....	45
2. Các thành phần tham gia dịch mã.....	48
2.1. mRNA mang trình tự mã hóa protein.....	49
2.2. Cấu trúc và chức năng của tRNA.....	49
3. Các giai đoạn của dịch mã.....	53
3.1. Hoạt hóa amino acid.....	53
3.2. Tổng hợp chuỗi polypeptide.....	56
4. Polyribosome.....	65
5. Tần số lỗi trong dịch mã.....	66
6. Sự hoàn thiện và vận chuyển protein sau dịch mã.....	67

7. Các kháng sinh và độc tố thường nhắm vào chu trình tổng hợp protein.....	74
PHẦN B: CÂU HỎI VÀ BÀI TẬP.....	78
I. CÂU HỎI VÀ BÀI TẬP CÓ HƯỚNG DẪN TRẢ LỜI.....	78
II. CÂU HỎI VÀ BÀI TẬP TỰ TRẢ LỜI.....	101
PHẦN C: KẾT LUẬN.....	109
TÀI LIỆU THAM KHẢO.....	110

CHUYÊN ĐỀ:
DÒNG THÔNG TIN DI TRUYỀN TỪ DNA ĐẾN PROTEIN:
PHIÊN MÃ VÀ DỊCH MÃ

ĐẶT VẤN ĐỀ

1. Lí do tiếp cận chuyên đề

Trong những thập kỷ gần đây, chúng ta đã trải qua một cuộc cách mạng kiến thức về những vấn đề liên quan đến quá trình lưu trữ, truyền đạt và biểu hiện của thông tin di truyền ở mức độ phân tử. Các kiến thức của sinh học phân tử cho phép chúng ta giải thích được mối quan hệ giữa cấu trúc và chức năng của các phân tử sinh học cũng như sự vận hành và kiểm soát các quá trình hóa sinh trong tế bào. Sinh học phân tử cũng ứng dụng cho nhiều lĩnh vực hơn là chỉ chăm sóc sức khỏe và đã tác động lớn đến các lĩnh vực khác như nông nghiệp, hành vi động vật, sự tiến hóa và vi sinh. Tâm điểm của sinh học phân tử là việc nghiên cứu các đại phân tử và các phức hệ đại phân tử của DNA, RNA và protein cùng các quá trình tái bản, phiên mã và dịch mã.

Các cơ chế di truyền ở cấp độ phân tử - chủ yếu là nhân đôi DNA, phiên mã và dịch mã - là những nội dung kiến thức nền tảng quan trọng được đưa vào chương trình giảng dạy cho học sinh THPT, đặc biệt là học sinh THPT chuyên. Trong phần cơ chế di truyền và biến dị, các nội dung về cơ chế di truyền phân tử tương đối khó, gây nhiều lúng túng cho học sinh ở các lớp chuyên trong quá trình nghiên cứu, học tập cũng như trong quá trình ôn luyện cho việc tham gia các kì thi. Đồng thời, các kiến thức về cơ chế di truyền phân tử cũng thường xuyên xuất hiện, làm cơ sở để trả lời các câu hỏi về cơ chế di truyền và biến dị trong đề thi chọn học sinh giỏi khu vực, cấp tỉnh/thành phố và đề thi Quốc gia, đa số đòi hỏi học sinh phải hiểu biết một cách có hệ thống và sâu sắc.

Hiện nay, đã có rất nhiều tài liệu viết về các cơ chế di truyền ở cấp độ phân tử nói riêng và cơ chế di truyền biến dị nói chung ở nhiều mức độ khác nhau, chủ yếu là các giáo trình giảng dạy tại các trường đại học, cao đẳng và một số tài liệu dành cho học sinh phổ thông nói chung và học sinh chuyên nói riêng. Tuy nhiên, qua thực tế giảng dạy các lớp chuyên sinh và tham gia bồi dưỡng học sinh giỏi tôi nhận thấy chính sự đa dạng trong tài liệu tham khảo và sự thiếu thống nhất ở các tài liệu này gây nhiều khó khăn cho việc lựa chọn tài liệu phục vụ cho việc học tập, ôn luyện của học sinh cũng như nghiên cứu, giảng dạy của giáo viên sao cho phù hợp.

Xuất phát từ thực tế đó, bằng cách tiếp cận riêng của bản thân, tôi biên soạn chuyên đề **“Dòng thông tin di truyền từ DNA đến protein: Phiên mã và dịch mã”** nhằm mục đích giúp các em học sinh tiếp cận các kiến thức có chiều sâu và có hệ thống hơn về phần này, qua đó các em có nền tảng tốt để theo học đội tuyển học sinh giỏi.

2. Mục đích của chuyên đề

- Hệ thống hóa một số kiến thức chuyên sâu về các giai đoạn của quá trình biểu hiện gene, truyền đạt thông tin di truyền từ DNA đến protein và các vấn đề liên quan.
- Định hướng tư duy, rèn luyện kĩ năng giải quyết vấn đề cho học sinh bằng các câu hỏi, bài tập thuộc nhiều cấp độ nhận thức về các cơ chế di truyền ở cấp độ phân tử và các vấn đề liên quan.

- Rèn luyện khả năng tự học, tự nghiên cứu tài liệu và khai thác kiến thức.
- Phát triển chuyên môn của bản thân, đồng thời cung cấp nguồn tư liệu tham khảo cho giáo viên và học sinh THPT chuyên trong giảng dạy và bồi dưỡng học sinh giỏi.

3. Đối tượng và phạm vi áp dụng

- Học sinh ở các lớp chuyên sinh và thành viên các đội tuyển học sinh giỏi các cấp.
- Giáo viên giảng dạy ở các trường THPT chuyên, giáo viên tham gia công tác bồi dưỡng học sinh giỏi.

4. Khái quát nội dung chuyên đề

Chuyên đề được biên soạn thành 2 phần:

Phần A: Cơ sở lý thuyết.

Phần này nghiên cứu các nội dung liên quan đến các cơ chế truyền đạt thông tin di truyền từ DNA đến protein là phiên mã và dịch mã.

Phần B: Câu hỏi, bài tập vận dụng.

Phần này gồm các câu hỏi, bài tập sưu tầm chủ yếu từ các đề thi học sinh giỏi Quốc gia, đề chọn học sinh giỏi tham gia đội tuyển dự thi IBO qua các năm và một số tài liệu khác cùng với hướng dẫn trả lời; đồng thời cung cấp thêm một số câu hỏi, bài tập tự trả lời giúp học sinh rèn luyện kỹ năng.

NỘI DUNG

PHẦN A: CƠ SỞ LÝ THUYẾT

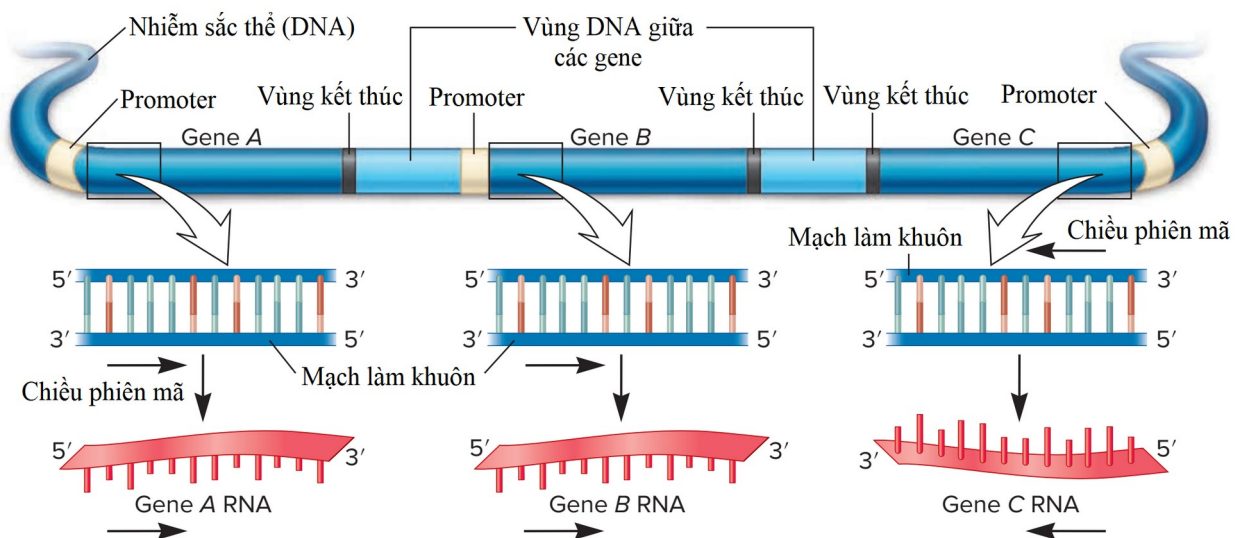
Cấu trúc, chức năng, sự sinh trưởng và sinh sản của một cơ thể phụ thuộc vào các thuộc tính protein có trong mỗi mô, tế bào của cơ thể đó. Một phân tử protein có thể gồm một hoặc một số chuỗi polypeptide; trong đó, trình tự các amino acid trong chuỗi polypeptide được qui định bởi trình tự nucleotide trong gene (DNA). Khi tế bào cần một loại protein nào đó, thông tin di truyền từ gene được “dịch mã” thành trình tự amino acid trong phân tử protein tương ứng. Quá trình truyền thông tin như vậy (còn gọi là sự biểu hiện của gene) được thực hiện qua hai giai đoạn là “phiên mã” và “dịch mã”. Trong đó, phiên mã là quá trình tổng hợp một phân tử RNA (trình tự các ribonucleotide) mạch đơn là bản phiên mã của gene (trình tự các deoxyribonucleotide), còn dịch mã là quá trình tổng hợp protein (trình tự các amino acid) dựa trên mạch khuôn là bản phiên mã (RNA). Không giống sao chép DNA vốn chỉ diễn ra một lần duy nhất trong mỗi chu trình tế bào, phiên mã và dịch mã diễn ra liên tục trong suốt chu trình tế bào (tuy có giảm ở pha M). Năm 1956 (3 năm sau ngày công bố mô hình xoắn kép DNA cùng Watson), Crick đưa ra khái niệm về “nguyên lý trung tâm” phản ánh hai giai đoạn của một quá trình biểu hiện gene là $DNA \rightarrow RNA \rightarrow protein$.

I. PHIÊN MÃ

1. Các thành phần cơ bản tham gia phiên mã

1.1. Khuôn mẫu của quá trình phiên mã

Hai sợi có trình tự base bổ sung nhau của DNA có vai trò khác nhau trong quá trình phiên mã. Sợi đóng vai trò như khuôn mẫu cho quá trình tổng hợp RNA được gọi là **mạch khuôn**. Sợi DNA bổ sung với sợi làm khuôn, hay **sợi không làm khuôn**, hoặc **sợi mã hóa**, giống nhau về trình tự base với RNA được phiên mã từ gene, với U trong RNA thay cho T trong DNA. Sợi mã hóa cho một gene cụ thể có thể nằm trong một trong hai sợi của một nhiễm sắc thể nhất định. Theo quy ước, các trình tự quy định điều hòa quá trình phiên mã được quy ước bởi các trình tự trong sợi mã hóa.



Hình 1.1. Sự phiên mã của ba gene khác nhau trên cùng một nhiễm sắc thể. RNA polymerase tổng hợp từng bản phiên mã RNA theo hướng 5' đến 3', trượt dọc theo mạch

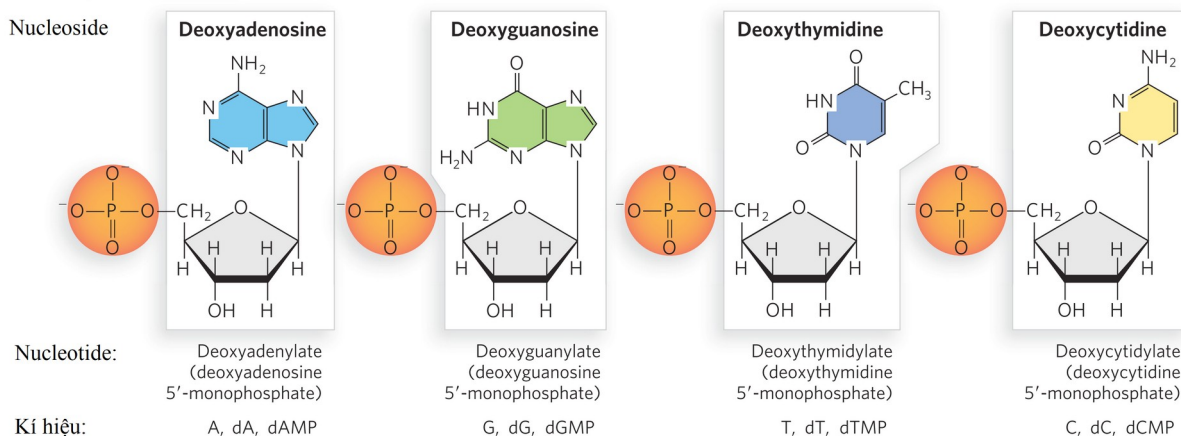
khuôn DNA theo hướng 3' đến 5'. Tuy nhiên, mạch nào được sử dụng làm sợi khuôn có sự khác nhau giữa các gene. Ví dụ, gene A và B sử dụng mạch phía dưới, nhưng gene C sử dụng mạch phía trên.

Trong trường hợp phiên mã của nhiều gene trong một nhiễm sắc thể, hướng phiên mã và chuỗi DNA được sử dụng làm khuôn sẽ khác nhau giữa các gene khác nhau. Hình 1.1 cho thấy ba gene liên kế nhau trong một nhiễm sắc thể. Các gene A và B được phiên mã từ trái sang phải, sử dụng mạch DNA phía dưới làm khuôn mẫu. Trong khi đó, gene C được phiên mã từ phải sang trái và sử dụng sợi DNA trên cùng làm khuôn mẫu. Lưu ý rằng trong cả ba trường hợp, sợi khuôn được đọc theo hướng 3' đến 5', và quá trình tổng hợp RNA xảy ra theo hướng 5' đến 3'. Trong mỗi gene nhất định, chỉ có một mạch khuôn được sử dụng để tổng hợp RNA.

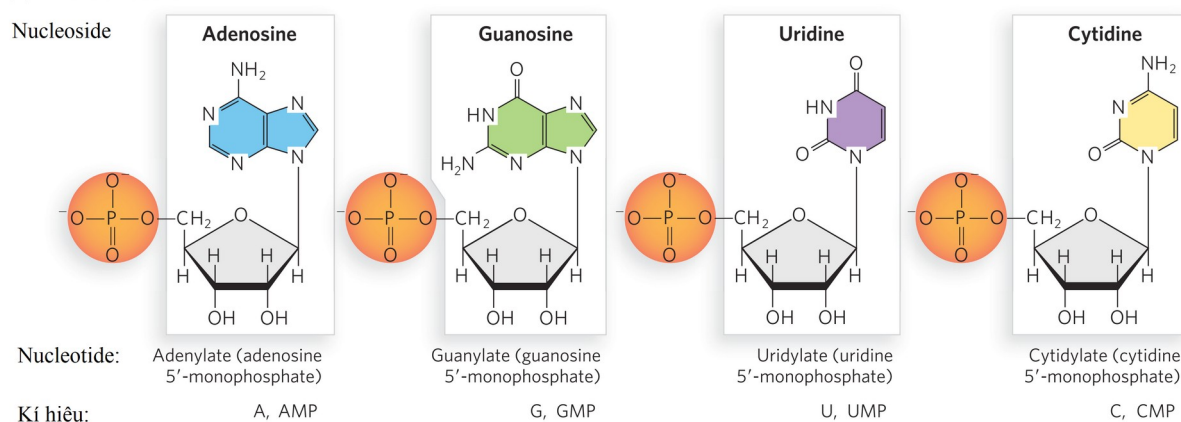
1.2. Các ribonucleotide triphosphate (rNTP)

Giống như các dNTP trong sao chép DNA, các rNTP (hay NTP) vừa là thành phần cấu trúc nên RNA, vừa là nguồn cung cấp năng lượng cho quá trình phiên mã. Cần phân biệt các nucleotide cấu tạo nên RNA và các nucleotide cấu tạo nên DNA. Chúng là 8 loại nucleotide khác nhau. Trong đó, cấu tạo nên RNA là các ribonucleotide có vị trí C-2' ở phần đường ribose là nhóm -OH; còn DNA được cấu tạo từ các deoxyribonucleotide có vị trí C-2' ở phần đường là nhóm -H.

(a) Deoxyribonucleotides



(b) Ribonucleotides



Hình 1.2. Các deoxyribonucleotide của DNA và các ribonucleotide của RNA.

1.3. Enzyme RNA polymerase

Enzyme RNA polymerase là enzyme trực tiếp xúc tác phản ứng trùng hợp RNA. Enzyme này có khả năng tự tách hai mạch đơn của phân tử DNA sợi kép, trượt dọc trên

một mạch và xúc tác phản ứng trùng hợp RNA. Cũng giống DNA polymerase, hoạt động xúc tác phản ứng của RNA polymerase được cung cấp năng lượng từ chính sự đứt gãy liên kết cao năng giữa hai nhóm phosphate của các dNTP; vì vậy, sự tổng hợp chỉ diễn ra theo chiều $5' \rightarrow 3'$ và RNA polymerase dịch chuyển trên mạch DNA trong phiên mã theo một chiều nhất định. Tuy nhiên, khác với DNA polymerase, RNA polymerase có khả năng tự khởi đầu phản ứng trùng hợp mà không cần đoạn mồi. Ở prokaryote, toàn bộ hệ gene được phiên mã bởi một loại RNA polymerase duy nhất; trong khi đó, ở eukaryote có ba loại (kí hiệu là RNA polymerase I, II và III).

Xét về cơ chế hóa học và hoạt động của enzyme, các bước của phiên mã rất giống với sao chép DNA. Cả hai quá trình đều là phản ứng trùng hợp các nucleotide xúc tác bởi enzyme; trong đó, một mạch nucleic acid mới được tổng hợp dựa trên một mạch DNA sẵn có. Tuy nhiên, giữa hai quá trình có một số khác biệt, bao gồm:

- RNA polymerase không cần đoạn mồi như DNA polymerase; nó có khả năng tự khởi đầu sự tổng hợp chuỗi polyribonucleotide (RNA). Tuy vậy, sự khởi đầu này chỉ có thể diễn ra từ những vị trí nhất định trên phân tử DNA, gọi là promoter.

- Để sao chép DNA, phức hệ sao chép gồm DNA polymerase liên kết đồng thời trên hai mạch của DNA sợi kép; việc tách hai mạch được thực hiện bởi một số protein đặc biệt. Trong khi đó, để phiên mã, RNA polymerase chỉ liên kết trên một mạch của phân tử DNA; nó tự giãn xoắn cấu trúc DNA xoắn kép, đồng thời xúc tác phản ứng phiên mã mà không cần các protein giãn xoắn và tách mạch đơn chuyên biệt.

- Nguyên liệu để RNA polymerase tổng hợp RNA là các ribonucleotide triphosphate (ATP, GTP, CTP và UTP), gọi chung là NTP; trong khi nguyên liệu để DNA polymerase tổng hợp DNA là các deoxyribonucleotide (dNTP). Ngoài ra, trong phiên mã, nucleotide được RNA polymerase dùng bổ sung vào chuỗi RNA để kết cặp với A là U, chứ không phải T như trường hợp của DNA polymerase. Như vậy, nếu mạch khuôn DNA là $3' - \text{ATACTGGAC} - 5'$ thì mạch RNA sẽ là $5' - \text{UAUGACCUG} - 3'$.

- Sản phẩm RNA được tạo ra từ phiên mã không duy trì liên kết hydrogen với mạch DNA làm khuôn. Thay vào đó, RNA polymerase “đẩy” mạch RNA mới tổng hợp ngay sau khi một vài ribonucleotide đã được bổ sung vào chuỗi RNA. Sự tách ra của RNA là cần thiết để sau đó nó có thể được dùng ngay làm khuôn tổng hợp protein (ở prokaryote) hoặc để biến đổi thành RNA trưởng thành sẵn sàng cho dịch mã (ở eukaryote). Cũng nhờ sự tách RNA xảy ra gần như ngay lập tức, một gene cùng lúc có thể được phiên mã bởi nhiều RNA polymerase để tạo nên nhiều bản phiên mã đồng thời. Nhờ vậy, tế bào có thể tổng hợp được một số lượng lớn protein từ một gene duy nhất trong thời gian ngắn.

- Tuy cũng có chức năng đọc sửa giống DNA polymerase, độ chính xác của các RNA polymerase trong việc ghép cặp các nucleotide theo nguyên tắc Chargaff là thấp hơn. Trung bình, tần số sai sót trong phiên mã là $10^{-4} - 10^{-5}$, còn tần số này trong sao chép DNA là $10^{-7} - 10^{-9}$.

1.4. Các trình tự điều hòa phiên mã

Đó là các trình tự nucleotide đặc thù trên DNA đánh dấu vị trí gene được bắt đầu và kết thúc phiên mã, hoặc là các trình tự điều khiển sự biểu hiện của gene. Đoạn trình tự DNA mà ở đó RNA polymerase gắn vào và khởi đầu phiên mã được gọi là **trình tự khởi**

đầu phiên mã (hay promoter). Đoạn trình tự DNA mà ở đó sự phiên mã của gene kết thúc được gọi là **trình tự (hay tín hiệu) kết thúc phiên mã (terminator).** Sự khởi đầu và kết thúc phiên mã ở prokaryote và eukaryote có những đặc điểm khác nhau. Nhìn chung, trong phiên mã, vì RNA polymerase chuyển dịch trên phân tử DNA theo một chiều, nên promoter thường nằm ngược dòng (tức là về phía đầu 5') của đoạn trình tự mã hóa ở mỗi gene; ngược lại, các terminator thường nằm xuôi dòng (về phía đầu 3') của đoạn trình tự mã hóa. Ngoài các trình tự nêu trên, sự phiên mã ở nhiều gene còn phụ thuộc vào các trình tự khác là vị trí liên kết của **các yếu tố hoạt hóa** (ví dụ: enhancer) hoặc của các protein điều hòa sự biểu hiện gene (như operator, attenuator, v.v...).

1.5. Các yếu tố điều hòa phiên mã

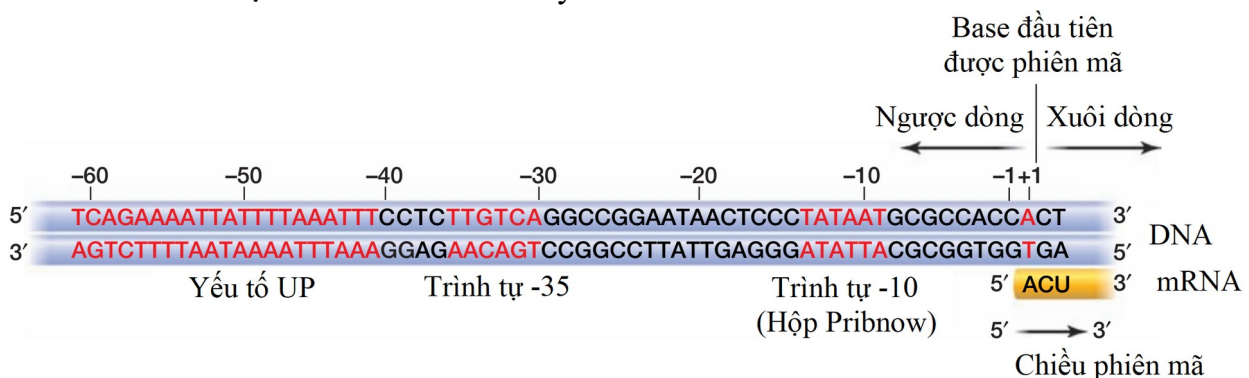
Sự phiên mã của phần lớn các gene ở cả prokaryote và eukaryote đều được điều khiển bởi nhiều protein khác nhau. Các protein này có thể là các protein hoạt hóa (điều hòa dương tính) hoặc ức chế phiên mã (điều hòa âm tính), hoặc là các yếu tố tham gia bộ máy phiên mã. Chúng được mã hóa bởi các gene khác trong hệ gene, nhưng tương tác với các trình tự điều hòa để điều khiển sự biểu hiện gene.

2. Phiên mã ở prokaryote

2.1. Promoter

Promoter là các đoạn nucleotide nằm ở đầu 5' trên mạch mã hóa (mạch không làm khuôn) trước điểm khởi đầu phiên mã (base thứ nhất, thường là adenine, kí hiệu +1), có vai trò là nơi cho RNA polymerase nhận biết, gắn vào thực hiện quá trình phiên mã. Đoạn promoter này có khoảng từ 20 - 200 cặp base.

Promoter ở prokaryote có hai đoạn DNA ổn định dài 6 bp có vai trò quyết định sự liên kết với RNA polymerase vào promoter được gọi là **trình tự liên ứng** (consensus sequence), đây là các trình tự bảo thủ cao trong quá trình tiến hóa, gồm trình tự 5'–TATAAT–3' xung quanh -10 gọi là **hộp TATA**, **hộp -10** hay **hộp Pribnow**¹ là điểm bám của đơn vị sigma của RNA polymerase và xung quanh vị trí -35 có trình tự 5'–TTGACA–3' là vị trí bám khác cho enzyme

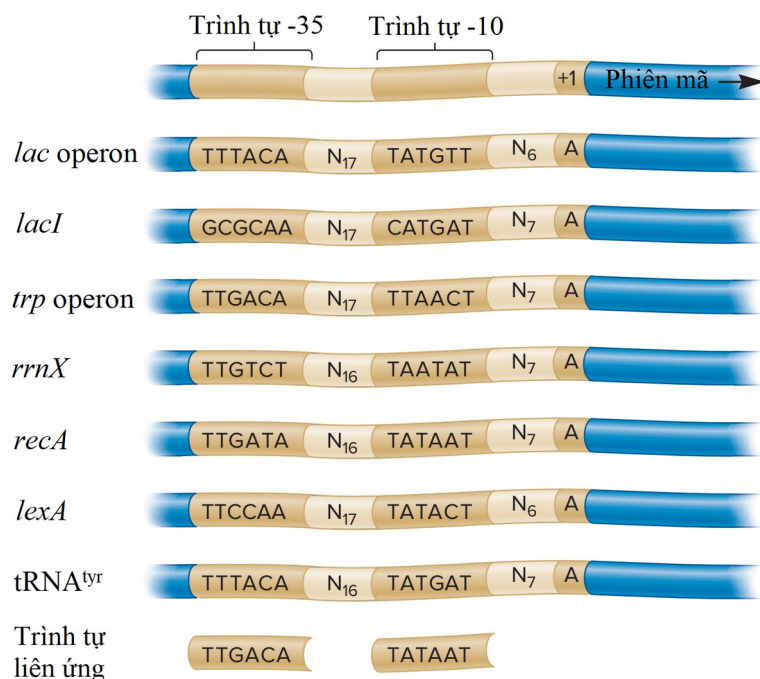


Hình 1.3. Các yếu tố cơ bản của promoter ở *Escherichia coli*. Kí hiệu "-" và "+" chỉ các vị trí nằm trước và sau vị trí bắt đầu phiên mã (+1), hay còn gọi là ngược dòng và xuôi dòng.

Trong thực tế, hiếm khi có 2 gene nào mà toàn bộ trình tự promoter của chúng giống hệt nhau (ngay cả các trình tự liên ứng -35 và -10 và đoạn giữa chúng). Vì vậy ái lực của RNA polymerase với vùng promoter của các gene khác nhau là khác nhau. Đây là một

¹ Được David Pribnow phát hiện lần đầu vào năm 1975

nguyên nhân giải thích tại sao các gene khác nhau thường được biểu hiện ở mức độ khác nhau. Nhìn chung, promoter càng có các trình tự liên ứng giống với các trình tự điển hình thì tốc độ phiên mã càng cao. Các đột biến thay thế base tại các hộp TATA, nghĩa là làm cho nó bớt giống với trình tự được bảo tồn (ví dụ, TATAAAT → TGTAAAT) do đó sẽ làm yếu khả năng phiên mã của promoter, gọi là **down mutation**. Ngược lại, các đột biến làm cho các trình tự promoter trở nên giống với các trình tự điều hòa (ví dụ, TATCTT → TATAAAT), sẽ làm mạnh khả năng phiên mã của promoter, gọi là **up mutation**.



Hình 1.4. Sự đa dạng của trình tự -35 và -10 trong các promoter khác nhau ở vi khuẩn. Hình này thể hiện các trình tự -35 và -10 với một mạch DNA được tìm thấy ở 7 promoter vi khuẩn và bacteriophage. Trình tự liên ứng điển hình được thể hiện bên dưới. Các vùng đệm gồm có số nucleotide giữa vùng -35 và -10 hay giữa vùng -10 và vị trí khởi đầu phiên mã. Chẳng hạn, N₁₇ nghĩa là có 17 nucleotide giữa cuối vùng -15 và bắt đầu vùng -10.

Ngoài các trình tự liên ứng, một số promoter hoạt động mạnh còn có một số trình tự khác giúp tăng cường sự tương tác giữa RNA polymerase và DNA, như promoter của các gene rRNA có **yếu tố UP (UP element, upstream element)** dài khoảng 20 bp xung quanh vị trí -50, giàu A và T, là vị trí có ái lực cao với đầu N của chuỗi α thuộc enzyme RNA polymerase do vậy RNA polymerase thường xuyên ở trạng thái liên kết với các promoter này.

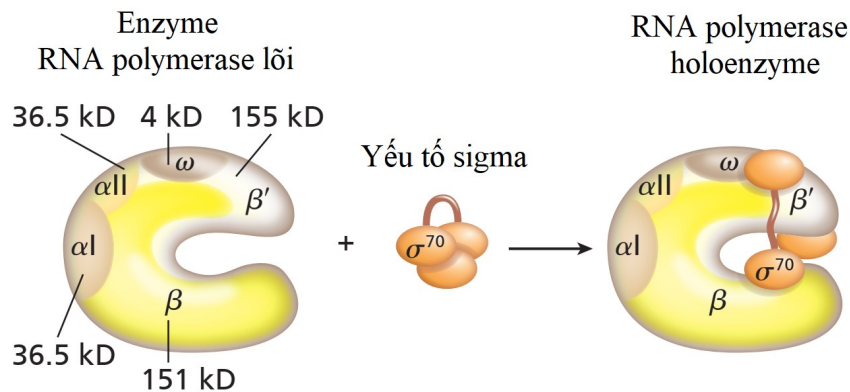
Một số gene, promoter thiếu một trong các trình tự liên ứng, nhưng lại được bổ sung bởi một trình tự khác để “bù đắp” cho trình tự thiếu. Chẳng hạn như **gene gal** ở *E. coli* thiếu trình tự -35 nhưng có đoạn -10 mở rộng có ái lực cao với yếu tố σ^{70} .

2.2. Các giai đoạn của phiên mã

2.2.1. Khởi đầu

Ở *E. coli*, **enzyme lõi** được cấu tạo bởi năm tiểu đơn vị, $\alpha_2\beta\beta'\omega$. Sự liên kết của tiểu đơn vị thứ sáu, **yếu tố sigma (σ)**, với enzyme lõi tạo thành RNA polymerase

holoenzyme. Các tiểu đơn vị khác nhau trong holoenzyme đóng các vai trò chức năng riêng biệt. Hai tiểu đơn vị α rất quan trọng trong việc lắp ráp chính xác holoenzyme và trong quá trình liên kết với DNA. Các tiểu đơn vị β và β' cũng cần thiết để liên kết với DNA, và chúng thực hiện xúc tác tổng hợp RNA. Tiểu đơn vị ω (omega) rất quan trọng đối với sự lắp ráp chính xác của enzyme lõi. Trọng lượng phân tử của RNA polymerase lõi với năm tiểu đơn vị là khoảng 390 kD (kiloDaltons), và với tiểu đơn vị sigma được thêm vào, holoenzyme có trọng lượng phân tử là 430 kD. Các holoenzyme cần để bắt đầu phiên mã; vai trò chính của yếu tố σ là nhận biết promoter. Các protein như yếu tố σ ảnh hưởng đến chức năng của RNA polymerase được gọi là các yếu tố phiên mã.



Hình 1.5. RNA polymerase lõi của vi khuẩn liên kết với tiểu đơn vị yếu tố σ hình thành holoenzyme có hoạt tính.

Các tiểu đơn vị σ khác nhau làm cho holoenzyme đặc hiệu với các promoter khác nhau, yếu tố σ^{70} là yếu tố phổ biến nhất ở vi khuẩn. Nó nhận biết các promoter của các “gene giữ nhà” (housekeeping genes), các sản phẩm protein của các gene được tế bào sử dụng liên tục. Do nhu cầu liên tục đối với các sản phẩm của chúng, các “gene giữ nhà” liên tục được biểu hiện. Các tiểu đơn vị σ^{54} và σ^{32} lần lượt nhận ra các promoter của các gene liên quan đến chuyển hóa nitrogen và các gene biểu hiện để phản ứng với stress của môi trường như sốc nhiệt, và chúng được sử dụng khi cần có tác động của các gene này. Yếu tố σ^{28} nhận biết các promoter của các gene cần thiết cho quá trình hướng hóa của vi khuẩn (cảm nhận hóa học và khả năng vận động).

Tính đặc hiệu của từng loại tiểu đơn vị sigma đối với các trình tự liên ứng của promoter khác nhau tạo ra các holoenzyme RNA polymerase có các đặc tính liên kết DNA khác nhau. Các nhà di truyền học vi sinh vật ước tính rằng mỗi tế bào *E. coli* chứa khoảng 3000 RNA polymerase holoenzyme tại bất kỳ thời điểm nào và mỗi loại trong số bốn loại tiểu đơn vị sigma có mặt ở một mức độ khác nhau giữa chúng. Bởi vì các tiểu đơn vị sigma dễ dàng gắn và tách ra khỏi các enzyme lõi để đáp ứng với những thay đổi của điều kiện môi trường, sinh vật có thể thay đổi các mô hình phiên mã của mình để điều chỉnh theo các điều kiện khác nhau.

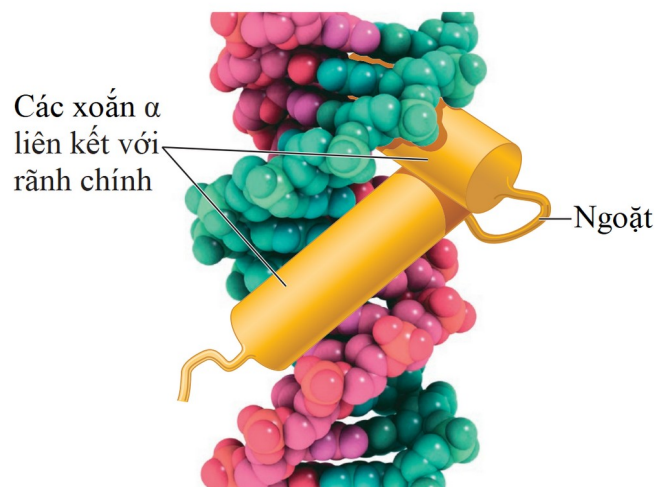
Bảng 1.1: Các tiểu đơn vị sigma của RNA polymerase ở *E. coli*.

Tiểu đơn vị	Khối lượng phân tử (kD)	Trình tự liên ứng		Promoter được nhận biết
		-35	-10	
σ^{28} (RpoF)	28	TAAA	GCCGATAA	Tổng hợp lông roi và vận động hướng hóa.

Tiêu đơn vị	Khối lượng phân tử (kD)	Trình tự liên ứng		Promoter được nhận biết
		-35	-10	
σ^{32} (RpoH)	32	CTTGAA	CCCCATTA	Các gene sốc nhiệt.
σ^{54} (RpoN)	54	CTGGPyAPyPu	TTGCA	Các gene cần cho chuyển hóa nitrogen và một số chức năng khác
σ^{70} (RpoD)	70	TTGACA	TATAAT	Các "gene giữ nhà" biểu hiện ở tất cả các tế bào đang sinh trưởng.

Py = pyrimidine; Pu = purine.

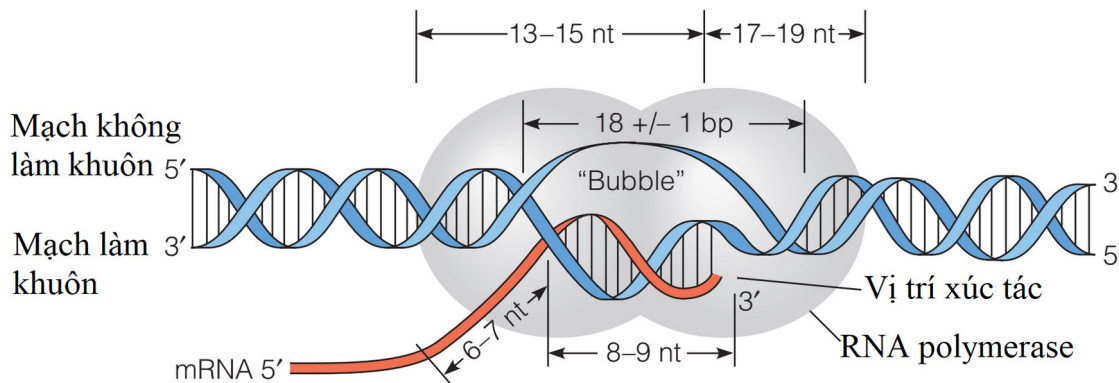
Sau khi RNA polymerase holoenzyme được tập hợp với sáu tiểu đơn vị, nó liên kết lỏng lẻo với DNA và sau đó trượt dọc DNA. Khi holoenzyme gặp một promoter, yếu tố σ nhận ra cả trình tự -35 và -10 . Protein yếu tố σ chứa một cấu trúc được gọi là **motif xoắn – ngoặt – xoắn (helix-turn-helix motif)** có thể liên kết chặt chẽ với các trình tự này. Các vòng xoắn alpha (α) trong protein phù hợp với rãnh chính của chuỗi xoắn kép DNA và hình thành liên kết hydrogen với các base. Liên kết hydrogen xảy ra giữa các nucleotide trong trình tự -35 và -10 của promoter và các chuỗi bên amino acid trong cấu trúc xoắn – ngoặt – xoắn của yếu tố σ .



Hình 1.6. Sự liên kết của protein yếu tố σ vào DNA của promoter. Một phần của protein chứa hai vòng xoắn α được nối với nhau bằng một ngoặt, được gọi là motif xoắn – ngoặt – xoắn. Hai vòng xoắn α của protein có thể nằm trong rãnh chính của DNA. Các chuỗi bên amino acid trong vòng xoắn α hình thành liên kết hydrogen với các base trong DNA.

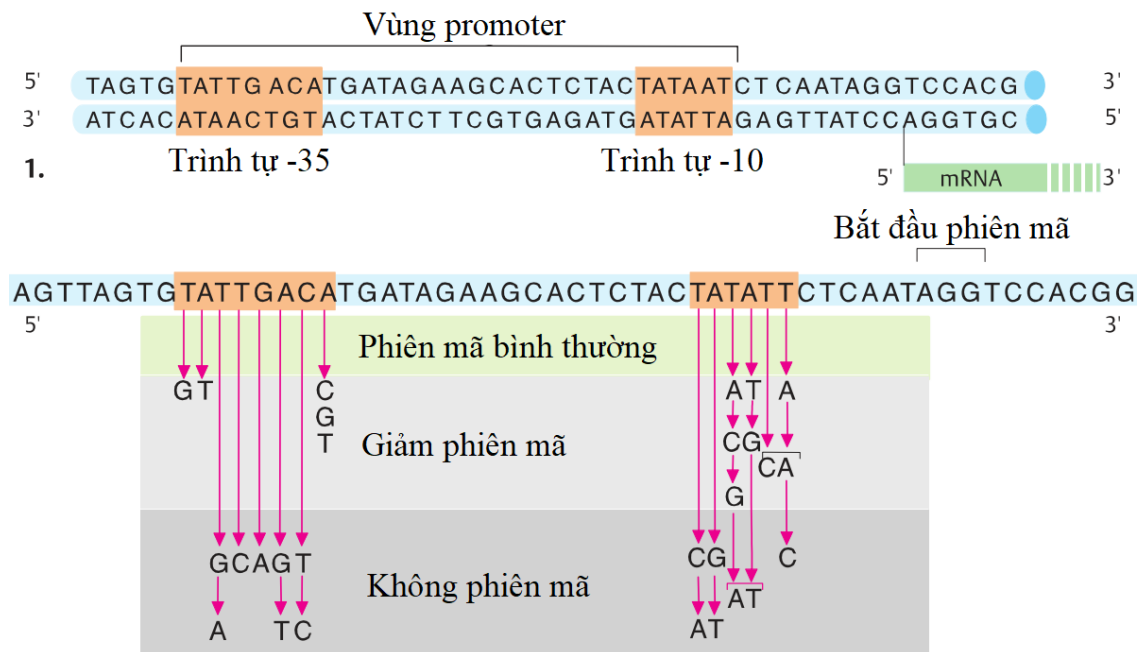
Như trong hình 1.9, quá trình phiên mã được bắt đầu khi σ yếu tố trong holoenzyme liên kết với promoter để tạo thành **phức hệ đóng**. Trong phức hệ đóng này, các sợi DNA trong promoter không bị tách rời. Để quá trình phiên mã bắt đầu, DNA sợi kép phải được tháo xoắn và biến tính một vùng khoảng 12-14 bp quanh vị trí khởi đầu phiên mã để tạo thành một **phức hệ mở** hay **bong bóng phiên mã**. Việc tháo xoắn này bắt đầu từ trình tự TATAAT tại vị trí -10 , chỉ chứa các cặp base A-T. Các cặp base A-T chỉ hình thành hai liên kết hydrogen, trong khi các cặp G-C tạo thành ba liên kết. Do đó, DNA ở vùng giàu

AT dễ dàng tách ra hơn vì phải phá vỡ ít liên kết hydrogen hơn. Một sợi RNA ngắn được tạo ra trong phức hợp mở, và sau đó yếu tố σ được giải phóng khỏi enzyme lõi. Yếu tố σ được giải phóng đánh dấu sự chuyển sang giai đoạn kéo dài của quá trình phiên mã. Lúc này, enzyme lõi có thể trượt dọc theo DNA để tổng hợp một chuỗi RNA.



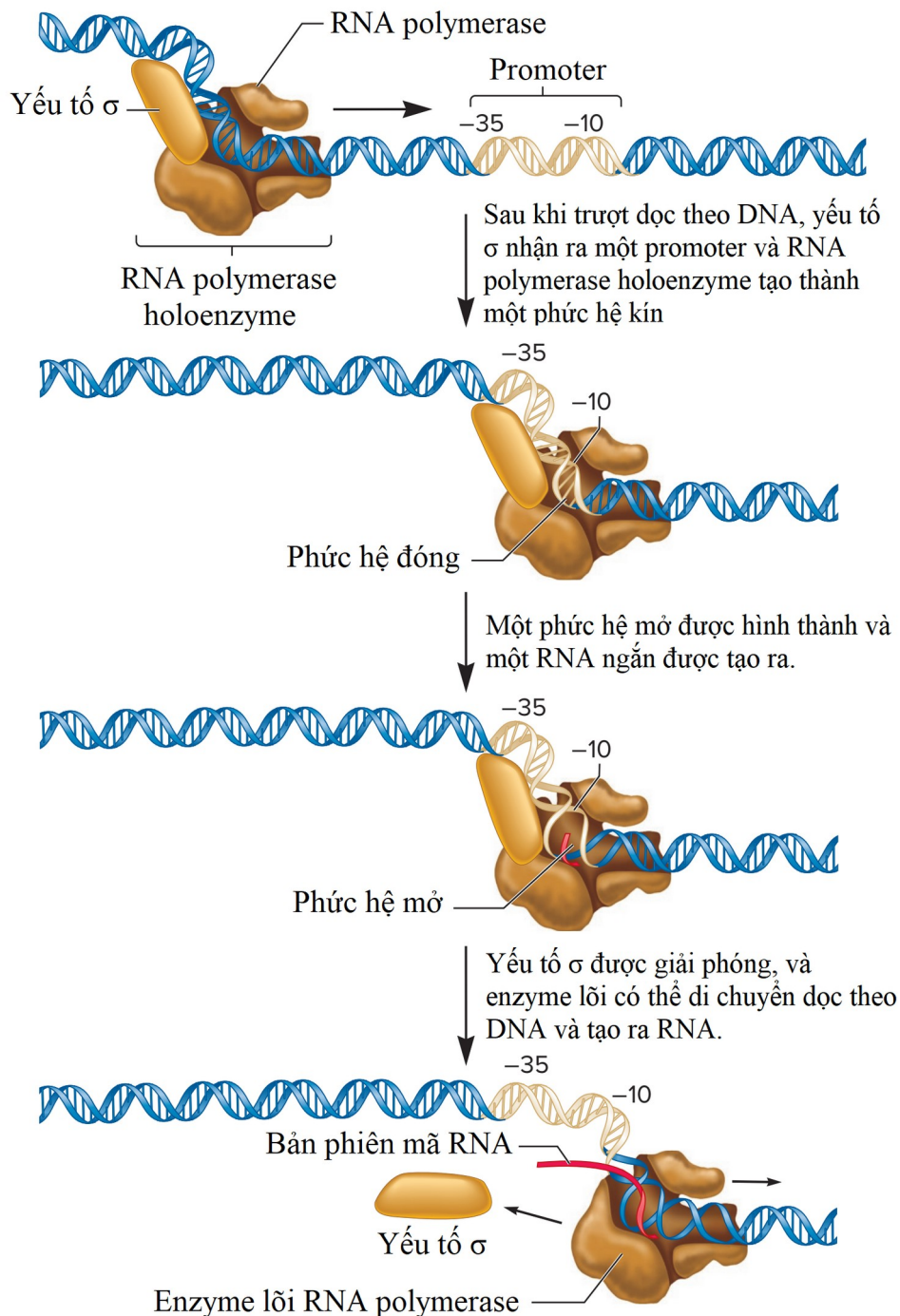
Hình 1.7. Bong bóng phiên mã. Độ dài của DNA tự do (tháo xoắn) và lai DNA-RNA ban đầu được ước tính từ khả năng phản ứng của các phức hợp phiên mã với các thuốc thử chẳng hạn như KMnO_4 – chất oxy hóa các base trong nucleic acid sợi đơn. Độ dài của DNA tiếp xúc với enzyme được xác định bằng cách in vết (footprinting – xác định vị trí liên kết của protein trên DNA). Sáu hoặc bảy nucleotide của RNA phía sau DNA lai được bảo vệ khỏi sự tấn công của ribonuclease bằng cách liên kết với enzyme. Nt = nucleotide.

Các nghiên cứu đột biến ở các promoter ở vi khuẩn *E. coli* cho thấy những biến đổi trong trình tự của trình tự -35 ảnh hưởng đến khả năng bám của RNA polymerase, sự thay đổi của hộp -10 làm ảnh hưởng đến việc chuyển từ phức hợp đóng sang phức hợp mở.



2. Ảnh hưởng của đột biến ở promoter lên tốc độ phiên mã

Hình 1.8. Các trình tự liên ứng ở vùng promoter và ảnh hưởng của đột biến ở promoter lên tốc độ phiên mã.



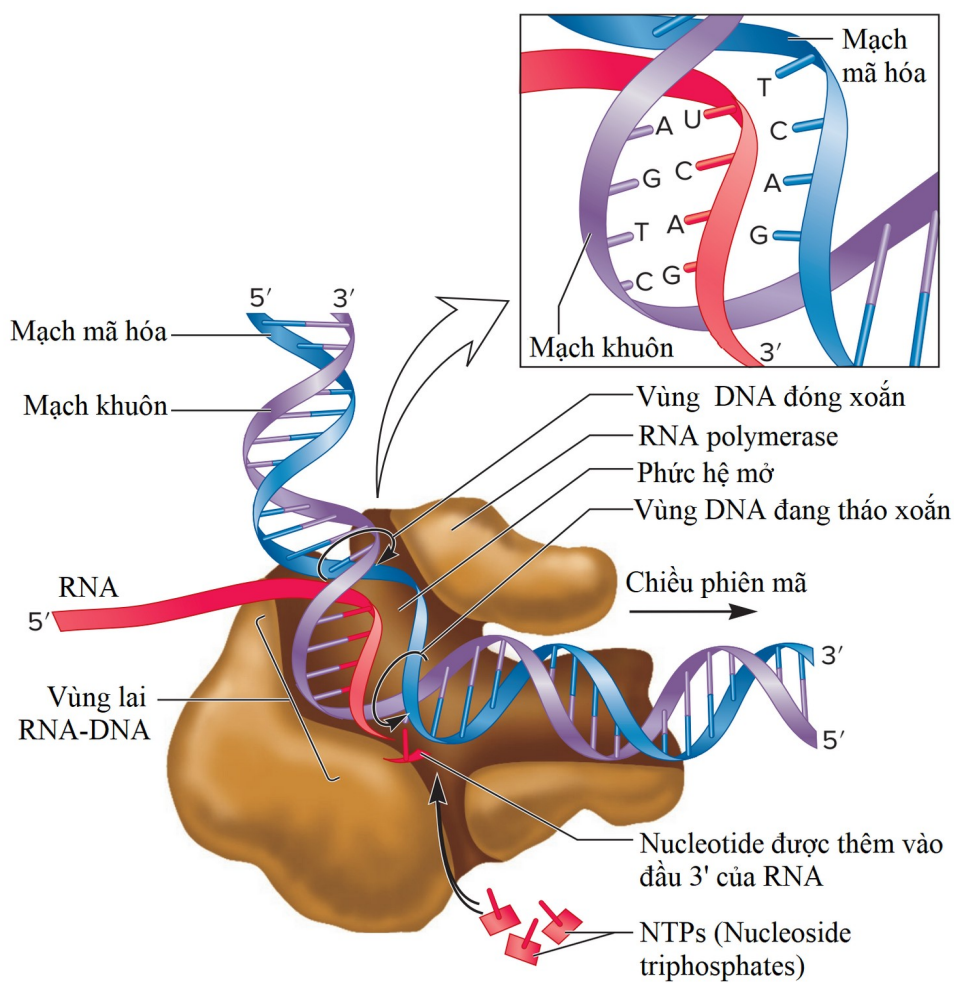
Hình 1.9. Giai đoạn khởi đầu phiên mã ở vi khuẩn. Tiểu đơn vị yếu tố σ của RNA polymerase holoenzyme nhận ra trình tự -35 và -10 của promoter. DNA tháo xoắn ở trình tự -10 để tạo thành một phức hợp mở, và một RNA ngắn được tạo thành. Sau đó, yếu tố σ phân ly khỏi holoenzyme, và enzyme lõi RNA polymerase có thể tiến hành dọc theo DNA, tổng hợp RNA và hình thành một phức hợp mở khi nó di chuyển.

Sự khởi đầu phiên mã ở vi khuẩn cổ (Archaea) có sự khác biệt với vi khuẩn. RNA polymerase của vi khuẩn cổ ngoài ba tiểu đơn vị lớn giống enzyme lõi của vi khuẩn còn có một số tiểu đơn vị khác và khá giống với RNA polymerase của eukaryote. Các promoter được nhận biết bởi protein bám TATA (TBP), và sự dính bám của RNA polymerase là nhờ vai trò trung gian của nhân tố phiên mã tương tự như nhân tố TFIIB của eukaryote, được gọi là TFB. Sự tương đồng của yếu tố phiên mã khác giữa vi khuẩn

cổ và eukaryote vẫn chưa được xác định, và những nghiên cứu về các trình tự hệ gene vi khuẩn cổ cho thấy có thể chúng không có mặt.

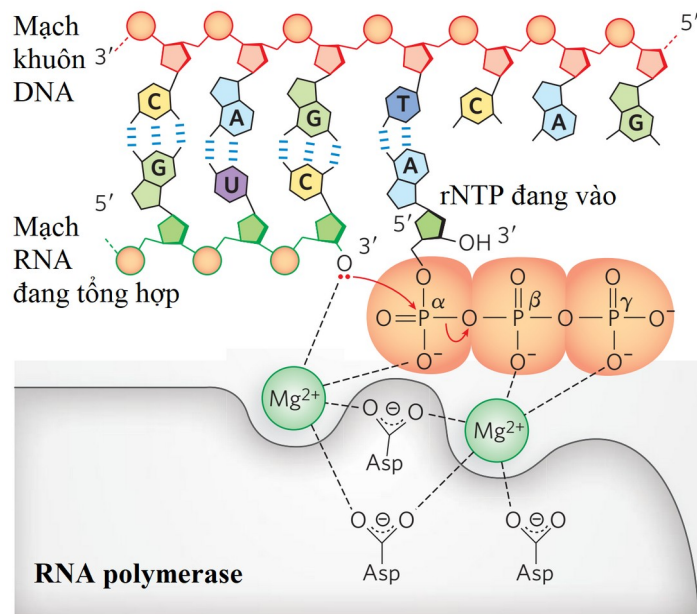
2.2.2. Kéo dài

Sau khi giai đoạn khởi đầu, phiên mã tổng hợp RNA tiến hành giai đoạn kéo dài. Trong giai đoạn kéo dài, RNA polymerase lõi di chuyển dọc theo, tháo xoắn kép DNA phía trước để lộ ra một đoạn DNA khuôn mẫu sợi đơn và duy trì phức hệ mở (bong bóng phiên mã) dài khoảng 17 bp. Phía sau vùng tháo xoắn (phức hệ mở), hai sợi DNA tái cấu trúc thành DNA sợi kép. Trong vùng phức hệ mở, khoảng 8-9 nucleotide tại đầu 3' của mạch RNA đang tổng hợp luôn liên kết với DNA mạch khuôn tạo thành vùng lai RNA-DNA tạm thời; phần còn lại của RNA mới được tổng hợp ra khỏi enzyme dưới dạng sợi đơn. Tốc độ tổng hợp RNA trung bình khoảng 43 nucleotide/giây, chậm hơn nhiều so với sao chép DNA – khoảng 1000 nucleotide/giây.



Hình 1.10. Quá trình tổng hợp bản sao mã RNA.

Về mặt hóa học của quá trình phiên mã bởi RNA polymerase tương tự như sự sao chép của DNA nhờ DNA polymerase. RNA polymerase luôn kéo dài mạch bằng cách xúc tác sự hình thành liên kết giữa nhóm 5'-PO₄²⁻ trên một nucleotide và nhóm 3'-OH trên nucleotide trước đó theo nguyên tắc bổ sung với mạch khuôn - kéo dài mạch theo hướng 5' đến 3'.



Hình 1.11. Cơ chế hóa học tổng hợp RNA. Sự bổ sung một rNTP vào mạch RNA đang tổng hợp là phản ứng phụ thuộc ion Mg^{2+} (Mg^{2+} -dependent reaction) tạo thành một liên kết phosphodiester có chiều $5' \rightarrow 3'$.

Tiểu đơn vị α trong enzyme lõi của RNA polymerase cũng đóng vai trò cần thiết cho việc lắp ráp và nhận biết promoter. RNA polymerase ở vi khuẩn có rãnh sâu ở giữa để có thể bám vào đoạn DNA dài 16 bp (15 bp ở eukaryote, ở nấm men có thể lớn hơn). Rãnh nhỏ hơn trên RNA polymerase có thể làm nhiệm vụ giữ sợi RNA đang được hình thành.

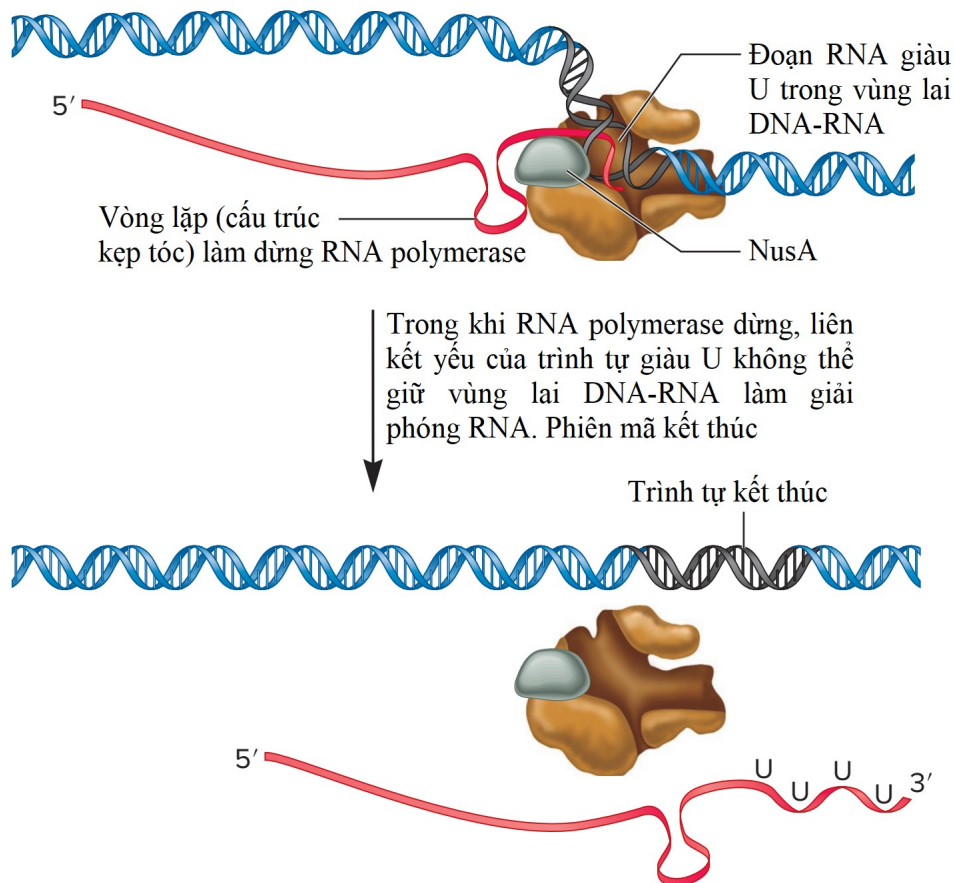
RNA polymerase có hai hoạt tính đọc sửa. Hoạt tính đọc sửa thứ nhất tương tự như quá trình đọc sửa bởi DNA polymerase, trong đó nucleotide được chèn không chính xác sẽ bị loại bỏ bởi enzyme ngược với phản ứng tổng hợp của nó, sau đó lùi lại một bước, và thay thế nucleotide không chính xác bằng nucleotide đúng. Ở hoạt tính đọc sửa thứ hai, enzyme di chuyển ngược lại một hoặc nhiều nucleotide và sau đó cắt bỏ RNA mang nucleotide sai tại vị trí đó trước khi tiếp tục tổng hợp RNA xuôi dòng.

2.2.3. Kết thúc

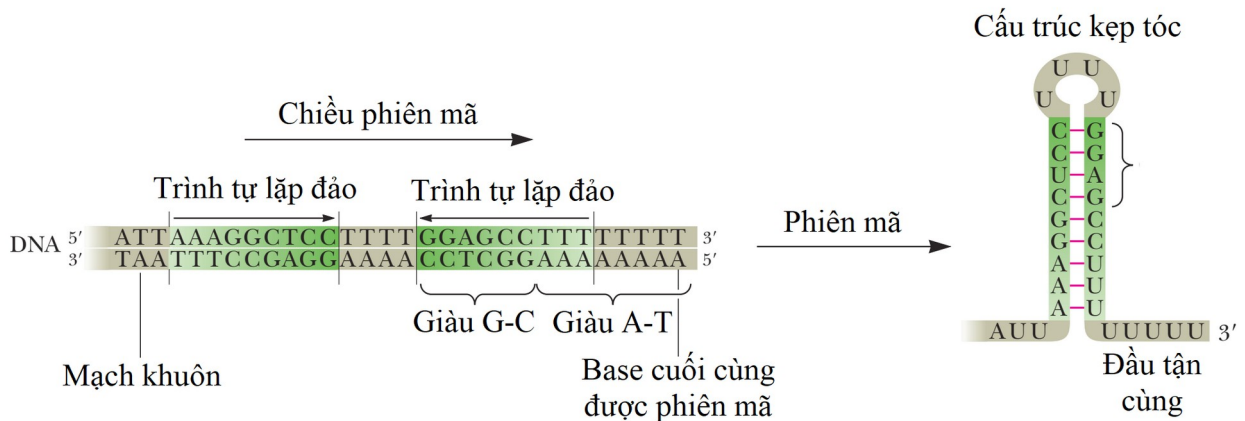
Trước khi kết thúc phiên mã, liên kết hydrogen giữa DNA và RNA trong phức hợp mở có tầm quan trọng chính trong việc ngăn chặn sự giải phóng RNA polymerase khỏi sợi khuôn. Sự kết thúc xảy ra khi vùng lai RNA-DNA ngắn này buộc phải tách ra, do đó giải phóng RNA polymerase cũng như bản phiên mã RNA mới được tạo ra. Ở *E. coli*, người ta đã xác định được hai cơ chế kết thúc phiên mã khác nhau. Đối với một số gene nhất định, một protein liên kết RNA được gọi là **ρ (rho)** chịu trách nhiệm kết thúc phiên mã, cơ chế này được gọi là kết thúc phụ thuộc ρ . Đối với các gene khác, sự kết thúc không cần sự tham gia của protein ρ , cơ chế này được gọi là sự kết thúc không phụ thuộc ρ .

- **Kết thúc phiên mã không phụ thuộc ρ .** Quá trình kết thúc không phụ thuộc ρ không đòi hỏi sự có mặt của protein ρ . Trong trường hợp này, trình tự kết thúc bao gồm hai trình tự nucleotide liên kế có chức năng bên trong RNA: Một là trình tự giàu uracil nằm ở đầu 3' của RNA tương ứng với trình tự 4-8 bp A-T trên gene. Trình tự thứ hai tiếp giáp với trình tự giàu uracil là trình tự bất cặp bổ sung chứa 15-20 nucleotide ngược

dòng điểm kết thúc phiên mã cho phép sự hình thành “cấu trúc kẹp tóc”. Như thể hiện trong Hình 1.12, sự hình thành của vòng lặp (cấu trúc kẹp tóc) làm cho RNA polymerase tạm dừng trong quá trình tổng hợp RNA của nó. Sự tạm dừng này được ổn định bởi các protein khác liên kết với RNA polymerase. Ví dụ, một protein được gọi là NusA, được liên kết với RNA polymerase, thúc đẩy quá trình tạm dừng ở các vòng lặp. Vào thời điểm chính xác mà RNA polymerase tạm dừng, trình tự giàu uracil trong bản sao RNA sẽ liên kết với sợi khuôn DNA. Liên kết hydrogen của RNA với DNA giữ RNA polymerase được kẹp vào DNA. Tuy nhiên, sự liên kết của trình tự giàu uracil này với sợi khuôn mẫu DNA là tương đối yếu, làm cho bản sao RNA phân tách một cách tự nhiên khỏi DNA và ngừng phiên mã. Bởi vì sự kết thúc độc lập với ρ không yêu cầu một protein (protein ρ) để giải phóng bản phiên mã RNA khỏi DNA, nó còn được gọi là sự **kết thúc nội tại**. Ở *E. coli*, khoảng một nửa số gene kết thúc phiên mã bởi sự kết thúc nội tại (không phụ thuộc ρ), và nửa số gene còn lại được kết thúc bởi protein ρ .



Hình 1.12. Quá trình kết thúc không phụ thuộc ρ (hoặc kết thúc nội tại). Khi RNA polymerase đến cuối gene, nó sẽ phiên mã một trình tự giàu uracil. Khi trình tự giàu uracil này được phiên mã, một vòng gốc hình thành ngay ngược dòng từ phức hợp mở. Sự hình thành của vòng gốc này làm cho RNA polymerase tạm dừng trong quá trình tổng hợp bản sao của nó. Quá trình tạm dừng này được ổn định bởi protein NusA, protein này liên kết gần vùng mà RNA thoát ra khỏi phức hợp mở. Trong khi nó đang tạm dừng, RNA trong phép lai RNA-DNA là một trình tự giàu uracil. Bởi vì liên kết hydrogen giữa U và A là tương tác tương đối yếu do đó bản phiên mã RNA và RNA polymerase phân li khỏi DNA.



Hình 1.13. Vị trí kết thúc với *E. coli trp operon* (*trp operon* mã hóa cho các enzyme của quá trình sinh tổng hợp tryptophan). Các trình tự lặp đảo tạo ra cấu trúc vòng lặp hay “kẹp tóc” kết thúc bằng một chuỗi các gốc U.

- **Kết thúc phiên mã phụ thuộc ρ .** Ngược lại với sự kết thúc nội tại phổ biến hơn, một số gene của vi khuẩn yêu cầu tác động của protein rho để liên kết với mRNA mới sinh và xúc tác sự phân tách mRNA khỏi RNA polymerase để chấm dứt phiên mã. Các gene mà phiên mã phụ thuộc vào rho có trình tự kết thúc khác với trình tự kết thúc trong các gene sử dụng trình tự kết thúc nội tại. Cấu trúc vòng lặp thường hình thành như một phần của phần cuối phụ thuộc rho, nhưng trình tự phần cuối phụ thuộc rho không có chuỗi trình tự uracil, vòng lặp này hình thành gần như ngay lập tức sau khi trình tự RNA được tổng hợp và nhanh chóng liên kết với RNA polymerase dẫn đến sự thay đổi cấu trúc khiến RNA polymerase tạm dừng trong quá trình tổng hợp RNA. Thay vào đó, một trình tự ngược dòng từ trình tự kết thúc dài khoảng 50 nucleotide giàu cytosine và nghèo guanine, được gọi là **vị trí *rut***, hoạt động như một vị trí nhận dạng cho sự liên kết của protein rho (ρ). Protein ρ có chức năng như một helicase chuyên hóa sử dụng năng lượng từ ATP để phân tách các vùng lai giữa RNA-DNA. Sau khi vị trí *rut* được tổng hợp trong RNA, protein ρ liên kết với RNA và di chuyển theo hướng của RNA polymerase.

Protein rho bao gồm sáu polypeptide giống nhau và có hai miền chức năng, cả hai đều được sử dụng trong quá trình kết thúc phiên mã gồm hai bước. Bước đầu tiên được bắt đầu khi protein rho được hoạt hóa bởi một phân tử ATP liên kết với một miền chức năng của rho. Protein rho được hoạt hóa sử dụng miền thứ hai của nó để liên kết với vị trí *rut* của phiên mã RNA. Sử dụng năng lượng có nguồn gốc từ ATP, rho sau đó di chuyển dọc theo mRNA theo hướng 3', cuối cùng bắt kịp RNA polymerase đã bị chậm lại gần trình tự kết thúc. Khi rho di chuyển, nó xúc tác sự phá vỡ các liên kết hydrogen giữa mRNA và sợi khuôn DNA. Sự phá vỡ liên kết giải phóng bản sao từ RNA polymerase và cảm ứng polymerase giải phóng khỏi DNA.

3. Phiên mã ở eukaryote

Sự phiên mã ở eukaryote nhìn chung là phức tạp hơn sự phiên mã ở prokaryote. Một phần bởi vì ở eukaryote có ba loại RNA polymerase khác nhau; ngoài ra, do sự khởi đầu phiên mã ở eukaryote thường cần sự tích hợp của nhiều loại tín hiệu hơn.

3.1. Các RNA polymerase ở eukaryote

Ở eukaryote, có 3 loại RNA polymerase tham gia phiên mã tổng hợp các loại phân tử RNA khác nhau:

- **RNA polymerase I** tập trung ở hạch nhân (nucleolus) và xúc tác phản ứng tổng hợp ba loại rRNA, gồm 28S, 18S và 5,8S.

- **RNA polymerase II** chỉ có trong chất nhân (dịch nhân) của tế bào; enzyme này trực tiếp xúc tác các phản ứng tổng hợp mRNA và một số RNA tham gia vào quá trình hoàn thiện RNA, bao gồm snRNA.

- **RNA polymerase III** cũng chỉ thấy trong chất nhân; enzyme này xúc tác phản ứng tổng hợp ba nhóm RNA, gồm các tRNA, RNA ribosome 5S (5S rRNA) và các snRNA không được tổng hợp bởi RNA polymerase II.

Bảng 1.2: Ba loại RNA polymerase nhân của sinh vật nhân chuẩn

Polymerase	RNA được phiên mã	Chức năng của RNA
RNA pol I	Tiền rRNA của 28S, 18S, 5,8S rRNA	Thành phần của ribosome, tổng hợp protein
RNA pol II	mRNA	Mã hóa protein
	snRNA	Cắt nối RNA
	siRNA	Kìm hãm nhờ nhiễm sắc chất, điều khiển dịch mã
	miRNA	Điều khiển dịch mã
RNA pol III	tRNA	Tổng hợp protein
	5S rRNA	Thành phần của ribosome, tổng hợp protein
	snRNA U6	Cắt nối RNA
	7S RNA	Hạt nhận biết tín hiệu (SRP) để đưa polypeptide vào lưới nội chất
	Các RNA ngắn và bền khác	Nhiều chức năng khác nhau, nhiều chức năng chưa biết đến

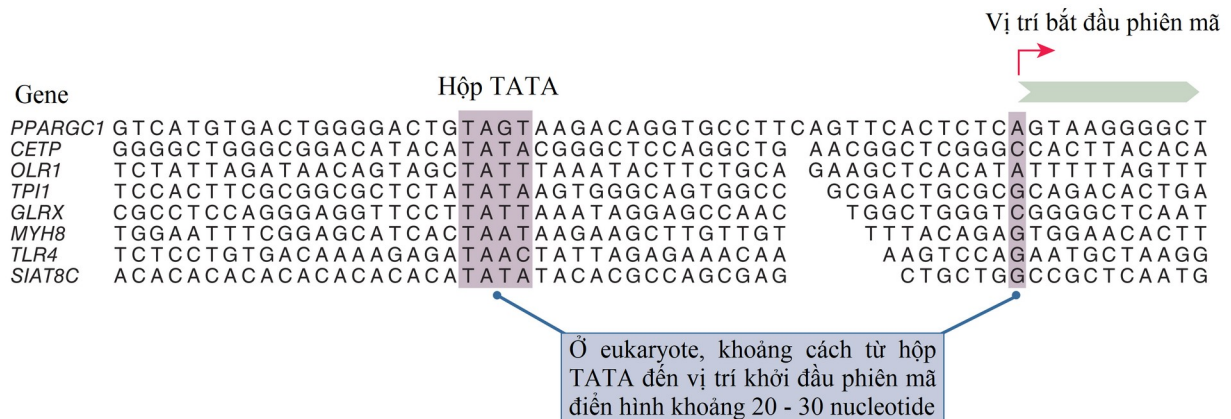
So với cấu trúc và chức năng của RNA polymerase ở *E. coli*, có thể nói đến nay cấu trúc và hoạt động chức năng của các RNA polymerase ở eukaryote còn chưa biết đầy đủ. Điều duy nhất có thể khẳng định là: các RNA polymerase ở eukaryote đều có cấu tạo phức tạp hơn từ rất nhiều tiểu phần khác nhau, cả ba đều chứa hai tiểu phần lớn và 10 – 14 tiểu phần nhỏ; nói cách khác, có nhiều gene cùng phối hợp mã hóa enzyme này. Chẳng hạn như RNA polymerase II của nấm men có 12 chuỗi polypeptide và có cấu trúc không gian kiểu chữ "U". Trong 12 chuỗi polypeptide là thành phần RNA polymerase II ở nấm men, có 5 chuỗi đồng thời là thành phần của RNA polymerase III ở sinh vật này. Nhìn chung, các enzyme RNA polymerase khác của nấm men cũng có cách tổ hợp tương tự. Ở thực vật chứa thêm hai RNA polymerase nhân là RNA polymerase IV và V có họ hàng gần với RNA polymerase II của chúng nhưng mang một tiểu phần lớn độc nhất vô nhị và một số tiểu phần nhỏ riêng, các RNA polymerase này tổng hợp các RNA can thiệp nhỏ (small interfering RNAs).

Bảng 1.3: Các tiểu đơn vị của các RNA polymerase holoenzyme.

Tiểu đơn vị ở vi khuẩn	Ở eukaryote		
	RNA polymerase I	RNA polymerase II	RNA polymerase III
β	RPA1	RPB1	RPC1
β'	RPA2	RPB2	RPC2
α	RPC5/RPC9	RPB3/RPB11	RPC5/RPC9
ω	RPB6	RPB6	RPB6
	[+ 9 tiểu đơn vị khác]	[+ 7 tiểu đơn vị khác]	[+ 11 tiểu đơn vị khác]

2.3.2. Các gene ở eukaryote có một promoter lõi và các yếu tố điều hòa

Để quá trình phiên mã xảy ra với tốc độ thích hợp, gene của sinh vật nhân thực có hai đặc điểm: một promoter lõi và các yếu tố điều hòa. **Promoter lõi** là một trình tự DNA tương đối ngắn cần thiết cho quá trình phiên mã diễn ra. Nó thường bao gồm một trình tự TATAAA được gọi là **hộp TATA (hộp Hogness hay hộp Goldberg-Hogness²)** và vị trí bắt đầu phiên mã. Hộp TATA, thường là khoảng 25 bp ngược dòng từ vị trí bắt đầu phiên mã, rất quan trọng trong việc xác định điểm bắt đầu chính xác cho quá trình phiên mã. Nếu nó bị thiếu trong promoter lõi, thì vị trí bắt đầu phiên mã sẽ không được xác định và phiên mã có thể bắt đầu ở nhiều vị trí khác nhau. Bản thân promoter lõi tạo ra một mức độ phiên mã thấp, đây được gọi là **phiên mã cơ bản (basal transcription)**.



Hình 1.14. Một số hộp TATA của người trong vùng promoter lõi gần vị trí bắt đầu phiên mã.

Các yếu tố điều hòa là các trình tự DNA ngăn ảnh hưởng đến khả năng nhận biết promoter lõi và bắt đầu quá trình phiên mã của RNA polymerase. Các yếu tố này được nhận biết bởi các yếu tố phiên mã - các protein ảnh hưởng đến tốc độ phiên mã. Có hai loại yếu tố điều hòa:

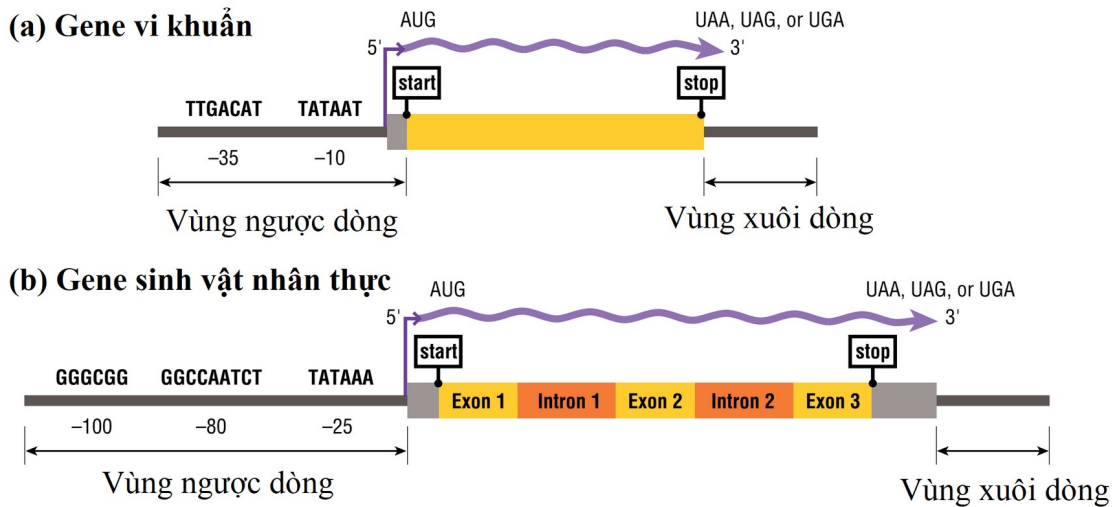
- Các trình tự hoạt hóa, được gọi là các **trình tự tăng cường (enhancer)**, cần thiết để kích thích phiên mã. Các nghiên cứu chỉ ra rằng enhancer thường dài khoảng 200 bp giống như các trình tự cận promoter và chứa nhiều vùng trình tự chức năng dài khoảng 6-10 bp. Trong trường hợp không có trình tự tăng cường, hầu hết các gene của sinh vật nhân thực có mức độ phiên mã cơ bản rất thấp.

² Được xác định vào năm 1978 bởi nhà hóa sinh người Mỹ David Hogness và nghiên cứu sinh của mình là Michael Goldberg.

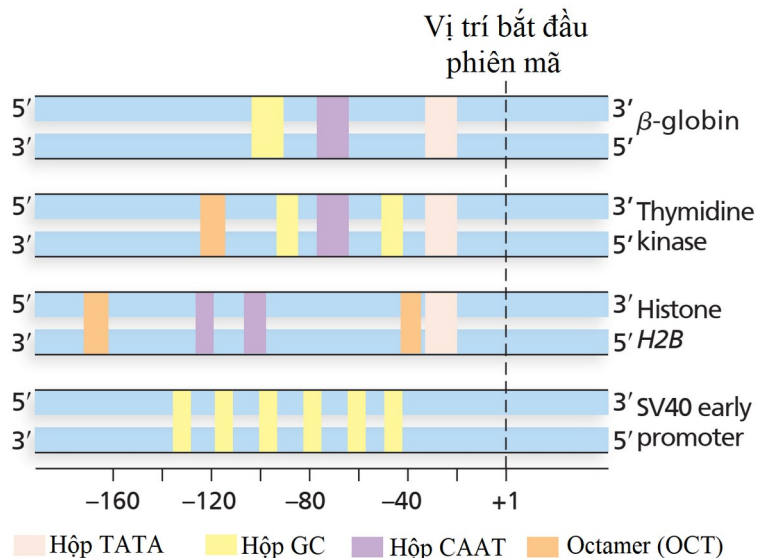
- Trong những điều kiện nhất định, có thể cần phải ngăn chặn sự phiên mã của một gene nhất định. Điều này xảy ra thông qua các **trình tự im lặng** hay **trình tự tắt (silencer)** - trình tự DNA được nhận biết bởi các yếu tố ức chế quá trình phiên mã.

Vị trí phổ biến cho các yếu tố điều hòa là vùng -50 đến -100. Tuy nhiên, vị trí của các yếu tố điều hòa khác nhau đáng kể giữa các gene khác nhau của sinh vật nhân thực. Những yếu tố này có thể ở xa promoter lõi nhưng lại ảnh hưởng mạnh đến khả năng bắt đầu phiên mã của RNA polymerase.

Ngoài ra, còn có **hộp CCAAT** nằm ngược dòng cách vị trí bắt đầu phiên mã khoảng 75-80 nucleotide. Hộp này ít phổ biến hơn hộp TATA, nó đóng vai trò điều hòa tốc độ phiên mã, tăng hiệu quả phiên mã. Vùng giàu G-C ngược dòng được gọi là **hộp giàu GC**, với trình tự liên ứng GGGCGG nằm ở vị trí -90, có tần suất nhỏ hơn tần suất của trình tự hộp CAAT. Ngoài ra, còn có các thành phần đặc hiệu khác.



Hình 1.15. Một gene điển hình với promoter của nó ở vi khuẩn và sinh vật nhân chuẩn. Vùng ngược dòng bao gồm hầu hết các tín hiệu cần thiết cho việc bắt đầu phiên mã. Khu vực xuôi dòng có thể liên quan đến các khu vực quy định khác. Mặc dù sơ đồ trình tự promoter của gene của sinh vật nhân chuẩn này hiển thị hộp TATA điển hình, nhưng các yếu tố trình tự của trình tự khởi động gene sinh vật nhân chuẩn rất khác nhau và chỉ một tập hợp con có các trình tự được hiển thị ở đây.



Hình 1.16. Một số ví dụ về tính đa dạng ở promoter của sinh vật nhân thực.

So sánh các promoter sinh vật nhân chuẩn cho thấy mức độ biến đổi cao về loại, số lượng và vị trí của các phần tử trình tự nhất trí (Hình 1.16). Một số promoter chứa cả ba trong số các trình tự liên ứng được xác định ở trên, một số khác chứa một hoặc hai trong số các trình tự này, một số không chứa trình tự liên ứng nào. Ví dụ, gene thymidine kinase chứa các hộp TATA, CAAT và GC cùng với trình tự octamer (OCT), được gọi là hộp OCT. Gene histone *H2B* chứa hai hộp OCT ngoài một hộp TATA và một cặp hộp CAAT. Tất cả các trình tự liên ứng này đóng vai trò quan trọng trong việc liên kết các yếu tố phiên mã.

Các trình tự DNA như hộp TATA, trình tự tăng cường (enhancer) và trình tự tắt (silencer) chỉ phát huy tác dụng của chúng trên một gene cụ thể. Chúng được gọi là các **yếu tố tác động *cis* (*cis*-acting element)**. Thuật ngữ *cis* xuất phát từ danh pháp hóa học có nghĩa là "bên cạnh." Các yếu tố tác động *cis*, mặc dù có thể ở xa promoter lõi, nhưng luôn được tìm thấy trong cùng một nhiễm sắc thể với các gene mà chúng quy định. Ngược lại, các yếu tố điều hòa phiên mã liên kết với các yếu tố như vậy được gọi là các **yếu tố tác động *trans* (*trans*-acting factor)**. Các yếu tố phiên mã điều hòa sự biểu hiện gene được mã hóa bởi các gene; các gene điều hòa mã hóa các yếu tố phiên mã có thể ở xa gene mà chúng kiểm soát, ngay cả trên một nhiễm sắc thể khác. Khi gene mã hóa yếu tố phiên mã được biểu hiện, protein của yếu tố phiên mã có thể khuếch tán vào nhân tế bào và liên kết với yếu tố tác động *cis* thích hợp của nó.

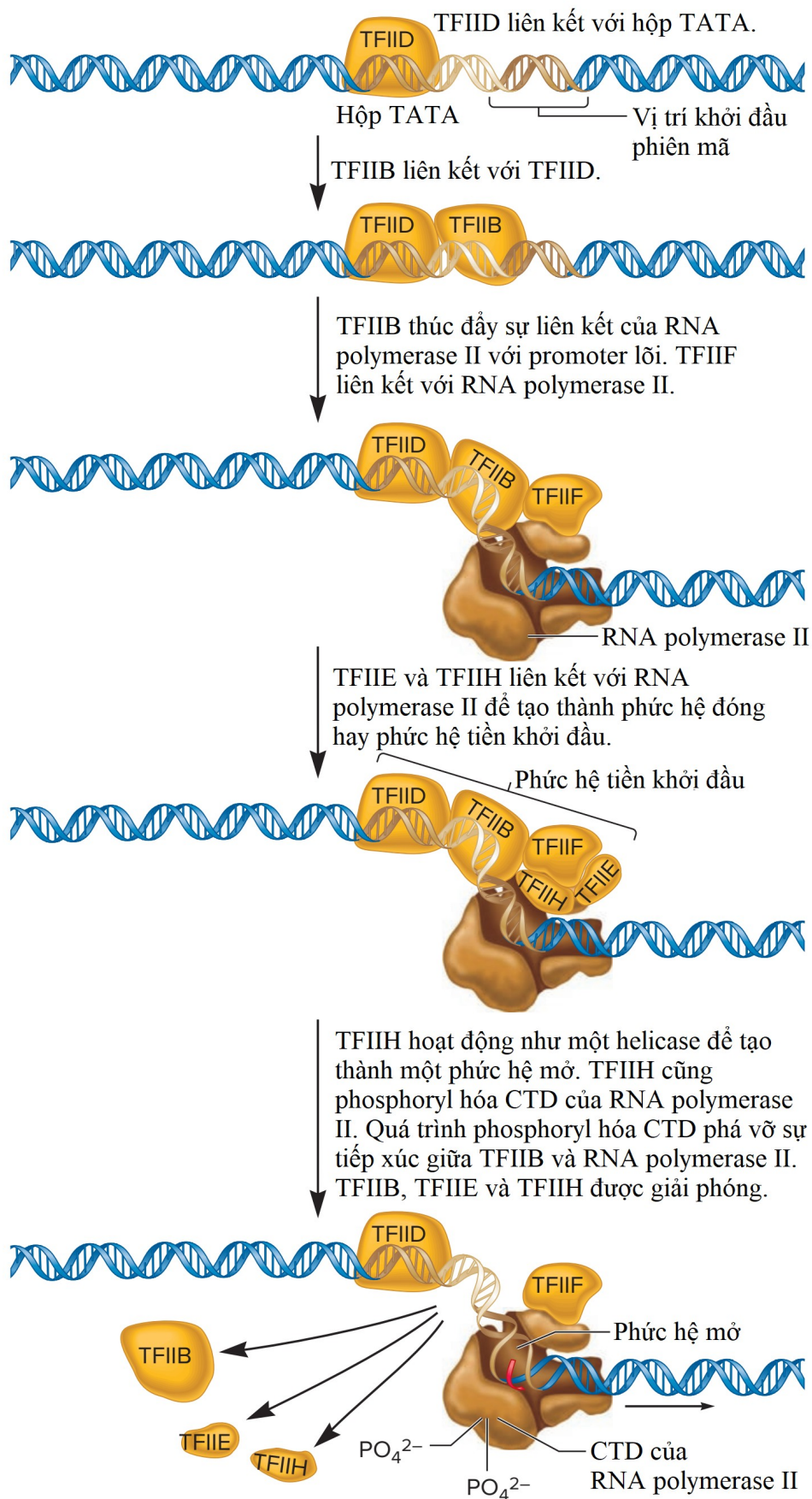
2.3.3. Khởi đầu phiên mã các gene mã hóa protein bởi RNA polymerase II

Bằng cách nghiên cứu quá trình phiên mã ở nhiều loài sinh vật nhân chuẩn, các nhà nghiên cứu đã phát hiện ra rằng ba loại protein cần thiết cho quá trình phiên mã cơ bản tại promoter lõi: RNA polymerase II, các yếu tố phiên mã chung và một phức hợp được gọi là yếu tố trung gian (mediator).

Năm loại protein khác nhau được gọi là các **yếu tố phiên mã chung (GTF)** luôn cần thiết cho RNA polymerase II để bắt đầu phiên mã các gene mã hóa protein. Trong quá trình lắp ráp của GTFs và RNA polymerase II tại hộp TATA, yếu tố phiên mã IID (TFIID) đầu tiên liên kết với hộp TATA và do đó đóng một vai trò quan trọng trong việc nhận biết promoter lõi. TFIID bao gồm một số tiểu đơn vị, bao gồm protein liên kết TATA (TBP), liên kết trực tiếp với hộp TATA và một số protein khác được gọi là các yếu tố liên kết với TBP (TAF). Sau khi TFIID liên kết với hộp TATA, nó sẽ liên kết với TFIIB. TFIIB thúc đẩy sự liên kết của RNA polymerase II và TFIIF. Cuối cùng, TFIIE và TFIIH liên kết với phức hợp. Điều này hoàn thành việc lắp ráp các protein để tạo thành một phức hệ đóng gọi là **phức hệ khởi đầu phiên mã hoàn chỉnh (complete transcription initiation complex)**, còn được gọi là **phức hệ tiền khởi đầu (PIC, preinitiation complex)**.

TFIIH đóng một vai trò quan trọng trong sự hình thành của phức hệ mở. TFIIH có một số tiểu đơn vị thực hiện các chức năng khác nhau. Một số tiểu đơn vị hoạt động như các helicase, phá vỡ các liên kết hydrogen giữa DNA sợi kép và cần thiết để tạo thành một phức hợp mở khởi đầu. Một tiểu đơn vị khác thủy phân ATP và phosphoryl hóa một miền trong RNA polymerase II được gọi là miền tận cùng carboxyl (CTD). Quá trình phosphoryl hóa CTD giải phóng phần tiếp xúc giữa RNA polymerase II và TFIIB. Tiếp

theo, TFIIB, TFIIIE và TFIIH phân ly, và RNA polymerase II tự do để tiến tới giai đoạn kéo dài của phiên mã.



Hình 1.17. Các bước dẫn đến sự hình thành của phức hợp mở trong khởi đầu phiên mã.

Trong ống nghiệm, khi các nhà nghiên cứu trộn TFIID, TFIIB, TFIIF, TFIIE, TFIIH, RNA polymerase II và trình tự DNA chứa hộp TATA và vị trí bắt đầu phiên mã lại với nhau, DNA sẽ được phiên mã thành RNA. Do đó, các thành phần này được gọi là **bộ máy phiên mã cơ bản**. Tuy nhiên, trong tế bào sống, có thêm các thành phần khác điều chỉnh quá trình phiên mã và cho phép nó tiến hành với tốc độ hợp lý.

Ngoài GTF và RNA polymerase II, một thành phần khác cần thiết cho quá trình phiên mã là phức hợp protein lớn gọi là yếu tố trung gian. Phức hợp này được phát hiện bởi Roger Kornberg và các đồng nghiệp vào năm 1990. Năm 2006, Kornberg được trao giải Nobel hóa học cho những nghiên cứu của ông về cơ sở phân tử của quá trình phiên mã ở sinh vật nhân chuẩn. Tên gọi yếu tố trung gian xuất phát từ quan sát rằng nó làm trung gian tương tác giữa RNA polymerase II và các yếu tố điều hòa phiên mã – các yếu tố liên kết với enhancer hoặc silencer. Nó đóng vai trò như bề mặt chung giữa RNA polymerase II và nhiều tín hiệu điều hòa đa dạng. Thành phần tiểu đơn vị của yếu tố trung gian khá phức tạp và thay đổi. Các tiểu đơn vị lõi tạo thành một phức hợp hình elip bao bọc một phần xung quanh RNA polymerase II. Bản thân chất trung gian có thể phosphoryl hóa CTD của RNA polymerase II và nó có thể điều chỉnh khả năng của TFIIH để phosphoryl hóa CTD. Do đó, nó đóng một vai trò quan trọng trong việc chuyển đổi giữa khởi đầu và kéo dài phiên mã.

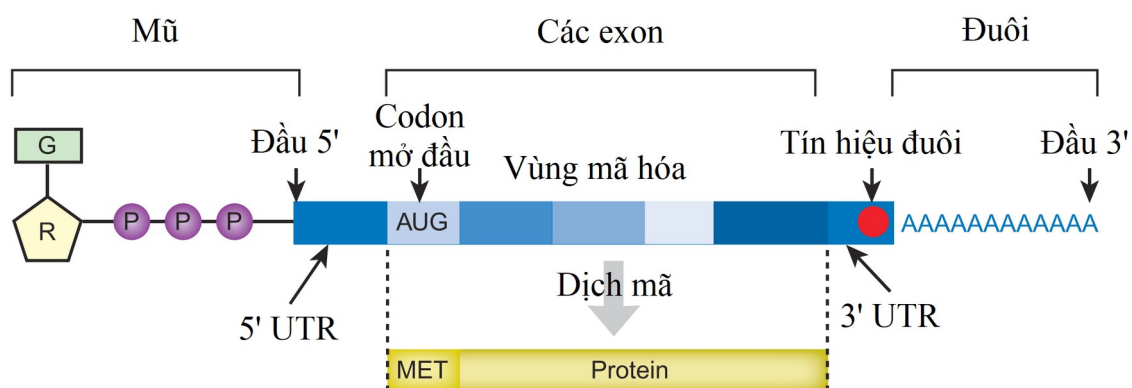
Bảng 1.4: Các protein cần cho phiên mã thông qua promoter lõi của các gene mã hóa protein ở eukaryote

Các protein	Chức năng
RNA polymerase II	Enzyme xúc tác sự liên kết của các nucleotide theo hướng 5' đến 3', sử dụng DNA làm khuôn mẫu. Hầu hết các protein RNA polymerase II của sinh vật nhân thực được cấu tạo bởi 12 tiểu đơn vị. Hai tiểu đơn vị lớn nhất có cấu trúc tương tự như tiểu đơn vị β và β' được tìm thấy trong <i>E. coli</i> RNA polymerase.
Các yếu tố phiên mã chung	
TFIID	Gồm protein bám TATA (TBP) và các yếu tố liên kết TBP khác (TAFs). Nhận biết hộp TATA của các promoter của gene mã hóa protein ở eukaryote.
TFIIB	Liên kết với TFIID và sau đó cho phép RNA polymerase II bám vào promoter lõi. Cũng thúc đẩy TFIIF liên kết.
TFIIF	Liên kết với RNA polymerase II và góp phần vào khả năng liên kết với TFIIB và promoter lõi. Cùng đóng vai trò trong khả năng liên kết với RNA polymerase II của TFIIE và TFIIH.
TFIIE	Đóng vai trò trong sự hình thành hoặc duy trì (hoặc cả hai) phức hệ mở bằng cách tạo điều kiện thuận lợi cho việc gắn TFIIH vào RNA polymerase II và điều chỉnh hoạt động của TFIIH.
TFIIH	Một protein đa đơn vị có nhiều vai trò. Đầu tiên, các tiểu đơn vị nhất định hoạt động như các helicase và thúc đẩy sự hình thành của phức hợp mở. Các tiểu đơn vị khác phosphoryl hóa miền tận cùng carboxyl (CTD) của RNA polymerase II, giải phóng sự

Các protein	Chức năng
	tương tác của nó với TFIIB, do đó cho phép RNA polymerase II tiến tới giai đoạn kéo dài.
Yếu tố trung gian	Một phức hợp đa đơn vị làm trung gian cho tác động của các yếu tố điều hòa phiên mã lên chức năng của RNA polymerase II. Mặc dù yếu tố trung gian thường có một số tiểu đơn vị lõi nhất định, nhưng nhiều tiểu đơn vị của nó có sự khác nhau, tùy thuộc vào loại tế bào và điều kiện môi trường. Khả năng của yếu tố trung gian ảnh hưởng đến chức năng của RNA polymerase II được cho là xảy ra thông qua CTD của RNA polymerase II. Yếu tố trung gian có thể ảnh hưởng đến khả năng phosphoryl hóa CTD của TFIIF, và bản thân các tiểu đơn vị trong yếu tố trung gian có khả năng phosphoryl hóa CTD. Bởi vì quá trình phosphoryl hóa CTD là cần thiết để giải phóng RNA polymerase II từ TFIIB, chất trung gian đóng một vai trò quan trọng trong khả năng RNA polymerase II chuyển từ giai đoạn khởi đầu sang giai đoạn kéo dài của phiên mã.

3.2. mRNA của eukaryote

Hình 4.5 minh họa cấu trúc chung của mRNA ở dạng hoàn thiện (sẵn sàng cho dịch mã) ở prokaryote và eukaryote. Theo đó, cấu trúc của mRNA có thể chia làm ba vùng chính. Ở tận cùng đầu 5' là một đoạn trình tự dẫn đầu, còn gọi là **vùng đầu 5' không được dịch mã (5'-UTR)** có chiều dài khác nhau ở các mRNA của các gene khác nhau. Theo sau 5'-UTR là vùng trình tự mã hóa; đây chính là trình tự nucleotide xác định trình tự các amino acid trên phân tử protein trong dịch mã. Theo sau vùng trình tự mã hóa là đoạn trình tự theo sau, hay còn gọi là **vùng đầu 3' không được dịch mã (3'-UTR)**; đoạn trình tự này qui định tính bền của mRNA và hiệu suất dịch mã.



Hình 1.18. Cấu trúc của một phân tử mRNA trưởng thành ở sinh vật nhân thực. Sau khi phiên mã, một mũ 7-methyl guanosine được thêm vào đầu 5' của mRNA, và đường ribose của nucleotide đầu tiên, và đôi khi thứ hai, được methyl hóa ở vị trí 2'. Đầu 3' của bản sao được polyadenyl hóa với việc bổ sung các gốc 200 - 250 A.

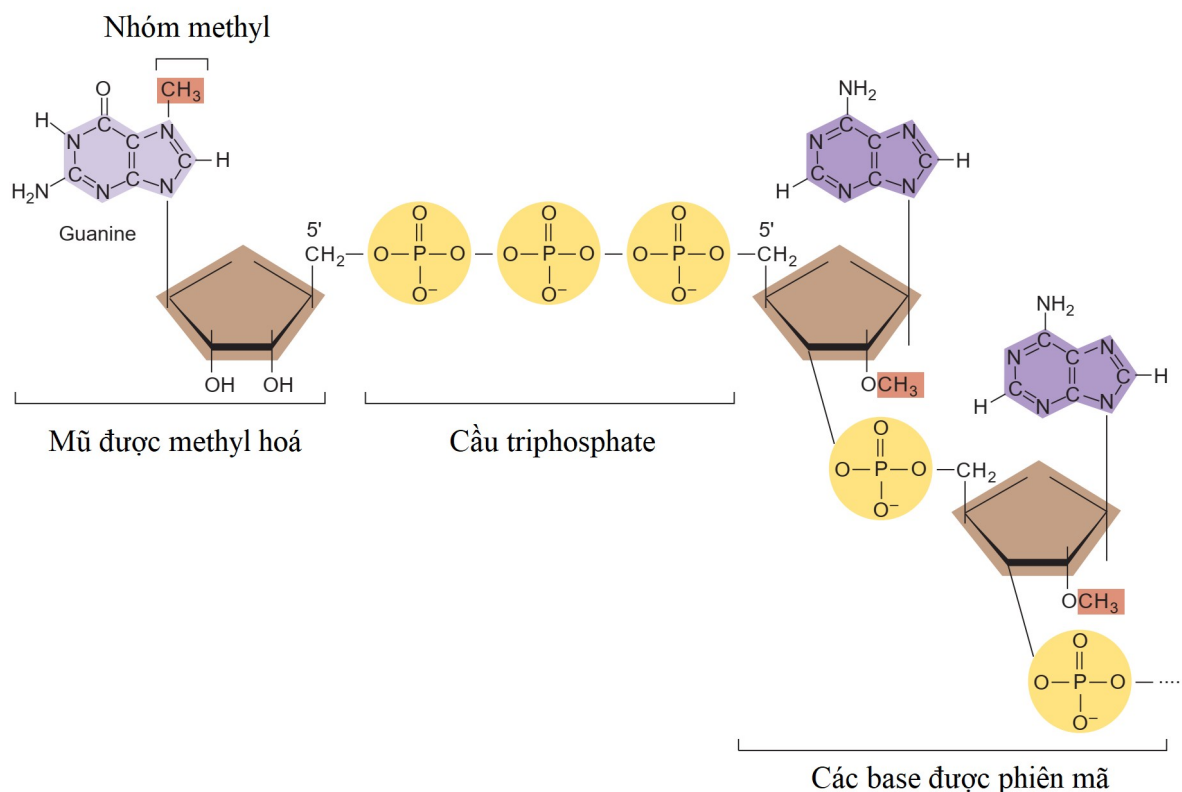
Sự tổng hợp mRNA có khác nhau giữa prokaryote và eukaryote. Ở prokaryote, phân tử mRNA mới được phiên mã thường được sử dụng ngay để tổng hợp protein. Ngoài ra, do vi khuẩn không có màng nhân nên các mRNA có thể được ribosome dùng ngay để

dịch mã, kể cả khi chúng chưa được phiên mã xong. Quá trình này được gọi là **sự kết cặp phiên mã - dịch mã** hay phiên mã và dịch xảy ra đồng thời. Ở eukaryote, phân tử mRNA sau phiên mã (tiền mRNA) phải trải qua một loạt các sự kiện biến đổi (quá trình hoàn thiện RNA) để hình thành nên phân tử mRNA hoàn thiện (còn gọi là mRNA thành thực, mRNA trưởng thành). Ngoài ra, ở eukaryote phân tử mRNA phải được vận chuyển từ nhân ra tế bào chất nơi có sự tập trung các ribosome thì sự dịch mã mới có thể bắt đầu. Như vậy, một phân tử mRNA ở eukaryote luôn được phiên mã hoàn chỉnh và được hoàn thiện trước dịch mã. Một khác biệt cơ bản nữa giữa mRNA của prokaryote và eukaryote là mRNA của prokaryote thường là **đa cistron (polycistronic mRNAs)**, nghĩa là chúng đồng thời chứa thông tin mã hóa của nhiều hơn một chuỗi polypeptide; trong khi đó, mRNA của eukaryote luôn là **đơn cistron (monocistronic mRNAs)**, nghĩa là nó chỉ mang thông tin mã hóa một chuỗi polypeptide duy nhất.

3.3. Biến đổi RNA sau phiên mã

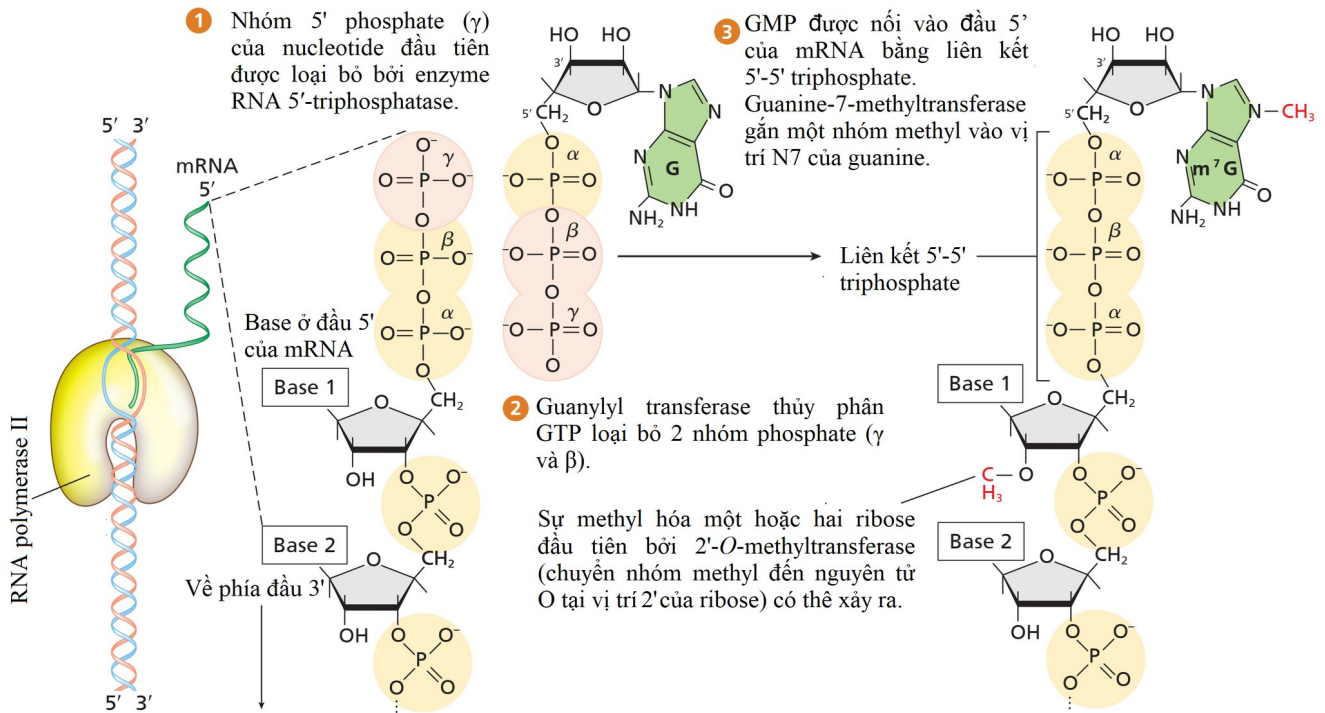
3.3.1. Sự cải biến đầu 5' và 3' của mRNA

Sự lắp mũ đầu 5' là bước đầu tiên của quá trình hoàn thiện mRNA. Một khi enzyme RNA polymerase II đã tổng hợp được một đoạn tiền mRNA dài khoảng 20 - 30 nucleotide, thì một enzyme "lắp mũ" sẽ bổ sung một nucleotide guanine được gắn nhóm methyl (-CH₃) tại vị trí số 7 của base guanosine (m⁷G) vào đầu 5' của chuỗi RNA qua liên kết 5'-5' (thay cho liên kết thông thường là 5'-3'). Nhờ cầu liên kết 5'-5', mũ m⁷G bảo vệ được mRNA trong suốt quá trình dịch mã tránh tác động của các exonuclease. Ngoài việc lắp mũ m⁷G, ở sinh vật nhân thật bậc cao, một hoặc hai nucleotide đầu tiên trong chuỗi được methyl hóa ở vị trí C-2' (hiện tượng này không gặp ở sinh vật nhân thật bậc thấp).



Hình 1.19. Cấu trúc của mũ được methyl hóa ở đầu 5' của mRNA nhân chuẩn.
Enzyme gắn mũ nối một G ngược với nucleotide đầu tiên của bản phiên mã sơ cấp thông

qua liên kết triphosphate. Sau đó, các enzyme methyl transferase sẽ thêm các nhóm methyl vào G này và vào một hoặc hai trong số các nucleotide được phiên mã đầu tiên từ khuôn mẫu DNA.



Hình 1.20. Gắn mũ đầu 5' của tiền mRNA. Một tiểu phần của enzyme guanylyl transferase (RNA 5'-triphosphatase) loại bỏ nhóm phosphate γ ra khỏi đầu 5' của chuỗi RNA mới sinh ❶. Một miền khác của tiểu phần này chuyển nhóm GMP từ GTP tới diphosphate 5' của sản phẩm phiên mã mới sinh, tạo ra cấu trúc dị thường 5'-5' triphosphate ❷. Ở bước cuối cùng các enzyme methyltransferase chuyển các nhóm methyl từ S-adenosylmethionine đến vị trí N₇ của guanine (enzyme guanine-7-methyltransferase) ❸ và các nguyên tử O ở vị trí 2' của các đường ribose đầu tiên hoặc thứ hai của mRNA mới sinh (enzyme 2'-O-methyltransferase).

Trong nhân, mũ gắn với một phức hợp protein được gọi là **CBC** (phức hợp gắn mũ, CAP-binding complex) giúp RNA được xử lý tốt và được chuyển ra ngoài. Mũ đầu 5' có một số chức năng, bao gồm (1) bảo vệ mRNA khỏi sự suy thoái nhanh chóng bởi sự phân giải của các enzyme phân hủy RNA (5'-exoribonuclease), (2) tạo điều kiện thuận lợi cho việc vận chuyển mRNA qua màng nhân, (3) tạo điều kiện thuận lợi cho quá cắt nối intron tiếp theo và (4) tăng cường hiệu quả dịch mã bằng cách định hướng ribosome trên mRNA.

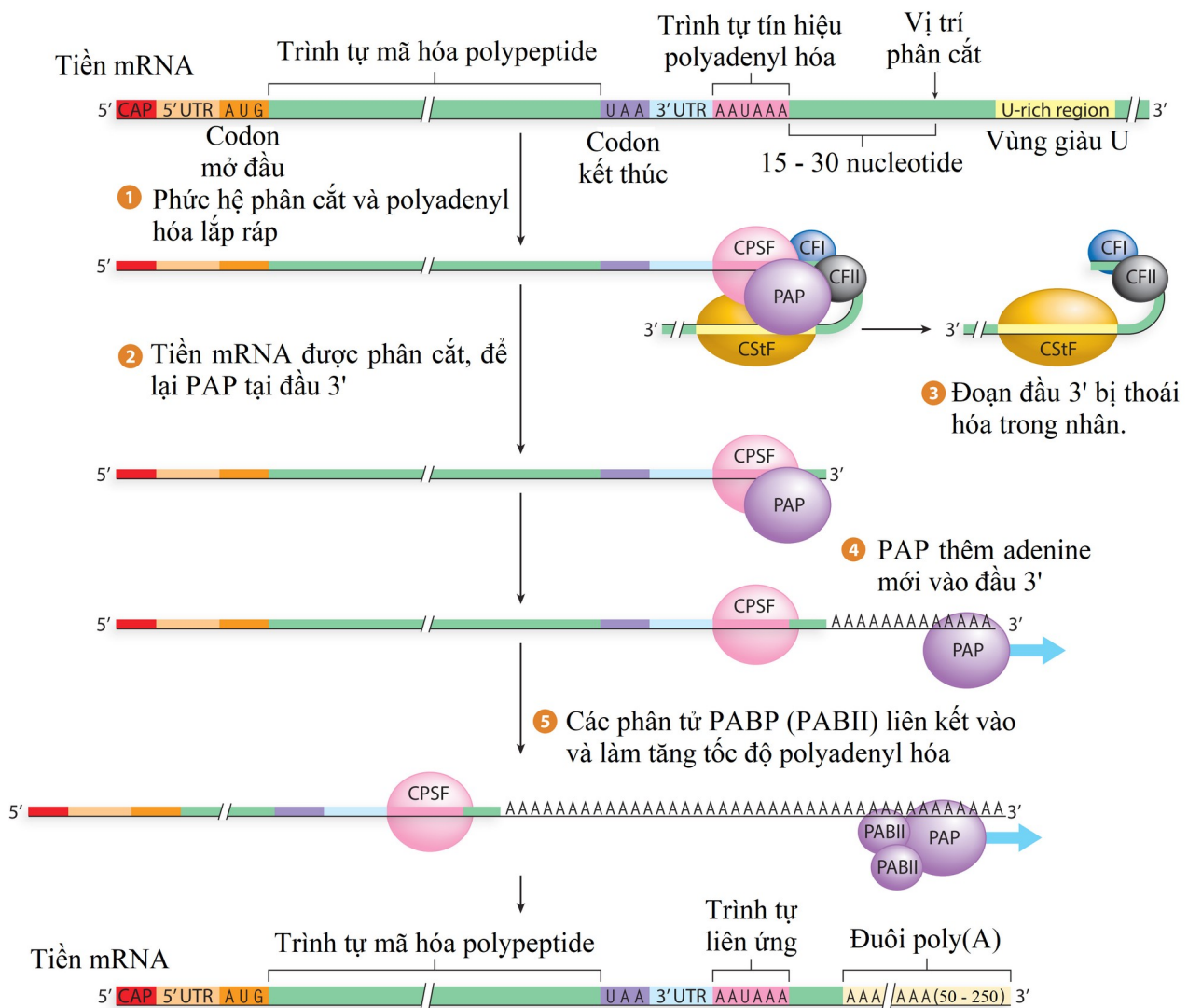
Các tiền mRNA trước khi được vận chuyển từ nhân ra tế bào chất đồng thời được biến đổi đầu 3' bằng việc **gắn đuôi poly(A)** - thường là một chuỗi liên tục gồm 200 - 250 A. Trên mạch DNA khuôn không có trình tự mã hóa đuôi poly(A) này, mà nó được bổ sung vào mRNA trong quá trình hoàn thiện. Đuôi poly(A) cần để mRNA được vận chuyển "an toàn" từ nhân ra tế bào chất. Trong tế bào chất, đuôi poly(A) ngoài việc giúp bảo vệ đầu 3' của mRNA khỏi sự tác động của các exonuclease, tạo điều kiện thuận lợi cho việc vận chuyển mRNA trưởng thành qua màng nhâ, còn có vai trò trong điều hòa tính ổn định của các phân tử mRNA và tăng hiệu suất dịch mã. Sự gắn đuôi poly(A) là

cần thiết cho sự dịch mã ở nhiều gene (một số mRNA ở eukaryote trong quá trình phát triển phôi không được dịch mã nếu thiếu đuôi poly(A)).

Sự kết thúc phiên mã và gắn đuôi poly(A) ở eukaryote có liên quan đến trình tự liên ứng nhận biết đuôi 5'-AAUAAA-3' nằm gần đầu 3' của mRNA. Thường thì trong quá trình phiên mã, RNA pol II tổng hợp phân tử tiền mRNA vượt qua vị trí này. Nhưng một loại exonuclease đặc biệt nhận ra vị trí này và tiến hành cắt phân tử RNA đang được kéo dài khoảng 10 đến 35 nucleotide kể từ một vị trí có đôi nucleotide 5'-CA-3'. Cách vị trí cắt một quãng về đầu 3' của mRNA là một đoạn trình tự giàu GU cũng có vai trò giúp nhận ra vị trí gắn đuôi poly(A), nhưng đoạn này sau đó được cắt bỏ.

Để gắn đuôi poly(A) vào mRNA, có sự tham gia của protein được gọi là **yếu tố đặc hiệu cắt rời và polyadenyl hóa (CPSF)** 360 kDa gồm 5 polypeptide khác nhau trước tiên tạo nên một phức không bền với tín hiệu poly(A) hóa AAUAAA ngược dòng. Sau đó ít nhất ba protein khác gắn vào phức hợp: một protein gồm ba tiểu phần khác nhau kích thước 200 kDa gọi là **yếu tố kích thích cắt rời (CStF)** tương tác với vùng giàu G/U; một protein bốn tiểu phần khác nhau kích thước 150 kDa gọi là **yếu tố cắt I (CFI)**; và một yếu tố cắt gồm hai tiểu phần khác nhau thứ hai gọi là **yếu tố cắt II (CFII)**. Một protein khoảng 150 kDa gọi là **symplekin** được cho là có vai trò giàn giáo để các yếu tố cắt/polyadenyl hóa tập hợp. Cuối cùng **poly(A) polymerase (PAP)** gắn vào toàn bộ phức này trước khi phản ứng cắt xảy ra. Điều kiện phải có PAP gắn vào giúp cặp đôi phản ứng cắt rời với polyadenyl hóa, vì thế đầu 3' tự do sinh ra sẽ được polyadenyl hóa nhanh chóng và thông tin quan trọng sẽ không bị mất đi trong trường hợp exonuclease phân hủy đầu 3' không được bảo vệ.

Tiếp theo sau phản ứng cắt rời trình tự xuôi dòng ở vị trí poly(A), polyadenyl hóa xảy ra theo hai giai đoạn. Quá trình gắn khoảng 12 nucleotide A đầu tiên xảy ra với tốc độ chậm, sau đó khoảng 200-250 A hoặc hơn được gắn thêm vào rất nhanh. Giai đoạn gắn nhanh đòi hỏi rất nhiều protein gắn poly(A) chứa motif RRM gắn vào. Protein này được gán cho tên gọi là PABPII để phân biệt với protein gắn poly(A) trong tế bào chất. PABPII gắn vào đuôi poly(A) ngăn do PAP tổng hợp, kích thích tốc độ polymer hóa các A tiếp theo cũng do PAP gắn vào, dẫn đến giai đoạn gắn nhanh của quá trình polyadenyl hóa. PABPII cũng đảm nhận vai trò phát tín hiệu cho PAP dừng phản ứng polymer hóa khi đuôi poly(A) đạt tới chiều dài 200-250 nucleotide, mặc dù cơ chế kiểm soát chiều dài đuôi vẫn chưa được làm sáng tỏ. Việc PABP gắn vào đuôi poly(A) đóng vai trò quan trọng cho sự vận chuyển RNA vào tế bào chất.

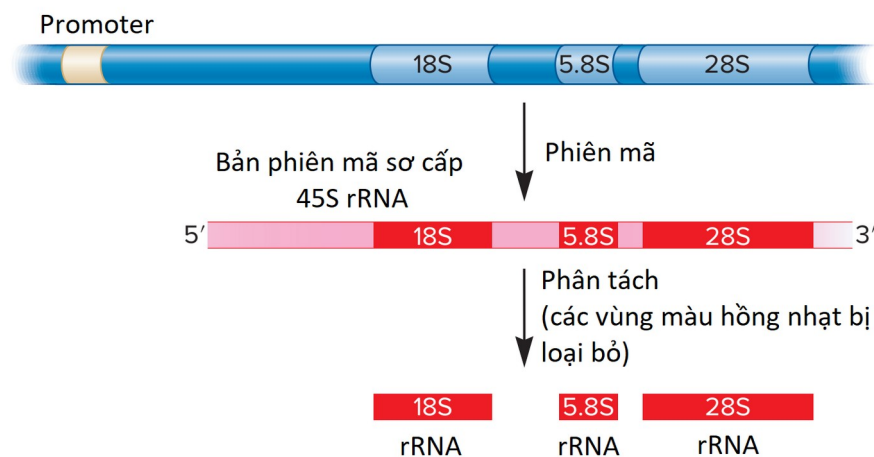


Hình 1.21. Mô hình cắt rời và polyadenyl hóa tiền mRNA. Yếu tố cắt rời và polyadenyl hóa đặc hiệu (CPSF) gắn vào tín hiệu poly(A) ngược dòng AAUAAA. CStF tương tác với trình tự giàu U hoặc GU xuôi dòng và với CPFS đang gắn, tạo thành cấu trúc vòng trong RNA; CFI và CFII gắn vào giúp ổn định phức này. Poly(A) polymerase gắn vào sẽ kích thích cắt bỏ ở vị trí poly(A), thường là 10 – 35 nucleotide về phía 3' của tín hiệu poly(A) ngược dòng. Các yếu tố cắt rời sau đó được giải phóng. Phần RNA xuôi dòng cũng được giải phóng và nhanh chóng bị phân hủy. PAP đang bám sau đó sẽ gắn thêm khoảng 12A ở tốc độ chậm vào nhóm 3' hydroxyl tạo ra bởi phản ứng cắt. Protein II gắn poly(A) (PABII) gắn vào đuôi poly(A) ngăn mới sinh sẽ làm tăng tốc độ gắn của PAP. Sau khi 200 – 250A đã được gắn thêm vào, PABP ra tín hiệu cho PAP dừng phản ứng.

Một số mRNA ở vi khuẩn cũng có đuôi poly(A) - dài khoảng 10-20 bp. Nhưng, điều đáng lưu ý là ở prokaryote, vai trò của đuôi poly(A) hoàn toàn khác biệt với ở eukaryote. Cụ thể, đuôi poly(A) lại thúc đẩy sự biến tính của mRNA ở prokaryote. Thực nghiệm cho thấy việc bổ sung đuôi poly(A) vào DNA lặp thể cũng thúc đẩy sự biến tính của nó. Đây cũng là một bằng chứng cho thấy biểu hiện của hệ gene ở các bào quan giống với các hệ gene vi khuẩn.

3.3.2. Sự hoàn thiện các tRNA và rRNA

Đối với một số gene, bản phiên mã RNA ban đầu được chế biến hoặc phân cắt thành các phần nhỏ hơn. Ví dụ, hình 1.22 cho thấy quá trình chế biến RNA ribosome (rRNA) của động vật có vú. Gene RNA của ribosome được phiên mã bởi RNA polymerase I, tạo ra một bản phiên mã sơ cấp dài được gọi là 45S rRNA. Sau khi tổng hợp rRNA 45S, sự phân cắt xảy ra tại một số điểm để tạo ra ba đoạn, được gọi là rRNA 18S, 5,8S và 28S. Các phân tử rRNA chức năng này đóng một vai trò quan trọng trong việc hình thành cấu trúc của ribosome. Ở eukaryote, sự phân cắt của rRNA 45S thành các rRNA nhỏ hơn và sự lắp ráp của các tiểu đơn vị ribosome xảy ra tại một vị trí trong nhân tế bào được gọi là **hạch nhân (nucleolus)**.



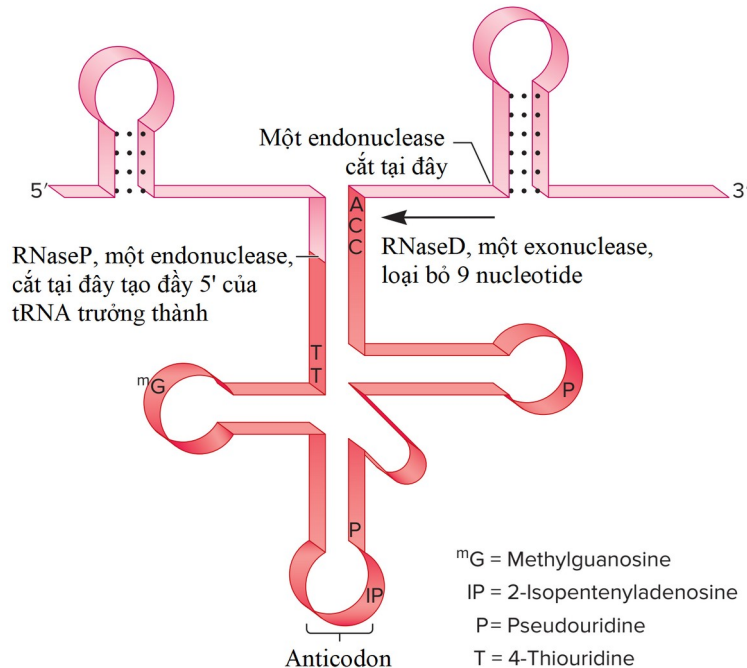
Hình 1.22. Quá trình chế biến (gia công) RNA ribosome ở sinh vật nhân thực. *Gene RNA của ribosome lớn được phiên mã thành một bản phiên mã sơ cấp dài 45S rRNA. Bản phiên mã này được phân cắt để tạo ra các phân tử rRNA 18S, 5,8S và 28S, các phân tử này liên kết với các tiểu đơn vị protein trong ribosome. Quá trình này xảy ra trong hạch nhân của tế bào.*

Ở vi khuẩn, có 3 loại rRNA, có hệ số lắng tương ứng là 5S, 16S và 23S, được hình thành từ một tiền RNA (pre-rRNA) duy nhất. Phân tử tiền RNA mang ba đoạn trình tự mã hóa các rRNA theo trật tự 16S-23S-5S xen giữa các đoạn nối không mã hóa (các đoạn nối không mã hóa này không được gọi là exon vì chúng nằm ngoài gene). Ở vi khuẩn, phân tử tiền rRNA còn mang trình tự mã hóa một số tRNA. Ở prokaryote cũng như ở eukaryote, hệ gene thường đồng thời có nhiều bản sao của gene mã hóa các rRNA.

Việc tạo ra các phân tử tRNA đòi hỏi quá trình xử lý thông qua các exonuclease và endonuclease. **Exonuclease** phân cắt liên kết cộng hóa trị giữa hai nucleotide ở một đầu của sợi. Một số exonuclease chỉ bắt đầu quá trình tiêu hóa này từ đầu 3', di chuyển theo hướng 3' đến 5', trong khi những exonuclease khác chỉ bắt đầu ở đầu 5' và phân tách nucleotide theo hướng 5' đến 3'. Trong khi đó, **endonuclease** phân cắt liên kết giữa hai nucleotide liền kề bên trong một sợi.

Giống như rRNA, tRNA cũng được tổng hợp dưới dạng các tiền tRNA lớn phải được phân cắt để tạo ra các tRNA có chức năng và trưởng thành vận chuyển các amino acid. Quá trình chế biến này đã được nghiên cứu rộng rãi ở *E. coli*. Các chi tiết của quá trình xử lý tRNA khác nhau giữa các tRNA khác nhau. Hình 1.23 cho thấy một ví dụ liên quan đến hoạt động của hai endonuclease và một exonuclease. Tiền tRNA được RNaseP, một endonuclease, nhận biết và cắt. Hoạt động của RNaseP tạo ra đúng đầu 5' của tRNA

trưởng thành. Một endonuclease khác sẽ phân cắt tRNA tiền thân để loại bỏ đoạn 170 nucleotide từ đầu 3'. Tiếp theo, một exonuclease, được gọi là RNaseD, liên kết với đầu 3' và phân cắt RNA theo hướng 3' → 5'. Tất cả các tRNA đều có trình tự CCA ở đầu 3' của chúng. Trong một số tRNA, trình tự CCA được mã hóa trong gene tRNA. Tuy nhiên, đối với hầu hết các tRNA ở sinh vật nhân sơ và sinh vật nhân chuẩn, trình tự CCA được thêm vào tRNA bởi một enzyme gọi là tRNA **nucleotidyltransferase**. Cuối cùng, một số base nhất định trong phân tử tRNA có thể bị biến đổi cộng hóa trị để thay đổi cấu trúc của chúng.



Hình 1.23. Quá trình chế biến phân tử tRNA tiền thân ở *E. coli*. *RnaseP* là một endonuclease cắt để tạo đầu 5' của tRNA trưởng thành. Ở đầu 3', một endonuclease khác cắt để loại bỏ một đoạn 170 nucleotid. Tiếp theo, endonuclease *RNaseD* loại bỏ 9 nucleotide ở đầu 3'. Với hầu hết tRNA, trình tự CCA ở đầu 3' được thêm vào bởi tRNA nucleotidyltransferase. Ngoài những bước này, một số base cũng được biến đổi. Một kiểu xử lý tRNA tiền thân tương tự xảy ra ở sinh vật nhân chuẩn.

Khi các nhà nghiên cứu nghiên cứu quá trình chế biến tRNA, họ đã phát hiện ra một số đặc điểm rất bất thường và thú vị, thay đổi cách các nhà sinh học nhìn nhận về hoạt động của chất xúc tác. RNaseP đã được phát hiện là một chất xúc tác chứa cả RNA và các tiểu đơn vị protein. Năm 1983, Sidney Altman và các đồng nghiệp đã phát hiện ra rằng phần RNA của RNaseP, có khả năng xúc tác để phân cắt tRNA tiền thân. RNaseP là một ví dụ về ribozyme, một phân tử RNA có hoạt tính xúc tác. Trước khi nghiên cứu RNaseP và việc xác định RNA tự cắt nối, các nhà hóa sinh đã tin chắc rằng chỉ có các phân tử protein mới có thể hoạt động như chất xúc tác sinh học.

3.3.3. Cắt nối RNA – các cơ chế cắt nối khác nhau loại bỏ các intron

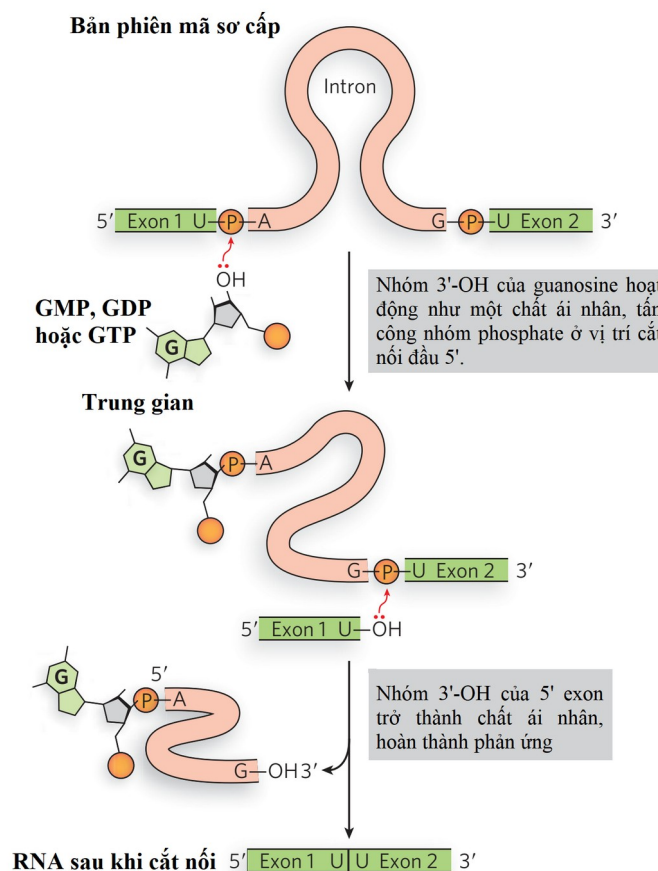
Các intron của eukaryote lần đầu tiên được phát hiện bằng cách so sánh trình tự base của các gene virus và các bản sao mRNA của chúng trong quá trình nhiễm virus adenovirus vào tế bào động vật có vú. Nghiên cứu này được thực hiện vào năm 1977 bởi hai nhóm do Philip Sharp và Richard Roberts đứng đầu. Một số nhóm nghiên cứu khác,

bao gồm cả nhóm của Pierre Chambon, Bert O'Malley và Philip Leder sau đó đã xác định được intron trong các gene mã hóa protein của eukaryote.

Kể từ khi phát hiện ban đầu về intron, nhiều nhóm nghiên cứu đã chỉ ra rằng hầu hết các gene mã hóa protein ở sinh vật nhân thực phức tạp đều chứa một hoặc nhiều intron. Ít phổ biến hơn, các intron có thể xuất hiện trong các gene tRNA và rRNA. Ở cấp độ phân tử, các cơ chế cắt nối RNA khác nhau đã được xác định.

3.3.3.1. Cơ chế tự cắt nối (self-splicing) ở intron nhóm I và II

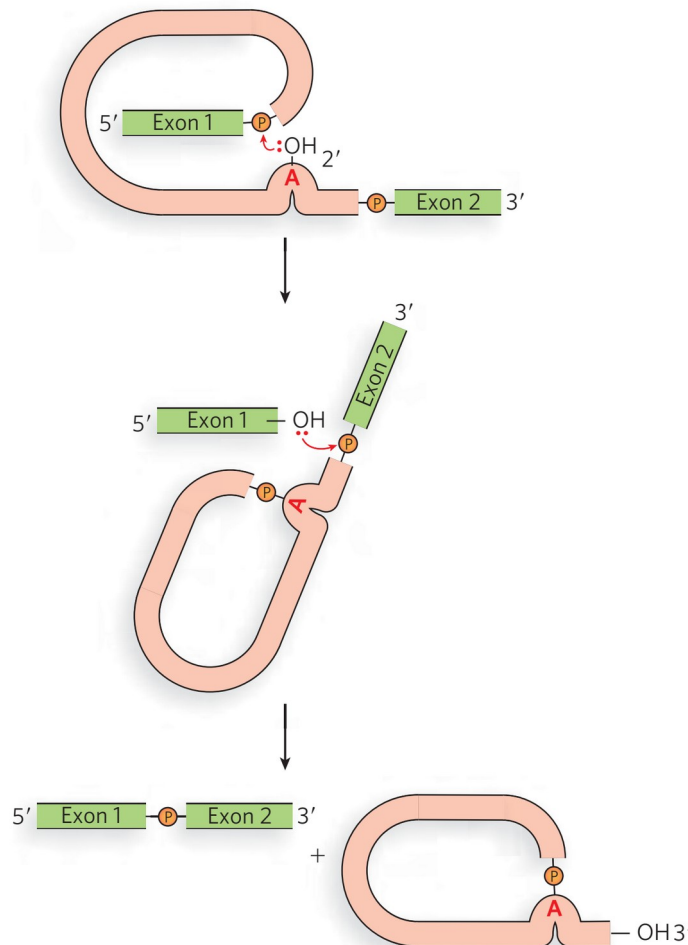
Sự cắt nối của các **intron nhóm I** và **nhóm II** xảy ra thông qua quá trình tự cắt nối – cắt nối mà không cần đến sự hỗ trợ của các chất xúc tác khác. Thay vào đó, RNA có chức năng như ribozyme của chính nó. Nhóm I và II khác nhau ở cách loại bỏ các intron và ghép nối các exon. Các intron nhóm I xuất hiện trong rRNA của *Tetrahymena* (một nguyên sinh vật) đã được nghiên cứu rộng rãi bởi Thomas Cech và các đồng nghiệp. Ở sinh vật này, quá trình nối liên quan đến sự liên kết của một guanosine vào vị trí liên kết guanosine bên trong intron (Hình 1.24). Guanosine này phá vỡ liên kết giữa exon đầu tiên và intron. Guanosine trở nên gắn vào đầu 5' của intron. Nhóm 3'-OH của exon 1 sau đó phá vỡ liên kết bên cạnh một nucleotide khác (trong ví dụ này là guanine, G) nằm ở ranh giới giữa phần cuối của intron và exon 2; exon 1 hình thành liên kết cộng hóa trị với đầu 5' của exon 2. Intron của RNA sau đó bị phân hủy. Trong ví dụ này, phân tử RNA có chức năng như ribozyme của chính nó, vì nó tự liên kết với nhau mà không cần đến sự hỗ trợ của protein xúc tác. Các intron nhóm I cũng được tìm thấy ở một vài gene mã hóa rRNA, mRNA, và các tRNA trong nhân, ti thể và lục lạp.



Hình 1.24. Cơ chế tự cắt nối của các intron nhóm I. Nhóm I sử dụng guanosine ngoại sinh (của một nucleotide guanine) để bắt đầu hai bước chuyển hóa trong quá trình tự

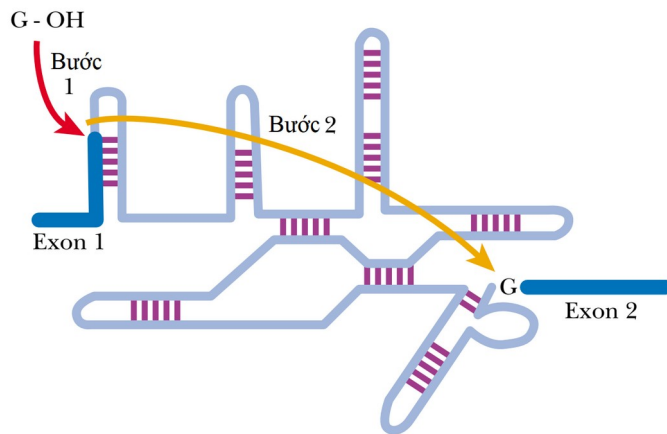
nối. Intron nhóm I được giải phóng dưới dạng cấu trúc thẳng (không tạo dạng vòng, hoặc thông lọng).

Với các intron nhóm II, tìm thấy trong bào quan (ti thể, lục lạp) của nấm, tảo, thực vật và một số trường hợp ở vi khuẩn, cũng là các intron tự cắt nối, ngoại trừ nhóm 2'-OH trên ribose trong nucleotide adenine (A) đã có trong intron bắt đầu quá trình xúc tác (Hình 1.25) hay nói cách khác nucleotide "tấn công" vào vị trí cắt là một adenine (A) có trong phân tử tiền RNA chứ không phải là nucleotide tự do trong môi trường. Theo thực nghiệm, quá trình tự nối nhóm I và II có thể xảy ra trong ống nghiệm mà không cần bổ sung bất kỳ protein nào. Tuy nhiên, trong tế bào sống, các protein được gọi là **maturase** thường tăng cường tỷ lệ cắt nối các intron nhóm I và II.

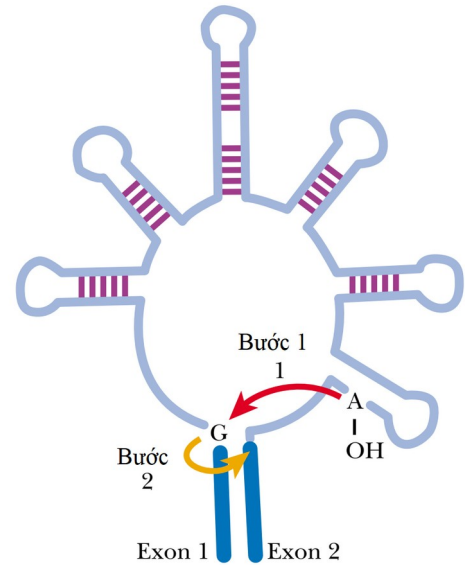


Hình 1.25. Cơ chế tự cắt nối của các intron nhóm II. Các intron nhóm II tự cắt nối bằng con đường tương tự sự cắt nối xúc tác bởi spliceosome, nucleotide "tấn công" vào vị trí cắt là một adenine (A) có trong phân tử tiền RNA chứ không phải là nucleotide tự do trong môi trường. Intron nhóm II được giải phóng dưới dạng cấu trúc thông lọng.

(a) Cơ chế tự cắt nối ở intron nhóm I



(b) Cơ chế tự cắt nối ở intron nhóm II



Hình 1.26. Phản ứng tự cắt nối của các intron nhóm I và II. Ở cả hai nhóm, các intron đều "gấp nếp" sao cho các đầu nối giữa các exon áp sát với nhau (ở đây chỉ minh họa sự kết cặp giữa các nucleotide qua liên kết hydrogen). **(a)** Ở intron nhóm I, một nucleotide G trong môi trường tấn công vào điểm cắt đầu 5' làm đứt intron khỏi exon; sau đó, đầu 3'-OH tự do của exon phản ứng với G tại điểm cắt đầu 3' (làm tách rời intron và nối hai exon). **(b)** Ở intron nhóm II, phản ứng cắt về cơ bản giống intron nhóm I, trừ đặc điểm là phản ứng cắt đầu tiên được thực hiện bởi A có trong chính tiền RNA.

3.3.3.2. Cắt nối tiền mRNA (pre-mRNA) bởi hoạt động của spliceosome

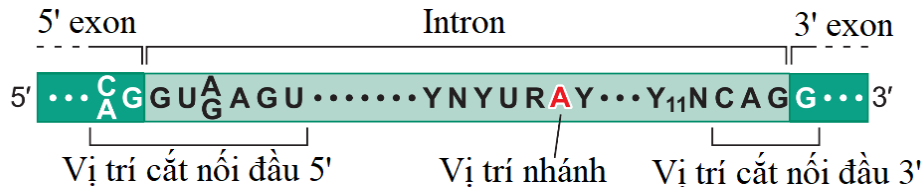
Ở sinh vật nhân thực, quá trình phiên mã của các gene mã hóa protein tạo ra một bản phiên mã dài được gọi là tiền mRNA (pre-mRNA), được tạo ra trong nhân. Tiền mRNA này thường bị thay đổi bằng cách nối và các sửa đổi khác trước khi đi ra khỏi nhân. Không giống như các intron nhóm I và II, có thể trải qua quá trình tự nối, quá trình cắt nối mRNA đòi hỏi sự hỗ trợ của một phức hợp được gọi là **spliceosome**. Spliceosome cần thiết để nhận biết các ranh giới của intron và loại bỏ nó một cách hợp lý.

Khác với sự cắt nối các intron nhóm I và II tương đối không phổ biến, sự cắt nối tiền mRNA là một hiện tượng phổ biến ở các sinh vật nhân chuẩn phức tạp. Ở động vật có vú và thực vật có hoa, hầu hết các gene mã hóa protein có ít nhất một intron có thể nằm ở bất kỳ vị trí nào trong gene. Một gene trung bình của con người có khoảng 8 intron. Trong một số trường hợp, một gene đơn lẻ có thể có một số lượng lớn các intron. Một ví dụ nổi bật là gene dystrophin của con người, khi bị đột biến, gây ra chứng loạn dưỡng cơ Duchenne, có 79 exon được xen kẽ bởi 78 intron.

Spliceosome là một phức hợp lớn cắt nối tiền mRNA ở sinh vật nhân chuẩn. Nó bao gồm năm tiểu đơn vị (U1, U2, U4, U5 và U6) được gọi là snRNP (đọc là "snurps"). Mỗi snRNP chứa RNA nhân kích thước nhỏ (small nuclear RNA) dài từ 100 – 300 nucleotide và một tập hợp các protein. Trong quá trình cắt nối, các tiểu đơn vị của một spliceosome thực hiện một số chức năng. Đầu tiên, các tiểu đơn vị của spliceosome liên kết với một trình tự intron và nhận biết chính xác ranh giới intron-exon. Ngoài ra, spliceosome phải giữ pre-mRNA ở cấu hình chính xác để đảm bảo sự liên kết của các exon với nhau. Và

cuối cùng, spliceosome xúc tác các phản ứng hóa học khiến intron bị loại bỏ và các exon được liên kết cộng hóa trị lại với nhau.

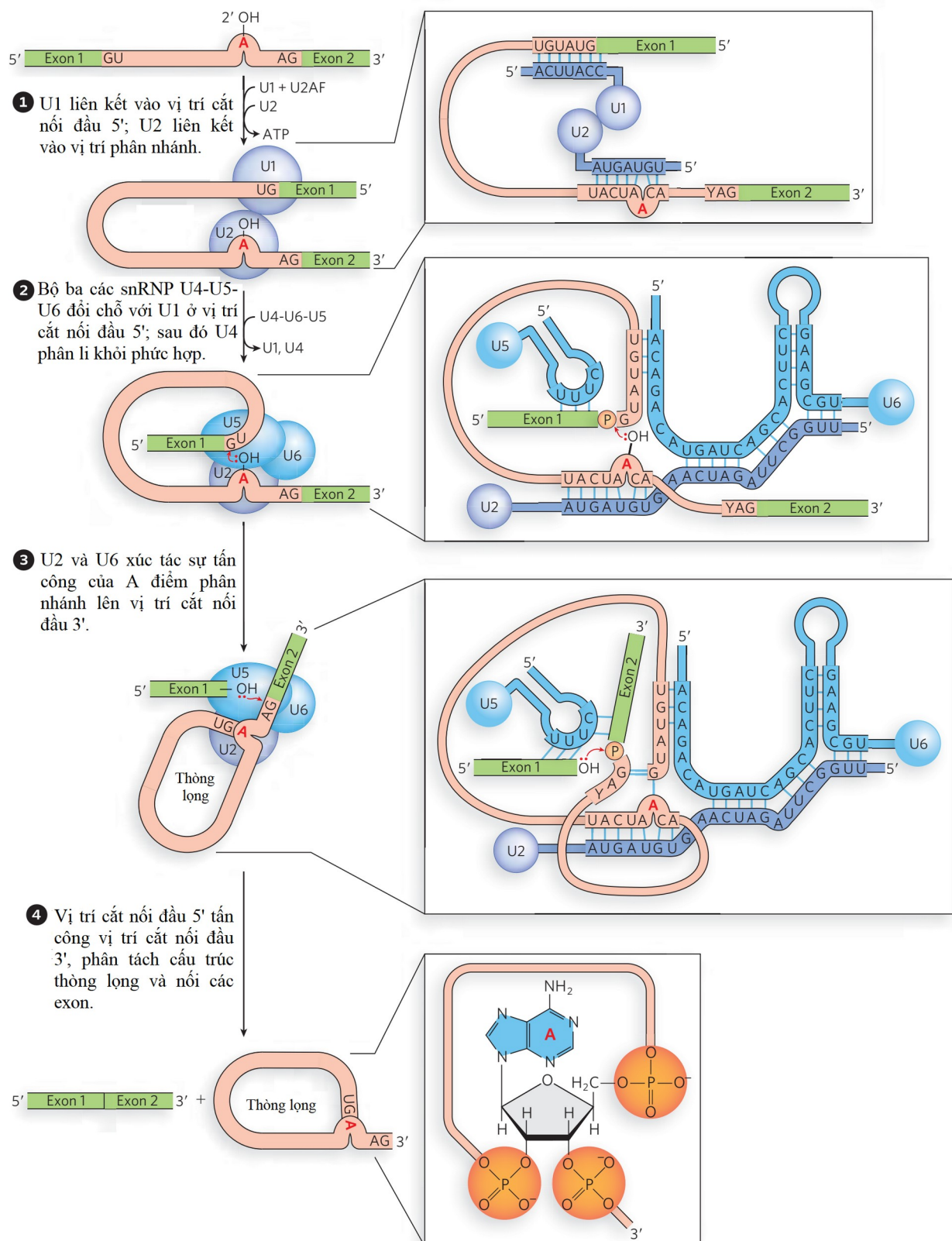
RNA intron được xác định bởi các trình tự đặc thù bên trong intron và ở ranh giới intron-exon. Có ba trình tự liên ứng nằm trên intron đóng vai trò quan trọng trong quá trình cắt nối là: vị trí cắt nối đầu 5' (chỗ tiếp giáp exon-intron, vị trí cho), vị trí phân nhánh là một trình tự giàu các pyrimidine bao quanh một nucleotide adenine ở gần đầu 3', và vị trí cắt nối đầu 3' (chỗ tiếp giáp intron-exon, vị trí nhận).



Hình 1.27. Các trình tự tại vị trí ranh giới intron - exon. Các trình tự liên ứng cho các vị trí cắt nối 5' và 3' (vị trí cho và vị trí nhận), và còn có A được bảo tồn vị trí nhánh (Y = any pyrimidine). Trong các trường hợp khác của trình tự liên ứng, hai base thay thế được ưu tiên tương tự, các base đó đều được chỉ ra ở vị trí đó. Trong hình này, các trình tự liên ứng được mô tả với người.

Cơ chế phân tử của quá trình cắt nối tiền mRNA được mô tả trong hình 1.28, snRNP U1 liên kết với vị trí cắt nối đầu 5'. Vị trí phân nhánh cũng được nhận dạng bởi BBP (branch-point binding protein, protein gắn vị trí phân nhánh) và U2AF. U2 snRNP đến thay thế BBP (nhờ U2AF trợ giúp). Tiếp theo là sự liên kết của một bộ ba các snRNP: một dimer U4/U6 kết hợp với U5. Intron vòng ra ngoài và hai exon được đưa lại gần nhau hơn. Vị trí cắt nối đầu 5' sau đó bị cắt, và đầu 5' của intron trở nên cộng hóa trị với nhóm 2'-OH của một nucleotide adenin đặc hiệu ở vị trí nhánh. U1 và U4 được giải phóng. Trong bước cuối cùng, vị trí cắt nối đầu 3' được cắt, và sau đó các exon được gắn cộng hóa trị với nhau. Ba snRNP gồm U2, U5 và U6 vẫn được gắn vào intron ở dạng thông lọng. Cuối cùng, intron bị phân hủy và các snRNP được sử dụng lại để nối các tiền mRNA khác.

Bằng chứng rằng các phân tử snRNA nhất định trong spliceosome đóng vai trò xúc tác trong việc loại bỏ các intron và ghép nối các exon. Nói cách khác, snRNA có thể hoạt động như ribozyme phân cắt RNA ở ranh giới exon-intron và nối các exon còn lại. Các nhà nghiên cứu đã suy đoán rằng các phân tử RNA trong U2 và U6 có thể có chức năng xúc tác này.



Hình 1.28. Sự lắp ráp spliceosome trên tiền mRNA liên quan đến sự bắt cặp base với các snRNA và cơ chế cắt bỏ intron – ghép nối exon thông qua spliceosome.

3.3.4. Một tiền mRNA (pre-mRNA) có thể được cắt bỏ intron và ghép nối exon theo những cách khác nhau

3.3.4.1. Cắt nối có lựa chọn hay cắt nối thay thế (Alternative splicing)

Khi lần đầu tiên được phát hiện, hiện tượng cắt nối giống như một quá trình lãng phí. Trong quá trình phiên mã, năng lượng được sử dụng để tổng hợp trình tự intron, sau đó

được loại bỏ bởi các spliceosome. Việc tạo ra và sau đó loại bỏ phần intron sử dụng một lượng năng lượng đáng kể. Quan sát này khiến nhiều nhà di truyền học tò mò, bởi vì chọn lọc tự nhiên có xu hướng loại bỏ các quá trình lãng phí. Do đó, các nhà nghiên cứu kỳ vọng phát hiện ra rằng sự cắt nối tiền mRNA có một hoặc nhiều vai trò sinh học quan trọng. Khi lần đầu tiên được phát hiện, hiện tượng cắt nối giống như một quá trình lãng phí. Trong quá trình phiên mã, năng lượng được sử dụng để tổng hợp trình tự intron, sau đó được loại bỏ bởi các spliceosome. Việc tạo ra và sau đó loại bỏ phần intron sử dụng một lượng năng lượng đáng kể. Quan sát này khiến nhiều nhà di truyền học tò mò, bởi vì chọn lọc tự nhiên có xu hướng loại bỏ các quá trình lãng phí. Do đó, các nhà nghiên cứu kỳ vọng phát hiện ra rằng sự cắt nối tiền mRNA có một hoặc nhiều vai trò sinh học quan trọng.

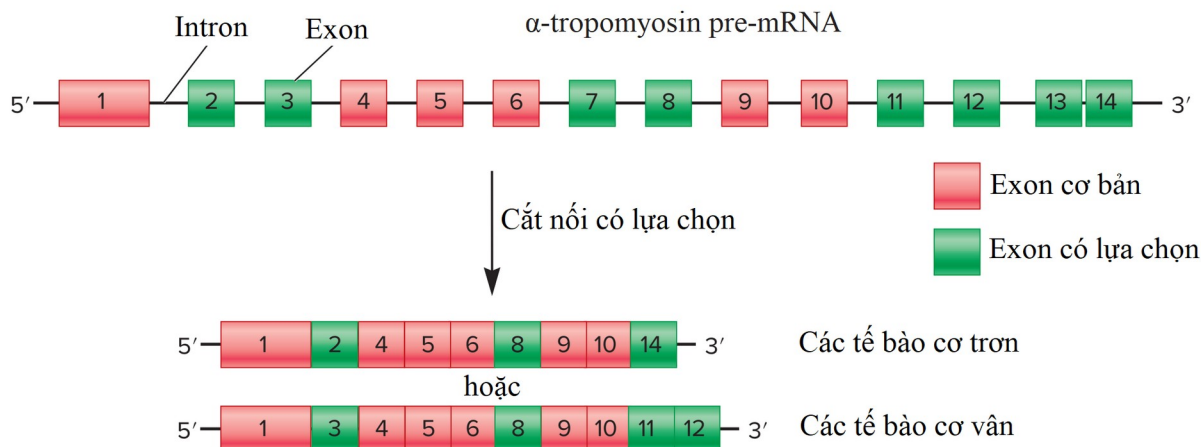
Trình tự các amino acid trong chuỗi polypeptide quyết định cấu trúc và chức năng của protein. Sự cắt nối có lựa chọn cho phép tạo ra hai hoặc nhiều polypeptide từ cùng một gene có sự khác biệt về trình tự amino acid của chúng, dẫn đến những thay đổi có thể có trong chức năng của chúng. Bởi vì việc ghép nối có lựa chọn cho phép hai hoặc nhiều trình tự polypeptide khác nhau có được từ một gene duy nhất, một số nhà di truyền học đã suy đoán rằng một lợi thế quan trọng của quá trình này là nó cho phép một sinh vật mang ít gene hơn trong bộ gene của nó.

Mức độ cắt nối và cắt nối có lựa chọn rất khác nhau giữa các loài khác nhau. Ví dụ: nấm men bánh mì (*S. cerevisiae*) chứa khoảng 6300 gene và khoảng 300 (tức là khoảng 5%) mã hóa tiền mRNA được cắt nối. Trong số này, chỉ một số ít được chứng minh là có cắt nối có lựa chọn. Do đó, ở sinh vật nhân chuẩn đơn bào này, sự cắt nối có lựa chọn không phải là cơ chế chính để tạo ra sự đa dạng của protein. Trong khi đó, các sinh vật đa bào phức tạp dường như dựa vào sự ghép nối có lựa chọn ở một mức độ lớn. Con người có khoảng 22.000 gene mã hóa protein khác nhau và hầu hết trong số này đều chứa một hoặc nhiều intron. Các ước tính gần đây cho thấy rằng khoảng 70% tất cả các tiền mRNA của con người được cắt nối có lựa chọn. Mức độ cắt nối có lựa chọn này cung cấp tiềm năng lớn hơn nhiều cho các tế bào người để tạo ra sự đa dạng protein.

Hình 1.29 là một ví dụ về cách nối thay thế cho một gene mã hóa một protein được gọi là α -tropomyosin, có chức năng điều hòa sự co cơ. Nó nằm dọc theo các sợi mảnh được tìm thấy trong các tế bào cơ trơn, chẳng hạn như trong tử cung và ruột non, và trong các tế bào cơ vân được tìm thấy trong cơ tim và cơ xương. Protein α -tropomyosin cũng được tổng hợp trong nhiều loại tế bào không phải tế bào hình nón nhưng với số lượng thấp hơn. Trong một sinh vật đa bào, các loại tế bào khác nhau phải điều chỉnh khả năng co bóp của chúng theo những cách khác nhau. Một cách có thể thực hiện được sự thay đổi chức năng này là bằng cách sản xuất các dạng α -tropomyosin khác nhau.

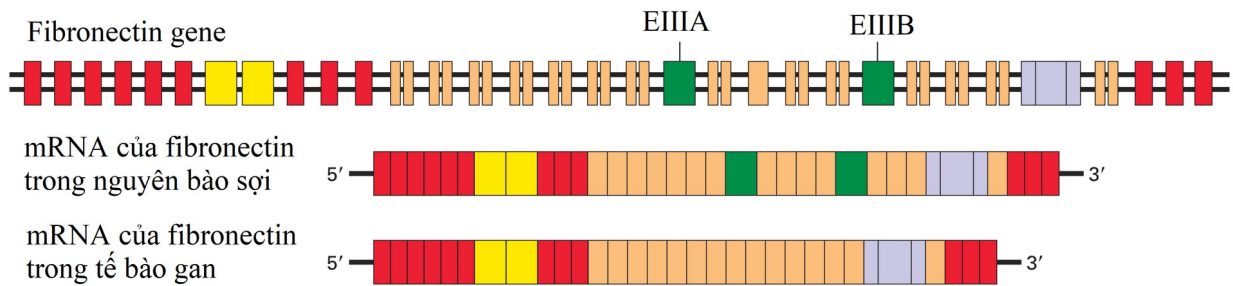
Cấu trúc intron-exon của tiền mRNA α -tropomyosin của chuột và hai cách thay thế để tiền mRNA có thể được nối được mô tả trong hình 1.29. Tiền mRNA chứa 14 exon, 6 trong số đó là các **exon cơ định** hay **exon cơ bản** (màu đỏ, **constitutive exon**), luôn được tìm thấy trong mRNA trưởng thành từ tất cả các loại tế bào. Có lẽ, các exon cơ định mã hóa các đoạn polypeptide của protein α -tropomyosin cần thiết cho cấu trúc và chức năng chung của nó. mRNA trưởng thành cũng chứa các exon khác, được gọi là **exon thay thế** hay **exon có lựa chọn** (màu xanh lá cây, **alternative exon**), thay đổi từ

loại tế bào này sang loại tế bào khác. Các trình tự polypeptide được mã hóa bởi các exon thay thế có thể thay đổi một cách tinh vi chức năng của α -tropomyosin để đáp ứng nhu cầu của loại tế bào mà nó được tìm thấy. Exon 2 mã hóa một đoạn của protein α -tropomyosin làm thay đổi chức năng của nó để phù hợp với tế bào cơ trơn. Trong khi đó, mRNA α -tropomyosin được tìm thấy trong tế bào cơ vân không bao gồm exon 2. Thay vào đó, mRNA này chứa exon 3, phù hợp hơn với loại tế bào đó.



Hình 1.29. Các cách khác nhau mà tiền mRNA α -tropomyosin của chuột có thể được nối. Phần trên của hình này mô tả cấu trúc của tiền mRNA α -tropomyosin của chuột. Các Exon được hiển thị dưới dạng các hộp màu và các intron là các đường màu đen kết nối. Phần dưới của hình mô tả mRNA trưởng thành của các tế bào cơ trơn và cơ vân. Các exon cơ bản (màu đỏ) được tìm thấy trong mRNA α -tropomyosin trưởng thành ở tất cả các loại tế bào. Các exon có lựa chọn cũng được tìm thấy trong mRNA nhưng chúng thay đổi từ loại tế bào này sang loại tế bào khác. Mặc dù exon 8 trong hình này được tìm thấy trong cả mRNA trưởng thành, nó được coi là exon có lựa chọn vì nó không được tìm thấy trong mRNA trưởng thành cho α -tropomyosin ở một số loại tế bào khác.

Fibronectin, protein đa miền của động vật có vú, là một ví dụ rõ nét khác về cắt nối có lựa chọn. Fibronectin là protein bám dính dài được tiết vào không gian ngoại bào. Fibronectin giúp gắn các protein khác lại với nhau. Việc gắn protein nào và ở đâu phụ thuộc vào miền nào của fibronectin cắt ghép với nhau. Gene mã hóa fibronectin chứa nhiều exon, nhóm lại thành từng vùng tương ứng với các miền protein đặc hiệu. Nguyên bào sợi tạo ra mRNA fibronectin chứa hai exon EIIIA và EIIIB. Hai exon này mã hóa trình tự amino acid mang khả năng gắn chặt với các protein trong màng bào tương của nguyên bào sợi. Kết quả là, đồng phân fibronectin giúp nguyên bào sợi dính với chất nền ngoại bào. Sản phẩm phiên mã bậc một của fibronectin (tiền mRNA) trong tế bào gan cắt nối có lựa chọn tạo ra protein không mang hai exon EIIIA và EIIIB. Vì vậy fibronectin do túi mật tiết vào máu không dính chặt với nguyên bào sợi hoặc hầu hết các loại tế bào khác. Điều này cho phép chúng lưu thông trong tuần hoàn. Tuy nhiên khi đông máu, miền gắn fibrin trong fibronectin của tế bào gan gắn với fibrin, sau đó tương tác với integrin trên màng của tiểu cầu đi qua, qua đó làm cục máu đông to thêm do chứa thêm tiểu cầu.



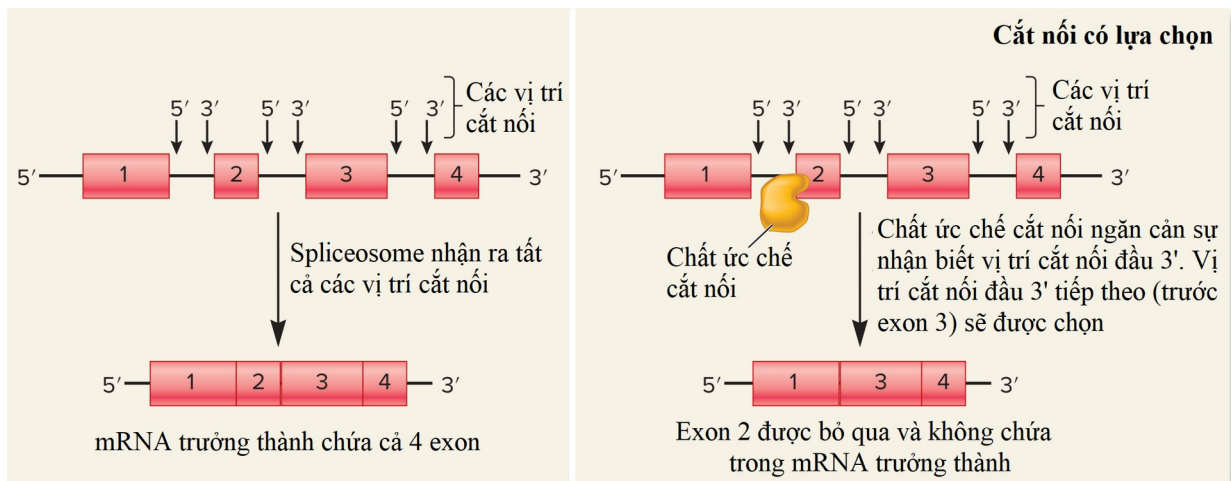
Hình 1.30. Cắt nối có lựa chọn ở tiền mRNA mã hóa fibronectin. Gene mã hóa fibronectin ~75 kb chứa nhiều exon, fibronectin cắt nối có lựa chọn tùy thuộc từng loại tế bào. Exon EIIIA và EIIIB (xanh lá cây) mã hóa các miền gắn với những protein đặc hiệu trên bề mặt nguyên bào sợi. mRNA của fibronectin trong nguyên bào sợi chứa cả EIIIA và EIIIB, trong khi những exon này bị loại bỏ khỏi mRNA fibronectin ở tế bào gan. Sơ đồ này không thể hiện intron (đường màu đen) theo đúng tỷ lệ; hầu hết trong số chúng lớn hơn nhiều so với bất kì exon nào.

Sự cắt nối có lựa chọn không phải là một sự kiện ngẫu nhiên. Thay vào đó, mô hình cắt nối cụ thể được quy định trong các tế bào nhất định nào. Cơ chế phân tử để điều chỉnh quá trình nối thay thế liên quan đến các protein được gọi là các **yếu tố cắt nối**, đóng vai trò quan trọng trong việc lựa chọn các vị trí cắt nối cụ thể. Các **protein SR** là một loại yếu tố cắt nối. Chúng chứa một miền ở đầu carboxyl của chúng rất giàu serines (S) và arginines (R) và tham gia vào quá trình nhận biết protein-protein. Chúng cũng chứa một miền liên kết RNA ở đầu tận amino.

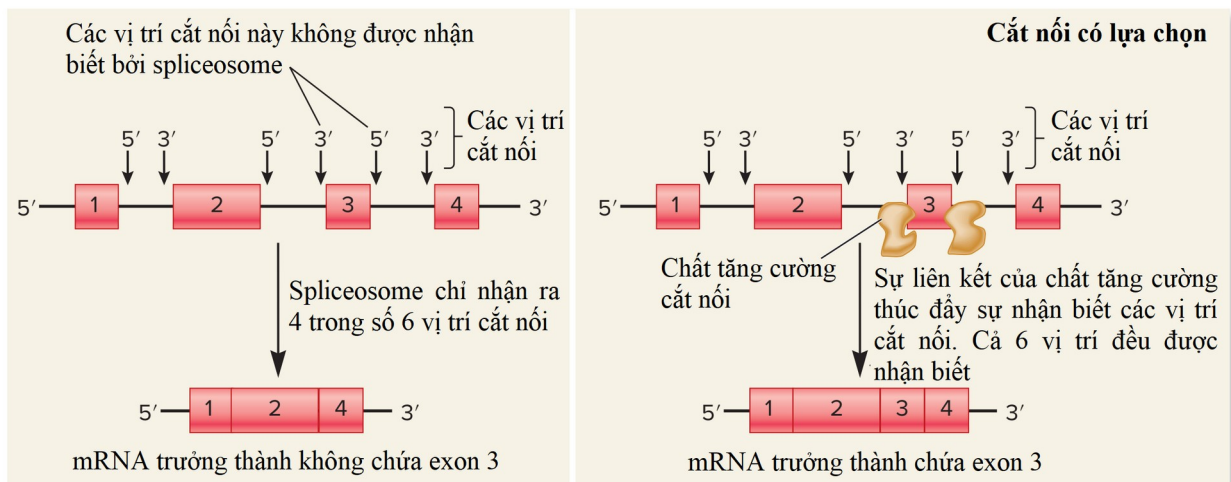
Các thành phần của một spliceosome nhận ra các vị trí cắt nối đầu 5' và 3' và sau đó loại bỏ intron xen vào. Chức năng của các yếu tố cắt nối là điều chỉnh khả năng của một spliceosome để chọn các vị trí cắt nối đầu 5' và 3'. Điều này có thể xảy ra theo hai cách:

- Một số yếu tố cắt nối hoạt động như **chất ức chế** ngăn cản khả năng nhận biết vị trí cắt nối. Trong hình 1.31a, một chất ức chế cắt nối liên kết với một vị trí cắt nối đầu 3' và ngăn không cho spliceosome nhận ra vị trí đó. Thay vào đó, spliceosome liên kết với vị trí cắt nối đầu 3' phù hợp tiếp theo. Cơ chế ức chế cắt nối làm cho exon 2 được cắt bỏ ra khỏi tiền mRNA và không có mặt trong mRNA trưởng thành, một sự kiện được gọi là **bỏ qua exon (exon skipping)**.

- Ngoài ra, các yếu tố nối khác giúp tăng cường khả năng nhận ra các vị trí cắt nối cụ thể của spliceosome. Trong hình 1.31b, các **chất tăng cường** cắt nối liên kết với các vị trí cắt nối 3' và 5' nằm hai bên exon 3, dẫn đến việc đưa exon 3 vào mRNA trưởng thành.



(a) Chất ức chế cắt nối



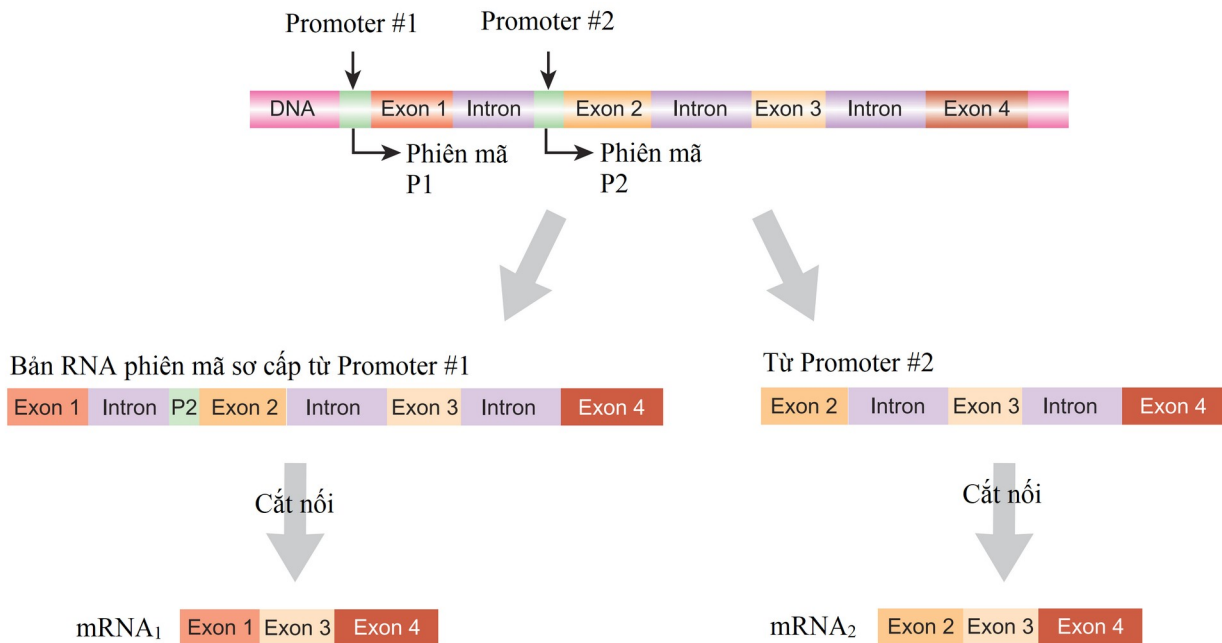
(b) Chất tăng cường cắt nối

Hình 1.31. Vai trò của các yếu tố cắt nối trong quá trình cắt nối có lựa chọn. (a) Các yếu tố cắt nối có thể hoạt động như chất ức chế để ngăn chặn việc nhận biết các vị trí cắt nối. Trong ví dụ này, sự hiện diện của một chất ức chế cắt nối khiến exon 2 bị bỏ qua và do đó mRNA trưởng thành không chứa exon này. **(b)** Các yếu tố nối khác có thể thúc đẩy khả năng nhận biết các vị trí cắt nối. Trong ví dụ này, các chất tăng cường cắt nối thúc đẩy việc nhận biết các vị trí nằm hai đầu exon 3, do đó làm cho nó bao gồm trong mRNA.

Sự cắt nối có lựa chọn xảy ra bởi vì mỗi loại tế bào có mức độ đặc trưng của nhiều loại yếu tố cắt nối. Hơn nữa, các yếu tố cắt nối có thể được điều chỉnh bởi sự liên kết của các phân tử tác động nhỏ, tương tác protein-protein và các biến đổi cộng hóa trị. Nhìn chung, sự khác biệt về thành phần của các yếu tố cắt nối và điều hòa hoạt động của chúng tạo cơ sở cho các kết quả của quá trình cắt nối có lựa chọn.

3.3.4.2. Chọn các promoter khác nhau

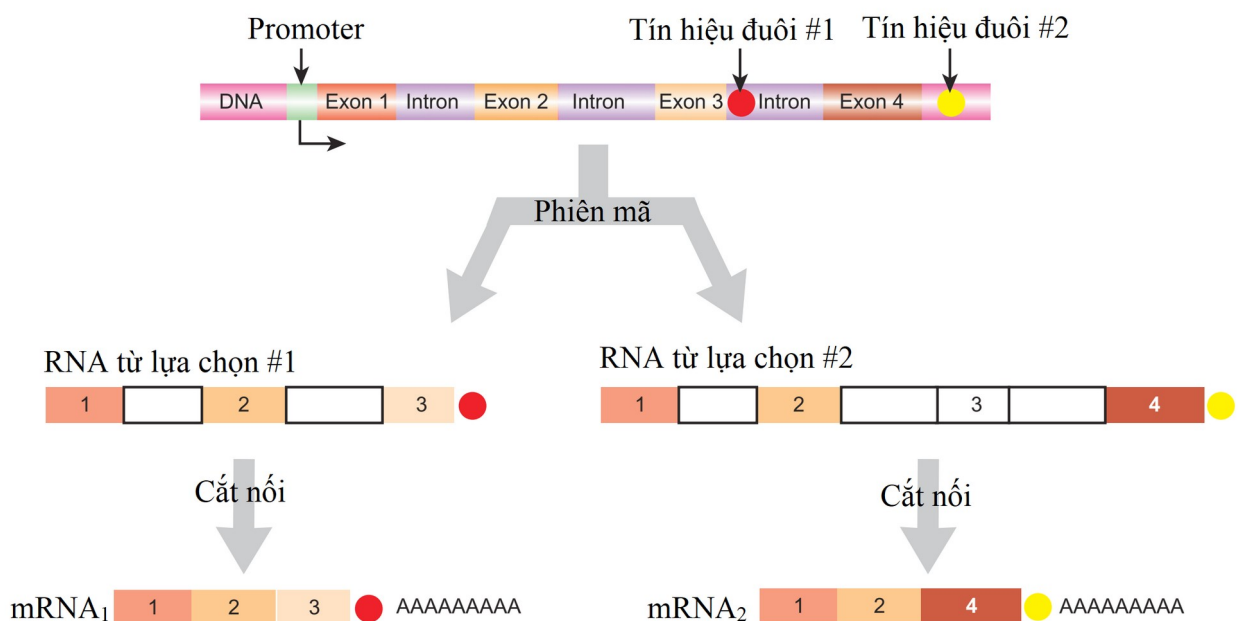
Theo cơ chế này, các tế bào khác nhau có các yếu tố phiên mã khác nhau. Các yếu tố phiên mã này nhận ra các promoter khác nhau (mặc dù chúng dùng chung một phần trình tự mã hóa).



Hình 1.32: Chọn các promoter khác nhau. DNA chứa hai trình promoter có thể tạo ra hai bản phiên mã sơ cấp. Nếu promoter #1 được sử dụng, thì đoạn chứa promoter #2 và exon #2 sẽ được tách ra. Nếu promoter #2 được sử dụng, thì exon #1 không có trong mRNA.

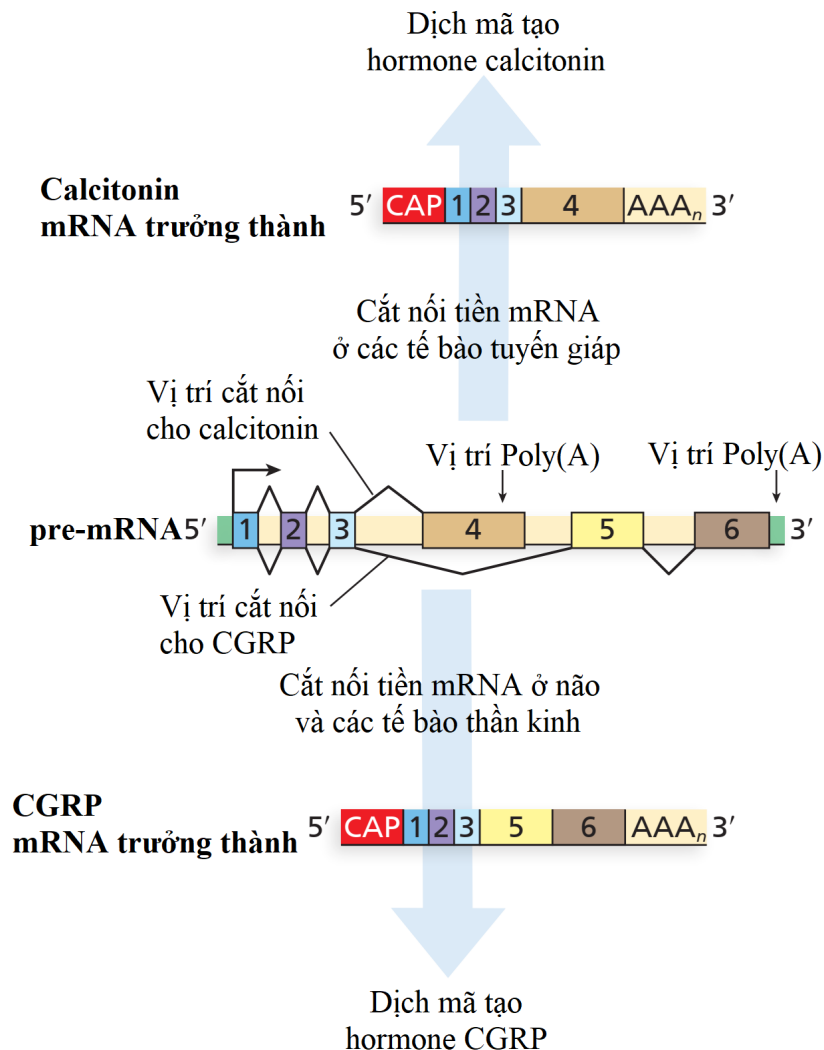
3.3.4.3. Chọn vị trí gắn đuôi poly(A) khác nhau

Theo cơ chế này, các tế bào khác nhau chọn vị trí gắn đuôi poly(A) khác nhau. Một số tế bào chọn tín hiệu gắn đuôi poly(A) sớm dẫn đến việc phân tử mRNA trưởng thành thiếu một số exon ở xa. Cơ chế này được dùng để sản sinh các kháng thể cùng nhận biết các phân tử kháng nguyên lạ xâm nhập vào tế bào, nhưng có phần đuôi khác nhau. Một loại được tiết vào máu, còn loại khác được gắn lên màng tế bào.



Hình 1.33: Chọn vị trí gắn đuôi poly(A) khác nhau. Trong DNA hiển thị có hai tín hiệu đuôi. Việc sử dụng tín hiệu đuôi đầu tiên dẫn đến một mRNA bao gồm exon 1, 2 và 3. Việc sử dụng tín hiệu đuôi #2 dẫn đến một mRNA chứa exon 1, 2 và 4.

Ở chuột, cả hai cơ chế - cắt nối có lựa chọn và lựa chọn vị trí gắn đuôi poly(A) khi một bản sao tiền mRNA được xử lý khác nhau để tạo ra hai loại hormone khác nhau: hormone điều hòa calcium là calcitonin ở tuyến giáp và peptide liên quan đến gene calcitonin (CGRP, calcitonin-gene-related peptide) trong não.



Hình 1.34. Cắt nối có lựa chọn và chọn vị trí gắn đuôi poly(A) ở tiền mRNA từ gene calcitonin của chuột. Calcitonin và peptide liên quan đến gene calcitonin (CGRP) có trình tự protein khác nhau nhưng được mã hóa bởi cùng một gene. Tiền mRNA chứa hai vị trí poly(A): một vị trí được sử dụng trong tuyến giáp và vị trí còn lại được sử dụng trong não. Trong tuyến giáp, exon 4 được giữ lại; trong não, exon 4 bị loại bỏ và exon 5 và 6 được giữ lại. Các mRNA thu được mã hóa các polypeptide khác nhau được xử lý thêm để tạo ra các sản phẩm hormone cuối cùng: calcitonin trong tuyến giáp và CGRP trong não.

3.3.4.4. Cắt nối chuyển chỗ intron (Trans-splicing)

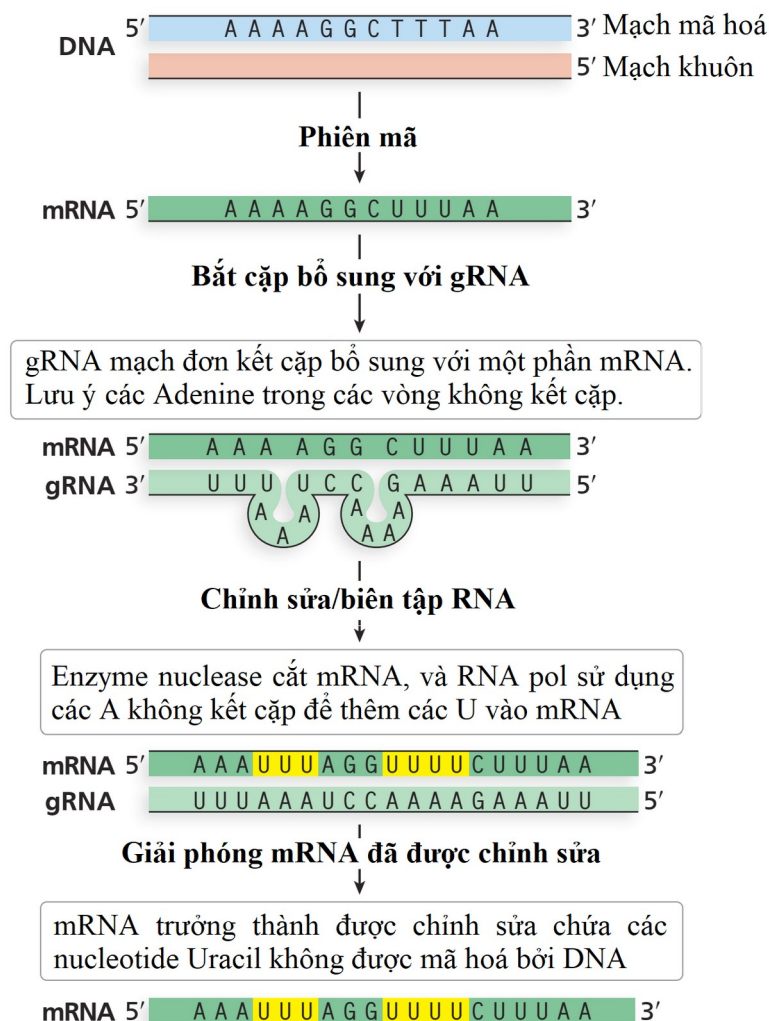
Trong cơ chế này, các exon có nguồn gốc từ các tiền mRNA khác nhau được cắt ra và tổ hợp lại trong phân tử RNA hoàn thiện. **Trùng mủi khoan (Trypanosomes)** là sinh vật nhân thật đơn bào gây bệnh "ốm ngủ" và một số bệnh nhiệt đới khác. Loài động vật nguyên sinh này có thể vượt qua sự giám sát của hệ miễn dịch của động vật có xương sống bằng cách thay đổi protein bề mặt tế bào dựa trên cơ chế "tráo trộn vị trí" các phần của gene. Ngoài ra, nó còn có thể tổ hợp các exon từ nhiều gene khác nhau. Mặc dù hiện

tượng này chưa từng gặp ở động vật có xương sống, nhưng cơ chế cắt chuyển chỗ intron cũng đã được tìm thấy ở giun tròn và trong hệ gene lục lạp ở thực vật.

3.3.5. *Chỉnh sửa trình tự RNA (RNA editing)*

Thuật ngữ **chỉnh sửa RNA** hay **biên tập RNA** đề cập đến quá trình tạo ra sự thay đổi trong trình tự nucleotide của phân tử RNA liên quan đến việc bổ sung hoặc xóa bỏ các nucleotide cụ thể hoặc chuyển đổi loại base này sang loại base khác, chẳng hạn như cytosine thành uracil. Trong trường hợp của mRNA, việc chỉnh sửa có thể có nhiều tác dụng khác nhau, chẳng hạn như tạo ra các codon mở đầu, tạo ra các codon kết thúc và thay đổi trình tự mã hóa cho một polypeptide.

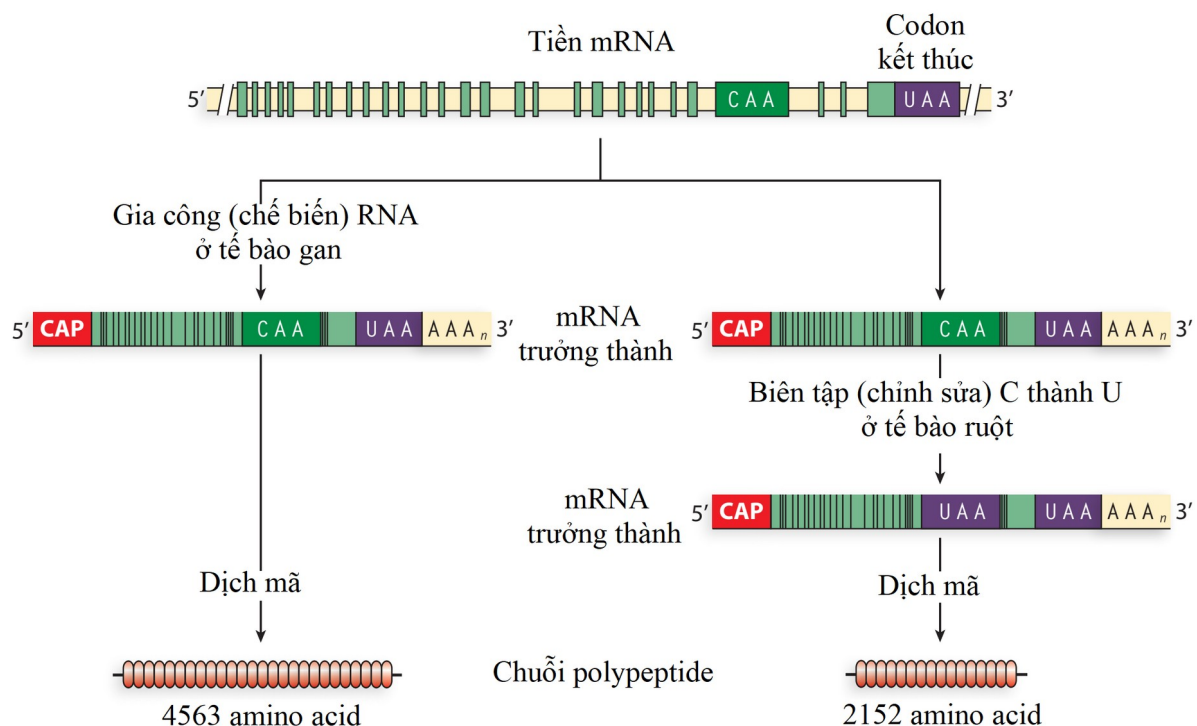
Hiện tượng chỉnh sửa RNA lần đầu tiên được phát hiện ở trùng mủi khoan (trypanosomes) – động vật nguyên sinh gây bệnh ngủ. Cũng như việc phát hiện ra sự cắt nối RNA, phát hiện ban đầu về chỉnh sửa RNA đã vấp phải sự hoài nghi lớn. Tuy nhiên, kể từ thời điểm đó, quá trình chỉnh sửa RNA đã được chứng minh là xảy ra ở nhiều sinh vật khác nhau và theo nhiều cách khác nhau, mặc dù ý nghĩa chức năng của nó chỉ mới xuất hiện từ từ (Bảng 1.5). Trong trường hợp cụ thể của trypanosomes, quá trình chỉnh sửa bao gồm việc thêm hoặc cắt bỏ một hoặc nhiều nucleotide uracil. Sự biên tập RNA ở một số gene của trùng mủi khoan là cần thiết bởi trình tự gene trên DNA không phải là các khung đọc mở (ORF) chính xác. Khi việc biên tập mRNA diễn ra sai sót, protein được tổng hợp "lệch" khung đọc và mất chức năng.



Hình 1.35. RNA hướng dẫn (gRNA) trực tiếp chỉnh sửa/biên tập mRNA

Các mRNA từ một số gene nhân ở sinh vật nhân chuẩn, một số gene ty thể thực vật và một số gene ty thể của trypanosome được chỉnh sửa bởi một RNA chuyên biệt gọi là **RNA hướng dẫn** hay **RNA dẫn đường (gRNA: guide RNA)**. Một phần của gRNA chứa trình tự bổ sung với vùng của mRNA mà nó chỉnh sửa. Với sự hỗ trợ của một phức hợp protein, một phần của gRNA bắt cặp với các nucleotide bổ sung của mRNA trước khi được chỉnh sửa và hoạt động như một khuôn mẫu để định hướng việc chèn (và đôi khi là xóa) uracil vào mRNA tại các vị trí đối diện với A trên gRNA (Hình 1.35). RNA hướng dẫn giải phóng mRNA đã chỉnh sửa sau khi chỉnh sửa xong. Protein được dịch mã từ mRNA đã chỉnh sửa có thể khác với protein được tạo ra từ bản phiên mã chưa chỉnh sửa.

Một ví dụ về biên tập RNA ở người là gene mã hóa apolipoprotein B. Protein này gồm 4563 amino acid, là một trong các protein lớn nhất đã từng biết đến nay. Apolipoprotein B có kích thước đầy đủ, kí hiệu apoB100, được tổng hợp ở tế bào gan trước khi được tiết vào máu. ApoB100 có vai trò "lắp ráp" các hạt lipoprotein mật độ thấp (LDL) và rất thấp (VLDL) làm nhiệm vụ vận chuyển lipid (kể cả cholesterol) trong cơ thể. Một phiên bản protein khác của gene này, kí hiệu là apoB48, chỉ gồm 2153 amino acid được tiết ra từ các tế bào ruột. ApoB48 có vai trò hấp thụ chất béo từ thức ăn. So với apoB100, apoB48 thiếu một đoạn trình tự amino acid là vị trí liên kết với thụ thể của LDL (gồm các amino acid từ 3129 đến 3532). Kết quả là các chất béo từ thức ăn được apoB48 vận chuyển chủ yếu được hấp thu ở gan, trong khi apoB100 có vai trò phân phối cholesterol vào các mô ngoại vi mà ở đó tế bào của chúng có thụ thể LDL.



Hình 1.36. Sự biên tập bản phiên mã mRNA của gene apolipoprotein B ở người. Một gene giống hệt nhau chứa 29 exon được tìm thấy trong tất cả các tế bào của động vật có vú và mRNA giống nhau được phiên mã trong tất cả các mô. Một phần của trình tự RNA thông tin này bao gồm codon số 2153 có trình tự CAA và được dịch mã dưới dạng glutamine trong apolipoprotein B ở gan, một loại protein bao gồm 4563 amino

acid. Tuy nhiên, trong các tế bào ruột, quá trình chỉnh sửa RNA sẽ thay đổi cytosine trong codon 2153 thành uracil, chuyển codon thành UAA. Sự thay đổi từ C → U này được tạo ra bởi quá trình chỉnh sửa RNA tạo ra mã kết thúc để ngăn chặn quá trình dịch mã sau khi lắp ráp 2152 amino acid đầu tiên của apolipoprotein B ở ruột.

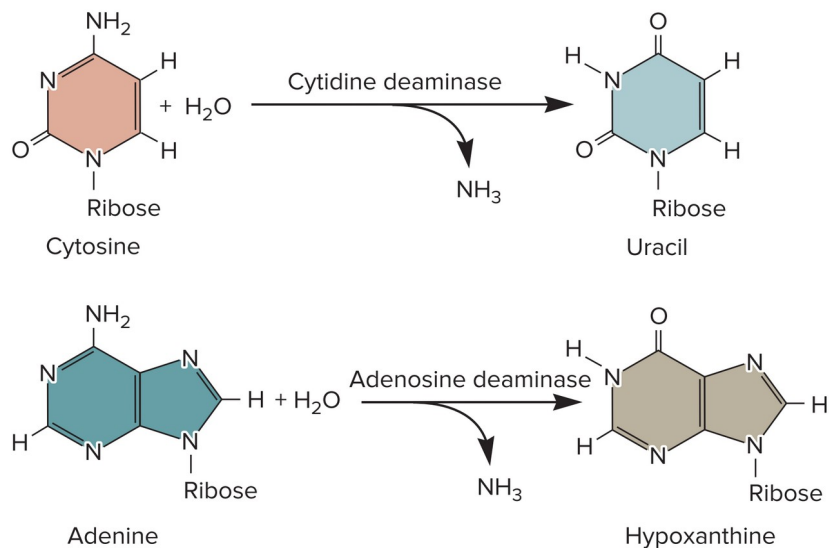
Hai phiên bản protein apoB48 và apoB100 đều được mã hóa bởi cùng một gene. Để tổng hợp được hai protein này, cần có sự biên tập lại RNA sau phiên mã. Sự biên tập RNA diễn ra như sau: bộ ba mã hóa Gln (tại vị trí 2154 của apoB100) tương ứng trên mRNA là CAA được biến đổi thành một mã bộ ba kết thúc UAA. Ở đây, việc biên tập được thực hiện bởi enzyme loại nhóm amin (-NH₂) của C dẫn đến việc biến đổi C → U. Quá trình loại amin cần một loạt protein hỗ trợ để đảm bảo phản ứng chỉ xảy ra đúng ở nucleotide cần thiết. Một điểm đáng chú ý trong biểu hiện gene apolipoprotein B là: mặc dù để tạo ra hai loại protein là apoB48 và apoB100 chỉ cần một gene duy nhất, nhưng để biên tập được RNA chính xác lại cần có nhiều protein hỗ trợ khác được mã hóa bởi nhiều gene khác. Như vậy, rõ ràng khi sử dụng hệ thống biên tập lại RNA tế bào cần bộ máy phức tạp và có lẽ cần chi phí năng lượng cao hơn nếu tế bào có hai gene mã hóa cho hai protein này. Tại sao chọn lọc lại duy trì cách thức "biên tập RNA" chứ không phải "hai gene" đến nay chưa rõ.

Bảng 1.5: Một số ví dụ về chỉnh sửa RNA

Sinh vật	Kiểu chỉnh sửa RNA	Tìm thấy ở
Trypanosomes (protozoa)	Thêm hoặc cắt bỏ một hoặc nhiều nucleotide U	Nhiều mRNA ti thể
Nấm nhầy	Thêm nucleotide C	Nhiều mRNA ti thể
Thực vật	Biến đổi C → U	Nhiều mRNA, tRNA, rRNA ti thể và lục lạp
Động vật	Biến đổi C → U	Apolipoprotein B mRNA và NFI mRNA, mã hóa cho protein ức chế khối u.
	Biến đổi A → I (inosine)	mRNA thụ thể glutamate và nhiều tRNA
<i>Drosophila</i>	Biến đổi A → I (inosine)	mRNA cho các protein kênh calcium và sodium.

Một cơ chế chỉnh sửa RNA phổ biến hơn liên quan đến sự thay đổi của loại base này sang loại base khác. Trong hình thức chỉnh sửa này, một base trong RNA bị khử amin, có nghĩa là một nhóm amin bị loại bỏ khỏi base. Khi cytosine bị khử amin, uracil được hình thành, và khi adenine bị khử amin, hypoxanthine được hình thành (Hình 1.37). Hypoxanthine được công nhận là guanine trong quá trình dịch mã.

Việc chỉnh sửa RNA liên quan đến sự thay thế C thành U và A thành H (hypoxanthine) phổ biến đến mức nào? Ở động vật không xương sống như *Drosophila*, các nhà nghiên cứu ước tính rằng 50-100 tiền mRNA được chỉnh sửa theo cách làm thay đổi trình tự mã hóa RNA. Ở động vật có vú, các tiền mRNA từ ít hơn 25 gene hiện được biết là đã được chỉnh sửa.



Hình 1.37. Chỉnh sửa RNA bằng cách khử amin. *Cytidine deaminase loại bỏ một nhóm amin khỏi cytosine, do đó tạo ra uracil. Adenosine deaminase loại bỏ một nhóm amin khỏi adenine để tạo ra hypoxanthine.*

II. DỊCH MÃ

Dịch mã là quá trình thông tin di truyền trong trình tự RNA được dùng để tổng hợp nên chuỗi amino acid tương ứng của phân tử protein. Đây có lẽ là quá trình tế bào sử dụng nhiều năng lượng nhất, đồng thời có tính "bảo thủ" nhất ở sinh vật (rất giống nhau ở mọi loài). Vi khuẩn trong pha tăng trưởng sử dụng đến 80% năng lượng tế bào và 50% lượng chất khô để tổng hợp protein. Để tổng hợp một phân tử protein cần sự phối hợp của trên 100 loại RNA và protein khác nhau. Dưới đây là một số nội dung cơ bản của dịch mã.

1. Bản chất của mã di truyền

1.1. Garrod đề xuất rằng một số gene mã hóa cho sự tạo thành các enzyme

Vào năm 1909, bác sĩ người Anh Archibald Garrod là người đầu tiên cho rằng các gene qui định kiểu hình thông qua các enzyme xúc tác các phản ứng diễn ra trong tế bào. Garrod dự đoán rằng các triệu chứng của một bệnh di truyền là kết quả của việc mất khả năng tổng hợp một enzyme nhất định nào đó ở người bệnh. Ông coi những bệnh như vậy là những **“rối loạn trao đổi chất bẩm sinh” (inborn error of metabolism)**. Garrod đã nêu ví dụ về một bệnh di truyền được gọi là alkapton niệu (alkaptonuria); ở những người mắc bệnh này, nước tiểu có màu đen do trong thành phần có alkapton là một chất chuyển màu sẫm khi tiếp xúc với không khí. Garrod cho rằng phần lớn mọi người đều có một enzyme giúp chuyển hóa alkapton (homogentisic acid), nhưng những người bị bệnh đã được di truyền gene mất khả năng tổng hợp enzyme này.

Garrod cũng có thể là một trong những người đầu tiên nhận ra các qui luật di truyền của Mendel có thể áp dụng cho người giống như với cây đậu Hà lan. Có thể nói nhận thức của Garrod đã đi trước thời đại, bởi vì các nghiên cứu được tiến hành sau đó hàng chục năm mới thực sự ủng hộ cho giả thiết của ông về việc mỗi gene điều khiển sự tổng hợp của một enzyme đặc thù. Các nhà hóa sinh học ngày càng tích lũy được nhiều bằng chứng cho thấy tế bào tiến hành tổng hợp và phân hủy phần lớn các chất hữu cơ thông

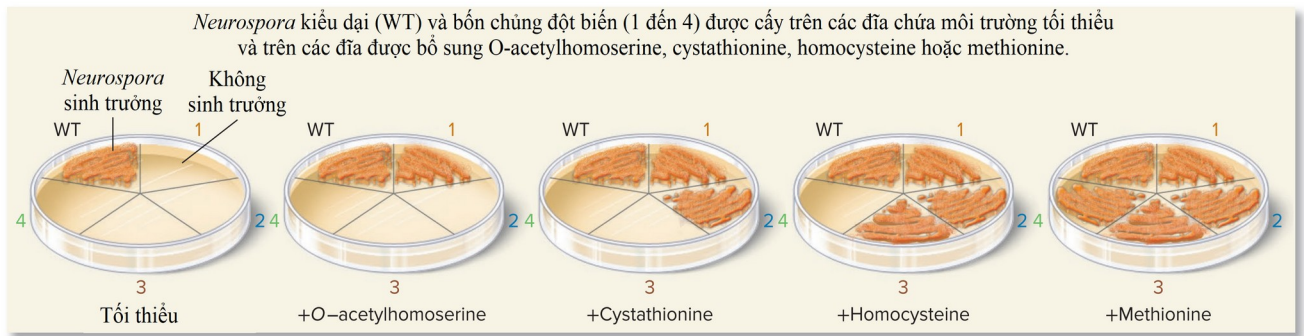
qua các con đường chuyển hóa, ở đó mỗi phản ứng hóa học đều được xúc tác bởi một enzyme đặc thù. Một ví dụ về con đường chuyển hóa như vậy là sự tổng hợp các sắc tố quy định màu mắt ở ruồi *Drosophila*. Vào khoảng những năm 1930, George Beadle và Boris Ephrussi dự đoán rằng ở ruồi *Drosophila*, mỗi một thể đột biến màu mắt đều có quá trình tổng hợp sắc tố bị ức chế tại một bước đặc thù nào đó, do thiếu sự tổng hợp enzym xúc tác bước phản ứng đó. Tuy vậy, vào thời điểm đó không có phản ứng nào cũng như enzyme có liên quan đến sự tổng hợp sắc tố quy định màu mắt ở ruồi giấm được biết đến.

1.2. Thí nghiệm của Beadle và Tatum với *Neurospora* dẫn đến đề xuất giả thuyết một gene - một enzyme

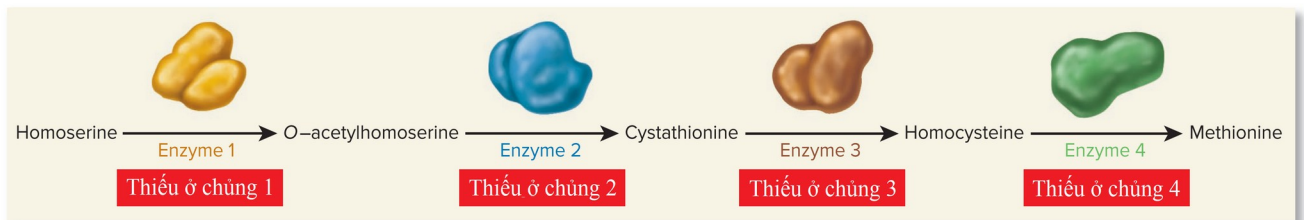
Một bước ngoặt trong việc làm sáng tỏ mối quan hệ giữa gene và enzyme đến sau đó vài năm khi Beadle và Edward Tatum nghiên cứu ở nấm men *Neurospora crassa*. Trên cơ sở các phương pháp gây tạo đột biến được tìm ra từ những năm 1920, các nhà khoa học đã dùng tia X “bắn phá” các chủng *Neurospora* để tạo nên các chủng đột biến có nhu cầu dinh dưỡng khác so với kiểu dại. Các chủng nấm men *Neurospora* kiểu dại có nhu cầu dinh dưỡng đơn giản. Chúng có thể dễ dàng sống trong môi trường thạch (agar) được bổ sung một số muối vô cơ, đường glucose và vitamin biotin. Từ **môi trường tối thiểu** này, các tế bào nấm men có thể dùng các con đường chuyển hóa của chúng để tạo nên tất cả các phân tử cần cho sự sinh trưởng và phát triển của mình. Beadle và Tatum đã xác định được nhiều chủng đột biến không có khả năng sống trên môi trường tối thiểu do nguyên nhân mất khả năng tổng hợp một hợp chất thiết yếu nào đó. Những chủng đột biến như vậy được gọi là đột biến khuyết dưỡng. Để có thể nuôi các chủng đột biến này, Beadle và Tatum phải nuôi chúng trong **môi trường đủ**, gồm các thành phần của môi trường tối thiểu, ngoài ra bổ sung thêm 20 loại amino acid và một số chất dinh dưỡng khác nữa. Trong môi trường đủ, mọi thể đột biến đều có khả năng sống dù chúng không có khả năng tổng hợp một chất nào đó.

Để phân tích đặc điểm của các dạng sai hỏng trao đổi chất ở các chủng đột biến khuyết dưỡng, Beadle và Tatum đã tiến hành lấy mẫu bằng cách nuôi chúng trong môi trường đủ, rồi phân phối chúng vào các ống đựng mẫu. Trong mỗi ống đựng mẫu, họ bổ sung môi trường tối thiểu, ngoài ra chỉ bổ sung thêm một chất dinh dưỡng nhất định (vốn khác nhau giữa môi trường đủ và môi trường tối thiểu). Chất bổ sung đặc thù cho phép nấm men đột biến có thể sinh trưởng sẽ cung cấp thông tin về kiểu sai hỏng chuyển hóa ở chủng nấm men đột biến. Ví dụ, nếu chủng đột biến được tìm thấy có khả năng phát triển trong môi trường bổ sung amino acid methionine, thì các nhà nghiên cứu kết luận rằng thể đột biến đó bị sai hỏng trong con đường chuyển hóa tổng hợp methionine so với chủng kiểu dại.

Beadle và Tatum sau đó tiếp tục xác định tính đặc thù của mỗi thể đột biến. Hình 2.1 minh họa cách họ dùng các phép thử tiếp theo để phân biệt ba thể đột biến khác nhau dù chúng đều là các đột biến khuyết dưỡng về methionine. Mỗi thể đột biến này đều cần một nhóm chất khác nhau dọc theo con đường sinh tổng hợp methionine gồm ba bước. Từ kết quả thí nghiệm, các nhà nghiên cứu cho rằng các thể đột biến đã bị ức chế ở các bước khác nhau của cùng con đường chuyển hóa trong đó mỗi thể đột biến thiếu một enzyme tương ứng với bước chuyển hóa bị ức chế.



(a) Sự sinh trưởng của các chủng trên các môi trường.



(b) Đơn giản hóa con đường sinh tổng hợp methionine.

Hình 2.1. Ví dụ về một thí nghiệm ủng hộ giả thuyết một gene - một enzyme của Beadle và Tatum. (a) Sự phát triển của các chủng loại hoang dã (WT) và đột biến trên môi trường sinh trưởng tối thiểu hoặc khi có O-acetylhomoserine, cystathionine, homocysteine hoặc methionine. (b) Một con đường đơn giản để sinh tổng hợp methionine. Lưu ý: Homoserine được tạo ra bởi *Neurospora* thông qua các enzyme và các phân tử tiền chất không được thảo luận trong thí nghiệm này.

Do trong nghiên cứu của Beadle và Tatum, các sai hỏng ở các thể đột biến đều liên quan đến một gene duy nhất, nên kết quả nghiên cứu của họ đã ủng hộ cho Giả thiết một gene - một enzyme mà chính hai nhà khoa học này đã đưa ra. Giả thiết “một gene - một enzyme” phát biểu rằng: chức năng của một gene là điều khiển sự tổng hợp một enzyme đặc thù. Giả thiết này sau đó tiếp tục được củng cố khi ngày càng có nhiều thể đột biến được xác định thiếu một enzyme đặc thù nào đó so với các dạng kiểu đại. Năm 1958, Beadle và Tatum được trao giải thưởng Nobel về “phát hiện của họ cho thấy các gene điều khiển các sự kiện hóa học xác định”.

1.3. Mã di truyền

Khả năng của mRNA được dịch mã thành một chuỗi amino acid cụ thể phụ thuộc vào mã di truyền. Trình tự các base trong phân tử mRNA cung cấp thông tin được mã hóa được đọc trong các nhóm ba nucleotide được gọi là codon (Hình 2.2). Trình tự của ba base trong hầu hết các codon quy định một amino acid cụ thể. Những codon này được gọi là **codon mã hóa/codon có nghĩa (sense codon)**. Ví dụ, codon AGC quy định serine amino acid. Codon AUG, quy định methionine, được sử dụng làm codon mở đầu; nó thường là codon đầu tiên bắt đầu chuỗi polypeptide. Mã hóa AUG cũng có thể được sử dụng để chỉ định các methionin bổ sung trong trình tự mã hóa. Cuối cùng, ba codon, UAA, UAG và UGA, được gọi là mã kết thúc, được sử dụng để kết thúc quá trình dịch. Các codon kết thúc còn được gọi là codon vô nghĩa.

		Base thứ hai				
		U	C	A	G	
Base thứ nhất (đầu 5' của codon)	U	UUU } Phe (F) UUC } UUA } Leu (L) UUG }	UCU } UCC } Ser (S) UCA } UCG }	UAU } Tyr (Y) UAC } UAA Stop UAG Stop	UGU } Cys (C) UGC } UGA Stop UGG Trp (W)	U C A G
	C	CUU } CUC } Leu (L) CUA } CUG }	CCU } CCC } Pro (P) CCA } CCG }	CAU } His (H) CAC } CAA } Gln (Q) CAG }	CGU } CGC } Arg (R) CGA } CGG }	U C A G
	A	AUU } AUC } Ile (I) AUA } AUG Met (M) or start	ACU } ACC } Thr (T) ACA } ACG }	AAU } Asn (N) AAC } AAA } Lys (K) AAG }	AGU } Ser (S) AGC } AGA } Arg (R) AGG }	U C A G
	G	GUU } GUC } Val (V) GUA } GUG }	GCU } GCC } Ala (A) GCA } GCG }	GAU } Asp (D) GAC } GAA } Glu (E) GAG }	GGU } GGC } Gly (G) GGA } GGG }	U C A G

Hình 2.2. Bảng mã di truyền với các codon trên mRNA. Các codon trên mRNA đọc theo hướng 5' → 3'. Codon AUG không chỉ mã hóa amino acid methionine (Met, hoặc M) mà còn có chức năng như một tín hiệu “bắt đầu” để ribosome bắt đầu dịch mã mRNA tại đó. Ba trong số 64 codon (UAG, UGA, UAA) có chức năng như tín hiệu “dừng”, đánh dấu nơi ribosome kết thúc quá trình dịch mã.

Từ bốn "kí tự" (A, T, G và C), hệ thống di truyền của các sinh vật có tổng cộng 64 bộ ba khác nhau. Hình 2.2 liệt kê tất cả các mã bộ ba và nghĩa phổ biến (amino acid được mã hóa) của chúng (các mã được viết theo chiều 5' → 3'). Trong 64 mã bộ ba, có 3 mã kết thúc và 61 bộ ba mã hóa cho 20 amino acid phổ biến. Điều này có nghĩa là một số amino acid được mã hóa bởi nhiều hơn một mã bộ ba. Hiện tượng này được gọi là **tính thoái hóa** của mã bộ ba. Các mã bộ ba được dùng để mã hóa cho cùng một amino acid được gọi là các mã bộ ba đồng nghĩa. Nhìn chung, các mã đồng nghĩa thường có hai nucleotide đầu giống nhau, nucleotide thứ ba hoặc là C hoặc là U, hoặc là A và G (thay thế nhau). Tuy vậy, nguyên tắc này không phải lúc nào cũng đúng, điển hình là Leu được mã hóa bởi các bộ ba UUA, UUG, CUU, CUC, CUA và CUG; hay Arg được mã hóa bởi các bộ ba CGU, CGC, CGA, CGG, AGA và AGG. Sự thoái hóa của mã bộ ba, đặc biệt ở vị trí thứ ba, giải thích cho hiện tượng tại sao tỉ lệ AT/GC ở phần lớn các loài là khác nhau, nhưng thành phần amino acid trong các protein của chúng lại rất giống nhau. Đồng thời, nhờ đặc tính thoái hóa này, sai sót dịch mã được giảm thiểu. Bởi vì, chúng ta biết rằng các liên kết G-C có lực liên kết lớn hơn so với A-U. Vậy nên, nếu hai nucleotide đầu tiên là G/C thì khả năng kết cặp sai trong dịch mã ở nucleotide thứ ba rất dễ xảy ra.

Việc sử dụng bất kỳ một nucleotide trong số 4 nucleotide ở vị trí thứ ba để mã hóa cho cùng một amino acid giúp giảm thiểu hậu quả đột biến.

Mặc dù có 61 codon có nghĩa nhưng trên thực tế trong mỗi tế bào prokaryote và eukaryote chỉ có khoảng 40 phân tử tRNA khác nhau. Lời giải đáp cho con số nhỏ hơn này nằm ở khả năng một codon đôi mã nhận ra nhiều hơn một codon tương ứng với một amino acid nhất định. Khả năng nhận biết mở rộng này tồn tại nhờ liên kết không chuẩn giữa các nitrogenous base trong **vị trí linh động** hay còn gọi là sự **"lỏng lẻo"** trong liên kết. Đó là vị trí thứ ba (đầu 3') trong codon mRNA và vị trí đầu tiên (đầu 5') tương ứng trong codon đôi mã của tRNA. Bảng 2.1 cho thấy, ngoài 4 nucleotide thông thường, vị trí đầu tiên (đầu 5') trên anticodon có thể mang một nucleotide biến đổi là I (inosine). I được biến đổi từ A bằng phản ứng loại nhóm amino khỏi vị trí C-6 (I thực chất là một nucleotide có phần đường là ribose và phần base là hypoxanthine). I có thể đồng thời tạo liên kết hydrogen với A, U hoặc C. Tại vị trí đầu tiên của anticodon, G có thể liên kết với C hoặc liên kết "lỏng lẻo" với U; ngược lại, U có thể liên kết với A hoặc "lỏng lẻo" với G. Khả năng liên kết "lỏng lẻo" trên đây của nucleotide thứ đầu tiên trong anticodon cho thấy thực tế một số tRNA có thể liên kết vào nhiều hơn một mã bộ ba trên mRNA (nhưng không bao giờ nhiều hơn 3 mã bộ ba khác nhau). Dựa trên các quy tắc "lỏng lẻo" này, chỉ có khoảng 30 loại tRNA khác nhau là cần thiết để phù hợp với 61 codon quy định amino acid. Trên thực tế, có 30 đến 40 loại tRNA có ở vi khuẩn và 41 đến 55 loại tRNA có ở sinh vật nhân chuẩn.

Bảng 2.1: Nguyên tắc kết cặp giữa anticodon và codon

Base tại vị trí đầu tiên (đầu 5') của anticodon trên tRNA	Base tại vị trí thứ ba (đầu 3') của codon trên mRNA
A	U
C	G
G	C hoặc U
U	A hoặc G
I	A, U, hoặc C

Tất cả các hệ thống sinh học (kể cả virus, vi khuẩn cho đến các sinh vật nhân thật) đều sử dụng mã di truyền là mã bộ ba và dùng một bảng mã chung (hình 2.2). Đặc điểm này được gọi là **tính phổ biến của mã di truyền**. Tính phổ biến của mã di truyền được biểu hiện đồng thời với các đặc tính chung của mã bộ ba là: (1) các mã bộ ba (trên mRNA) luôn được đọc theo chiều 5' → 3'; (2) kể từ mã bắt đầu (của ORF), các mã bộ ba luôn được đọc liên tục không ngắt quãng cho đến khi gặp mã kết thúc (nói cách khác là các mã bộ ba được đọc trong các ORF cố định). Đây là những bằng chứng quan trọng cho thấy các dạng sống trên Trái đất đều có nguồn gốc chung. Ngoài ra, cũng nhờ các đặc tính này, các nghiên cứu về kỹ thuật di truyền hiện nay có thể tiến hành tái tổ hợp DNA (hoặc các gene) giữa bất cứ hai đối tượng sinh vật nào (kể từ virus cho đến thú và người).

Bên cạnh tính phổ biến của hệ thống mã di truyền nói trên, các nghiên cứu gần đây cho thấy một vài chệch hướng mà hầu hết là xảy ra trong các bộ gene ty thể. Một vài

thay đổi thỉnh thoảng cũng xảy ra ở các bộ gene vi khuẩn và bộ gene nhân của các eukaryote, nhưng thường thì liên quan với các codon kết thúc. Sự phân bố phát sinh chủng loại của các thay đổi này chỉ ra rằng mã di truyền vẫn còn tiến hóa.

Bảng 2.2: Những biến thể mã di truyền đã biết

Codon	Mã phổ biến	Mã biến thể	Xảy ra ở
AUA	Isoleucine	Methionine	Ty thể nấm men và động vật có vú
UGA	Kết thúc	Tryptophan	Ty thể động vật có vú
CUU, CUA, CUC, CUG	Leucine	Threonine	Ty thể nấm men
AGA, AGG	Arginine	Kết thúc	Động vật nguyên sinh có nhung mao; ty thể nấm men và động vật có vú;
UAA, UAG	Kết thúc	Glutamine	Động vật nguyên sinh có nhung mao (<i>Paramecium</i>); <i>Acetabularia</i> (một loại thực vật đơn bào)
UGA	Kết thúc	Selenocysteine	Được tìm thấy ở một số gene ở vi khuẩn, vi khuẩn cổ và eukaryotes
UAC	Kết thúc	Pyrrolysine	Được tìm thấy ở một số gene ở vi khuẩn cổ sinh methane
UGA	Kết thúc	Cysteine	<i>Euplotes</i> (một loại trùng tiêm mao)

2. Các thành phần tham gia dịch mã

Bộ máy tổng hợp protein gồm bốn thành phần cơ bản là mRNA, tRNA, enzyme aminoacyl-tRNA synthetase và ribosome. Khi kết hợp với nhau, các thành phần của bộ máy dịch mã làm nhiệm vụ phiên dịch ngôn ngữ dưới dạng trình tự của 4 loại nucleotide gồm A, G, T và C ở dạng mã bộ ba được viết trên phân tử DNA thành ngôn ngữ của 20 loại amino acid phổ biến trên phân tử protein; trong đó mỗi một mã bộ ba (codon) qui định một amino acid và được đọc liên tục trong vùng mã hóa (ORF) của mRNA hoàn thiện. Trong quá trình dịch mã, mRNA được dùng làm khuôn để tổng hợp protein; như vậy, phân tử này có vai trò cung cấp thông tin cho bộ máy dịch mã. tRNA là “giao diện” vừa để các mã bộ ba của mRNA dính vào, vừa để các amino acid tương ứng với các mã bộ ba của mRNA có thể liên kết vào chuỗi polypeptide. Vai trò của tRNA có thể ví như “người phiên dịch” giữa hai “ngôn ngữ” DNA và protein. Enzyme aminoacyl-tRNA synthetase xúc tác phản ứng hình thành liên kết cộng hóa trị đặc hiệu giữa amino acid với phân tử tRNA vận chuyển; chính enzyme này đảm bảo việc dịch mã chính xác qua hoạt động lựa chọn đúng các phân tử tRNA phù hợp với mỗi loại amino acid. Thành phần cuối cùng của bộ máy dịch mã là ribosome có khối lượng phân tử lớn (hàng ngàn kDa) được cấu tạo từ protein và rRNA. Ribosome làm nhiệm vụ “điều phối” việc nhận ra chính xác các mã bộ ba trên mRNA của mỗi tRNA, đồng thời thúc đẩy phản ứng hình thành liên kết peptide.

2.1. mRNA mang trình tự mã hóa protein

Trong quá trình dịch mã, bộ máy dịch mã có nhiệm vụ giải mã vùng mã hóa của mRNA. Vùng mã hóa protein của mỗi phân tử mRNA hoàn thiện, còn được gọi là **khung đọc mở (kí hiệu là ORF)**, là một trình tự nucleotide liên tiếp (không "ngắt quãng") được đọc liên tục (không "dấu phẩy") từ một mã mở đầu (thường là AUG) và kết thúc ở một trong các mã kết thúc (thường là UAA, UAG và UGA). Cần phân biệt trình tự bắt đầu và kết thúc của một ORF thường tương ứng với các điểm bắt đầu và kết thúc dịch mã, còn các trình tự khởi đầu và kết thúc của mRNA (không kể đuôi polyA) tương ứng với điểm bắt đầu và kết thúc phiên mã. Dịch mã bắt đầu từ điểm bắt đầu của một ORF và đọc liên tục mỗi lần một mã bộ ba cho đến khi gặp mã kết thúc của ORF đó. Các mã bắt đầu và kết thúc của một ORF được gọi tắt tương ứng là mã bắt đầu và mã kết thúc dịch mã.

Ở prokaryote, mã bắt đầu dịch mã thường là 5'-AUG-3', nhưng đôi khi là 5'-GUG-3' hay 5'-UUG-3'. Tế bào eukaryote (trừ hệ gene ti thể và Lạp thể) luôn dùng mã bắt đầu là 5'-AUG-3'. Mã bắt đầu có vai trò rất quan trọng, vì nó đồng thời thực hiện hai chức năng. Chức năng thứ nhất là xác định amino acid đầu tiên của chuỗi polypeptide, thường là Met (mặc dù ở một số protein, amino acid này được cắt bỏ khi hoàn thiện). Chức năng thứ hai là xác định điểm bắt đầu một ORF và toàn bộ các bộ ba còn lại trong ORF. Mọi đột biến liên quan đến mã bắt đầu đều làm mất hoàn toàn nghĩa của ORF và dẫn đến việc protein không được tổng hợp. Về nguyên tắc, trên một mạch DNA có thể có tối đa 3 ORF khác nhau; nếu xét trên cả hai mạch DNA thì sẽ có tối đa 6 ORF khác nhau. Đây là cơ sở của hiện tượng các ORF có thể nằm gối lên nhau (mặc dù hiện tượng này đến nay mới gặp ở virus) và có thể được đọc theo cả hai chiều (tương ứng với hai mạch) của phân tử DNA. Tuy nhiên, một khi sự dịch mã đã bắt đầu thì mỗi mã bộ ba được dịch mã tiếp theo luôn nằm kề sát mã bộ ba phía trước. Nói cách khác, nếu ribosome đã xác định một mã bắt đầu, thì mã bộ ba đó xác định tất cả các bộ ba theo sau của quá trình dịch mã đó.

Phần lớn sinh vật sử dụng 3 mã kết thúc là 5'UAA3', 5'UAG3' và 5'UGA3'. Các mã kết thúc vừa xác định ranh giới tận cùng của một ORF, vừa là tín hiệu kết thúc dịch mã.

2.2. Cấu trúc và chức năng của tRNA

Mặc dù có nhiều loại tRNA (trên 300 loại đã được phát hiện), nhưng mỗi tRNA chỉ đính kết với một loại amino acid đặc trưng duy nhất và nhận biết một hoặc một số mã bộ ba nhất định (trên mRNA, phần lớn tRNA nhận biết được nhiều hơn một mã bộ ba). Kích thước của phần lớn tRNA là khoảng 75 – 95 nucleotide. Mặc dù, trình tự chi tiết của các tRNA là khác nhau, nhưng chúng đều có các đặc điểm chung sau:

- Các tRNA chứa nhiều ribonucleotide được biến đổi sau phiên mã, chẳng hạn như pseudouridine (ψ), dihydrouridine (DHU), hypoxanthine, thymine (T), methylguanine (mG); mặc dù các base biến đổi này không có vai trò thiết yếu trong hoạt động của tRNA, nhưng nếu thiếu các base biến đổi này tế bào thường sinh trưởng kém (riêng hypoxanthine có vai trò quan trọng trong nhận biết mã bộ ba của một số tRNA);

- Các tRNA đều có cấu tạo gồm các đoạn trình tự xoắn kép (hình thành bởi liên kết hydro giữa các phân trình tự khác nhau của phân tử) xen kẽ với các đoạn trình tự mạch đơn và cấu hình điển hình của chúng giống một chiếc lá gồm 4 thùy chính, đó là các thùy

tiếp nhận amino acid, thùy ψ (hay ψ U), thùy D (DHU) và thùy đôi mã. Ngoài ra còn có một thùy biến đổi có kích thước và trình tự nucleotide khác nhau ở các tRNA khác nhau. Có thể mô tả đặc điểm cơ bản của mỗi thùy như sau:

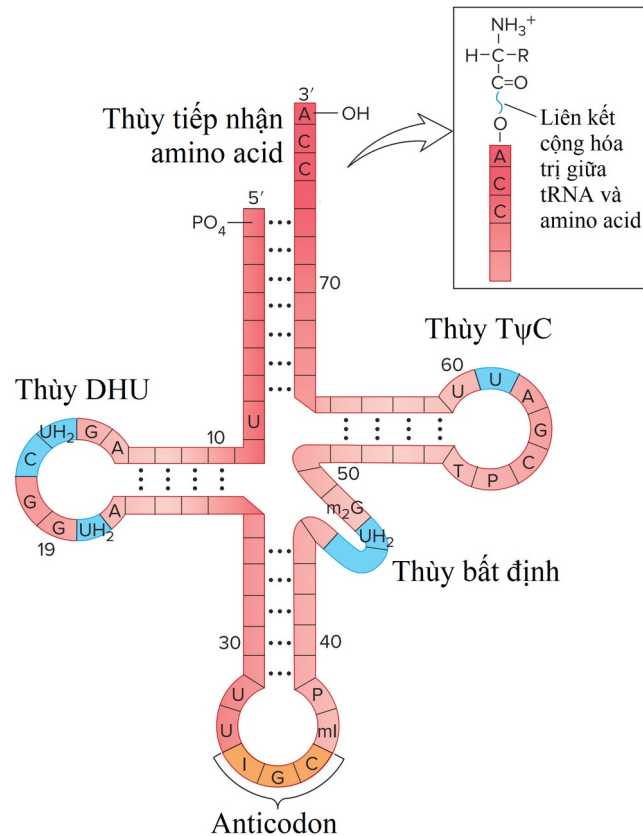
+ *Thùy tiếp nhận amino acid* được hình thành bởi sự kết cặp của các đoạn trình tự ở hai đầu 5' và 3' của phân tử tRNA. Tất cả tRNA đều có trình tự đầu 3' tận cùng thuộc thùy tiếp nhận amino acid là 5'-CCA-3'. Đây chính là vị trí tRNA liên kết với amino acid trong quá trình dịch mã (nhờ xúc tác bởi aminoacyl-tRNA synthetase);

+ *Thùy T ψ C* có đặc điểm đặc trưng là có một base uridine được biến đổi thành ψ , và base này thường được tìm thấy trong trình tự 5'-T ψ CC-3'.

+ *Thùy DHU* có đặc điểm đặc trưng là chứa base dihydrouridine (DHU).

+ *Thùy đôi mã* mang mã bộ ba đối mã (anticodon) tương ứng với mã bộ ba (codon) trên phân tử mRNA. Mã bộ ba đối mã luôn liền kề với một purine (A hoặc G) ở đầu 3' và một uracil (U) ở đầu 5' trên phân tử tRNA.

+ *Thùy bất định* ở giữa thùy đôi mã và thùy T ψ C, có chiều dài từ 3 đến 21 nucleotide.

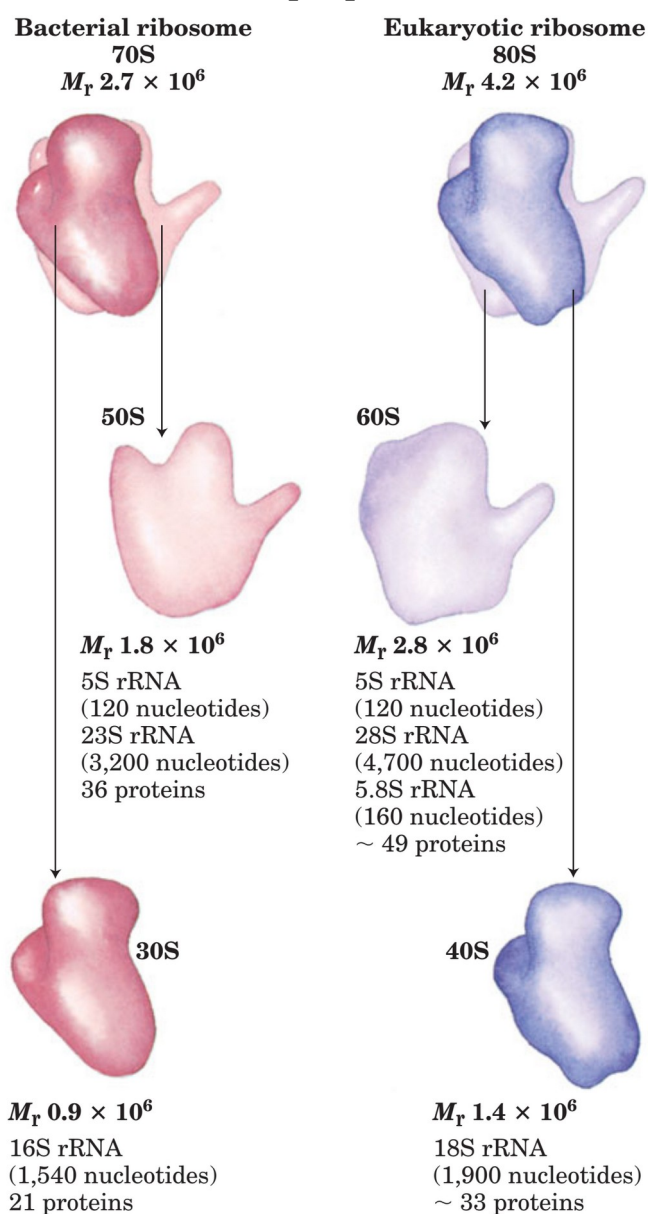


Hình 2.3. Cấu trúc bậc hai của tRNA. Trong tất cả các tRNA, các nucleotide ở đầu 3' chứa trình tự CCA. Một số vị trí có thể có thêm nucleotide không được tìm thấy trong tất cả các phân tử tRNA. Các trang vị trí biến đổi này được hiển thị bằng màu xanh lam. Hình này cũng cho thấy vị trí của một số base biến đổi đặc trưng được tìm thấy trong tRNA mang alanine của nấm men. Các base bị biến đổi: I = inosine, mI = methylinosine, T = ribothymidine, UH₂ = dihydrouridine, m₂G = dimethylguanosine, và P = pseudouridine.

tRNA khi đã liên kết với amino acid được gọi là **tRNA đã nạp amino acid**, còn các tRNA chưa liên kết với amino acid được gọi là **tRNA chưa nạp amino acid**. Quá trình “nạp amino acid” dùng năng lượng từ ATP để hình thành liên kết acyl giữa nhóm -

COOH của amino acid với nhóm 2'- hoặc 3'-OH của nucleotide A ở tận cùng đầu 3' của tRNA. Đáng chú ý là liên kết acyl của tRNA sau khi nạp amino acid là một liên kết cao năng. Điều này rất quan trọng, bởi năng lượng tự do giải phóng từ liên kết cao năng này trong quá trình kéo dài chuỗi polypeptide chính là động lực hình thành liên kết peptide giữa các amino acid.

2.2.2. Cấu trúc của ribosome và sự lắp ráp

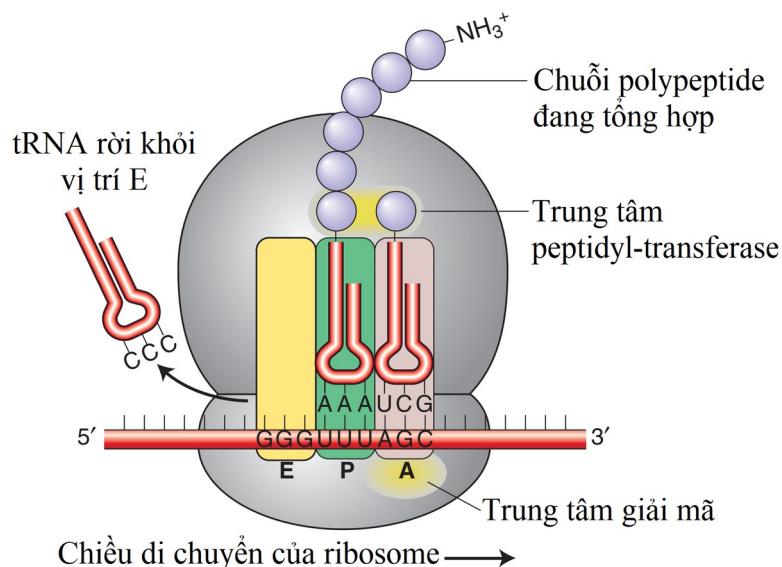


Hình 2.4. Thành phần phân tử và khối lượng (M_r) của các ribosome ở vi khuẩn và eukaryote. Các tiểu đơn vị ribosome được xác định bằng các giá trị S (đơn vị Svedberg), hệ số lắng đề cập đến tốc độ lắng của chúng trong máy ly tâm. Các giá trị S không nhất thiết phải cộng khi các tiểu đơn vị được kết hợp, vì tốc độ lắng bị ảnh hưởng bởi hình dạng cũng như khối lượng.

Ribosome được cấu tạo từ nhiều đại phân tử và có vai trò điều khiển quá trình tổng hợp protein. So với các bộ máy sao chép DNA và phiên mã, thì ribosome có cấu trúc phức tạp và lớn hơn nhiều. Về cấu trúc, ribosome ở tất cả các giới sinh vật, dù hoạt động trong tế bào chất hay trong các bào quan, đều được cấu tạo gồm hai tiểu phần (gọi là tiểu phần lớn và tiểu phần nhỏ). Tiểu phần lớn mang vùng xúc tác hình thành liên kết peptide,

gọi là **trung tâm peptidyl transferase**. Tiểu phần nhỏ mang vùng có vai trò đối chiếu bộ ba đối mã của tRNA (đã nạp amino acid) với mã bộ ba trên mRNA nên được gọi là **trung tâm giải mã**. Cả tiểu phần lớn và tiểu phần nhỏ, cũng như các rRNA cấu tạo nên chúng được phân biệt và gọi tên dựa trên hệ số lắng trong li tâm và được đọc là S (Svedberg). Ribosome của prokaryote có hệ số lắng là 70S được cấu tạo từ một tiểu phần lớn 50S và một tiểu phần nhỏ 30S (hệ số lắng S không chỉ phụ thuộc vào khối lượng, mà còn phụ thuộc vào kích thước, hình dạng... ; nên hệ số lắng của một phức hệ thường không bằng tổng hệ số lắng thành phần). Ribosome của eukaryote có hệ số lắng 80S, được cấu tạo từ một tiểu phần lớn 60S và một tiểu phần nhỏ 40S.

Để xúc tác phản ứng hình thành liên kết peptide, mỗi ribosome phải đồng thời có ít nhất hai vị trí liên kết tRNA. Trong thực tế, mỗi ribosome có ba vị trí liên kết tRNA, được gọi là các vị trí A, P và E (Hình 2.5). **Vị trí A** là nơi đính kết với tRNA đã nạp amino acid; **vị trí P** là nơi đính kết với tRNA mang amino acid đã liên kết vào chuỗi polypeptide đang tổng hợp (peptidyl-tRNA); **vị trí E** là nơi tRNA giải phóng khỏi chuỗi polypeptide và ribosome. Mỗi vị trí liên kết của tRNA đều nằm trên vùng giáp nối (gối nhau) giữa tiểu phần lớn và tiểu phần nhỏ của ribosome. Nhờ vậy, tRNA có thể đồng thời tiếp cận được “trung tâm peptidyl transferase” của tiểu phần lớn, vừa tiếp cận được “trung tâm giải mã” của tiểu phần nhỏ. Trong khi thùy mang amino acid (đầu 3') của tRNA đính vào tiểu phần lớn, thì thùy đối mã của nó đính vào tiểu phần nhỏ.



Hình 2.5. Các vị trí chính trong sự tương tác của ribosome với các thành phần khác trong quá trình kéo dài dịch mã.

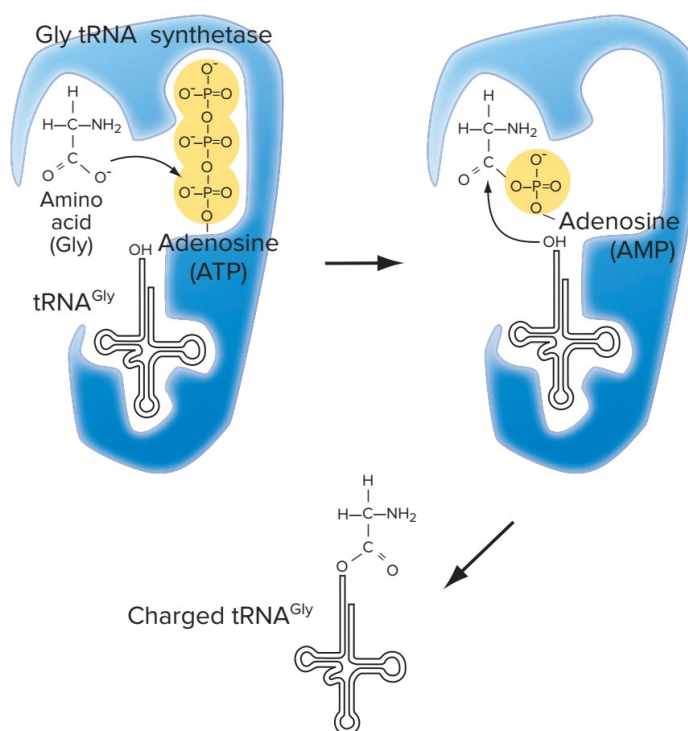
Để tiến hành dịch mã, cấu hình phân tử của ribosome hình thành nhiều đường dẫn nhỏ ra và vào. Phân tử mRNA đi vào và ra khỏi “trung tâm giải mã” qua một khe dẫn hẹp chỉ vừa đủ rộng cho một phân tử RNA dạng mạch đơn chui qua. Điều này đảm bảo cho việc dịch mã diễn ra khi các mối tương tác nội phân tử giữa các ribonucleotide không thể hình thành. ở giữa khe dẫn hẹp đó là vùng tiếp xúc tRNA. Điều đáng chú ý là kích thước vùng này vừa đủ chỗ cho đúng hai mã bộ ba (tương ứng với mã bộ ba đối mã của hai tRNA) liên kết vào, đảm bảo mã di truyền được đọc theo đúng mã bộ ba. Tiểu phần lớn của ribosome có một khe dẫn thứ hai là đường dẫn ra của chuỗi polypeptide mới tổng hợp. Cũng giống như khe dẫn mRNA ở tiểu phần nhỏ, khe dẫn chuỗi

polypeptide ở tiểu phần lớn chỉ vừa đủ rộng để chuỗi polypeptide đi ra và hạn chế sự tương tác giữa các amino acid trong phân tử protein. Trong ribosome, các chuỗi polypeptide chỉ hình thành được cấu trúc bậc hai dạng xoắn α , chứ không hình thành được dạng mặt phẳng β hay các cấu trúc bậc cao hơn. Vì vậy, cấu hình của protein ở dạng hoạt động chức năng chỉ hình thành sau khi chúng rời ribosome.

3. Các giai đoạn của dịch mã

3.1. Hoạt hóa amino acid

Để tổng hợp chuỗi polypeptide có trình tự được xác định bởi mRNA, phải đáp ứng hai yêu cầu hóa học: (1) nhóm carboxyl của mỗi amino acid phải được hoạt hóa để tạo điều kiện thuận lợi cho việc hình thành liên kết peptide và (2) một mối liên kết phải được thiết lập giữa mỗi amino acid mới và thông tin trong mRNA mã hóa nó. Cả hai điều này đều được thỏa mãn bằng cách gắn cộng hóa trị của amino acid vào tRNA trước khi tổng hợp protein. Việc gắn mỗi amino acid với tRNA tương ứng của nó là rất quan trọng. Phản ứng này diễn ra trong tế bào chất, không diễn ra trên ribosome. Mỗi amino acid trong số 20 amino acid được liên kết cộng hóa trị với một tRNA đặc hiệu với năng lượng từ sự thủy phân ATP, được xúc tác bởi các enzyme hoạt hóa phụ thuộc Mg^{2+} được gọi là các **aminoacyl-tRNA synthetase**. Khi được gắn vào một amino acid, một tRNA được gọi là được “ **nạp** ” (**tRNA charged**).



Hình 2.6. Các aminoacyl-tRNA synthetase gắn các amino acid vào tRNA. Enzyme có các vị trí nhận biết một amino acid, các tRNA tương ứng và ATP. Đầu tiên synthetase hoạt hoá amino acid, tạo thành AMP-amino acid. Sau đó, enzyme này chuyển nhóm carboxyl của amino acid từ AMP sang nhóm hydroxyl ($-OH$) của ribose ở đầu 3' của tRNA, tạo ra tRNA được "nạp" amino acid.

Quá trình nạp tRNA đòi hỏi sự hình thành một liên kết acyl giữa nhóm carboxyl của amino acid và đầu 2'- hoặc 3'-hydroxyl tự do của tRNA. Tất cả các tRNA đều có chung

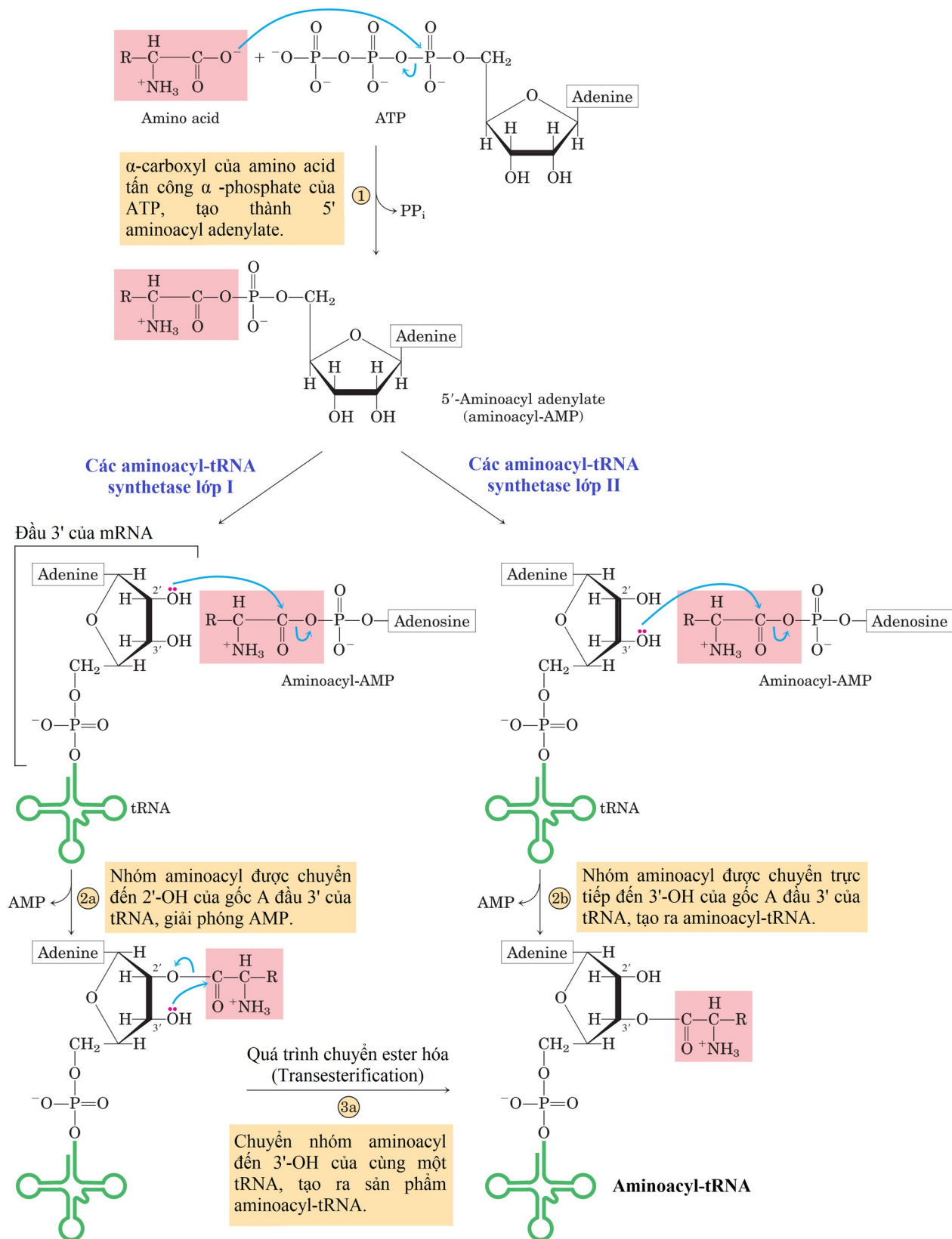
cấu trúc, bao gồm ba hoặc bốn nhánh và trình tự ở đầu CCA ở đầu 3'. Phản ứng acyl hóa dẫn đến việc gắn amino acid vào đầu 3' adenosine.

Aminoacyl-tRNA synthetase phải hoạt hóa amino acid trước khi nó được gắn vào tRNA. Phản ứng này xảy ra theo hai bước trong vị trí hoạt động của enzyme. Cấu trúc của tất cả các aminoacyl-tRNA synthetase của *E. coli* đã được xác định. Các nhà nghiên cứu đã chia chúng thành hai lớp dựa trên sự khác biệt đáng kể về cấu trúc bậc 1 và bậc 3 và cơ chế phản ứng; hai lớp này giống nhau ở tất cả các sinh vật.

Bảng 2.3: Hai lớp aminoacyl-tRNA synthetase

Lớp I		Lớp II	
Arg	Leu	Ala	Lys
Cys	Met	Asn	Phe
Gln	Trp	Asp	Pro
Glu	Tyr	Gly	Ser
Ile	Val	His	Thr

Trong bước **adenylyl hóa (adenylylation)**, một chất trung gian liên kết với enzyme, 5'-aminoacyl adenylate (5'-aminoacyl-AMP), hình thành khi nhóm carboxyl của amino acid phản ứng với nhóm α -phosphoryl của ATP để tạo thành nên một liên kết photphoanhydride I, với sự chuyển vị của pyrophosphate. Trong bước **nạp amino acid vào tRNA** tiếp theo, nhóm aminoacyl được chuyển từ aminoacyl-AMP liên kết với enzyme đến tRNA đặc hiệu của nó. Amino acid có thể được chuyển sang 2'-OH hoặc 3'-OH của adenosine đầu 3' của tRNA, tùy thuộc vào loại aminoacyl-tRNA synthetase. Các enzyme synthetase lớp I gắn amino acid vào 2'-OH và các enzyme synthetase lớp II gắn amino acid với 3'-OH. Trong con đường lớp I, liên kết aminoacyl ester di chuyển đến vị trí 3'-OH một cách tự phát bằng phản ứng transesterification (phản ứng trao đổi giữa rượu và ester).



Hình 2.7. Liên kết các amino acid vào tRNA bằng aminoacyl-tRNA synthetase.
 Bước ❶: bước adenyl hóa, amino acid được liên kết với adenylate, tạo thành aminoacyl-AMP. Bước ❷: amino acid được chuyển từ AMP sang tRNA theo một trong hai con đường, được xúc tác bởi enzyme aminoacyl-tRNA synthetase lớp I hoặc lớp II. Nhóm aminoacyl được chuyển đến 2'-OH của gốc A đầu 3' của tRNA, giải phóng AMP. Bước ❸: Nhóm aminoacyl sau khi được các enzyme lớp I chuyển đến 2'-OH của A đầu 3'

sẽ được tiếp tục chuyển đến 3'-OH của gốc A ở chính tRNA đó tạo thành aminoacyl-tRNA.

3.2. Tổng hợp chuỗi polypeptide

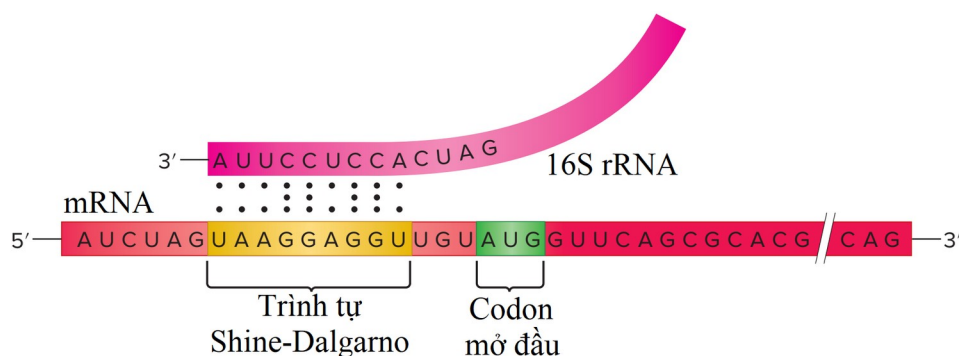
3.2.1. Khởi đầu dịch mã

Quá trình bắt đầu dịch mã ở tất cả các sinh vật bắt đầu khi tiểu đơn vị nhỏ của ribosome liên kết gần đầu 5' của mRNA và xác định trình tự codon bắt đầu. Trong giai đoạn tiếp theo, **tRNA khởi đầu** - tRNA mang amino acid đầu tiên của chuỗi polypeptide, liên kết với codon mở đầu trên mRNA. Cuối cùng, tiểu đơn vị lớn tham gia liên kết với tiểu đơn vị nhỏ để tạo thành ribosome hoàn chỉnh. Trong giai đoạn này, **các protein yếu tố khởi đầu** giúp kiểm soát sự hình thành và liên kết của ribosome với tRNA khởi đầu, và guanosine triphosphate (GTP) cung cấp năng lượng. Các tRNA được sử dụng trong quá trình dịch mã đều mang một amino acid đặc hiệu và được xác định là các tRNA đã liên kết với amino acid (tRNA được " nạp " amino acid).

3.2.1.1 Khởi đầu dịch mã ở vi khuẩn

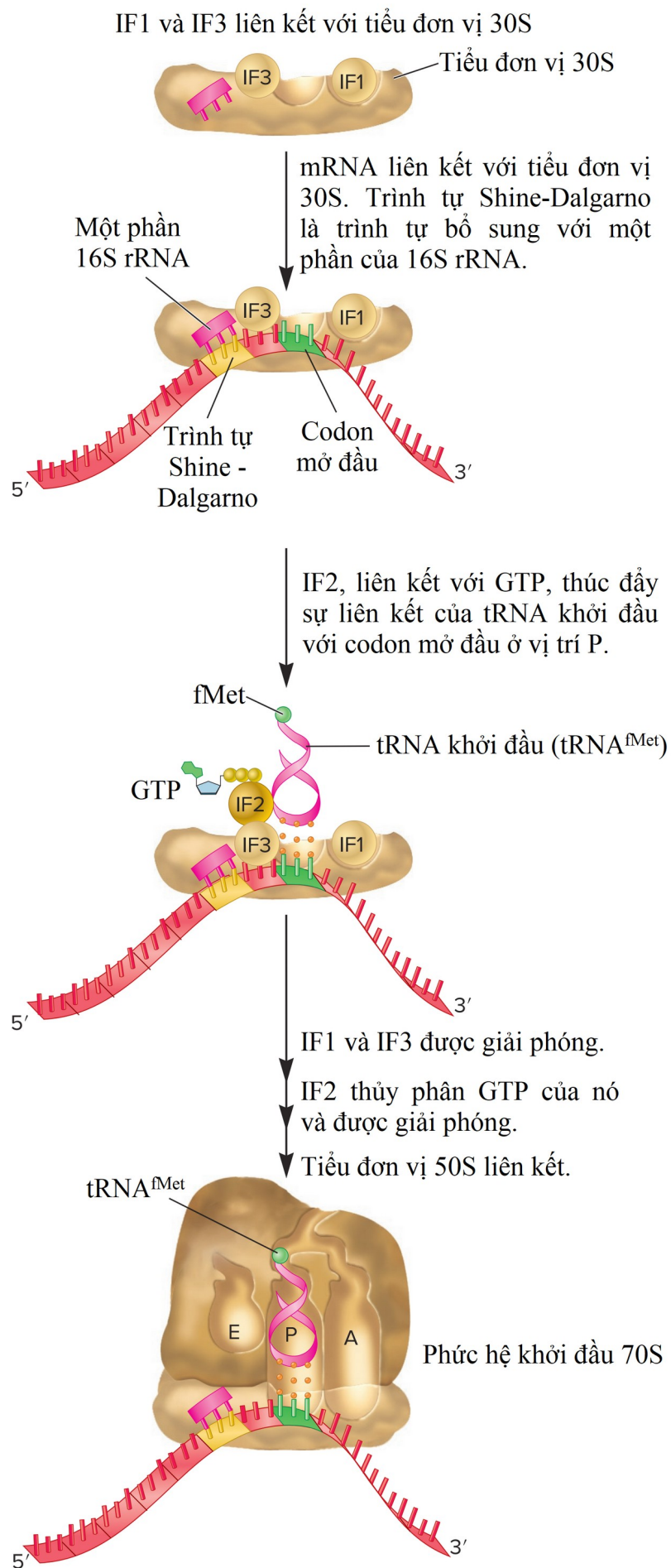
Hình 2.9 mô tả giai đoạn bắt đầu dịch mã ở vi khuẩn, trong đó các tiểu đơn vị mRNA, tRNA^{fMet} và ribosome kết hợp với nhau để tạo thành một phức hợp khởi đầu. Sự hình thành phức hợp này cần có sự tham gia của 3 yếu tố khởi đầu: IF1, IF2, IF3. Đầu tiên, IF1 và IF3 liên kết với tiểu đơn vị 30S của ribosome, điều này ngăn cản sự liên kết của tiểu đơn vị 50S. Tiếp theo, mRNA liên kết với tiểu đơn vị 30S. Sự liên kết này được tạo điều kiện thuận lợi bởi trình tự chín nucleotide trong mRNA của vi khuẩn được gọi là **trình tự Shine-Dalgarno**³ còn gọi là **trình tự bám ribosome (RBS, ribosome binding site)**. Trình tự Shine-Dalgarno tạo điều kiện thuận lợi cho việc liên kết mRNA với ribosome bằng cách bổ sung với một chuỗi ngắn trong rRNA 16S, giúp thúc đẩy liên kết hydrogen của mRNA với tiểu đơn vị 30S.

Trình tự Shine-Dalgarno là một trình tự giàu purine gồm khoảng sáu nucleotide nằm ngược dòng của codon bắt đầu từ 3 đến 9 nucleotide. Một đoạn giàu pyrimidine bổ sung chứa trình tự UCCUCC được tìm thấy gần đầu 3' của rRNA 16S, và nó bắt cặp với trình tự Shine-Dalgarno để định vị mRNA trên tiểu đơn vị 30S. Trình tự Shine – Dalgarno là một ví dụ khác về trình tự liên ứng, trình tự nucleotide chính xác và vị trí chính xác của trình tự Shine – Dalgarno thay đổi rất ít giữa các mRNA.



Hình 2.8. Vị trí của trình tự Shine-Dalgarno và codon mở đầu ở mRNA vi khuẩn. Trình tự Shine-Dalgarno bổ sung với một trình tự trong 16S rRNA, thúc đẩy sự khởi đầu dịch mã.

³ Theo tên của các nhà nghiên cứu người Úc là John Shine và Lynn Dalgarno, người đã nhận ra nó.



Hình 2.9. Giai đoạn khởi đầu của dịch mã ở vi khuẩn

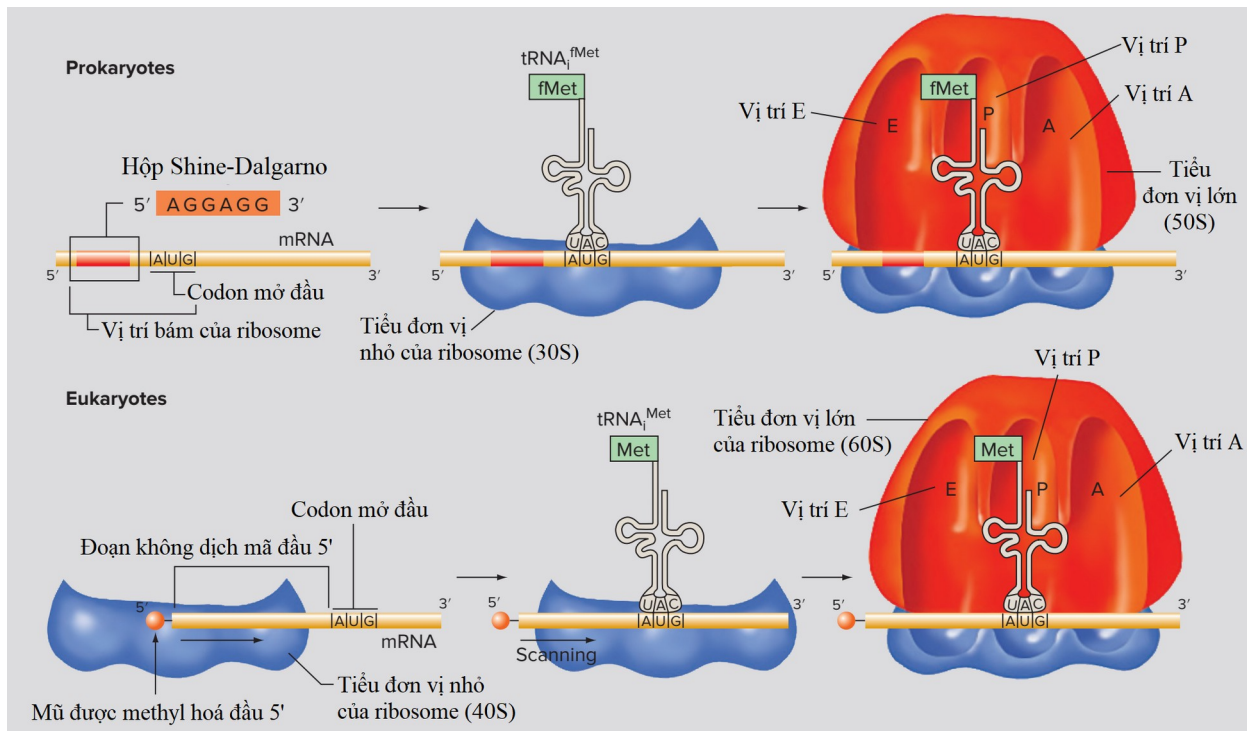
Bước tiếp theo đòi hỏi IF2, có GTP liên kết với nó. IF2 thúc đẩy sự liên kết của tRNA khởi đầu (tRNA^{fMet}) với mRNA đã được liên kết với tiểu đơn vị 30S. tRNA^{fMet} liên kết với codon bắt đầu, thường là một vài nucleotide xuôi dòng từ trình tự Shine-Dalgarno. Mã mở đầu thường là AUG, nhưng trong một số trường hợp, nó có thể là GUG hoặc UUG. Ngay cả khi codon bắt đầu là GUG (thường mã hóa valine) hoặc UUG (thường mã hóa leucine), thì amino acid đầu tiên trong polypeptide vẫn là formylmethionine vì chỉ một tRNA^{fMet} mới có thể bắt đầu dịch mã. Trong hoặc sau khi dịch mã toàn bộ polypeptide, nhóm formyl hoặc toàn bộ N-formylmethionine có thể bị loại bỏ. Do đó, một số polypeptide có thể không có N-formylmethionine là amino acid đầu tiên của chúng. Như đã lưu ý trong Hình 2.9, tRNA^{fMet} liên kết với vị trí P trên ribosome. IF1 được cho là chiếm một phần của vị trí A, do đó ngăn cản sự liên kết của tRNA^{fMet} với vị trí A trong quá trình khởi đầu.

Sau khi mRNA và tRNA^{fMet} liên kết với tiểu đơn vị 30S, IF1 và IF3 được giải phóng, sau đó IF2 thủy phân GTP của nó và cũng được giải phóng. Điều này cho phép tiểu đơn vị 50S ribosome liên kết với tiểu đơn vị 30S tạo thành ribosome hoàn chỉnh. Sau khi dịch mã hoàn thành, sự liên kết IF1 là cần thiết để phân ly các tiểu đơn vị 50S và 30S của ribosome để tiểu đơn vị 30S có thể bắt đầu lại với một phân tử mRNA khác.

3.2.1.2. Khởi đầu dịch mã ở eukaryote

Ở eukaryote, sự lắp ráp của phức hợp khởi đầu mang những điểm tương tự như ở vi khuẩn. Tuy nhiên, cần có thêm các yếu tố khác cho quá trình khởi đầu gọi là eIF để phân biệt chúng với các yếu tố khởi đầu của vi khuẩn (IF). tRNA khởi đầu ở sinh vật nhân chuẩn mang methionine chứ không phải N-formylmethionine, như ở vi khuẩn. Một yếu tố khởi đầu sinh vật nhân chuẩn, eIF2, liên kết trực tiếp với tRNA^{Met} để thúc đẩy nó vào tiểu đơn vị 40S của ribosome. mRNA của sinh vật nhân chuẩn không có trình tự Shine-Dalgarno. mRNA được nhận dạng bởi eIF4, là một phức hợp đa protein nhận biết mũ 7-methylguanosine ở đầu 5' của mRNA. Sau đó eIF4 tạo điều kiện cho sự liên kết của đầu 5' của mRNA với tiểu đơn vị 40S của ribosome.

Việc xác định codon mở đầu AUG chính xác ở sinh vật nhân chuẩn khác với ở vi khuẩn. Sau khi liên kết ban đầu của mRNA với ribosome, bước tiếp theo là xác định vị trí một codon mở đầu AUG nằm ở đâu đó xuôi dòng từ mũ 7-methylguanosine. Năm 1986, Marilyn Kozak đề xuất rằng ribosome bắt đầu ở đầu 5' và sau đó quét dọc theo mRNA theo hướng 3' để tìm kiếm codon mở đầu AUG. Trong nhiều trường hợp, nhưng không phải tất cả, ribosome sử dụng codon AUG đầu tiên mà nó gặp làm codon mở đầu. Khi một codon mở đầu được xác định, eIF5 gây ra việc giải phóng các yếu tố khởi đầu khác, cho phép tiểu đơn vị lớn 60S kết hợp với tiểu đơn vị bé 40S tạo thành ribosome hoàn chỉnh.



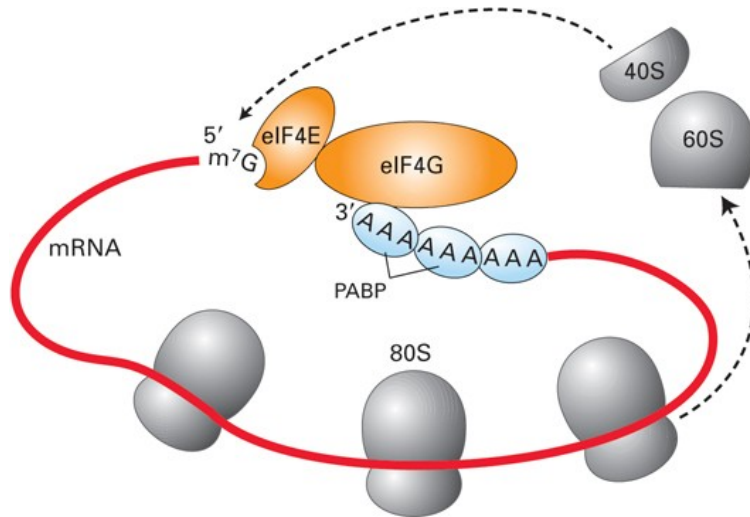
Hình 2.10. Khởi đầu dịch mã ở prokaryote và eukaryote.

Bằng cách phân tích trình tự của nhiều mRNA sinh vật nhân chuẩn, các nhà nghiên cứu đã phát hiện ra rằng không phải tất cả các codon AUG gần đầu 5' của mRNA đều có thể hoạt động như các codon mở đầu. Trong một số trường hợp, ribosome quét đi qua codon AUG đầu tiên và chọn một AUG ở xa hơn mRNA. Trình tự các base xung quanh codon AUG đóng một vai trò quan trọng trong việc xác định xem nó có được chọn làm codon mở đầu bởi ribosome quét hay không. Trình tự liên ứng để nhận dạng codon mở đầu tối ưu ở sinh vật nhân chuẩn phức tạp, chẳng hạn như động vật có xương sống và thực vật có mạch, như sau:

						Codon mở đầu			
G	C	C	(A/G)	C	C	A	U	G	G
-6	-5	-4	-3	-2	-1	+1	+2	+3	+4

Ngoài codon AUG, guanin ở vị trí +4 và purine, tốt nhất là adenin, ở vị trí -3 là những vị trí quan trọng nhất để lựa chọn codon bắt đầu. Các quy tắc này để khởi đầu dịch mã tối ưu được gọi là **quy tắc Kozak (Kozak's rules)**. Trình tự liên ứng 5'-ACCAUGG-3' (codon mở đầu được in đậm) được gọi là **trình tự Kozak**, theo tên của Marilyn Kozak, người đã phát hiện ra nó năm 1978.

Một điểm đặc biệt trong dịch mã ở eukaryote là các yếu tố khởi đầu dịch mã không chỉ tương tác với phần đầu 5' của mRNA, mà chúng còn tương tác với phần đuôi poly(A) ở đầu 3' (cầu liên kết 5' – 3') làm mRNA trong dịch mã có dạng vòng tròn (hình 2.11). Sự tương tác này xảy ra giữa yếu tố **eIF4F** (gồm eIF4A, eIF4E, và eIF4G) và **protein liên kết đuôi poly(A) (PABP)**. Chính vì cấu trúc mRNA ở eukaryote trong dịch mã có dạng vòng tròn, nên đuôi polyA mặc dù ở xa lại có ảnh hưởng trực tiếp đến sự khởi đầu dịch mã. Tác động này là do phần lớn các ribosome vừa được giải phóng khi dịch mã kết thúc sẽ có xu hướng được dùng lại ngay cho một chu kỳ dịch mã mới của chính mRNA đó.



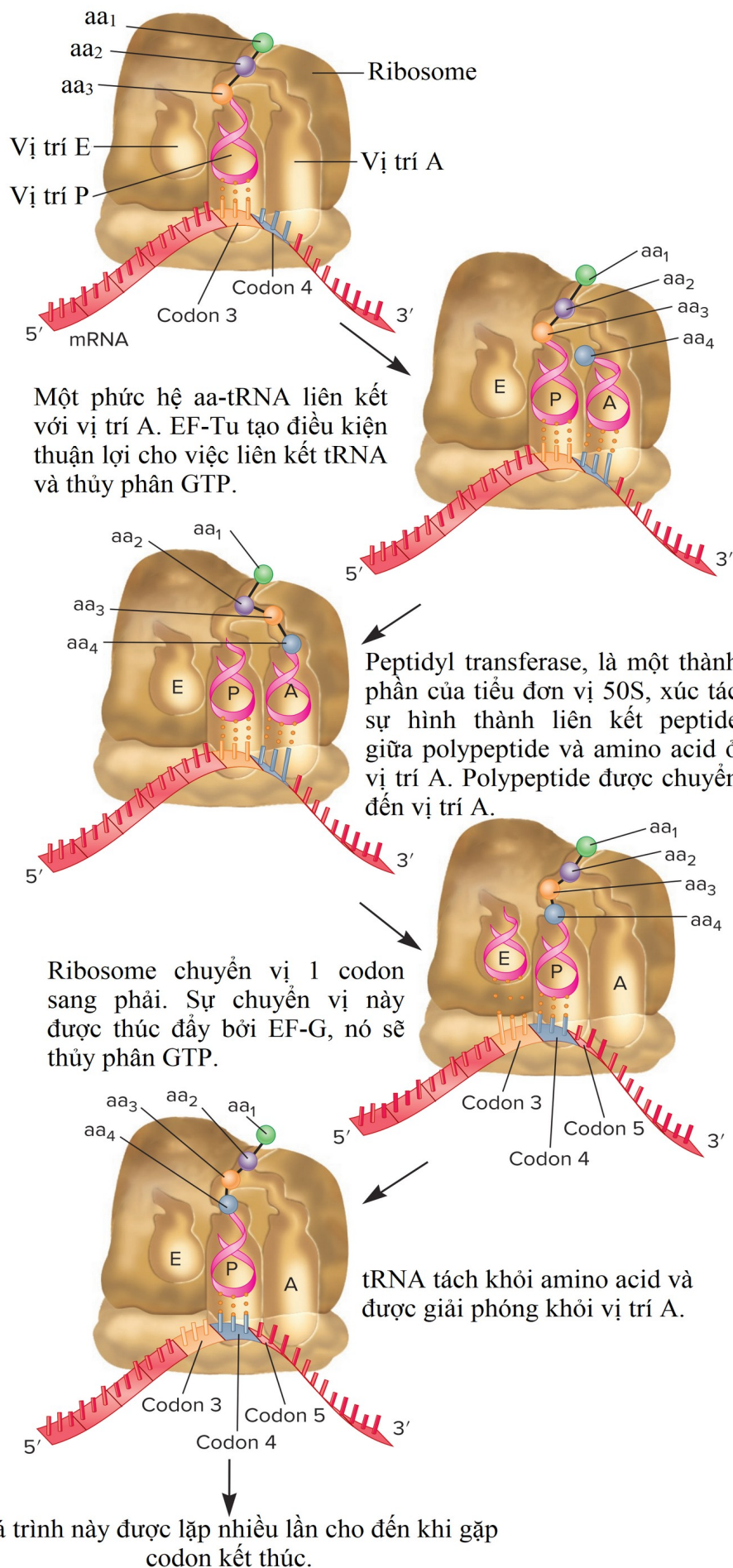
Hình 2.11. Cấu trúc vòng của mRNA làm tăng hiệu quả dịch mã. mRNA của eukaryote hình thành cấu trúc vòng nhờ những tương tác của ba protein chính gắn tại đầu 3' và 5' là PABP, eIF4E và eIF4G. Khi một ribosome hoàn thành dịch mã và tách khỏi đầu 3', các tiểu phần đã phân tách có thể nhanh chóng tìm thấy mũ 5' (m^7G) và đuôi poly(A) gắn PABP gần đó và khởi đầu vòng tổng hợp kế tiếp.

3.2.2. Kéo dài dịch mã

Trong giai đoạn kéo dài của quá trình dịch mã, từng amino acid được thêm vào một polypeptide đang phát triển. Mặc dù quá trình này khá phức tạp, nhưng nó diễn ra với tốc độ đáng kể. Trong điều kiện tế bào bình thường, một polypeptide có thể kéo dài với tốc độ 15–20 amino acid mỗi giây ở vi khuẩn và 2–6 amino acid mỗi giây ở sinh vật nhân chuẩn.

3.2.2.1. Liên kết của tRNA đã gắn amino acid (aminoacyl-tRNA) vào vị trí A của ribosome

Để bắt đầu kéo dài, tRNA mang một amino acid mới đến ribosome để nó có thể được gắn vào phần cuối của polypeptide đang phát triển. Ở phần trên Hình 13.17, mô tả quá trình dịch mã của vi khuẩn, một polypeptide ngắn được gắn vào tRNA nằm ở vị trí P của ribosome. Một tRNA mang một amino acid duy nhất sẽ liên kết với vị trí A. Sự gắn kết này xảy ra do anticodon trong tRNA bổ sung với codon trong mRNA. Sự thủy phân GTP bởi yếu tố kéo dài EF-Tu (hay Tu) cung cấp năng lượng cho sự liên kết của tRNA với vị trí A. Khi aminoacyl - tRNA – Tu-GTP thích hợp được liên kết với codon mRNA, GTP bị thủy phân và phức hợp Tu-GDP được giải phóng, để lại aa-tRNA mới đến nằm ở vị trí A của ribosome. Ngoài ra, 16S rRNA, là một thành phần của tiểu đơn vị nhỏ của ribosome 30S, đóng một vai trò quan trọng bằng cách đảm bảo sự nhận biết chính xác giữa mRNA và tRNA. 16S rRNA có thể phát hiện khi một tRNA không chính xác được liên kết tại vị trí A và sẽ ngăn chặn sự kéo dài cho đến khi tRNA được ghép nối sai được giải phóng khỏi vị trí A. Hiện tượng này, được gọi là **chức năng giải mã (decoding function)** của ribosome, rất quan trọng trong việc duy trì tính chính xác cao của quá trình dịch mã mRNA. Một amino acid không chính xác được kết hợp vào một polypeptide đang phát triển với tỉ lệ khoảng 1 sai sót trên 10.000 amino acid, hoặc 10^{-4} .



Hình 2.12. Giai đoạn kéo dài dịch mã ở vi khuẩn

3.2.2.2. Sự hình thành liên kết peptide

Bước tiếp theo của quá trình kéo dài là một phản ứng được gọi là **chuyển peptidyl (peptidyl transfer)** – polypeptide bị loại bỏ khỏi tRNA ở vị trí P và chuyển đến amino acid ở vị trí A (các tRNA ở vị trí A lúc này gọi là peptidyl-tRNA). Sự chuyển này đi kèm với sự hình thành liên kết peptide giữa amino acid ở vị trí A và polypeptide, kéo dài polypeptide bằng một amino acid. Phản ứng chuyển peptidyl được xúc tác bởi một thành phần của tiểu đơn vị 50S được gọi là **peptidyl transferase**, bao gồm một số protein và rRNA. Điều thú vị là, dựa trên cấu trúc tinh thể của tiểu đơn vị 50S, Thomas Steitz, Peter Moore và các đồng nghiệp của họ đã kết luận rằng 23S rRNA (một thành phần của peptidyl transferase, không phải là protein ribosome) xúc tác sự hình thành liên kết giữa các amino acid liên kế xảy ra trong quá trình chuyển peptidyl. Trong trường hợp này, rRNA hay ribosome là một ribozyme.

3.2.2.3 Ribosome di chuyển dọc theo mRNA hay sự chuyển vị (translocation)

Sau khi phản ứng chuyển peptidyl hoàn tất, ribosome di chuyển sang codon tiếp theo trong mRNA. Điều này sẽ di chuyển các tRNA tại các vị trí P và A đến các vị trí E và P tương ứng. Cuối cùng, tRNA đã tách khỏi amino acid sẽ thoát ra khỏi vị trí E. Tại thời điểm này, tRNA gắn amino acid có thể đi vào vị trí A trống và một chuỗi các bước sẽ thêm amino acid tiếp theo vào polypeptide đang phát triển.

3.2.3. Kết thúc dịch mã

Giai đoạn cuối cùng của quá trình dịch mã, được gọi là kết thúc, xảy ra khi đến codon kết thúc trên mRNA. Ở hầu hết các loài, ba codon kết thúc là UAA, UAG và UGA. Các codon kết thúc không được nhận dạng bởi tRNA có trình tự bổ sung. Thay vào đó, chúng được nhận biết bởi các protein được gọi là các **yếu tố giải phóng (RF)**. Các yếu tố giải phóng có thể liên kết đặc hiệu với trình tự codon kết thúc ở vị trí A của ribosome. Hoạt động xúc tác của các RF giải phóng polypeptide liên kết với tRNA tại vị trí P. Sự giải phóng polypeptide gây ra sự đẩy RF ra khỏi vị trí P và dẫn đến sự phân tách các tiểu đơn vị của ribosome. Ở vi khuẩn, RF1 nhận ra UAA và UAG, và RF2 nhận biết UAA và UGA. Yếu tố giải phóng thứ ba, RF3, cũng cần thiết. Ở sinh vật nhân chuẩn, chỉ một yếu tố giải phóng duy nhất là eRF1 nhận ra cả ba mã kết thúc và eRF3 cũng được cần để kết thúc dịch mã.

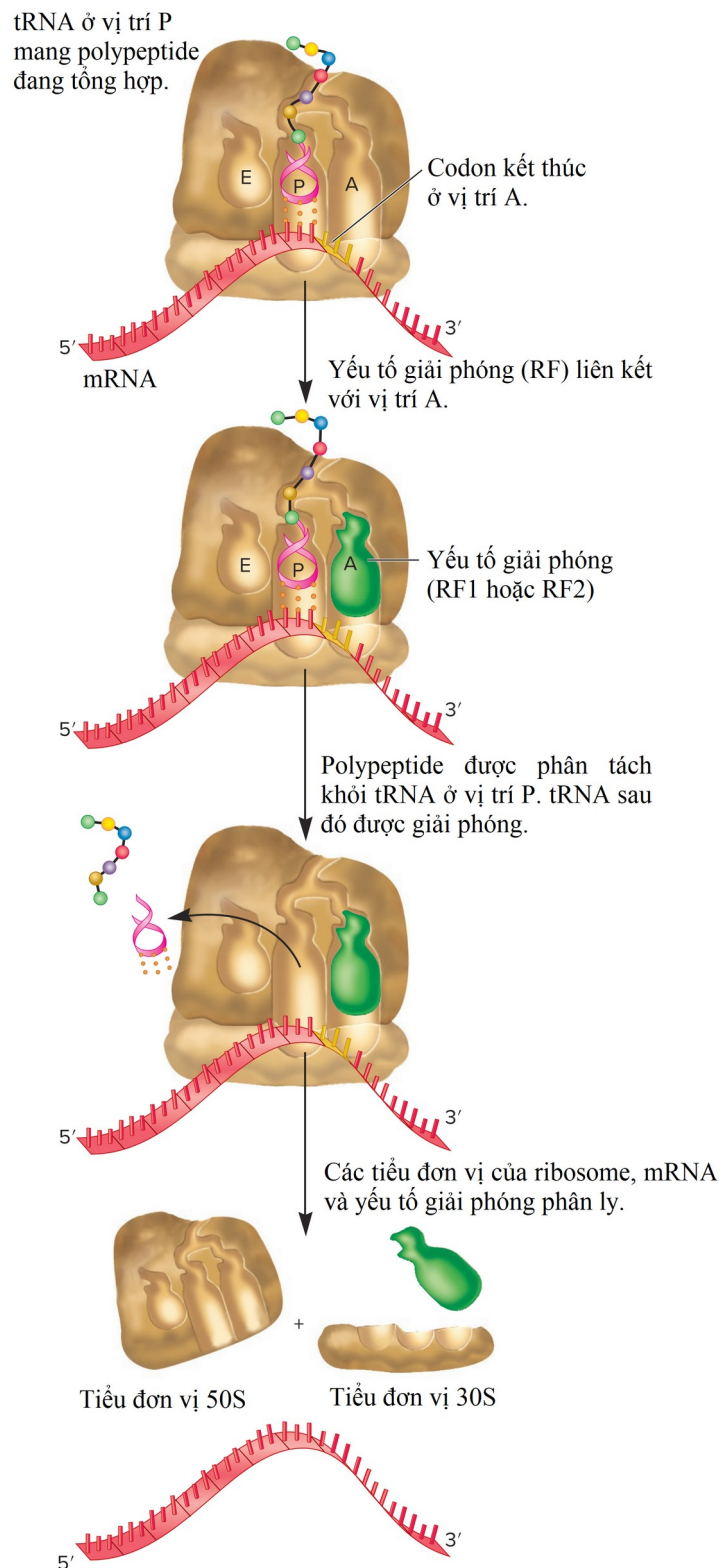
Bảng 2.4: So sánh đơn giản các yếu tố protein dịch mã ở vi khuẩn và eukaryote

Các yếu tố ở vi khuẩn	Các yếu tố ở eukaryote*	Chức năng
Các yếu tố khởi đầu		
	eIF4	Nhận biết mũ 7-methylguanosine đầu 5' của mRNA và tạo điều kiện thuận lợi cho việc liên kết mRNA với tiểu đơn vị bé của ribosome.
IF1, IF3	eIF1, eIF3, eIF6	Ngăn chặn sự liên kết giữa các tiểu đơn vị bé và lớn của ribosome và tạo điều kiện cho sự phân ly của chúng.
IF2	eIF2	Thúc đẩy sự liên kết của tRNA khởi đầu với tiểu đơn vị bé của ribosome.

Các yếu tố ở vi khuẩn	Các yếu tố ở eukaryote*	Chức năng
	eIF5	Giúp phân tách các yếu tố khởi đầu, cho phép hai tiểu đơn vị ribosome kết hợp.
Các yếu tố kéo dài		
EF-Tu	eEF1 α	Liên quan đến sự liên kết của các tRNA vào vị trí A của ribosome.
EF-Ts	eEF1 $\beta\gamma$	Các yếu tố trao đổi nucleotide cần thiết cho hoạt động của EF-Tu và eEF1 α .
EF-G	eEF2	Cần cho sự chuyển vị của ribosome.
Các yếu tố giải phóng		
RF1, RF2	eRF1	Nhận biết codon kết thúc và kích hoạt sự phân cắt chuỗi polypeptide khỏi tRNA.
RF3	eRF3	Các GTPase cũng liên quan đến sự kết thúc.
* Các yếu tố dịch mã của sinh vật nhân thực thường được cấu tạo từ nhiều protein.		

Sau khi chuỗi polypeptide và các yếu tố RF đã rời khỏi ribosome, thì mRNA vẫn còn gắn với ribosome, đồng thời có hai tRNA tự do đang ở các vị trí E và P. Để tham gia được vào các chu kỳ dịch mã mới, các tRNA phải được giải phóng khỏi ribosome, rồi hai tiểu phần của ribosome phải tách nhau ra. Các sự kiện này được gọi là sự phục hồi ribosome.

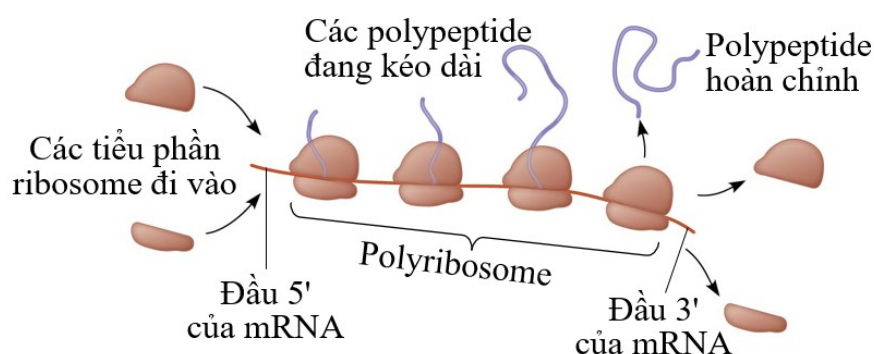
Ở prokaryote, có các yếu tố phục hồi ribosome, kí hiệu là RRF. Các RRF tương tác với các yếu tố EF-G và RF3 theo kiểu giống tRNA để phục hồi ribosome. Cụ thể, sau khi chuỗi polypeptide rời khỏi ribosome, RRF liên kết vào vị trí A (đang trống). Do RRF "giống" tRNA, EF-G liên kết vào một đầu của RRF. Như một phản ứng kéo dài chuỗi, EF-G thúc đẩy sự giải phóng các tRNA đã tách amino acid khỏi ribosome. RF3 lúc này có vai trò giải phóng mRNA và tách hai tiểu phần của ribosome khỏi nhau. Kết quả là tất cả các phân tử tRNA, tiểu phần nhỏ của ribosome (liên kết với RF3) và tiểu phần lớn của ribosome đều tách khỏi nhau và khỏi mRNA. Tất cả các thành phần này đều có thể được dùng lại cho các chu kỳ tổng hợp protein mới.



Hình 2.13. Giai đoạn kết thúc dịch mã ở vi khuẩn. Trong bước đầu tiên, RF1 hoặc RF2 liên kết với codon kết thúc tại vị trí A và RF3 (không được hiển thị) liên kết tại một vị trí khác trên ribosome. Sau khi RF1 (hoặc RF2) và RF3 liên kết, liên kết giữa polypeptide và tRNA bị thủy phân. Sau đó, polypeptide và tRNA được giải phóng khỏi ribosome. Bước cuối cùng trong quá trình kết thúc dịch mã là việc tách rời các tiểu đơn vị của ribosome, mRNA và các yếu tố giải phóng.

4. Polyribosome

Các nhà sinh học tế bào ước tính rằng mỗi tế bào vi khuẩn chứa khoảng 20.000 ribosome, chiếm gần một phần tư khối lượng của tế bào. Số lượng ribosome trên mỗi tế bào nhân thực là thay đổi, nhưng nó cũng lên tới hàng chục nghìn. Một ribosome có thể tổng hợp một chuỗi polypeptide kích thước trung bình trong vòng dưới một phút. Tuy vậy, thường thì sự dịch mã một phân tử mRNA được tiến hành đồng thời bởi nhiều ribosome khác nhau; nghĩa là, một phân tử mRNA có thể được dùng làm khuôn để cùng lúc tạo nên nhiều bản sao của một chuỗi polypeptide. Một khi ribosome đã vượt qua codon bắt đầu, thì một ribosome thứ hai có thể đính kết được vào mRNA; kết quả là một loạt ribosome nối tiếp nhau trượt dọc trên phân tử mRNA. Một chuỗi gồm nhiều ribosome như vậy được gọi là **polyribosome** (hay **polysome**); nó có thể được quan sát thấy bằng kính hiển vi điện tử. Cấu trúc polyribosome có cả ở các tế bào vi khuẩn và tế bào sinh vật nhân thật. Chúng giúp cho tế bào trong một thời gian rất ngắn có thể tổng hợp được nhiều bản sao của một chuỗi polypeptide nhất định.



Hình 2.14. Cấu trúc và hoạt động của ribosome. Mỗi phân tử mRNA thường được dịch mã đồng thời bởi một số ribosome tập hợp thành cụm được gọi là polyribosome.

Trong quá trình hình thành một polysome của sinh vật nhân chuẩn, ribosome đầu tiên tải lên mRNA phải đối mặt với nhiệm vụ khó khăn nhất trong “vòng dịch mã tiên phong” của nó. mRNA đến từ nhân mang nhiều protein: một số trong số này còn sót lại từ các quá trình cắt nối và polyadenylation; các protein liên kết với mRNA khác giúp dẫn mRNA ra khỏi nhân và bảo vệ nó khỏi bị phá hủy. Nhưng hầu như không có chỗ cho bản thân mRNA nằm giữa hai tiểu đơn vị của ribosome, vì vậy những protein này phải được loại bỏ khi mRNA luồn qua ribosome đầu tiên. Những protein này sẽ sớm được thay thế bằng những protein khác cần thiết cho quá trình dịch mã.

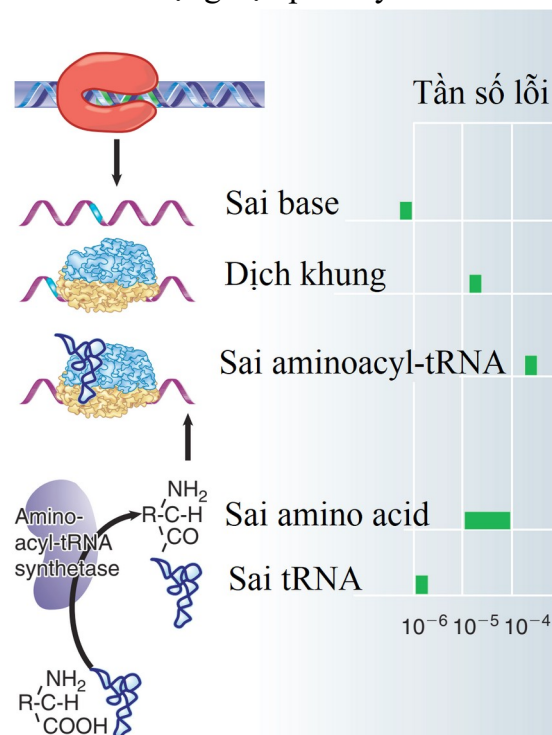
Bảng 2.5: So sánh dịch mã của vi khuẩn và eukaryote

Điểm phân biệt	Vi khuẩn	Eukaryote
Thành phần ribosome	Các ribosome 70S = tiểu đơn vị 30S (21 protein + 1 rRNA) + tiểu đơn vị 50S (34 protein + 2 rRNA).	Các ribosome 80S = tiểu đơn vị 40S (33 protein + 1 rRNA) + tiểu đơn vị 60S (49 protein + 3 rRNA).
Amino acid khởi đầu	<i>N</i> -formyl methionine	Methionine
tRNA khởi đầu	tRNA ^{fMet}	tRNA ^{Met}

Điểm phân biệt	Vi khuẩn	Eukaryote
Sự hình thành phức hợp khởi đầu	Đòi hỏi IF1, IF2 và IF3	Đòi hỏi nhiều yếu tố khởi đầu hơn so với vi khuẩn
Sự liên kết khởi đầu của mRNA với ribosome	Đòi hỏi trình tự Shine-Dalgarno	Đòi hỏi mũ 7-methylguanosine đầu 5'
Lựa chọn codon mở đầu	AUG, GUG, hoặc UUG nằm sau (xuôi dòng) từ trình tự Shine-Dalgarno	Theo quy tắc Kozak (trình tự Kozak)
Tốc độ kéo dài	Thường 15 – 20 amino acids trên giây	Thường 2 – 6 amino acids trên giây
Kết thúc	Đòi hỏi RF1, RF2, và RF3	Đòi hỏi eRF1 và eRF3
Đồng thời với phiên mã	Có	Không

5. Tần số lỗi trong dịch mã

Ribosome có thể mắc hai loại lỗi trong quá trình dịch mã. Nó có thể gây ra dịch khung hình bằng cách bỏ qua một base khi nó đọc mRNA (hoặc, theo hướng ngược lại, bằng cách đọc một base hai lần - một lần là base cuối cùng của một codon và sau đó là base đầu tiên của codon tiếp theo, hoặc hai lần trong cùng một codon). Những lỗi này rất hiếm, xảy ra với tần số $\sim 10^{-5}$. Hoặc, nó có thể cho phép một aminoacyl-tRNA không chính xác bắt cặp (không đúng) với một codon, do đó amino acid sai được kết hợp. Đây có lẽ là lỗi phổ biến nhất trong quá trình dịch mã, xảy ra với tần số $\sim 5 \times 10^{-4}$. Nó được kiểm soát bởi cấu trúc ribosome và động học phân ly.



Hình 2.15. Lỗi xảy ra ở tỷ lệ từ 10^{-6} đến 5×10^{-4} ở các giai đoạn dịch mã khác nhau.

Một aminoacyl-tRNA synthetase có thể mắc hai loại lỗi: nó có thể đặt sai amino acid trên tRNA của nó hoặc nó có thể nạp amino acid của nó bằng tRNA sai. Sự kết hợp sai amino acid phổ biến hơn, có thể là do tRNA cung cấp một bề mặt lớn hơn mà enzyme có thể tạo ra nhiều tiếp xúc hơn để đảm bảo tính đặc hiệu. Các aminoacyl-tRNA synthetase có các cơ chế cụ thể để sửa lỗi trước khi một tRNA nạp sai được giải phóng.

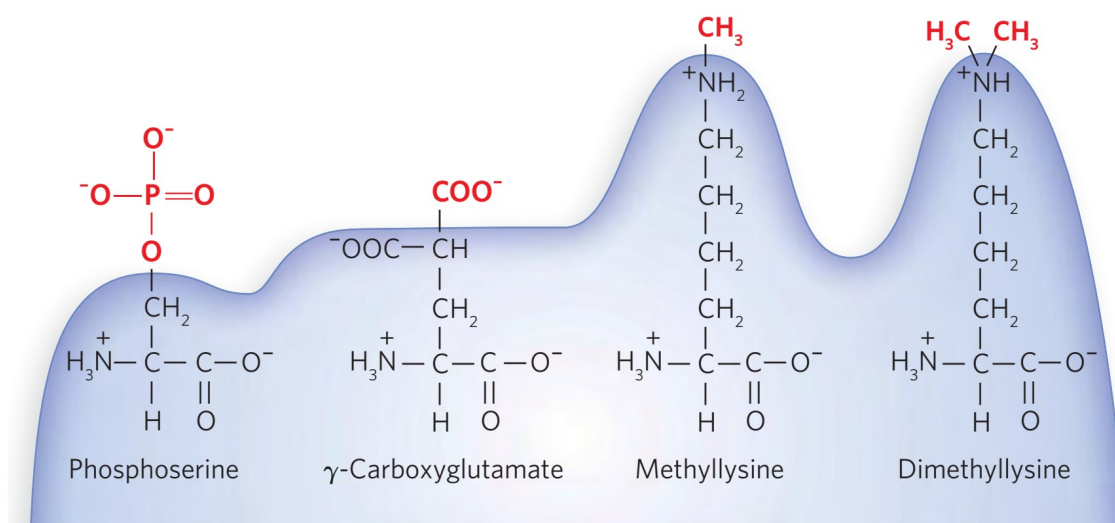
6. Sự hoàn thiện và vận chuyển protein sau dịch mã

6.1. Biến đổi chuỗi polypeptide sau dịch mã

Giống như các bản phiên mã RNA, thường được sửa đổi sau khi chúng được tổng hợp, cấu trúc của chuỗi polypeptide có thể thay đổi khi hoàn thành quá trình dịch mã thông qua một số quá trình khác nhau được gọi chung là **biến đổi sau dịch mã (posttranslational modification)**. Các biến đổi sau dịch mã xảy ra và chúng rất quan trọng đối với khả năng hoạt động của sản phẩm protein cuối cùng. Cấu trúc của chuỗi polypeptide có thể thay đổi khi hoàn thành quá trình dịch mã thông qua một số quá trình khác nhau.

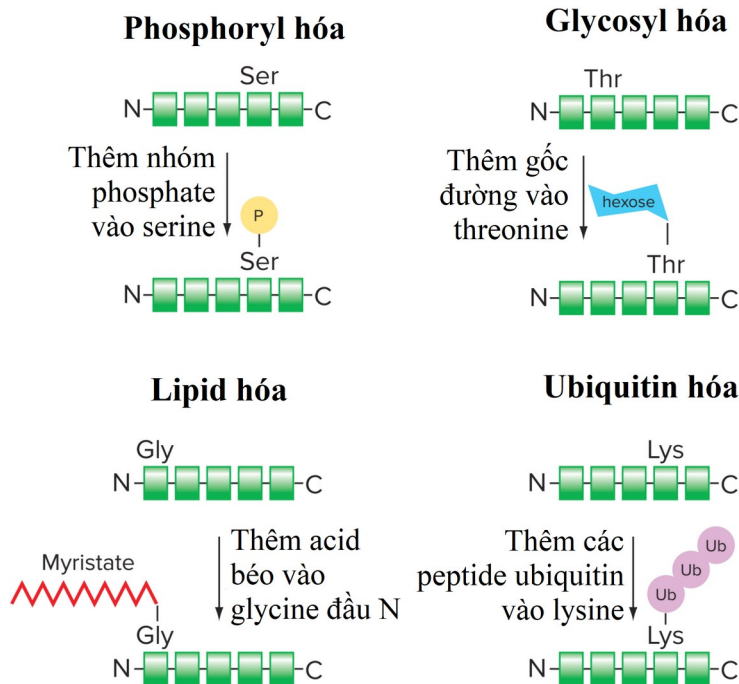
(1) *Amino acid đầu N thường bị loại bỏ hoặc biến đổi.* Ví dụ, nhóm formyl hoặc toàn bộ gốc formylmethionine trong polypeptide của vi khuẩn thường được loại bỏ bằng enzyme. Trong chuỗi polypeptide của sinh vật nhân chuẩn, nhóm amino của gốc methionine ban đầu thường bị loại bỏ và nhóm amino của gốc đầu N có thể bị biến đổi (acetyl hóa).

(2) *Các gốc amino acid riêng lẻ đôi khi được sửa đổi.* Ví dụ, phosphate có thể được thêm vào các nhóm hydroxyl của một số amino acid, chẳng hạn như tyrosine. Những biến đổi như thế này tạo ra các gốc tích điện âm có thể hình thành liên kết ion với các phân tử khác hoặc thay đổi cấu tạo cục bộ của protein. Quá trình phosphoryl hóa cực kỳ quan trọng trong việc điều chỉnh nhiều hoạt động của tế bào và là kết quả của hoạt động của các enzyme gọi là **kinase**. Tại các gốc amino acid khác, nhóm methyl hoặc nhóm acetyl có thể được thêm vào nhờ enzyme, điều này có thể ảnh hưởng đến chức năng của chuỗi polypeptide được biến đổi.



Hình 2.16. Các gốc amino acid được biến đổi. Các gốc amino acid có thể được phosphoryl hóa, carboxyl hóa hoặc methyl hóa để thay đổi chức năng của protein.

(3) *Các chuỗi phụ carbohydrate đôi khi được gắn thêm vào.* Chúng được thêm vào theo phương thức cộng hóa trị, tạo ra các glycoprotein, một loại quan trọng của các phân tử bề mặt tế bào, chẳng hạn như những phân tử xác định các kháng nguyên trong hệ thống nhóm máu ABO ở người. Ở glycoprotein, chuỗi carbohydrate gắn với nhóm -OH của serine hoặc threonine được gọi là oligosaccharide liên kết O, và chuỗi carbohydrate gắn với N. Oligosaccharide liên kết N chứa nhiều nhánh, phổ biến hơn, lớn hơn và phức tạp hơn.

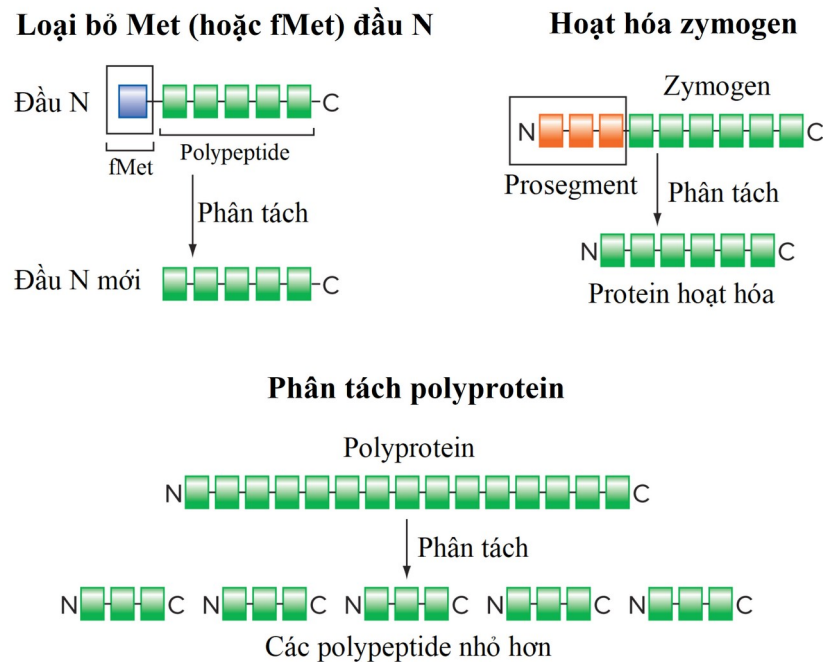


Hình 2.17. Gia công sau dịch mã có thể sửa đổi cấu trúc chuỗi polypeptide bằng cách thêm các nhóm chức. Các enzyme thêm nhiều nhóm chức vào các amino acid đặc hiệu.

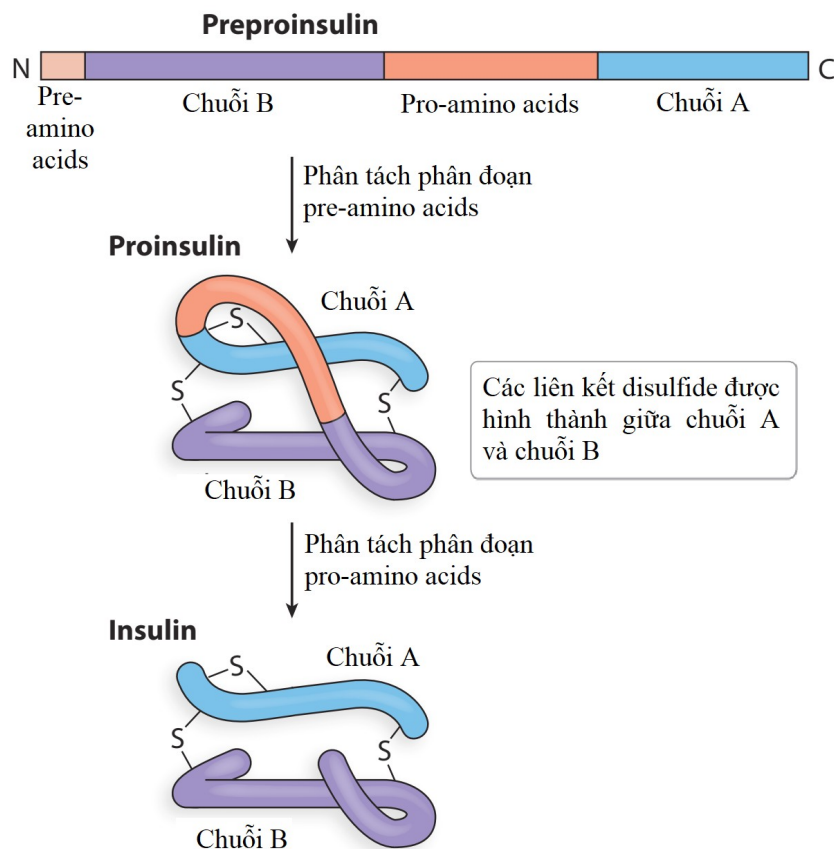
(4) *Chuỗi polypeptide có thể được phân cắt.* Sự phân cắt cũng có thể tạo ra một số polypeptide nhỏ hơn từ một sản phẩm dịch mã lớn hơn. Trong trường hợp này, polypeptide lớn hơn được tạo ra trước khi nó bị phân cắt thường được gọi là **polyprotein**. Ngoài ra, một số protein được tổng hợp ở dạng không hoạt động được gọi là các **zymogen** được hoạt hóa bằng cách phân cắt bằng enzyme loại bỏ đoạn **prosegment** đầu N.

Việc sản xuất hormone insulin, tạo điều kiện thuận lợi cho việc vận chuyển glucose vào tế bào, bao gồm hai bước biến đổi sau dịch mã loại bỏ các đoạn của chuỗi polypeptide ban đầu. Sản phẩm polypeptide được dịch mã từ gene insulin được gọi là preproinsulin. Đó là một protein không hoạt động có chứa một phân đoạn dẫn đầu, được gọi là phân đoạn pre-amino acid (pre-amino acid segment), ở đầu N và một phân đoạn kết nối, được gọi là pro-amino acid (pre-amino acid segment), phân tách phân đoạn chuỗi A và phân đoạn chuỗi B, hai phần chức năng của polypeptide. Trong quá trình xử lý tiền proinsulin sau dịch mã, các pre-amino acids của chuỗi tín hiệu bị loại bỏ, sau khi polypeptide được vận chuyển qua màng tế bào, để tạo thành tiền insulin. Ba liên kết disulfide hình thành bên trong và giữa các phân đoạn chuỗi A và chuỗi B, sau đó là sự phân cắt polypeptide loại bỏ phân đoạn pro-amino acid. Kết quả là một phân tử insulin

chức năng bao gồm 20 amino acid trong phân đoạn chuỗi A và 31 amino acid trong phân đoạn chuỗi B.



Hình 2.18. Gia công sau dịch mã có thể sửa đổi cấu trúc chuỗi polypeptide bằng cách phân tách bởi enzyme. Sự phân tách bởi enzyme có thể loại bỏ một amino acid, phân tách polyprotein, hoặc hoạt hóa zymogen.



Hình 2.19. Ví dụ về biến đổi chuỗi polypeptide sau dịch mã bằng cách phân cắt chuỗi polypeptide. Việc sản xuất hormone insulin bao gồm hai bước biến đổi sau dịch mã loại bỏ các đoạn của chuỗi polypeptide ban đầu (preproinsulin).

(5) **Các peptide đích có thể bị loại bỏ.** Protein thường hoạt động ở những vị trí cụ thể của tế bào như màng sinh chất hoặc một bào quan cụ thể. Trong một quá trình được gọi là **hướng đích protein (protein targeting)**, các protein được hướng đến các đích thích hợp của chúng bằng các trình tự ngắn bên trong (dài 3 - 7 amin acid) được gọi là **peptide đích (target peptides)**, có chức năng giống như “mã bưu điện” cho tế bào. Ví dụ, một peptide đích được gọi là *trình tự tín hiệu (signal sequence)* hướng các protein được dành để tiết ra khỏi tế bào hoặc gắn vào màng đến mạng lưới nội chất hạt (RER), nơi chúng có thể được đưa vào màng hoặc chuyển vị trí vào trong xoang (lumen) cùng lúc chúng đang được tổng hợp bởi các ribosome. Tương tự như vậy, một số protein được hướng đến ty thể bởi một peptide đích được gọi là *tín hiệu hướng đích ty thể (mitochondrial targeting signal)*. Các peptide đích thường được tách ra bằng enzyme sau khi protein đã được chuyển đến vị trí thích hợp của nó.

(6) **Các chuỗi polipeptide gấp nếp thường tạo phức với các nhóm ngoại.** Bậc 3 và bậc 4 của cấu trúc protein thường bao gồm và phụ thuộc vào các yếu tố phi protein được gọi là *các nhóm ngoại (prosthetic groups)*, thường là các vitamin, kim loại hoặc các phân tử chứa kim loại. Bốn nhóm heme chứa sắt có trong huyết sắc tố là một ví dụ điển hình.

6.2. Cuộn gấp và cuộn gấp sai protein

Từ lâu, người ta đã nghĩ rằng sự cuộn gấp của protein là một quá trình tự phát, theo đó một phân tử mạch thẳng thoát khỏi ribosome đạt được cấu trúc ba chiều, ổn định về mặt nhiệt động chỉ dựa trên các đặc tính hóa học kết hợp vốn có trong trình tự amino acid. Trên thực tế đây thực sự là trường hợp của nhiều protein, giờ đây rõ ràng là quá trình chuyển đổi về hình dạng như vậy bắt đầu trước khi polypeptide thoát ra khỏi ribosome. Được gọi là quá trình **cuộn gấp và dịch mã đồng thời (cotranslational folding)**, quá trình này bắt đầu trong đường hầm thoát peptide của ribosome trong quá trình kéo dài.

Tuy nhiên, người ta đã chứng minh rằng đối với các protein khác, việc gấp lại chính xác phụ thuộc vào các thành viên của một họ phân tử gọi là **chaperones**. Bản thân các chaperone là các protein (đôi khi được gọi là các *phân tử chaperone* hoặc các *chaperonin*) làm trung gian cho quá trình gấp nếp theo một trong hai cách. Một số chaperone chỉ đơn giản là liên kết các polypeptide gấp nếp để ngăn chặn sự hình thành của các hình dạng không chính xác. Một số khác đóng vai trò tích cực hơn trong việc cuộn gấp polypeptide bằng cách sử dụng năng lượng ATP để đảm bảo hình dạng phù hợp. Mặc dù các chaperone có thể liên kết với protein, giống như các enzyme, nhưng chúng không trở thành một phần của sản phẩm cuối cùng.

Ở sinh vật nhân thực, chaperone ban đầu được phát hiện ở *Drosophila*, nơi chúng được gọi là các **protein sốc nhiệt (heat-shock proteins)** phản ánh thực tế rằng quá trình gấp protein bị ảnh hưởng bởi nhiệt. Các protein sốc nhiệt được biểu hiện để đáp ứng với nhiệt độ cao nhằm đảm bảo sự gấp nếp thích hợp của protein trong những điều kiện này. Ngày nay chúng ta biết rằng chaperone có mặt trong tất cả các sinh vật và thậm chí được tìm thấy trong ty thể và lục lạp.

Ngay cả khi có chaperone, việc cuộn gấp nhầm vẫn có thể xảy ra và tồn tại thêm một hệ thống “kiểm soát chất lượng”. Các protein gấp sai ở vi khuẩn thường được phân giải

bởi các protease phụ thuộc ATP. Ở sinh vật nhân thực, các protein gấp sai thường được “gắn thẻ” bằng cách gắn cộng hoá trị với một protein nhỏ gọi là **ubiquitin**. Các enzyme được gọi **ubiquitin ligase** nhận ra các protein bị gấp sai và xúc tác cho sự gắn kết của các phân tử ubiquitin. Sau khi một protein được gắn với một số gốc ubiquitin này, nó sẽ trở thành cơ chất cho **proteasome**, một phức hợp protein lớn với hoạt tính protease giải phóng ubiquitin và làm thoái hoá protein bị gấp sai. Ngoài việc loại bỏ các protein bị gấp sai, sự phân hủy protein qua trung gian ubiquitin bởi proteasome đóng một vai trò quan trọng trong quá trình điều hòa sau dịch mã ở sinh vật nhân chuẩn.

Cuộn gập protein là một quá trình cực kỳ quan trọng, không chỉ bởi vì các protein gấp sai có thể không hoạt động, mà còn vì các protein gấp không đúng cách có thể tích tụ và gây bất lợi cho tế bào và các sinh vật chứa chúng. Ví dụ, một nhóm các rối loạn não có thể lây truyền ở động vật có vú – **bệnh scrapie** ở cừu, **bệnh não xốp ở bò (bovine spongiform encephalopathy, BSE) (bệnh bò điên)** ở gia súc và **bệnh Creutzfeldt-Jakob (CJD)** ở người - được gây ra bởi sự hiện diện trong não của tập hợp các protein bị cuộn gập sai được gọi là **prion**. Protein bị gấp sai (được gọi là PrP^{Sc}) là một phiên bản thay đổi của protein bình thường (được gọi là PrP^{C}) được tổng hợp trong các tế bào thần kinh và được tìm thấy trong não của tất cả các động vật có vú trưởng thành. Sự khác biệt giữa PrP^{C} và PrP^{Sc} nằm ở cấu trúc protein bậc 2 của chúng. PrP^{C} bình thường, không lây nhiễm gấp lại thành một protein có cấu trúc chủ yếu là xoắn α , trong khi đó PrP^{Sc} lây nhiễm gấp lại thành một protein có số lượng lớn hơn cấu trúc phiên gập nếp β . Khi một phân tử PrP^{Sc} bất thường tiếp xúc với một phân tử PrP^{C} , protein bình thường sẽ cuộn lại thành cấu trúc bất thường. Quá trình này tiếp tục như một phản ứng dây chuyền, với những kết quả có khả năng tàn phá cao - sự hình thành các cụm prion cuối cùng sẽ phá hủy não bộ. Do đó, nhóm rối loạn này có thể được coi là bệnh của cấu trúc protein bậc 2.

Hiện nay, nhiều phòng thí nghiệm đang nghiên cứu về sự gập nếp và gập nếp sai của protein, đặc biệt là liên quan đến di truyền học. Nhiều chứng rối loạn di truyền ở người là do các protein bị sai lệch tạo thành các khối/cụm kết tập. Bệnh thiếu máu hồng cầu hình liềm, là một trường hợp điển hình, trong đó các chuỗi β của hemoglobin bị thay đổi do thay đổi một amino acid duy nhất, khiến các phân tử tập hợp lại trong hồng cầu, gây hậu quả nghiêm trọng. Một dạng bệnh Creutzfeldt-Jakob di truyền trội trên nhiễm sắc thể thường được biết đến trong đó đột biến làm thay đổi trình tự amino acid PrP, dẫn đến sự hình thành prion. Các bệnh thoái hóa thần kinh tiến triển khác nhau như **bệnh Huntington**, **bệnh Alzheimer** và **bệnh Parkinson** cũng liên quan đến sự hình thành các tập hợp protein bất thường trong não. Bệnh Huntington được di truyền như một đặc điểm trội trên nhiễm sắc thể thường, trong khi các thành phần di truyền ít được xác định rõ ràng hơn có liên quan đến bệnh Alzheimer và Parkinson.

6.3. Phân loại protein trong tế bào

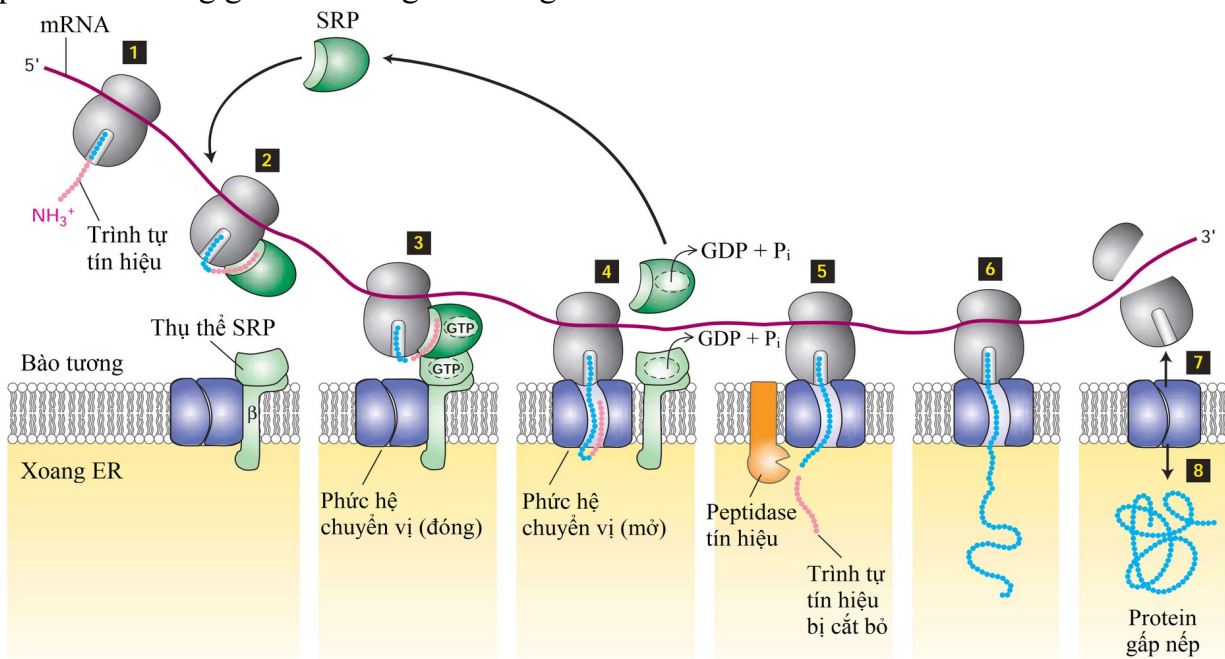
Ở vi khuẩn và sinh vật nhân thực, một số protein có thể được tiết ra khỏi tế bào; và ở sinh vật nhân chuẩn, một số protein khác phải được đặt trong các xoang khác nhau của tế bào, chẳng hạn như nhân, ti thể, lục lạp và lysosome. Việc sắp xếp các protein vào các xoang thích hợp của chúng nằm dưới sự kiểm soát di truyền, trong đó các trình tự “tín hiệu” hay trình tự “dẫn đầu” đặc hiệu trên các protein sẽ hướng chúng đến các bào quan

chính xác. Tương tự, ở vi khuẩn, một số protein nhất định trở nên khu trú trong màng và những protein khác được tiết ra.

Các protein tiết ở tế bào nhân thực được chuyển qua lưới nội chất (ER) và bộ máy Golgi. Vào năm 1975, Günther Blobel, B. Dobberstein và các đồng nghiệp đã phát hiện ra rằng các protein được tiết ra và các protein khác được sắp xếp bởi Golgi ban đầu có chứa thêm các amino acid ở đầu amin. Công trình của Blobel đã dẫn đến **giả thuyết tín hiệu**, trong đó nói rằng các protein được sắp xếp bởi Golgi liên kết với ER bằng một phần mở rộng đầu amin kỵ nước (**trình tự tín hiệu**) tới màng mà sau đó bị loại bỏ và phân hủy. Blobel đã giành giải Nobel Sinh lý học hoặc Y học năm 1999 cho công trình này.

Trình tự tín hiệu của protein hướng đến ER bao gồm khoảng 15 đến 30 amino acid đầu N. Khi trình tự tín hiệu được tạo ra bằng cách dịch mã và tiếp xúc trên bề mặt ribosome, một **hạt nhận biết tín hiệu** (SRP, một phức hợp RNA-protein) trong tế bào chất sẽ liên kết với trình tự và ngăn chặn quá trình dịch mã mRNA tiếp theo cho đến khi phức hợp chuỗi polypeptide đang phát triển – SRP – ribosome – mRNA tới và liên kết với ER. SRP liên kết với **thụ thể SRP** trong màng ER, gây ra sự liên kết chắc chắn của ribosome với ER, giải phóng SRP và tiếp tục dịch mã. Polypeptide đang phát triển kéo dài qua màng ER vào không gian bên trong của ER.

Một khi trình tự tín hiệu được đưa hoàn toàn vào không gian xoang bên trong của ER, nó sẽ bị loại bỏ khỏi polypeptide bởi enzyme **peptidase tín hiệu**. Khi polypeptide hoàn chỉnh nằm hoàn toàn trong không gian các kênh ER, nó thường được sửa đổi bằng cách bổ sung các nhóm cacbohydrate cụ thể để tạo ra glycoprotein. Các glycoprotein sau đó được chuyển trong các túi đến bộ máy Golgi, nơi mà hầu hết quá trình phân loại diễn ra. Ví dụ, các protein dự trữ tiết ra được đóng gói thành các túi dự trữ bài tiết, các túi này sẽ di chuyển lên bề mặt tế bào, nơi chúng kết hợp với màng sinh chất và giải phóng các protein đã đóng gói của chúng ra bên ngoài tế bào.



Hình 2.20. Cơ chế chuyển vị đưa protein vào ER kèm dịch mã. Bước ①, ②: sau khi thò ra khỏi ribosome, trình tự tín hiệu ER gắn với hạt nhận biết tín hiệu (SRP). Bước ③:

SRP đưa phức hệ ribosome – polypeptide đang sinh tới thụ thể SRP nằm trên màng ER. GTP gắn với SRP và thụ thể SRP làm bền tương tác này. Bước ④: phức hệ ribosome – polypeptide đang sinh chuyển tới translocon (một phức hệ protein tạo thành kênh dẫn qua màng ER) hay phức hệ chuyển vị làm mở kênh dẫn chuyển vị này và đưa trình tự tín hiệu cùng như phân đoạn nối liền với nó vào lỗ trung tâm. Sau khi tách khỏi translocon, SRP và thụ thể SRP thủy phân GTP gắn với chúng để sẵn sàng gắn với chuỗi polypeptide khác. Bước ⑤: Khi kéo dài, chuỗi polypeptide xuyên qua kênh dẫn translocon vào xoang ER, nơi trình tự tín hiệu bị peptidase tín hiệu cắt và nhanh chóng phân hủy. Bước ⑥: Chuỗi polypeptide tiếp tục kéo dài khi mRNA dịch mã tới đầu 3'. Vì ribosome gắn với translocon nên chuỗi đang tổng hợp bị đẩy qua translocon vào xoang ER. Bước ⑦, ⑧: Khi quá trình dịch mã hoàn tất, ribosome được giải phóng, phần còn lại của protein được kéo vào xoang ER, translocon đóng lại và protein thu được cấu hình gấp nếp tự nhiên. Trong ER và sau đó là trong Golgi, Protein này tiếp tục được biến đổi và hoàn thiện. Cuối cùng các nang vận chuyển sẽ vận chuyển nó đến màng nguyên sinh chất và tiến hành xuất bào.

Hầu hết protein của ty thể và lục lạp do gene trong nhân mã hoá, tổng hợp trên ribosome của bào tương và tiến tới các bào quan này theo cơ chế hấp thụ sau dịch mã. Mọi thông tin cần thiết để đưa tiền protein từ bào tương tới chất nền ty thể hoặc chất nền lục lạp nằm trong trình tự tín hiệu hay trình tự hướng đích (targeting sequence) đầu N. Vì ty thể và lục lạp chứa hai lớp màng và nhiều khoảng trống giữa các lớp nên để protein hướng tới vị trí chính xác thường cần hai trình tự hướng đích và hai hệ thống chuyển vị gắn màng: một hướng protein tới bào quan và một đưa protein này tới đúng ngăn phụ hay màng bào quan. Sau khi protein đã hấp thụ, trình tự này bị protease trong chất nền cắt bỏ.

Mọi protein của peroxisome đều tổng hợp trên ribosome trong bào tương và tích hợp vào bào quan sau khi dịch mã. Hầu hết protein chất nền peroxisome chứa trình tự hướng đích PTS1 đầu C; một số ít chứa trình tự hướng đích PTS2 đầu N. Cả hai trình tự hướng đích này đều không bị cắt bỏ sau khi protein được hấp thụ.

Protein nhập vào hay xuất khỏi nhân mang một trình tự amino acid đặc hiệu có chức năng giống tín hiệu hướng nhân (NLS, nuclear-localization signal) hoặc tín hiệu xuất nhân (NES, nuclear-export signal). Protein chỉ nằm trong nhân mang NLS nhưng không mang NES, trong khi protein chuyển dịch qua lại giữa nhân và bào tương mang cả hai tín hiệu. Hiện đã xác định được nhiều loại NES và NLS khác nhau. Các nhà khoa học cho rằng mỗi loại tín hiệu vận chuyển nhân tương tác với một protein vận chuyển vào nhân đặc hiệu thuộc một họ các protein tương đồng

Bảng 2.6: Trình tự hướng hấp thụ đưa protein từ bào tương tới bào quan

Bào quan đích	Vị trí của trình tự trong protein	Cắt bỏ trình tự	Bản chất của trình tự
ER (xoang)	Đầu N	Có	Vùng lõi gồm 6 – 12 amino acid kỵ nước, phía trước đó thường là một hoặc nhiều amino acid có tính base (Arg, Lys).

Bào quan đích	Vị trí của trình tự trong protein	Cắt bỏ trình tự	Bản chất của trình tự
Ty thể (chất nền)	Đầu N	Có	Xoắn lưỡng tính, dài 20 – 50 amino acid, với các amino acid tích điện dương (Arg và Lys) ở một phía và các amino acid kỵ nước ở phía kia.
Lục lạp (stroma)	Đầu N	Có	Không có motif chung; thường giàu Ser, Thr và các amino acid kỵ nước nhỏ, ít chứa Glu, Asp.
Peroxisome (chất nền)	Đầu C (hầu hết protein); đầu N (một số protein)	Không	Tín hiệu PTS1 (Ser-Lys-Leu) tại tận cùng đầu C; tín hiệu PTS2 tại đầu N.
Nhân (chất nhân)	Khác nhau	Không	Nhiều loại khác nhau; motif chung chứa một đoạn ngắn giàu Lys và Arg

7. Các kháng sinh và độc tố thường nhắm vào chu trình tổng hợp protein

Nhiều bệnh khác nhau ảnh hưởng đến con người và vật nuôi là do vi khuẩn gây bệnh gây ra. **Kháng sinh** là bất kỳ chất nào được tạo ra bởi một vi sinh vật có tác dụng ức chế sự phát triển của các vi sinh vật khác, chẳng hạn như vi khuẩn gây bệnh. Hầu hết các chất kháng sinh là các phân tử hữu cơ nhỏ, với khối lượng nhỏ hơn 2000 dalton (Da). Trong một số trường hợp, thuốc kháng sinh phát huy tác dụng vì chúng ức chế hoặc cản trở quá trình dịch của vi khuẩn. Do các thành phần của dịch mã có phần khác nhau giữa vi khuẩn và sinh vật nhân thực, nên một số kháng sinh ức chế quá trình dịch mã của vi khuẩn mà không ảnh hưởng đến quá trình dịch mã ở sinh vật nhân thực. Do đó, chúng có thể được sử dụng để điều trị các bệnh nhiễm trùng do vi khuẩn ở người và vật nuôi.

Puromycin, được tạo ra bởi nấm mốc *Streptomyces alboniger*, là một trong những loại kháng sinh ức chế được hiểu biết rõ nhất. Cấu trúc của nó rất giống với đầu 3' của aminoacyl-tRNA, vì vậy puromycin có thể liên kết với vị trí A của ribosome và tham gia vào sự hình thành liên kết peptide, tạo ra peptidyl-puromycin. Tuy nhiên, do puromycin chỉ giống với đầu 3' của tRNA nên nó không tham gia vào quá trình chuyển vị và tách khỏi ribosome ngay sau khi được liên kết với đầu C của peptide. Điều này làm sớm dừng tổng hợp protein.

Tetracycline ức chế cả ribosome của vi khuẩn và sinh vật nhân chuẩn. Chúng liên kết với rRNA 16S (hoặc 18S) của tiểu đơn vị nhỏ và chặn sự gắn kết của aminoacyl-tRNA. Mặc dù ức chế cả hai loại ribosome, nhưng tetracycline ức chế vi khuẩn tốt hơn do vi khuẩn tích cực hấp thụ chúng, trong khi các tế bào nhân thực tích cực đào thải chúng. **Chloramphenicol** ức chế quá trình tổng hợp protein bởi các ribosome của vi khuẩn (và ty thể và lục lạp) bằng cách ngăn chặn quá trình chuyển peptidyl; nó không ảnh hưởng đến quá trình tổng hợp protein tế bào ở sinh vật nhân chuẩn. Ngược lại, **cycloheximide** ngăn chặn peptidyl transferase của ribosome sinh vật nhân thực 80S nhưng không ngăn chặn ribosome 70S của vi khuẩn (hoặc ty thể và lục lạp).

Streptomycin, một trisaccharide, gây độc sai mã di truyền (ở vi khuẩn) ở nồng độ tương đối thấp và ức chế sự khởi đầu ở nồng độ cao hơn.

Fusidic acid là một dẫn xuất steroid bám vào yếu tố kéo dài dịch mã của prokaryote EF-G. Khi có mặt của fusidic acid, EF-G cùng với GDP bám của nó, bị đông cứng trên ribosome. Fusidic acid cũng ức chế các yếu tố kéo dài EF2 eukaryote tương ứng; tuy nhiên trong thực nghiệm, các tế bào động vật không bị ảnh hưởng khi sử dụng kháng sinh này.

Một số chất ức chế tổng hợp protein khác đáng chú ý vì độc tính của chúng đối với con người và các động vật có vú khác. **Độc tố bạch hầu** là một enzyme xúc tác quá trình ADP-ribosyl hóa (ADP-ribosylation: bổ sung một hoặc nhiều gốc ADP-ribose vào protein) của gốc diphthamide (một histidine biến đổi) của yếu tố kéo dài ở sinh vật nhân thực eEF2, do đó làm bất hoạt nó. Độc tố được tiết ra bởi vi khuẩn *Corynebacterium diphtheriae*. Những người bị nhiễm bệnh bị sốt, ớn lạnh, sưng cổ và nhịp tim nhanh. Mặc dù phổ biến trong lịch sử, bệnh bạch hầu phần lớn đã được xóa sổ ở các nước phát triển thông qua tiêm chủng rộng rãi. Trong số các độc tố do thực vật tạo ra, **ricin**, một loại protein cực độc của hạt thầu dầu, làm bất hoạt tiểu đơn vị 60S của ribosome nhân chuẩn bằng cách khử một adenosine đặc hiệu trong rRNA 23S. Bảng 2.4 liệt kê một số kháng sinh và độc tố nhắm vào dịch mã, hậu quả chức năng của chúng, và một số ứng dụng lâm sàng và các ứng dụng khác.

Bảng 2.7: Một số kháng sinh và độc tố có thể ức chế dịch mã

Lớp	Ví dụ	Đích	Cơ chế tác động
Ức chế khởi đầu dịch mã			
Edeine	Edeine	Vi khuẩn và tế bào nhân thực	Liên kết các vị trí P và E của tiểu đơn vị nhỏ, làm mất ổn định tương tác giữa mRNA và chất khởi đầu aminoacyl-tRNA (ngăn cản sự gắn kết aminoacyl-tRNA của chất khởi đầu)
Kasugamycin	Kasugamycin	Vi khuẩn và nấm; độc tính thấp với các sinh vật nhân chuẩn bậc cao hơn	Liên kết các vị trí P và E của tiểu đơn vị nhỏ, làm mất ổn định tương tác giữa mRNA và aminoacyl-tRNA khởi đầu (ngăn cản sự liên kết của aminoacyl-tRNA khởi đầu)
Orthosomycins	Avilamycin, Evernimicin	Vi khuẩn	Liên kết tiểu đơn vị 50S, ngăn chặn sự liên kết với phức hợp khởi đầu
Pactamycin	Pactamycin	Vi khuẩn và tế bào nhân thực	Ức chế sự khởi đầu và kéo dài; cơ chế chưa biết rõ
Ức chế kéo dài dịch mã			

Lớp	Ví dụ	Đích	Cơ chế tác động
Aminoglycosides	Gentamicin, Hygromycin B, Kanamycin, Neomycin, Naromomycin	Vì khuẩn và tế bào nhân thực	Thúc đẩy các lỗi trong giải mã bằng cách ổn định các sự bất cặp codon-anticodon không chính xác
Aminoglycosides	Streptomycin	Vì khuẩn	Thúc đẩy lỗi giải mã bằng cách thay đổi tốc độ hoạt hoá GTPase trong EF-Tu; cản trở sự kết cặp bình thường giữa aminoacyl-tRNA và codon, điều này gây ra việc đọc sai, từ đó tạo ra các protein bất thường.
Amphenicols	Azidamfenicol, Chloramphenicol , Florfenicol, Thiamphenico	Vì khuẩn	Liên kết vào vị trí A, ức chế phản ứng peptidyl transferase (ngăn chặn sự kéo dài bằng cách hoạt động như chất ức chế cạnh tranh của peptidyl transferase)
Enacyloxins và kirromycins	Aurodox, Azdimucin, Delvomycin, Efrotomycin, Heneicomycin, Kirromycin, Mocimycin,	Vì khuẩn	Dừng phức hợp tam nguyên (eIF2–GTP–Met-tRNA ^{Met}) trên ribosome bằng cách ngăn chặn sự thay đổi về hình dạng ở EF-Tu
Lincosamides	Clindamycin, Lincomycin	Vì khuẩn	Liên kết với vị trí A
Macrolides	Azithromycin, Carbomycin, Clarithromycin, Erythromycin, Roxithromycin, Spiramycin, Telithromycin, Tylosin	Vì khuẩn	Liên kết đường hầm thoát ra của ribosome (ribosomal exit tunnel), ngăn chặn sự thoát ra của polypeptide từ ribosome.
Oxazolidinones	Cycloserine, Linezolid, Ranbezolid	Vì khuẩn	Liên kết với vị trí A, cơ chế chính xác chưa được biết rõ
Pleuromutilin	Retapamulin, Tiamulin,	Vì khuẩn	Ức chế khởi đầu và phản ứng peptidyl transferase bằng cách liên kết với các vị trí A và P

Lớp	Ví dụ	Đích	Cơ chế tác động
	Valnemulin		cùng một lúc
Sparsomycin	Sparsomycin	Vì khuẩn và tế bào nhân thực	Ức chế tRNA liên kết với vị trí A và tăng cường liên kết tRNA với vị trí P
Streptogramin	Dalfopristin, Pristinamycin, Quinupristin, Streptogramin A, Streptogramin B	Vì khuẩn	Liên kết các vị trí liên kề, riêng biệt trên trung tâm peptidyl transferase và hoạt động phối hợp để chặn cả hai vị trí A và P trong quá trình khởi đầu và kéo dài
Tetracyclines	Chlortetracycline, Demeclocycline, Doxycycline, Lymecycline, Meclocycline, Methacycline, Minocycline, Oxytetracycline, Rolitetracycline, Tetracycline	Vì khuẩn	Liên kết với tiểu đơn vị 30S, ngăn phức hợp tam nguyên liên kết với ribosome (ngăn chặn sự kéo dài bằng cách ức chế sự liên kết của aminoacyl-tRNA với ribosome)
Ức chế sự chuyển vị (sự di chuyển của ribosome trên mRNA)			
Aminoglycosides	Spectinomycin	Vì khuẩn	Ức chế chuyển vị bằng cách ổn định một chất trung gian
Fusidic acid (steroid)	Fusidic acid	Vì khuẩn	Ngăn chặn sự phân ly EF-G bằng cách liên kết với EF-G-GTP trong phức hợp với ribosome
Ricin (protein)	Ricin	Vì khuẩn và tế bào nhân thực	Loại bỏ 23S rRNA, có thể làm gián đoạn hoạt tính kích thích của GTPase
Thiopeptides	Micrococcin, Nosiheptide, Thiostrepton	Vì khuẩn	Liên kết với vị trí A, ức chế liên kết tRNA-IF-Tu và liên kết EF-G (ức chế khởi đầu và chuyển vị)
Tuberactinomycin s	Capreomycin, Enviomycin, Spectinomycin, Viomycin	Vì khuẩn	Ức chế chuyển vị bằng cách ổn định một chất trung gian

PHẦN B: CÂU HỎI VÀ BÀI TẬP

I. CÂU HỎI VÀ BÀI TẬP CÓ HƯỚNG DẪN TRẢ LỜI

Câu 1:

Dưới đây là trình tự DNA sợi kép của vi khuẩn (*E. coli*) mã hóa cho một loại protein giả định. Cả hai mạch được thể hiện, sợi phía trên đọc theo chiều 5' đến 3' từ trái sang phải, trong khi sợi dưới cùng đọc theo chiều 5' đến 3' từ phải sang trái. Các nucleotide được đánh số từ 1 đến 100. Quá trình phiên mã bắt đầu từ vị trí cặp base C/G (sợi trên/sợi dưới) màu đỏ và được gạch dưới và RNA polymerase tiến hành từ trái sang phải dọc theo DNA.

```

      1           20           40
5' -GTGTCCGTCTAATATTGTGAGATGTTATATTCCCGCCGTCAACACCATCAA-3'
      - - - - - + - - - - - + - - - - - + - - - - - + - - - - -
3' -CACAGGCAGATTATAA CACTCTACAATATAGGGCGGCAGTTGTGGTAGTT-5'

      60           80           100
5' -ACAGGATTAATCGCCTGCTGGGGCAAAGGCGGTGAAGGTAAAGGTGTTGCC-3'
      - - - - - + - - - - - + - - - - - + - - - - - + - - - - -
3' -TGTCTTATTAGCGGACGACCCCGTTTCCGCCACTTCCATTTCACACACGG-5'

```

- Mạch nào của DNA được dùng làm khuôn để phiên mã (mạch phía trên hay phía dưới)?
- Vị trí của promoter so với vị trí bắt đầu phiên mã?
- 15 nucleotide đầu tiên của mRNA thu được là gì? Cho biết đầu 5' và 3' của mRNA.
- 5 amino acid đầu tiên được dịch mã từ mRNA thu được là gì? Cho biết đầu tận cùng amino (NH_3^+) và carboxyl (COO^-) của protein.
- Các nucleotide được gạch chân **TAA** (được chỉ định bằng màu xanh lam) có mã hóa một codon kết thúc cho tổng hợp protein không? Giải thích ngắn gọn.
Xét các tình huống trong các phần f - h một cách độc lập với nhau.
- Một đột biến xảy ra dẫn đến việc chèn thêm một cặp gốc G/C (mạch trên/mạch dưới) ngay sau cặp base 11 (được in đậm). Đột biến thêm này sẽ có ảnh hưởng gì đến bản phiên mã mRNA và protein thu được?
- Một đột biến khác dẫn đến việc thay thế cặp base T/A ở vị trí 30 (được in đậm và gạch chân) bằng một cặp base G/C. Đột biến này sẽ ảnh hưởng như thế nào đến trình tự của protein được tạo ra?
- Đột biến thứ ba xảy ra dẫn đến sự thay thế cặp base C/G ở vị trí 42 (được in nghiêng đậm) thành cặp base T/A. Đột biến này sẽ ảnh hưởng như thế nào đến trình tự của protein được tạo ra?

Hướng dẫn trả lời

- Mạch phía dưới.
- Promoter sẽ nằm bên trái vị trí bắt đầu phiên mã.
- 5' CUAAUAUUGUGAGAU 3'
- N Met-Leu-Tyr-Pro-Ala C
- Không. TAA được gạch chân không được đọc là TAA do khung đọc. Trình tự GATAAT tạo thành các codon: GAU AAU.

- f) mRNA sẽ dài hơn một nucleotide, nhưng vì phần thêm vào xảy ra trước codon khởi đầu nên protein không thay đổi.
- g) Codon UAU mã hóa Tyr, nhưng bây giờ nó là UAG, một codon kết thúc. Protein bị cắt ngắn.
- h) Codon AAC mã hóa Asn, bây giờ là AAU, cũng mã hóa Asn. Protein không thay đổi.

Câu 2:

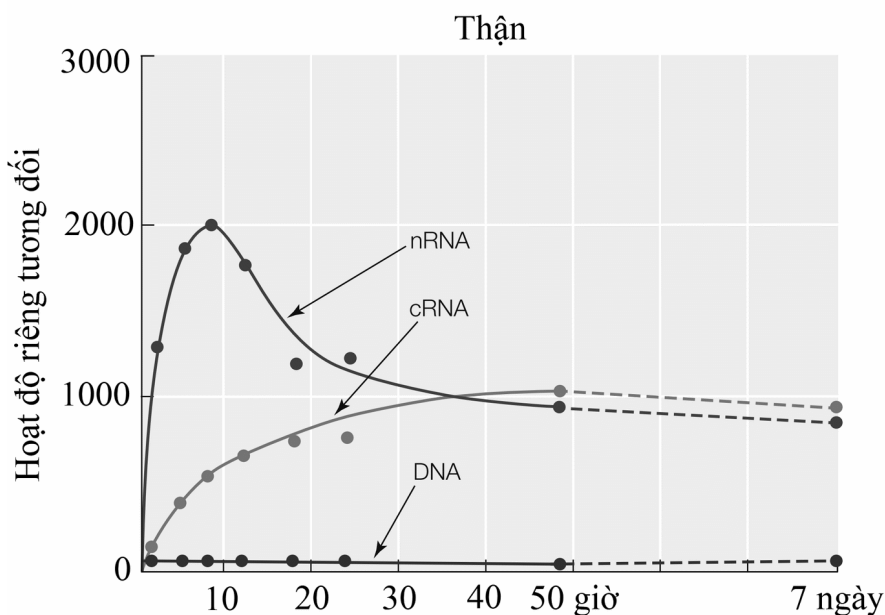
Chức năng của RNA polymerase và poly(A) polymerase giống nhau và khác nhau thế nào?

Hướng dẫn trả lời

Cả 2 enzyme đều gắn ribonucleotide vào chuỗi RNA đang kéo dài. PolyA polymerase không cần DNA làm khuôn và bổ sung một loại ribonucleotide duy nhất là Adenine (tới khoảng 200A) vào đầu 3' tận cùng của mRNA. Trong khi đó, RNA polymerase cần DNA làm mạch khuôn, bổ sung cả 4 loại ribonucleotide A, T, G và C vào mạch mRNA đang mở rộng theo nguyên tắc bổ sung.

Câu 3:

Các con thỏ được tiêm [^{32}P]phosphate; DNA, RNA nhân (nRNA) và RNA tế bào chất (cRNA) được tách chiết từ thận và các hoạt độ riêng của chúng (độ phóng xạ/mg nucleic acid) đã được xác định sau 2, 5, 8, 12, 18, 24, 48 và 168 giờ sử dụng đồng vị phóng xạ.



Hình C3

- a) Quá trình nào đã được nghiên cứu trong thí nghiệm này?
- b) Em có thể kết luận gì về quá trình phân chia của tế bào thận trong thời gian sinh nhân?
- c) Có thể rút ra kết luận gì từ các đường cong RNA nhân (nRNA) và RNA tế bào chất (cRNA)?
- d) Tại sao các hoạt động cụ thể của nRNA và cRNA không giảm đáng kể sau 2 ngày đánh dấu?

Hướng dẫn trả lời

- a) Nghiên cứu sao chép DNA, tổng hợp RNA và chế biến RNA.

- b) Phosphate phóng xạ được sử dụng/gắn vào các phân tử RNA mới được tổng hợp, nhưng không có trong DNA. Do đó, các tế bào thận không phân chia như hầu hết các tế bào trong cơ thể.
- c) Quá trình tổng hợp các phân tử RNA diễn ra trong nhân, sau đó chúng được vận chuyển ra tế bào chất.
- d) Phần lớn các phân tử RNA tế bào chất là rRNA và thời gian bán hủy của chúng khá dài.

Câu 4:

Một đột biến được tìm thấy trong gene mã hóa tRNA. Loại allele kiểu dại đã tạo ra một tRNA nhận ra codon GAA và được nạp amino acid là glutamic acid (Glu). tRNA đột biến vẫn được nạp Glu, nhưng anticodon bị đột biến sao cho nó nhận ra codon của triplet TAA. Điều này sẽ có ảnh hưởng gì đến quá trình dịch mã trong các tế bào này?

Hướng dẫn trả lời

Điều quan trọng là phải nhận ra rằng, trong các tế bào bình thường, TAA là mã kết thúc. Trong các tế bào đột biến, mã kết thúc TAA đôi khi sẽ liên kết tRNA đột biến được nạp Glu. Trong trường hợp này, quá trình tổng hợp polypeptide sẽ tiếp tục vượt qua codon kết thúc và protein được tạo ra sẽ dài hơn bình thường.

Câu 5:

Mô tả thí nghiệm ủng hộ giả thuyết thành phần rRNA chứ không phải protein của ribosome có hoạt tính xúc tác của peptidyl transferase.

Hướng dẫn trả lời

Thiết kế thí nghiệm đơn giản gồm: xử lý ribosome với protease (để loại bỏ thành phần protein) và kiểm tra phần ribosome còn lại còn hoạt tính xúc tác hay không. Thực tế thí nghiệm cho thấy: khi phần lớn protein đã được loại bỏ, liên kết peptide vẫn được hình thành. Các nghiên cứu sinh học cấu trúc sau này cũng ủng hộ giả thuyết thành phần rRNA của ribosome xúc tác phản ứng peptide khi không có amino acid nào có mặt ở vị trí hoạt động vốn có kích thước rất nhỏ (~18 Å).

Câu 6:

Tại sao qua ức chế dịch mã mà các chất kháng sinh lại "tiêu diệt" đặc hiệu được các tế bào vi khuẩn mà ít ảnh hưởng các tế bào sinh vật nhân thực?

Hướng dẫn trả lời

- Chất kháng sinh hoạt động theo cách ức chế một trong các bước dịch mã bằng cách liên kết (phong bế) vị trí đặc hiệu trên ribosome hoặc EF-Tu hoặc EF-G. Sự ức chế một bước dẫn đến dừng cả quá trình. Dịch mã là quá trình thiết yếu để tế bào sống sót.

- Thành phần protein và rRNA của ribosome của vi khuẩn và sinh vật nhân thực là khác nhau đáng kể. Vì vậy, chất kháng sinh liên kết đặc hiệu vào các thành phần bộ máy dịch mã của vi khuẩn, song không ảnh hưởng đến bộ máy dịch mã của sinh vật nhân thực.

Câu 7:

Có nhiều phương pháp cho phép bạn đưa gene nguyên vẹn của sinh vật nhân thực, ví dụ, đưa gene triose phosphate isomerase, vào tế bào nhân sơ.

- a) Bạn có mong đợi gene này được phiên mã chính xác bởi RNA polymerase của sinh vật nhân sơ không? Trong tình huống ngược lại, trong đó một gene nguyên vẹn của sinh vật nhân sơ được đưa vào một tế bào nhân thực; liệu nó có được phiên mã chính xác bởi một phức hợp phiên mã của sinh vật nhân thực không?
- b) Giả sử rằng, trong một trường hợp hiếm hoi, một gene triose phosphate isomerase điển hình của sinh vật nhân chuẩn chứa các trình tự chính xác để cho phép phiên mã chính xác trong tế bào nhân sơ. Liệu RNA này sẽ được dịch đúng cách để tạo ra enzyme nguyên vẹn có chức năng?

Hướng dẫn trả lời

- a) - Không. Rất khó có khả năng trình khởi động của gene của sinh vật nhân chuẩn sẽ chứa các trình tự chính xác ở đúng vị trí để cho phép khởi đầu chính xác bởi các RNA polymerase của sinh vật nhân sơ.
- Tương tự như vậy, rất khó có khả năng trình khởi động của gene prokaryote sẽ chứa trình tự chính xác ở đúng vị trí để cho phép khởi đầu chính xác bởi RNA polymerase II.
- b) Không. Một gene isomerase triose phosphate điển hình của sinh vật nhân chuẩn có chứa các đoạn intron. Tế bào nhân sơ không chứa spliceosome và do đó sẽ không thể xử lý chính xác bản phiên mã chính. Do đó, quá trình dịch mã RNA sẽ tạo ra một đoạn protein bất thường.

Câu 8:

Bảng sau chứa thông tin trình tự DNA do Marilyn Kozak (1987) tổng hợp. Dữ liệu bao gồm tỷ lệ phần trăm của A, C, G và T ở mỗi vị trí trong số 12 nucleotide trước codon mở đầu trong 699 gene từ các loài động vật có xương sống khác nhau và nucleotide đầu tiên sau codon khởi đầu. Codon mở đầu (MĐ) chiếm các vị trí từ +1 đến +3 và nucleotide +4 xuất hiện ngay sau codon mở đầu đầu.

Bảng C8.

Vị trí	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	[MĐ] +4
Tỉ lệ A	23	26	25	23	19	23	17	18	25	61	27	15	[AUG] 23
Tỉ lệ C	35	35	35	26	39	37	19	39	53	2	49	55	[AUG] 16
Tỉ lệ G	23	21	22	33	23	20	44	23	15	36	13	21	[AUG] 46
Tỉ lệ T	19	18	18	18	19	20	20	20	7	1	11	9	[AUG] 15

Sử dụng dữ liệu để xác định trình tự liên ứng của 13 nucleotide (-12 đến -1 và +4) xung quanh codon mở đầu trong gene của động vật có xương sống.

Hướng dẫn trả lời

Trình tự liên ứng là **CCCGCCGCCACCAUGG**.

Vị trí	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	[MĐ] +4
Tỉ lệ A	23	26	25	23	19	23	17	18	25	61	27	15	[AUG] 23
Tỉ lệ C	35	35	35	26	39	37	19	39	53	2	49	55	[AUG] 16
Tỉ lệ G	23	21	22	33	23	20	44	23	15	36	13	21	[AUG] 46
Tỉ lệ T	19	18	18	18	19	20	20	20	7	1	11	9	[AUG] 15
Trình tự liên ứng	C	C	C	G	C	C	G	C	C	A	C	C	AUGG

Câu 9:

Isoginkgetin là một hóa chất thấm vào tế bào được phân lập từ cây Bạch quả (*Ginkgo biloba*) có tác dụng liên kết và ức chế snRNP (ribonucleoprotein hạt nhân nhỏ).

- a) Bạn dự đoán những loại vấn đề nào trong các tế bào được điều trị bằng isoginkgetin?
- b) Đây là vấn đề nghiêm trọng nhất đối với tế bào *E. coli*, tế bào nấm men hay tế bào người? Tại sao?

Hướng dẫn trả lời

- a) - Isoginkgetin ức chế snRNP bằng cách liên kết với chúng sẽ ảnh hưởng đến sự cắt nối của tiền mRNA để tạo ra mRNA trưởng thành.
 - Quá trình cắt nối của chuỗi mRNA mới sinh có thể có hai loại: tự cắt nối hoặc cắt nối qua trung gian spliceosome. Trong quá trình tự cắt nối, các intron tự ghép nối và không có protein nào tham gia vào phương pháp ghép nối này. Quá trình cắt nối qua trung gian spliceosome liên quan đến một số phân tử RNA ngắn được gọi là RNA nhân nhỏ (có tên là U1, U2, v.v.) các phân tử RNA ngắn này được liên kết với snRNP để hình thành phức hợp spliceosomal. phức hợp spliceosome thực hiện cắt nối
 - Khi các snRNP bị ức chế, tiền mRNA không trải qua quá trình ghép nối để loại bỏ các intron → các intron được đưa vào mRNA trưởng thành → sự dịch mã của các nucleotide thuộc intron → protein có thêm các amino acid có thể được hình thành và có thể dài hơn protein bình thường, protein này sẽ bị rối loạn chức năng và sẽ ảnh hưởng đến các chức năng liên quan như chuyển hóa tế bào.
 - (Có thể nói đơn giản: snRNP có vai trò quan trọng trong quá trình cắt nối có lựa chọn của tiền mRNA của sinh vật nhân chuẩn có chứa intron. Nếu snRNP bị ức chế → tiền mRNA này sẽ được ghép nối không chính xác → mất protein thường được mã hóa và có thể là quá trình tổng hợp các sản phẩm protein bất thường từ các mRNA chưa được chế biến).
- b) - Quá trình cắt nối RNA không xảy ra trong tế bào vi khuẩn, vì vậy các tế bào này sẽ không bị ảnh hưởng bởi isoginkgetin.
 - Tuy nhiên, tế bào nấm men và người, cả hai đều sử dụng quá trình ghép nối trong quá trình chế biến RNA do hệ gene chứa intron, vì vậy việc chế biến sẽ gặp vấn đề đối với cả hai.

Câu 10:

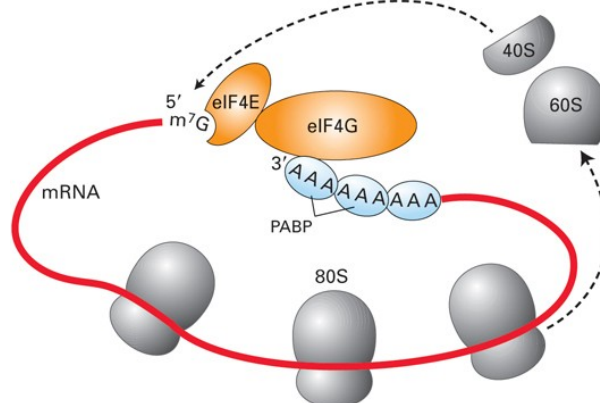
Không giống như DNA polymerase, RNA polymerase không có hoạt động đọc sửa (proofreading activity). Giải thích tại sao thiếu hoạt động đọc sửa lại không có hại cho tế bào.

Hướng dẫn trả lời

- Việc thiếu hoạt tính đọc sửa ở RNA polymerase làm cho tỷ lệ lỗi phiên mã lớn hơn tỷ lệ lỗi sao chép DNA. Tuy nhiên, các phân tử RNA khiếm khuyết được tạo ra không có khả năng ảnh hưởng đến khả năng tồn tại của tế bào vì hầu hết các bản sao RNA được tổng hợp từ một gene nhất định là bình thường.
- Trong trường hợp mRNA bị lỗi, số lượng protein bị lỗi chỉ là một tỷ lệ nhỏ trong tổng số protein được tổng hợp. Ngoài ra, những lỗi mắc phải trong quá trình phiên mã sẽ nhanh chóng được loại bỏ vì hầu hết các phân tử mRNA đều có thời gian bán hủy ngắn.

Câu 11:

Hình C11 dưới đây mô tả mRNA của eukaryote hình thành cấu trúc vòng trong quá trình khởi đầu dịch mã nhờ những tương tác của ba protein gắn tại đầu 3' và 5' là protein liên kết đuôi poly(A) (PABP), eIF4E và eIF4G.



Hình C11

Hãy cho biết sự hình thành cấu trúc vòng của mRNA trong dịch mã ở eukaryote có ý nghĩa như thế nào? Đột biến gene mã hóa protein gắn đuôi poly(A) (PABP) tác động như thế nào đến quá trình dịch mã.

Hướng dẫn trả lời

- Cấu trúc vòng của mRNA làm tăng hiệu quả dịch mã: Khi một ribosome hoàn thành dịch mã và tách khỏi đầu 3', các tiểu phần đã phân tách có thể nhanh chóng tìm thấy mũ 5' (m^7G) và đuôi poly(A) gắn PABP gần đó và khởi đầu vòng tổng hợp kế tiếp.

- Vì protein liên kết poly(A) có liên quan đến việc tăng hiệu quả dịch mã, một đột biến trong protein liên kết poly(A) sẽ gây ra quá trình dịch mã kém hiệu quả hơn. Các polyribosome trong tế bào có đột biến như vậy sẽ không chứa các cấu trúc tròn của mRNA trong quá trình dịch mã vì thiếu protein liên kết poly(A) sẽ loại bỏ vị trí liên kết 3' với eIF4G.

Câu 12:

A, B, C là các loại thuốc phổ biến được dùng để điều trị bệnh nhiễm trùng do vi khuẩn hoặc virus. Đặc điểm của các loại thuốc như sau được thể hiện ở Bảng C12. Từ cơ chế tác động của mỗi loại thuốc, hãy cho biết:

Bảng C12

Thuốc	Cơ chế tác động
Thuốc A	Liên kết vào phía trước vị trí A trên tiểu đơn vị 50S của ribosome
Thuốc B	Là một chất tương tự guanine
Thuốc C	Ức chế DNA gyrase

- Bước nào trong quá trình tổng hợp protein bị ức chế bởi mỗi loại thuốc? Giải thích.
- Thuốc nào có hiệu quả hơn để chống lại vi khuẩn? Giải thích.

Hướng dẫn trả lời

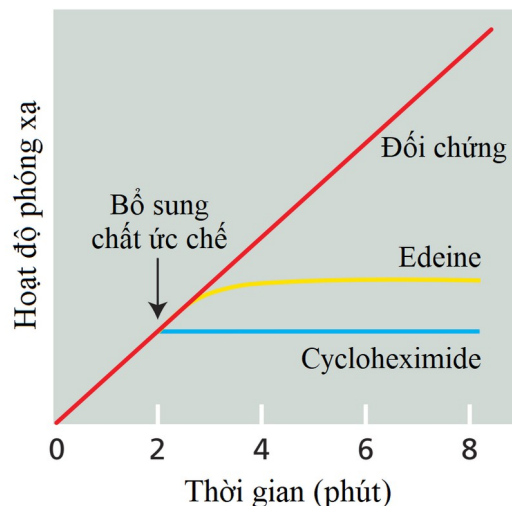
- **Thuốc A** ức chế chuyển vị aminoacyl bằng cách can thiệp vào việc chuyển tRNA từ vị trí A sang vị trí P. Do đó, kéo dài chuỗi polypeptide bị ức chế.
- **Thuốc B**: một chất tương tự guanine, có khả năng liên kết với polymerase của virus và ức chế tổng hợp DNA của virus.

- **Thuốc C** ức chế DNA gyrase, enzyme giúp giải toả lực căng DNA trước chạc sao chép. Do đó, quá trình tổng hợp RNA từ DNA (phiên mã) và bắt đầu dịch mã (tổng hợp protein) bị ức chế.

- b) Thuốc A chống vi khuẩn hiệu quả hơn vì nó hoạt động bằng cách liên kết với vị trí A trên tiểu đơn vị 50S của ribosome vi khuẩn và ức chế sự kéo dài của chuỗi polypeptid. Do đó, nó sẽ không gây ra bất kỳ ảnh hưởng nào đến cơ thể người vì ribosome của sinh vật nhân thực là loại 80S bao gồm hai tiểu phần 60S và 40S trong đó ribosome của vi khuẩn là loại 70S gồm hai tiểu phần là 50S và 30S.

Câu 13:

Kháng sinh edenie có khả năng ức chế tổng hợp protein nhưng không ảnh hưởng đến sự tổng hợp DNA hoặc RNA. Khi bổ sung edenie vào dịch ly giải hồng cầu lưới, người ta thấy quá trình tổng hợp bị ức chế sau một thời gian ngắn như hình bên. Ngược lại, xicloheximide ngay lập tức làm dừng sự tổng hợp protein. Khi ly tâm dịch ly giải hồng cầu lưới có edenie, người ta thấy không tồn tại polyribosome sau khi sự tổng hợp protein bị ức chế, thay vào đó mRNA lại liên kết với một ribosome 40S không bình thường – chứa một lượng tương ứng tiểu đơn vị ribosome và tRNA khởi đầu.



Hình C13. Tác động của các chất ức chế edeine và cycloheximide đối với quá trình tổng hợp protein trong ly giải hồng cầu lưới.

- Edeine ức chế bước nào trong quá trình tổng hợp protein? Giải thích.
- Tại sao có khoảng trễ giữa thời điểm bắt đầu bổ sung edenie và khi protein hoàn toàn bị ngừng tổng hợp? Xác định độ dài khoảng trễ này?
- Cơ chế ức chế tổng hợp protein của xiclohexamide khác gì so với edenie? Nếu bổ sung xiclohexamide vào cùng thời điểm bổ sung edenie thì có xảy ra sự biến mất của polyribosome không? Giải thích.

Hướng dẫn trả lời

- Edeine ngăn chặn tiểu phần lớn ribosome (60S) với phức hệ gồm tiểu phần nhỏ (40S), mRNA và tRNA khởi đầu. Vì sau khi bị ức chế, mRNA chỉ có thể liên kết với tiểu phần 40S và tRNA khởi đầu. Hoặc ức chế sự hình thành phức hệ khởi đầu phiên mã.
- Vì quá trình kéo dài chuỗi polypeptide không bị ảnh hưởng nên các ribosome đã liên kết thành công vẫn có thể tiếp tục hoàn thành quá trình dịch mã.

- Độ dài khoảng trễ khoảng 1 phút.
- c) - Xiclohexamide ngay lập tức làm ngừng sự gia tăng hoạt độ phóng xạ → ức chế kéo dài chuỗi polypeptide (hoặc ức chế sự di chuyển của ribosome).
- Không. Vì xiclohexamide là các ribosome trong polyribosome bị “đóng băng” (ức chế trượt trên mRNA) chứ không làm chúng tách rời khỏi mRNA và không thể liên kết lại như edenie.

Câu 14:

Nhiều loại thuốc kháng sinh thường gặp hiện nay có hiệu quả ức chế sự tổng hợp protein trong tế bào vi khuẩn. Bảng C25 cho biết cơ chế hoạt động của một số loại kháng sinh.

Bảng C14

Kháng sinh	Cơ chế hoạt động
Streptomycin	Liên kết không thuận nghịch với tiểu đơn vị 30S của ribosome
Chloramphenicol	Ức chế hoạt tính peptidyl transferase của ribosome
Tetracycline	Ức chế aminoacyl - tRNA gắn vào vị trí A của ribosome
Erythromycin	Liên kết với tiểu đơn vị 50S, ngăn cản sự hình thành cấu trúc ribosome hoàn chỉnh (70S)
Kasugamycin	Ức chế sự liên kết của tRNAMet vào ribosome
Thiostrepton	Ức chế sự chuyển vị của ribosome bằng cách ức chế EF-G

Sử dụng thông tin được cung cấp trong bảng trên, hãy cho biết

- a) Giai đoạn nào của quá trình dịch mã ở vi khuẩn (mở đầu, kéo dài hay kết thúc) bị ức chế bởi mỗi loại kháng sinh?
- b) Để ức chế sự tổng hợp của protein vi khuẩn, các nhà khoa học còn có thể thiết kế những loại thuốc kháng sinh có cơ chế hoạt động như thế nào (khác với những loại kháng sinh được nhắc đến ở bảng trên)?

Hướng dẫn trả lời

- a) Giai đoạn của quá trình dịch mã ở vi khuẩn bị ức chế bởi mỗi loại kháng sinh:
 - **Streptomycin:** ức chế giai đoạn khởi đầu dịch mã. Do: streptomycin liên kết không thuận nghịch với tiểu đơn vị 30S của ribosome - ngăn cản tiểu đơn vị 30S liên kết với mRNA → ức chế hình thành phức hệ khởi đầu
 - **Chloramphenicol:** ức chế giai đoạn kéo dài chuỗi polypeptide. Do: chloramphenicol ức chế hoạt tính peptidyl transferase của ribosome (peptidyl transferase là hoạt tính enzyme có vai trò xúc tác hình thành liên kết peptide giữa đầu N của amino acid ở vị trí A với đầu C của amino acid ở vị trí P) → liên kết peptide giữa các amino acid không được hình thành giai đoạn kéo dài chuỗi 1 polypeptide bị ức chế.
 - **Tetracycline:** ức chế giai đoạn kéo dài chuỗi polypeptide. Do: ức chế sự liên kết của aminoacyl - tRNA vào vị trí A trên ribosome sẽ ngăn cản các amino acid đi vào ribosome (vì amino acid được vận chuyển bởi tRNA tương ứng) → các amino acid không được lắp ráp vào chuỗi polypeptide - giai đoạn kéo dài chuỗi polypeptide bị ức chế.
 - **Erythromycin:** ức chế giai đoạn khởi đầu dịch mã. Do: Erythromycin liên kết với tiểu đơn vị 50S, khiến cho tiểu đơn vị này không thể tham gia vào cấu trúc ribosome

hoàn chỉnh cùng với tiểu đơn vị 30S → ribosome hoàn chỉnh không được hình thành → không thể hình thành phức hệ khởi đầu dịch mã.

- **Kasugamycin:** ức chế giai đoạn khởi đầu dịch mã. Do: Formylmethionine (fMet) là amino acid khởi đầu ở sinh vật nhân thực. Suy ra Kasugamycin có tác dụng ức chế sự khởi đầu dịch mã, vì chất kháng sinh này ngăn cản hình thành liên kết giữa fMet-tRNA và ribosome.

- **Thiostrepton:** ức chế giai đoạn kéo dài chuỗi polypeptide. Do: sự chuyển vị của ribosome là cần thiết để làm trống vị trí A, dịch chuyển tRNA không mang amino acid sang vị trí E (vị trí thoát) và đưa tRNA liên kết chuỗi polypeptide sang vị trí P. Nhờ đó, quá trình dịch mã có thể xảy ra liên tục. Khi ngăn cản sự chuyển vị thì quá trình dịch mã sẽ bị ngừng trệ, chuỗi polypeptide không được kéo dài giai đoạn kéo dài chuỗi bị ức chế.

b) Những cơ chế hoạt động khác của kháng sinh gây ức chế tổng hợp protein vi khuẩn:

- Ngăn cản quá trình cải biến sau dịch mã của protein (phosphoryl hóa, pupyl hóa, thêm bớt hoặc chỉnh sửa các amino acid, cuộn gập thành các cấu trúc bậc cao, lắp ráp với các protein khác, ...).

- Ngăn cản sự di chuyển (hoặc ức chế sự hình thành) của hệ polyribosome chậm quá trình dịch mã.

Câu 15:

Một nhà nghiên cứu phân lập các biến thể đột biến của các yếu tố dịch mã ở vi khuẩn: IF-2, EF-Tu và EF-G. Trong mỗi trường hợp, đột biến cho phép gập đúng protein và liên kết với GTP nhưng không cho phép thủy phân GTP. Ở giai đoạn nào dịch mã sẽ bị chặn bởi mỗi protein đột biến?

Hướng dẫn trả lời

- IF-2: Ribosome 70S sẽ hình thành, nhưng các yếu tố khởi đầu sẽ không được giải phóng và quá trình kéo dài không thể bắt đầu.

- EF-Tu: Aminoacyl-tRNA thứ hai sẽ liên kết với vị trí A của ribosome, nhưng không hình thành liên kết peptide nào.

- EF-G: Liên kết peptide đầu tiên sẽ hình thành, nhưng ribosome sẽ không di chuyển dọc theo mRNA (chuyển vị trí) để rời khỏi vị trí A để liên kết với EF-Tu-tRNA mới.

Câu 16:

Khi nghiên cứu chức năng của gene mã hóa thụ thể yếu tố phát triển đặc biệt ở người, nhà nghiên cứu thấy rằng hai loại protein được tổng hợp từ gene này. Protein lớn có chứa miền xuyên màng có chức năng nhận biết các yếu tố phát triển tại bề mặt tế bào, kích thích con đường truyền tín hiệu xuôi dòng đặc hiệu. Ngược lại, protein nhỏ tương đồng được tế bào tiết và gắn với yếu tố phát triển tuần hoàn trong máu, do đó kìm hãm con đường truyền tín hiệu xuôi dòng. Giải thích vì sao tế bào có chức năng tổng hợp những protein có chức năng khác biệt này.

Hướng dẫn trả lời

Một giải thích đơn giản là protein chứa miền lớn hơn, kéo dài qua màng và protein nhỏ, được tiết ra ngoài được mã hóa bởi cùng một gene nhưng được chế biến (cắt nối) theo những cách khác nhau. Cụ thể, exon cuối cùng của gene có thể chứa thông tin cho

miền xuyên màng, và trong protein tiết ra nhỏ hơn, exon này có thể bị bỏ qua trong quá trình cắt nối.

Câu 17: Đề HSG Quốc gia năm 2012

Nêu những điểm khác nhau cơ bản của quá trình phiên mã ở sinh vật nhân thực với sinh vật nhân sơ. Những điểm khác nhau này có ý nghĩa gì cho sinh vật nhân sơ và sinh vật nhân thực?

Hướng dẫn trả lời

a) Khác biệt

Điểm khác biệt	Nhân sơ	Nhân thực
Enzyme	Chỉ cần một loại enzyme tổng hợp 3 loại RNA (rRNA, mRNA, tRNA).	Cần 3 loại enzyme khác nhau tổng hợp 3 loại RNA (RNA polymerase I tổng hợp rRNA; RNA polymerase II tổng hợp mRNA; RNA polymerase III tổng hợp tRNA).
Đơn vị phiên mã	Một đơn vị phiên mã gồm nhiều gene (một gene điều hoà, một vùng điều hoà điều khiển sự phiên mã của cả một nhóm gene gọi là operon)	Một đơn vị phiên mã chỉ gồm một gene (một gene điều hoà, một vùng điều hoà điều khiển sự phiên mã của một gene).
Hoàn thiện mRNA	RNA tổng hợp ra được dùng để dịch mã ngay mà không cần biến đổi.	RNA tổng hợp ra cần phải được cắt bỏ intron và nối các exon lại với nhau để tạo ra mRNA; ngoài ra, còn gắn thêm mũ 7 methylguanosine ở đầu 5' và đuôi poly(A) ở đầu 3' của mRNA.

b) Ý nghĩa:

- Đối với sinh vật nhân sơ: Giúp tiết kiệm năng lượng và thời gian cho các quá trình phiên, dịch mã diễn ra nhanh hơn (phiên mã và dịch mã xảy ra gần như đồng thời), góp phần làm cho nhân sơ có thể sinh sản nhanh.
- Đối với sinh vật nhân thực: Việc gắn mũ và đuôi poly(A) có tác dụng kích thích mRNA đi ra tế bào chất để dịch mã và tránh khỏi sự phân huỷ của một số enzyme, là tín hiệu để cho ribosome nhận biết gắn vào mRNA để dịch mã và tạo ra sự ổn định lâu dài hơn trong tế bào. Việc cắt bỏ intron và nối exon có thể tạo ra các mRNA trưởng thành khác nhau, từ đó qua dịch mã tạo ra được các chuỗi polypeptide khác nhau để cấu trúc lên các loại protein khác nhau.

Câu 18: Đề chọn đội tuyển dự thi IBO năm 2008

Hãy nêu sự khác biệt giữa mRNA đã thành thực và tiền mRNA trong quá trình phiên mã ở sinh vật nhân thật (eukaryote).

Hướng dẫn trả lời

Tiền mRNA	mRNA thành thực
Mới phiên mã từ DNA, nằm trong nhân	Là sản phẩm của quá trình chế biến tiền mRNA đã hoặc chuẩn bị được vận chuyển

Tiền mRNA	mRNA thành thực
	ra tế bào chất
Kích thước dài bởi mang cả exon và intron	Kích thước ngắn bởi chỉ mang các exon trong vùng mã hóa (nếu không tính đuôi polyA).
Không có phần đầu 3' và 5' được cải biến	Có “mũ” 7-methylguanosine ở đầu 5' và đuôi poly(A) ở đầu 3'.
Thường ít khi có kích thước hoàn chỉnh, bởi sự cắt intron có thể xảy ra ngay khi phiên mã chưa kết thúc	Có chiều dài hoàn chỉnh từ khi được vận chuyển từ nhân ra tế bào chất cho đến khi kết thúc dịch mã
Là sản phẩm từ đó hình thành nên mRNA thành thực (một phân tử tiền mRNA có thể tạo nên một số phân tử mRNA thành thực khác nhau)	Là khuôn tổng hợp nên phân tử protein (ở sinh vật nhân thực, thường một phân tử mRNA thành thực được dùng để tổng hợp một chuỗi polypeptide duy nhất)

Câu 19: Đề HSG Quốc gia năm 2015

- Nêu vai trò của các exon trong gene phân mảnh. Sau khi các intron bị cắt bỏ thì trật tự và số lượng của exon trong mRNA trưởng thành sẽ như thế nào?
- Đột biến điểm ở intron có ảnh hưởng đến exon không? Giải thích.

Hướng dẫn trả lời

- Vai trò của exon trong gene phân mảnh là mã hóa các amino acid để cấu trúc nên chuỗi polipeptide và mã hóa phân tử RNA. Trong vùng mã hóa amino acid, mỗi exon quy định một miền cấu trúc biểu hiện chức năng của protein.
- Số lượng và trình tự các exon:
+ Về trật tự: sau khi các intron bị cắt bỏ thì trật tự sắp xếp của các exon trong mRNA trưởng thành có thể bị xáo trộn, tuy nhiên thường giữ nguyên như trật tự vốn có trên gene. Các vị trí của exon đầu (ở đầu 5') và cuối (ở đầu 3') thường không thay đổi.
+ Về số lượng: một vài exon có thể bị loại bỏ do cơ chế điều hòa hoạt động của gene. Ví dụ, gene mã hóa troponin T gồm 5 exon mã hóa cho 2 loại protein cơ mà mRNA trưởng thành khác nhau, trong đó dạng 1 không có exon 4, còn dạng 2 không có exon 3.
- Nếu đột biến intron là đột biến nguyên khung thì không ảnh hưởng đến exon, còn nếu là đột biến dịch khung thì có thể làm biến đổi intron thành trình tự mã hóa amino acid, bổ sung thêm trình tự nucleotide mã hóa amino acid vào các exon, làm cho chuỗi peptide dài ra khi được tổng hợp sẽ có hại cho cơ thể sinh vật.

Câu 20: Đề chọn đội tuyển dự thi IBO năm 2012

Các nhà khoa học cho rằng một số intron có chức năng điều hoà hoạt động gene theo một trong 2 cách sau đây: (1) intron của gene trực tiếp tham gia điều hoà hoạt động gene và (2) intron trong RNA sơ cấp tham gia điều hoà hoạt động gene. Hãy giải thích cơ chế điều hoà hoạt động gene của intron trong 2 cách nêu trên.

Hướng dẫn trả lời

- Intron của gene có thể chứa các trình tự tăng cường, khi nó liên kết với các yếu tố phiên mã sẽ làm tăng ái lực của RNA polymerase với promoter và do vậy sẽ làm tăng cường mức độ phiên mã của gene.

- Nếu intron trong RNA sơ cấp có chức năng điều hoà hoạt động gene thì chỉ có thể theo cơ chế nó sẽ liên kết bổ sung được với một trình tự của promoter và do vậy ngăn cản quá trình phiên mã của gene.

Câu 21: Đề chọn đội tuyển dự thi IBO năm 2016

EF-Tu là một yếu tố kéo dài với GTP tham gia giai đoạn kéo dài chuỗi pôlipeptide ở tế bào nhân sơ. EF-Tu gắn với tất cả các phức hợp aminoacyl -tRNA (aa-tRNA) với ái lực gần như nhau để đưa chúng đến ribosome với tần xuất giống nhau. Sau đây là kết quả thí nghiệm xác định sự liên kết của EF-Tu và phức hợp aminoacyl-tRNA bắt cặp chính xác và không chính xác.

Phức hợp aminoacyl-tRNA	Hệ số phân ly (nM)
Ala-tRNA ^{Ala}	6,2
Gln-tRNA ^{Ala}	0,05
Gln -tRNA ^{Gln}	4,4
Ala-tRNA ^{Gln}	260

- Dựa vào số liệu trên hãy giải thích vì sao hệ thống nhận biết tRNA- EF-Tu có thể ngăn ngừa sự ghép sai amino acid trong quá trình dịch mã?
- Hãy chỉ ra vai trò của EF-Tu trong quá trình dịch mã.

Hướng dẫn trả lời

- Phức hợp aminoacyl-tRNA bắt cặp chính xác (Ala-tRNA^{Ala} và Gln-tRNA^{Gln}) có ái lực gần như nhau với EF-Tu và được chuyển đến vị trí A trên ribosome.
 - Phức hợp bắt cặp không chính xác Ala-tRNA^{Gln} gắn với EF-Tu lỏng lẻo hơn nhiều và sẽ phân ly với EF-Tu trước khi tiến đến ribosome.
 - Phức hợp Gln-tRNA^{Ala} gắn chặt với EF-Tu làm cho EF-Tu không tách được khỏi chúng tại ribosome.
 - Do đó, dù ái lực gắn kết cao hay thấp hơn đều ảnh hưởng đến hoạt động của EF-Tu và làm giảm tốc độ gắn vào vị trí A trên ribosome của phức hợp aminoacyl-tRNA bắt cặp sai.
- Vai trò của EF-Tu giúp sự bắt cặp chính xác bộ ba đối mã của tRNA với bộ ba mã hóa của mRNA.
 - Sự thủy phân GTP gắn với EF-Tu khi có sự cặp đôi chính xác tạo cấu hình phù hợp cho sự tương tác giữa codon - anticodon và đảm bảo cho sự hình thành liên kết peptide xảy ra tiếp theo.

Câu 22: Đề Olympic chuyên KHTN năm 2019

- Nêu vai trò của enzyme aminoacyl-tRNA synthetase và enzyme peptidyl transferase trong quá trình dịch mã ở sinh vật nhân thực.
- Giải thích tại sao trong khởi đầu dịch mã ở sinh vật nhân thực cần sự đính kết của phức hợp tRNA-Met vào tiểu phần nhỏ của ribosome trước khi liên kết với mRNA, nhưng ở sinh vật nhân sơ thì phức hợp tRNA-fMet lại liên kết với mRNA sau khi tiểu phần nhỏ của ribosome đã liên kết với mRNA?

Hướng dẫn trả lời

- a) - Enzyme aminoacyl-tRNA synthetase: Xúc tác cho sự hình thành liên kết cộng hóa trị chính xác giữa tRNA và amino acid tương ứng qua một phản ứng được thúc đẩy bởi sự thủy phân ATP. Trung tâm xúc tác của mỗi loại enzyme chỉ phù hợp cho một sự kết cặp đặc thù giữa một loại amino acid với tRNA. Có 20 loại synthetase khác nhau, mỗi loại dành cho một loại amino acid, mỗi enzyme synthetase có thể liên kết với nhiều tRNA khác nhau cùng mã hóa cho một loại amino acid.
- Enzyme peptidyl transferase: Là một phần của tiểu phần lớn ribosome, có vai trò xúc tác cho sự tạo thành các liên kết peptide giữa các amino acid trong quá trình tổng hợp chuỗi polypeptide.
- b) - Trong khởi đầu dịch mã ở sinh vật nhân thực: Tiểu phần nhỏ của ribosome luôn liên kết với tRNA_i trước khi nó được huy động tới liên kết vào mũ 5'm⁷G. Sau đó, tiểu phần nhỏ cùng phức hợp tRNA-Met sẽ trượt dọc phân tử mRNA theo chiều 5' – 3' cho đến khi gặp mã 5'AUG 3' đầu tiên (trong phần lớn trường hợp nó nhận bộ ba này làm bộ ba mở đầu). → Nếu không có sự kết dính của phức hợp tRNA-Met vào ribosome trước thì ribosome không thể xác định được vị trí của bộ ba khởi đầu dịch mã (AUG).
- Trong khởi đầu dịch mã ở sinh vật nhân sơ: Tiểu phần nhỏ của ribosome được huy động đến và liên kết với mRNA thông qua tương tác Chargaff giữa trình tự trên rRNA 16S với trình tự Shine-Dalgarno. Sau đó, tiểu phần nhỏ thường được lắp ráp vào vị trí RBS (Ribosome binding site) sao cho vị trí P của ribosome tương ứng với mã mở đầu AUG. Sau khi tiểu phần nhỏ của ribosome gắn vào đúng vị trí thì tRNA đã nạp amino acid fMet mới đi vào vị trí P trong ribosome → ở sinh vật nhân sơ tiểu phần nhỏ không cần gắn trước với tRNA_i.

Câu 23: Đề HSG Quốc gia năm 2018

Dưới đây là một đoạn trình tự nucleotide thuộc vùng mã hóa của một gene quy định chuỗi polypeptide có 300 amino acid, mang bộ 3 tương ứng mã mở đầu và chưa xác định được các đầu tận cùng (3' hoặc 5') của đoạn gene này.

A G A T G T A G T A X G G A A T T G A T X X A G T A A G T X A T T X
T X T A X A T X A T G X X T T A A X T A G G T X A T T X A G T A A G

- a) Dựa vào trình tự nucleotide của đoạn gene trên, hãy nêu cách xác định sợi làm khuôn cho quá trình phiên mã của gene. Viết trình tự nucleotide của đoạn mRNA được phiên mã từ đoạn gene này với các đầu tận cùng (3' hoặc 5') và viết kí hiệu +1 để xác định bộ 3 mở đầu dịch mã.
- b) Không thay đổi các viết thứ tự các nucleotide, hãy viết lại trình tự nucleotide của đoạn gene đã cho và bổ sung mũ tên chỉ chiều phiên mã, kí hiệu vị trí tương ứng mã mở đầu (+1) và các đầu tận cùng (3' hoặc 5') trên 2 sợi của đoạn gene này.

Hướng dẫn trả lời

- a) Đây là đoạn thuộc vùng mã hóa và mang bộ ba mở đầu dịch mã nên mRNA tương ứng được tạo thành phải mang bộ 3 mở đầu AUG

- Nếu sợi bên dưới làm khuôn cho phiên mã, trình tự mRNA là :

AGAUGUAGUAXGGAAUUGAUXXAGUAAGUXAUUX

Nếu phiên mã theo chiều từ trái sang phải hoặc từ phải sang trái mRNA đều xuất hiện bộ 3 mở đầu AUG (như gạch chân). Nhưng kế tiếp hoặc gần kế tiếp các bộ 3 AUG này đều là các bộ ba kết thúc dịch mã (UGA) → đây không phải là trình tự mRNA đúng của gene này.

- Sợi bên trên là sợi làm khuôn cho phiên mã. Khi đó mRNA có trình tự là:

UXUAXXAUXAUGXXUUAAXUAGGUXAUUXAGUAAAG

Phiên mã từ trái sang phải hoặc từ phải sang trái, mRNA đều có bộ 3 mở đầu AUG. Nhưng nếu theo chiều từ trái sang phải, cách bộ 3 AUG 1 bộ 3 khác là bộ 3 kết thúc UAA → Không thỏa mãn.

⇒ Do đó, sợi bên trên làm khuôn cho phiên mã, theo chiều từ phải sang trái

- Trình tự và các đầu tận cùng của đoạn mRNA, vị trí bộ 3 mở đầu là:

5' GAAUGAXUUAXUGGAUXAAUUXXGUAXUAXAUXU 3'
+1

Hoặc cũng có thể viết :

3' UXUAXAUXAUGXXUUAAXUAGGUXAUUXAGUAAAG 5'
+1

Vẽ mũi tên chỉ chiều phiên mã và viết các đầu tận cùng (3' và 5')

5' A G A T G T A G T A X G G A A T T G A T X X A G T A A G T X A T T X 3'
3' T X T A X A T X A T G X X T T A A X T A G G T X A T T X A G T A A G 5'
← +1

Câu 24: Đề HSG Quốc gia năm 2019

Để nghiên cứu tốc độ tích lũy đột biến thay thế nucleotide trên gene, các nhà khoa học đã so sánh trình tự nucleotide ở vùng đầu (chứa trình tự nucleotide mã hóa tín hiệu nhận biết và tiến hành dịch mã của ribosome) của 149 gene của *E. coli*. Một phần kết quả nghiên cứu được thể hiện trên hình C24.

<i>lacI</i>	T	G	G	T	G	A	A	T	G	T	G	A	A	A	C	C	A
<i>recA</i>	G	A	A	T	A	A	A	A	T	G	G	C	T	A	T	C	
<i>galR</i>	G	T	A	T	T	T	C	A	T	G	G	C	G	A	C	C	
<i>metJ</i>	A	G	T	A	T	C	T	C	A	T	G	G	C	T	G	A	A
<i>lexA</i>	G	G	G	G	C	G	G	A	A	T	G	A	A	A	G	C	G
<i>trpR</i>	G	A	C	A	T	A	T	T	A	T	G	G	C	C	C	A	A
	5'	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	3'

Hình C19.

- Mạch DNA của các gene trên hình là mạch khuôn hay mạch không làm khuôn trong quá trình phiên mã? Giải thích.
- Hãy viết trình tự một bộ ba mã hóa bảo thủ nhất trong đoạn trình tự ở hình trên. Chức năng của chúng là gì? Tại sao chúng được bảo tồn trong quá trình tiến hóa?

Hướng dẫn trả lời

- Mạch DNA trên hình là **mạch không làm khuôn**.
- Giải thích: Mạch DNA trên hình có chiều 5' → 3'; điểm khởi đầu dịch mã có bộ ba mở đầu (ATG) → Đây là mạch không làm khuôn.
- Trình tự bộ ba mã hóa bảo thủ nhất trong đoạn trình tự là **ATG**, tại **vị trí 0-1-2**.

- Chức năng: Đây là bộ ba mở đầu có chức năng **mở đầu** cho quá trình tổng hợp chuỗi polypeptide và **mã hóa cho amino acid Methionine**.

- Giải thích: Nói chung, các đột biến xảy ra trên các nucleotide của DNA là **ngẫu nhiên**, nên **tần số đột biến** tại các nucleotide khác nhau là **tương đương nhau**.

- Bộ ba **ATG** là **tín hiệu nhận biết** cho các yếu tố **khởi đầu quá trình dịch mã**. Các đột biến tại bộ ba này ảnh hưởng nghiêm trọng đến quá trình dịch mã → bị **CLTN đào thải** → được **bảo tồn trong tiến hóa**.

Câu 25: Đề chọn đội tuyển IBO năm 2015

Nêu những điểm giống và khác nhau cơ bản giữa quá trình khởi đầu phiên mã ở sinh vật nhân sơ với quá trình khởi đầu phiên mã ở sinh vật nhân thực, không tính đến sự khác biệt về quá trình giãn xoắn của nhiễm sắc thể trước khi khởi đầu phiên mã.

Hướng dẫn trả lời

a) Giống nhau: Điều cần có sự tương tác của các loại protein điều hòa nhất định với vùng điều hòa của gene thì quá trình khởi đầu phiên mã mới xảy ra.

b) Khác nhau:

- Ở sinh vật nhân sơ, chỉ có một loại phân tử RNA polymerase phiên mã cho tất cả các loại gene trong khi đó ở sinh vật nhân thực có 3 loại RNA polymerase khác nhau.

- Ở sinh vật nhân sơ, RNA polymerase có thể trực tiếp nhận ra promoter và khởi đầu phiên mã mà không cần có sự trợ giúp của các protein riêng biệt là các yếu tố khởi đầu phiên mã như ở sinh vật nhân thực. Bản thân RNA polymerase của sinh vật nhân sơ chứa một tiểu phân protein được gọi là yếu tố sigma có khả năng nhận biết và tách DNA thành 2 mạch đơn để RNA khởi đầu phiên mã.

- Ở sinh vật nhân thực, quá trình khởi đầu phiên mã chỉ xảy ra sau khi đã có hàng loạt các yếu tố khởi đầu phiên mã liên kết với promoter hoặc vùng enhancer và nhờ đó làm tăng ái lực của RNA polymerase với promoter để khởi đầu phiên mã.

Câu 26: Đề chọn đội tuyển IBO năm 2020

Các gene ở vi khuẩn *E. coli* được khởi động phiên mã nhờ RNA polymerase nhận biết và liên kết vào các hộp -10 (5'-TATAAT-3') và -35 (5'-TTGACA-3') trong vùng khởi động của gene. Một gene có sản phẩm phiên mã chứa 2 nucleotide đầu tiên là 5'-AG-3', đồng thời có trình tự vùng khởi động như sau:



Do mỗi mạch của phân tử DNA sợi kép đều có thể làm khuôn phiên mã, nên sự phiên mã có thể diễn ra theo một trong hai chiều ① hoặc ② như ở hình trên. Hãy trả lời các câu hỏi dưới đây và giải thích:

a) Điểm khởi đầu phiên mã và các hộp -10 và -35 tương ứng với các vị trí nucleotide nào?

b) Chiều phiên mã với trình tự khởi động nêu trên theo chiều ① hay chiều ②?

c) Mạch trình tự 5'→ 3' ở trên là mạch làm khuôn phiên mã hay mạch mã hóa?

d) RNA polymerase nhận biết hộp -10 và -35 trên phân tử DNA sợi kép bằng cách nào?

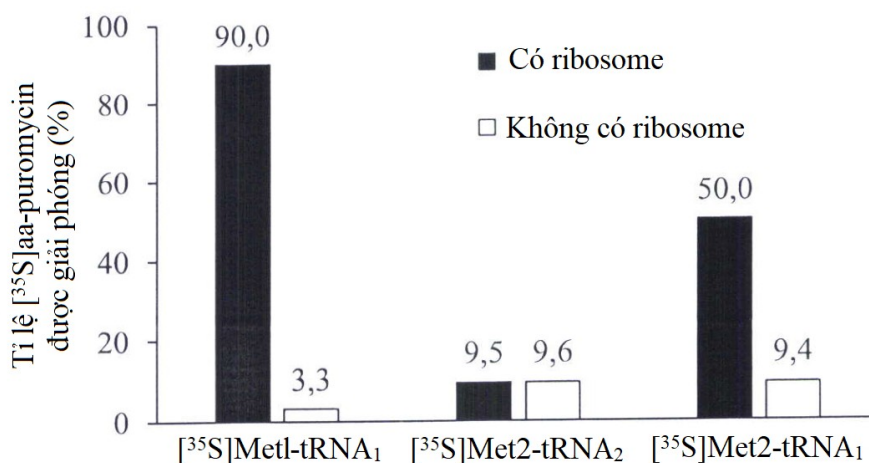
Hướng dẫn trả lời

- a) Điểm khởi động phiên mã sẽ là vị trí A trong cặp nucleotide 5'-AG-3' của mạch mã hóa cách ngược dòng khoảng 10 nucleotide tới hộp 5'-TATAAT-3' và khoảng 35 nucleotide tới hộp 5'-TTGACA-3' trên mạch mã hóa;
Hoặc vị trí C trong cặp nucleotide 5'-CT-3' cách xuôi dòng khoảng 10 nucleotide tới hộp 5'-ATTATA-3' và khoảng 35 nucleotide tới hộp 5'-TGTCGA-3' trên mạch làm khuôn (theo nguyên tắc bổ sung)
Từ trình tự trên có thể nhận thấy Vị trí khởi động phiên mã là vị trí số 6; hộp -10 từ vị trí 14 đến vị trí 19; hộp -35 từ vị trí 38 đến vị trí 43.
- b) Theo chiều ②, vì các hộp -10 và -35 nằm phía phải (nhìn từ ngoài vào) so với vị trí khởi đầu phiên mã.
- c) Mạch trình tự ở trên là mạch làm khuôn phiên mã, như đã được giải thích ở ý a).
- d) RNA polymerase của *E. coli* (holoenzyme) chứa yếu tố σ^{70} dùng 2 vòng xoắn α (đường kính mỗi vòng xoắn 20 Å) cài tương ứng vào hộp -10 và -35 là 2 rãnh lớn (rãnh chính) liền kề trên phân tử DNA sợi kép (có đường kính 22 Å). Các gốc cho và nhận liên kết hidro ở rãnh chính có tính đặc hiệu trình tự nucleotide, trong khi ở rãnh phụ (rãnh nhỏ) thì không.

Câu 27: Đề HSG Quốc gia năm 2023

Puromycin là kháng sinh có cấu trúc gần giống với phức hợp Tyr-tRNA. Puromycin được sử dụng để xác định aa-tRNA đang ở vùng P hay vùng A trên ribosome. Nếu aa-tRNA có ở vùng P của ribosome mà có puromycin ở vùng A thì aa-tRNA hình thành liên kết peptide với puromycin, tạo thành phức hợp aa-puromycin. Sau đó các phức hợp aa-puromycin và tRNA rời khỏi ribosome. Nếu aa-tRNA có ở vùng A thì nó liên kết chặt với ribosome, không tạo phức hợp aa-puromycin và aa-tRNA không rời khỏi ribosome. Ngoài ra, puromycin cũng có thể hình thành liên kết peptide với các aa-tRNA trong điều kiện không có ribosome nhưng với tốc độ chậm hơn khi ở trong ribosome.

Để xác định vị trí trên ribosome mà tRNA đi vào, người ta đánh dấu phóng xạ ^{35}S (kí hiệu [^{35}S]) amino acid methionine và một dẫn xuất của nó. Hai loại amino acid này được kí hiệu ngẫu nhiên là [^{35}S]Met1 và [^{35}S]Met2. [^{35}S]Met1 và [^{35}S]Met2 được sử dụng để tạo ba phức hợp [^{35}S]Met1-tRNA₁, [^{35}S]Met2-tRNA₂, [^{35}S]Met2-tRNA₁ (gọi chung là [^{35}S]aa-tRNA) có thể tham gia dịch mã. Sau đó, tiến hành các thí nghiệm trộn từng phức hợp [^{35}S]Met1-tRNA₁, hoặc [^{35}S]Met2-tRNA₁, hoặc [^{35}S]Met2-tRNA₂, với dung dịch gồm ribosome, puromycin và đoạn ba nucleotide AUG. Các thí nghiệm đối chứng tương tự các thí nghiệm trên nhưng không có ribosome. Sau 15 phút, xác định tỉ lệ (%) [^{35}S]aa-puromycin được giải phóng so với tổng số [^{35}S]aa-tRNA ban đầu, thu được kết quả ở Hình C22. Biết rằng các tRNA₁ và tRNA₂ có antiodon khớp với codon AUG.



Hình C22

- Puromycin ảnh hưởng tới quá trình dịch mã như thế nào? Giải thích.
- Giải thích tại sao đánh dấu phóng xạ ³⁵S ở [³⁵S]Met1, [³⁵S]Met2 trong thí nghiệm này.
- Dựa vào kết quả thí nghiệm, hãy cho biết phức hợp nào ([³⁵S]Met1-tRNA₁, [³⁵S]Met2-tRNA₂, [³⁵S]Met2-tRNA₁) được đưa vào vùng P của ribosome. Giải thích.

Hướng dẫn trả lời

- Puromycin làm kết thúc quá trình dịch mã. Giải thích: Puromycin có cấu trúc tương tự Tyr-tRNA nên có thể vào vùng A của ribosome đang tham gia dịch mã. Khi có mặt tại vùng A, puromycin hình thành liên kết peptide với đoạn pôlipeptide ở vùng P (đang gắn với tRNA) tạo nên phức hợp peptide-puromycin/peptidil-puromycin và phức hợp peptide-puromycin rời khỏi ribosome. Vì vậy, quá trình tổng hợp chuỗi pôlipeptide không tiếp tục xảy ra.
 - Trong thí nghiệm này, cần đánh dấu phóng xạ ³⁵S ở [³⁵S]Met1, [³⁵S]Met2 là vì gốc R của methionine có S và để xác định mức độ hình thành [³⁵S]Met1-puromycin hoặc [³⁵S]Met2-puromycin.
 - [³⁵S]Met1-tRNA và [³⁵S]Met2-tRNA, được đưa vào vùng P của ribosome.
 - Nếu aa-tRNA đặt vào vùng A, nó được giữ ở ribosome, ngăn cản ribosome xúc tác cho phản ứng hình thành liên kết peptide. Vì vậy tỉ lệ [³⁵S]aa-puromycin được giải phóng tương đương với tỉ lệ [³⁵S]aa-puromycin được giải phóng khi không có mặt của ribosome.
 - Ở thí nghiệm sử dụng [³⁵S]Met2-tRNA2 có dùng ribosome, tỉ lệ [³⁵S]aa-puromycin được giải phóng tương tự với thí nghiệm sử dụng [³⁵S]Met2-tRNA, không có ribosome. Điều này chứng tỏ [³⁵S]Met2-tRNA, được đặt vào vùng A.
 - Nếu aa-tRNA đặt vào vùng P, ribosome xúc tác phản ứng tạo liên kết peptide với puromycin, tạo phức aa-puromycin, sau đó phức này rời khỏi ribosome. Vì vậy ribosome có thể tham gia xúc tác/tổng hợp phản ứng hình thành liên kết peptide tiếp theo. Vì vậy, khi có ribosome, tỉ lệ [³⁵S]aa-puromycin được giải phóng cao hơn tỉ lệ [³⁵S]aa-puromycin được giải khi không có ribosome.
- Ở thí nghiệm sử dụng [³⁵S]Met1-tRNA, hoặc [³⁵S]Met2-tRNA, có ribosome [³⁵S]Met1-tRNA, tỉ lệ [³⁵S]aa-puromycin giải phóng tương ứng là 90% và 50% cao hơn với thí nghiệm sử dụng cùng phức hợp [³⁵S]Met1-tRNA (3,3%) hoặc [³⁵S]Met2-tRNA₁

(9,4%) khi không có ribosome, chứng tỏ [³⁵S]Met1-tRNA, và [³⁵S]Met2-tRNA đưa vào vùng P.

Câu 28:

Nội dung của một giả thuyết mới được phát biểu như sau: “Mã di truyền bộ ba được sử dụng phổ biến trong toàn sinh giới hiện nay thực chất được tiến hóa từ “mã di truyền bộ đôi” trong quá khứ”. Những người ủng hộ giả thuyết này cho rằng: độ đa dạng của các loại amino acid tham gia tổng hợp protein trong quá khứ là thấp hơn so với các loại amino acid thường gặp hiện nay, do đó các tế bào trong quá khứ chỉ cần mã bộ đôi là đủ để mã hóa cho toàn bộ các amino acid.

- a) Hãy cho biết có thể tìm thấy những bằng chứng nào ủng hộ cho giả thuyết “mã bộ đôi”?
- b) Ala, Val, Gly, Asp và Glu là những amino acid xuất hiện sớm nhất trong quá trình tiến hóa. Hãy quan sát bảng mã di truyền và chỉ ra điểm chung giữa các codon mã hóa cho 5 amino acid này? Trong những codon đó, cho biết nucleotide nào có vai trò phân biệt amino acid được mã hoá?
- c) Khi so sánh protein nguyên thủy với protein hiện đại, người ta nhận thấy cysteine, tyrosine, và phenylalanine là các amino acid xuất hiện tương đối muộn trong quá trình tiến hóa. Ngoài ra, mỗi amino acid trong ba amino acid trên đều chỉ được mã hóa bởi hai codon, trong khi những amino acid khác xuất hiện sớm hơn thì được mã hóa bởi nhiều codon hơn. Dựa vào thông tin vừa được cung cấp, ta có thể rút ra được những nhận xét gì về sự hình thành bộ mã di truyền?

Hướng dẫn trả lời

- a) Các bộ ba mã hoá cho cùng 1 amino acid thường sẽ giống nhau về hai nucleotide đầu tiên (tính từ đầu 5'), và chủ yếu khác nhau ở nucleotide cuối cùng. Chứng tỏ 2 nucleotide đầu tiên trong các codon đã được tiến hoá chọn lọc và ổn định từ rất sớm, trong khi nucleotide thứ ba được tiến hoá muộn hơn. Do đó, với điều kiện số lượng các loại amino acid trong quá khứ là thấp hơn so với hiện nay, ta có thể ra rằng mã bộ đôi đã từng được sử dụng vào thời kì nguyên thủy.
- b) Cả 5 amino acid Ala, Val, Gly, Asp, Glu đều được mã hoá bởi các bộ ba có dạng GNN (N là nucleotide bất kì). Trong đó, những codon khác loại (mã hoá cho các amino acid khác nhau) thì phân biệt với nhau ở nucleotide thứ hai tính từ đầu 5').
- c) Dễ dàng nhận thấy 5 amino acid xuất hiện sớm trong quá trình tiến hoá đều được quy định bởi những codon có dạng GNN; trong khi các amino acid hình thành muộn hơn (gọi tắt là amino acid 'muộn') lại được mã hóa bởi codon có dạng U(U, A, G)(U, C). Điều đó cho thấy các amino acid 'muộn' đã trải qua một quá trình tiến hoá dài, diễn ra với tốc độ chậm và xảy ra nhiều đột biến. Câu hỏi được đặt ra ở đây là: những codon hiện nay mã hóa cho các amino acid 'muộn' đã từng tồn tại trước đó với vai trò mã hóa cho các amino acid 'sớm, hay chỉ xuất hiện khi các amino acid "muộn được hình thành lần đầu trong sinh giới? Nếu như trường hợp đầu tiên là đúng thì sẽ lý giải được lý do vì sao số lượng các codon là có giới hạn, và số lượng các codon mã hóa cho những amino acid 'muộn" là ít hơn so với số lượng các codon mã hóa cho những amino acid sớm".

Cần chú ý thêm rằng những codon mã hóa cho các amino acid muộn và các bộ ba kết thúc đều có nucleotide khởi đầu là U. Do đó, một câu hỏi lại tiếp tục được đặt ra là: liệu các codon kết thúc có từng được dùng để mã hóa cho các amino acid 'muộn hay không? Nếu có, sự thay đổi chức năng của codon mã hóa sẽ ít gây gián đoạn cho quá trình sinh tổng hợp protein hơn so với hiện tượng được mô tả trong trường hợp đầu tiên được nhắc đến ở trên.

Câu 29:

Xét một exon đặc biệt thuộc gene mã hoá cho thụ thể hormone thần kinh ở thú; exon này tham gia vào quá trình tổng hợp hai loại protein khác nhau là XLas và ALEX. Chuỗi ký tự dưới đây là trình tự đầu 5' của exon đó khi được tách ra từ hệ gene chuột:

5'-gtcccaaccatgccaccgatcttcgcctgcttctgaagATGCGGGCCCAG-3'

Biết rằng phần chữ in thường sẽ tương ứng với đoạn exon khởi đầu mã hóa cho protein XLas; phần chữ in hoa sẽ tương ứng với đoạn exon khởi đầu mã hóa cho protein ALEX.

- Hãy cho biết trình tự nucleotide của phân tử RNA được phiên mã từ đoạn DNA trên.
- Xác định codon mở đầu của phân đoạn mã hóa cho protein XLas?
- Xác định codon mở đầu của phân đoạn mã hóa cho protein ALEX? Quá trình dịch mã của protein ALEX và protein XLas có xảy ra trên cùng 1 khung đọc hay không?
- Dựa vào đoạn DNA trên, hãy cho biết trình tự các amino acid của protein ALEX và protein XLas. Những codon gối lên nhau của hai khung đọc mở có mã hóa cho cùng 1 loại amino acid hay không?
- Hiện tượng một exon cùng tham gia mã hóa cho 2 loại protein khác nhau có những ưu điểm và nhược điểm gì? Giải thích.

Hướng dẫn trả lời

- 5'-gucccaaccaugeccaccgaucuuccgccugcuucugaagAUGCGGGCCCAG-3'
- 5'-gucccaacc**aug**eccaccgaucuuccgccugcuucugaagAUGCGGGCCCAG-3'
- 5'-gucccaaccaugeccaccgaucuuccgccugcuucugaag**AUG**CGGGCCCAG-3'

Quá trình dịch mã của protein ALEX và protein XLas xảy ra trên hai khung đọc khác nhau. Các bộ ba mã hóa của hai khung đọc này chỉ bị gối lên nhau (chứ không trùng nhau).

- XLas: Met-Pro-Thr-Asp-Leu-Pro - Pro-Ala-Ser-Glu-Asp-Ala-Gly-Pro
ALEX: Met-Arg- Ala – Gln

So sánh 2 đoạn trình tự trên, ta nhận thấy: những bộ ba thuộc vùng gối lên nhau của 2 khung đọc mở không mã hóa cho cùng một amino acid.

- Lợi ích: Thứ nhất, giúp tiết kiệm tối đa lượng vật chất di truyền cần phải sử dụng, qua đó làm giảm kích thước hệ gene nhưng vẫn đảm bảo duy trì đủ các chức năng cần thiết. Thứ hai, các đoạn gene gối lên nhau giúp tối ưu hóa thời gian và nguyên liệu cần thiết để thực hiện quá trình điều hòa biểu hiện gene ở mức độ phiên mã và dịch mã.

Tác hại: Khi một đột biến xảy ra ở một gene thuộc cụm các gene gối lên nhau, đột biến đó có thể làm ảnh hưởng đến hàng loạt những gene khác. Khi đó, thể đột biến nhiều khả năng sẽ bị tử vong.

Câu 30:

- a) Khi dịch mã phân tử mRNA được tạo thành từ nhiều đơn vị lặp tetranucleotide UUAC thì chỉ thu được ba loại amino acid. Tuy nhiên, khi dịch mã phân tử mRNA được tạo thành từ nhiều lần lặp lại đoạn tetranucleotide UAUC thì lại thu được bốn loại amino acid khác nhau. Hãy giải thích hiện tượng này?
- b) Khi dịch mã phân tử mRNA được tạo thành từ nhiều đơn vị lặp dinucleotide AC thì chỉ thu được hai loại amino acid là threonine và histidine. Khi dịch mã phân tử mRNA được tạo thành từ nhiều đơn vị lặp trinucleotide CAA thì lại thu được ba loại amino acid là glutamine, asparagine và threonine. Hãy xác định bộ ba nucleotide mã hóa cho amino acid Threonine.
- c) Trong một thí nghiệm dịch mã *in vitro*, các nhà khoa học đã tiến hành dịch mã những phân tử mRNA được tạo thành từ hai đơn vị lặp khác nhau và quan sát thành phần phân tử các loại amino acid thu được. Kết quả thí nghiệm được thể hiện trong **bảng C30** dưới đây:

Bảng C30.

Đơn vị lặp trên mRNA	Các loại codon trên mRNA	Amino acid xuất hiện trên chuỗi polypeptide
AG	AGA, GAG	Arg, Glu
AAG	AGA, AAG, GAA	Lys, Arg, Glu

Biết rằng AGG là bộ ba mã hóa Arginine (Arg). Hãy cho biết mỗi codon được liệt kê ở trên mã hóa cho loại amino acid nào?

Hướng dẫn trả lời

- a) Trên phân tử mRNA 5' - UUACUUACUUAC... - 3' có xuất hiện các codon như sau (tùy thuộc vào từng khung đọc được sử dụng trong quá trình dịch mã): UUA, CUU = Leu, UAC = Tyr, ACU = Thr.
Trong những codon trên, có hai codon mã hóa cho amino acid leucine (tính thoái hóa của mã di truyền), do đó chỉ thu được ba loại amino acid khác nhau. Trên phân tử mRNA 5' - UAUCUAUCUAUC... - 3' có xuất hiện các codon như sau (tùy thuộc vào từng khung đọc được sử dụng trong quá trình dịch mã): UAU = Tyr, AUC = Ile, UCU = Ser, CUA = Leu.
Trong trường hợp này, mỗi codon xuất hiện trên mRNA chỉ mã hóa cho một loại amino acid duy nhất, do đó có thể thu được bốn loại amino acid khác nhau.
- b) - Dịch mã phân tử mRNA 5' - ACAC... - 3' sẽ tạo ra 2 loại amino acid do trên phân tử mRNA này có 2 bộ ba mã hóa là ACA và CAC.
- Dịch mã phân tử mRNA 5' - CAACAA... - 3' sẽ tạo ra 3 loại amino acid do trên phân tử mRNA này có 3 bộ ba mã hóa là CAA, AAC và ACA.
- Trên cả hai phân tử mRNA được đề cập đến đều xuất hiện chung một bộ ba mã hóa là ACA, mà kết quả dịch mã hai loại mRNA đều thu được amino acid threonine. Suy ra ACA là bộ ba mã hóa cho amino acid threonine.
- c) - Dựa vào tính thoái hóa của bộ mã di truyền và hiện tượng "bắt cặp linh động ở nucleotide thứ ba", ta suy ra được nhận xét sau: các bộ ba cùng mã hóa cho một loại amino acid thường chỉ khác nhau ở nucleotide thứ ba.

- Ta đã biết AGG là bộ ba mã hóa cho amino acid arginine — bộ ba AGA cũng mã hóa cho amino acid arginine - bộ ba GAG mã hóa cho acid glutamic (dựa vào phân tử mRNA 5' AGAGAG ... 3') — bộ ba GAA cũng sẽ mã hóa cho acid glutamic.
- Còn lại là bộ ba AAG mã hóa cho amino acid lysine.

Câu 31:

Trọng lượng phân tử trung bình của protein được mã hóa trong bộ gene (genome) của người là khoảng 50.000 dalton. Một số protein lớn hơn rất nhiều so với mức trung bình này. Ví dụ, một protein gọi là titin, được tạo ra bởi các tế bào cơ, có trọng lượng phân tử là 3.000.000 dalton.

- Hãy ước tính một tế bào cơ sẽ mất bao lâu để dịch mã một mRNA mã hóa cho một protein có trọng lượng phân tử trung bình và một mRNA mã hóa cho titin. Khối lượng phân tử trung bình của amino acid là khoảng 110 dalton. Giả sử rằng tốc độ dịch mã là hai amino acid mỗi giây.
- Nếu các nucleotide trong phần mã hóa của mRNA chiếm 5% tổng số nucleotide được phiên mã thì một tế bào cơ sẽ mất bao lâu để phiên mã một gene mã hóa cho một protein có trọng lượng phân tử trung bình so với gene titin. Giả sử rằng tốc độ phiên mã là 20 nucleotide/giây.

Hướng dẫn trả lời

- Vì một protein trung bình chứa khoảng 455 amino acid $[(50,000 \text{ d/protein}) \times (\text{amino acid}/110 \text{ d})]$, nên một tế bào cơ sẽ mất khoảng **3,8 phút** để tạo ra nó $[(455 \text{ amino acids}) \times (\text{giây}/2 \text{ amino acids}) \times (\text{phút}/60 \text{ giây})]$.
- Vì titin có kích thước gấp 60 lần kích thước của một protein trung bình nên tế bào cơ sẽ cần **3,8 giờ** để tạo ra nó $[(3.000.000 \text{ Da}/110 \text{ Da}) / (2 \text{ amino acid}/\text{giây} \times 3600 \text{ giây}) \sim 3,8 \text{ giờ}]$. $[(3,000,000 \text{ d/titin}) \times (\text{amino acid}/110 \text{ d}) \times (\text{giây}/2 \text{ amino acids}) \times (\text{giờ}/3600 \text{ giây})]$
- Một tế bào cơ sẽ mất khoảng 23 phút để phiên mã một gene trung bình và 23 giờ để phiên mã gene titin. Đối với protein trung bình, 455 amino acid tương ứng với 1365 nucleotide của RNA $[(3 \text{ nt/codon}) \times (455 \text{ codon})]$. Cho rằng 5% bản phiên mã ban đầu được chuyển đổi thành mRNA, bản phiên mã ban đầu là $2,7 \times 10^4$ nucleotide $(1365 \text{ nt} \times 20)$, sẽ cần khoảng 23 phút để phiên mã $[(2,73 \times 10^4 \text{ nt}) \times (\text{giây}/20 \text{ nt}) \times (\text{phút}/60 \text{ giây})]$. Vì titin lớn gấp 60 lần nên tế bào cơ sẽ cần khoảng 23 giờ để phiên mã nó.

Câu 32:

Quá trình tổng hợp protein dùng bốn liên kết phosphate cao năng cho mỗi amino acid được thêm vào chuỗi polypeptide. Phiên mã dùng hai liên kết phosphate cao năng cho mỗi nucleotide được thêm vào chuỗi polynucleotide. Hãy tính xem có bao nhiêu phân tử protein sẽ được tạo ra từ một mRNA tại thời điểm mà chi phí năng lượng cho quá trình dịch mã bằng với chi phí năng lượng cho quá trình phiên mã. Giả sử rằng các nucleotide trong phần mã hóa của mRNA chiếm 5% tổng số được phiên mã.

Hướng dẫn trả lời

Chi phí năng lượng cho quá trình dịch mã và phiên mã sẽ bằng nhau khi **30 phân tử protein được tạo ra từ một mRNA**.

Quá trình tổng hợp protein đòi hỏi bốn liên kết phosphate cao năng trên mỗi codon (trên ba nucleotide). Quá trình phiên mã tiêu tốn 6 liên kết phosphate cao năng để tạo ra một codon, nhưng cũng tiêu tốn gấp 19 lần năng lượng tổng hợp RNA sẽ bị loại bỏ (95%). Do đó, quá trình phiên mã tiêu tốn tổng cộng 120 liên kết phosphate cao năng trên mỗi codon ($6 + 114$), so với 4 liên kết trên mỗi codon để dịch mã. Tỷ lệ chi phí năng lượng cho mỗi codon ($120/4 = 30$) xác định số lượng phân tử protein sẽ được tạo ra khi chi phí năng lượng của quá trình dịch mã phù hợp với quá trình phiên mã. Vì hầu hết các mRNA được sử dụng để tạo ra hàng trăm đến hàng nghìn protein nên quá trình dịch mã tiêu tốn một phần năng lượng của tế bào cao hơn nhiều so với quá trình phiên mã.

Câu 33:

Cả sinh vật nhân sơ và sinh vật nhân chuẩn đều bảo vệ chống lại sự nguy hiểm của việc dịch mã các mRNA bị đứt gãy. Những mối nguy hiểm nào mà việc dịch mã các mRNA bị đứt gãy này có thể gây ra cho tế bào?

Hướng dẫn trả lời

Một mRNA bị ngắn đi do đứt gãy hay sai hỏng khi được dịch mã sẽ tạo ra một loại protein bị cắt ngắn có thể gây hại cho tế bào.

- Một đoạn protein có thể giữ lại/vẫn duy trì một số chức năng của toàn bộ protein, ví dụ, cho phép nó liên kết với protein mục tiêu nhưng giữ nó trong một phức hợp không hiệu quả.

- Ngoài ra, một đoạn protein có thể bộc lộ các bề mặt liên kết mới, bất thường cho phép nó liên kết với các phân tử mới, can thiệp vào chức năng của chúng.

- Cuối cùng, việc loại bỏ một phần protein có thể loại bỏ phần protein thường kiểm soát hoạt động của nó. Trong trường hợp này, protein bị cắt ngắn có thể bị khóa ở trạng thái hoạt động, gây hậu quả nghiêm trọng cho tế bào.

Câu 34:

Một số thí nghiệm đã được tiến hành để thu thập thông tin về cách thức ribosome của sinh vật nhân thực nhận ra codon khởi đầu AUG. Trong một thí nghiệm, gene mã hóa tRNA mở đầu mang methionine ($\text{tRNA}_i^{\text{Met}}$) đã được xác định vị trí và thay đổi. Các nucleotide xác định anticodon trên $\text{tRNA}_i^{\text{Met}}$ đã bị đột biến sao cho anticodon trong tRNA là 5'-CCA-3' thay vì 5'-CAU-3'. Khi gene đột biến này được chuyển vào một tế bào nhân chuẩn, quá trình tổng hợp protein diễn ra, nhưng các protein được tạo ra bất thường. Một số protein được tổng hợp có chứa thêm amino acid và một số khác chứa ít amino acid hơn bình thường.

- Những kết quả này cho thấy gì về cách ribosome nhận ra điểm bắt đầu dịch mã trong các tế bào nhân chuẩn? Giải thích.
- Nếu thí nghiệm tương tự được tiến hành trên tế bào vi khuẩn, bạn sẽ mong đợi kết quả gì?
- Giải thích tại sao một số protein chứa thêm amino acid trong khi một số khác chứa ít amino acid hơn bình thường.

Hướng dẫn trả lời

- Do đột biến anticodon thành 5'-CCA-3' từ 5'-CAU-3' trên $\text{tRNA}_i^{\text{Met}}$, tRNA khởi đầu giờ đây sẽ nhận ra codon 5'-UGG-3', mà thường chỉ mã hóa cho Trp.

- Nếu quá trình dịch mã bắt đầu bởi ribosome ở sinh vật nhân thực xảy ra bằng cách gắn vào đầu 5' của mRNA, sau đó quét/dò tìm (scan), thì codon 5'-UGG-3' đầu tiên được tRNA^{Met} đột biến nhận ra sẽ là vị trí bắt đầu dịch mã.

- Nếu codon 5'-UGG-3' đầu tiên xuất hiện trước codon 5'-AUG-3' bình thường, thì một loại protein có chứa thêm amino acid có thể được tạo ra. Nếu codon 5'-UGG-3' đầu tiên xuất hiện sau 5'-AUG-3' bình thường, thì một protein ngắn hơn sẽ được tạo ra.

- Các protein bị ngắn hơn cũng có thể được tạo ra do 5'-UGG-3' đầu tiên nằm ngoài khung của trình tự mã hóa thông thường. Nếu điều này xảy ra, thì rất có thể một codon kết thúc sẽ gặp phải trước khi kết thúc trình tự mã hóa bình thường và sẽ kết thúc quá trình dịch mã.

→ Dữ liệu gợi ý rằng quá trình bắt đầu dịch mã diễn ra bằng cách quét ribosome để tìm trình tự bắt đầu thích hợp.

- b) Rất ít hoặc không có sự tổng hợp protein. Sự bắt đầu dịch mã ở vi khuẩn đòi hỏi RNA 16S của tiểu đơn vị nhỏ của ribosome tương tác với trình tự Shine-Dalgarno. Sự tương tác này phục vụ để sắp xếp ribosome trên codon khởi đầu. Nếu anticodon đã bị thay đổi đến mức không thể nhận ra codon khởi đầu, thì quá trình tổng hợp protein sẽ không diễn ra.

Câu 34:

Các kết quả sau đây thu được trong các nghiên cứu ban đầu về quá trình dịch mã của các protein tiết. Dựa trên những gì chúng ta biết về quá trình này, hãy giải thích lý do tại sao từng kết quả được quan sát thấy.

- a) Một hệ thống dịch mã *in vitro* chỉ bao gồm mRNA và ribosome đã tạo ra các protein tiết lớn hơn protein chính protein đó khi được dịch mã trong tế bào.
- b) Một hệ thống tương tự nhưng bổ sung thêm các microsome (túi có màng bao bọc được hình thành chủ yếu từ màng lưới nội chất bị phá vỡ nhưng một số phát sinh từ màng sinh chất) tạo ra các protein tiết có kích thước giống hệt với các protein được tìm thấy trong một tế bào.
- c) Khi các microsome được thêm vào sau quá trình dịch mã *in vitro*, các protein được tổng hợp lại lớn hơn so với các protein được tạo ra trong tế bào.

Hướng dẫn trả lời

- a) Khi không có màng ER, toàn bộ protein được dịch mã và chuỗi tín hiệu ER vẫn còn trên protein.
- b) Khi quá trình dịch mã xảy ra với sự có mặt của các microsome chứa ER, protein sẽ được dịch vào lòng của các microsome. Theo quy trình này, chuỗi tín hiệu được phân cắt để tạo ra một protein nhỏ hơn.
- c) Dịch mã và chuyển vị trên màng ER là các quá trình đồng thời. Nếu chúng không xảy ra đồng thời, thì protein không được nhập đúng cách vào ER nơi chuỗi tín hiệu có thể bị cắt (mặc dù có một số ví dụ về chuyển vị sau dịch mã).

II. CÂU HỎI VÀ BÀI TẬP TỰ TRẢ LỜI

Câu 1:

RNA polymerase của vi khuẩn kéo dài RNA với tốc độ 70 nucleotide mỗi giây và mỗi phức hợp phiên mã bao phủ 70 bp DNA.

- Số lượng phân tử RNA tối đa có thể được tạo ra trong một phút từ một gene có kích thước 6000 bp là bao nhiêu? (Giả sử rằng tốc độ khởi đầu không bị giới hạn.)
- Số phức hợp phiên mã tối đa có thể liên kết với gene này tại một thời điểm là bao nhiêu?

Câu 2:

Bộ gene của *E. coli* có kích thước xấp xỉ 4600 kb và chứa khoảng 4000 gene. Bộ gen của động vật có vú có kích thước xấp xỉ 33×10^6 bp và chứa tối đa 30.000 gene. Trung bình một gene ở *E. coli* dài 1000 bp.

- Tính phần trăm DNA của vi khuẩn *E. coli* không được phiên mã.
- Mặc dù nhiều gene của động vật có vú lớn hơn gene của vi khuẩn, nhưng hầu hết các sản phẩm của gene ở động vật có vú đều có cùng kích thước với các sản phẩm của gene của vi khuẩn. Tính tỷ lệ phần trăm DNA trong exon trong bộ gene của động vật có vú.

Câu 3:

Một ribosome của vi khuẩn có thể tổng hợp khoảng 20 liên kết peptide mỗi phút. Nếu protein vi khuẩn trung bình dài khoảng 260 amino acid, thì có bao nhiêu protein mà các ribosome trong tế bào *E. coli* có thể tổng hợp trong 20 phút nếu tất cả các ribosome đều hoạt động với tốc độ tối đa? Biết rằng tế bào vi khuẩn trung bình có khoảng 15.000 ribosome.

Câu 4:

Bộ gene của vi khuẩn thường chứa nhiều bản sao gene cho rRNA. Chúng được phiên mã rất hiệu quả để tạo ra một lượng lớn rRNA để lắp ráp thành ribosome. Ngược lại, các gen mã hóa protein ribosome chỉ hiện diện dưới dạng các bản sao đơn lẻ. Giải thích sự khác nhau về số bản sao của các gene rRNA và protein ribosome.

Câu 5:

Tại sao sự điều hòa phiên mã thường liên quan đến các trình tự khởi đầu phiên mã (promoter) và tương tác giữa các trình tự này trên DNA với các loại protein?

Câu 6:

Phản ứng hình thành liên kết peptide được coi là phản ứng trùng ngưng, nhưng sao nó cũng được gọi là phản ứng loại nước?

Câu 7:

Để kiểm tra có đúng bộ ba 5'-GUG-3' mã hóa cho Valine không, bạn phải thiết kế trình tự mRNA gồm cặp nucleotide lặp lại nào? Khi ribosome dịch mã trình tự lặp lại cặp nucleotide đó, loại amino acid nào khác cũng đồng thời được tạo ra trong cuối polypeptide? Giải thích.

Câu 8:

- a) Một đặc điểm đáng chú ý của mã di truyền là các amino acid có tính chất hóa học giống nhau thường có các codon giống nhau. Ví dụ, các codon có U hoặc C là nucleotide thứ hai, có xu hướng xác định các amino acid kỵ nước. Bạn có thể đề xuất một lời giải thích khả thi cho hiện tượng này về sự tiến hóa ban đầu của bộ máy tổng hợp protein không?
- b) Các mRNA đa cistron (polycistronic mRNAs) là phổ biến ở sinh vật nhân sơ nhưng cực kỳ hiếm ở sinh vật nhân chuẩn. Mô tả sự khác biệt chính trong quá trình tổng hợp protein làm cơ sở cho quan sát này.

Câu 9:

Nấm mũ tử thần, *Amanita phalloides*, chứa một số chất nguy hiểm, bao gồm α -amanitin gây chết người. Độc tố này ngăn chặn sự kéo dài RNA ở người ăn phải nấm bằng cách liên kết với RNA polymerase II của sinh vật nhân chuẩn với ái lực rất cao; nó gây chết người ở nồng độ thấp tới 10^{-8} M. Phản ứng ban đầu khi ăn phải nấm là đau dạ dày (do một số chất độc khác gây ra). Những triệu chứng này biến mất, nhưng khoảng 48 giờ sau, người ăn nấm chết, thường là do rối loạn chức năng gan. Suy đoán về lý do tại sao phải mất nhiều thời gian như vậy để α -amanitin gây chết người.

Câu 10:

Trong những ngày đầu nghiên cứu về ribosome, trước khi vai trò chính xác của ribosome được làm rõ, một nhà nghiên cứu đã đưa ra quan sát sau đây. Cô ấy có thể tìm thấy, trong các thí nghiệm lắng đọng trên dịch ly giải vi khuẩn, không chỉ các hạt 30S, 50S và 70S mà còn cả một số hạt lắng đọng ở khoảng 100S và 130S. Khi cô ấy xử lý hỗn hợp như vậy bằng EDTA, mọi thứ phân tách thành các hạt 30S và 50S. Khi thêm các ion hóa trị hai, cô ấy có thể lấy lại các hạt 70S, nhưng không bao giờ lấy lại được các hạt 100S hoặc 130S.

- a) Đề xuất những gì các hạt 100S và 130S có thể đại diện, dựa trên kiến thức hiện tại về tổng hợp protein. Nhà nghiên cứu đã bỏ lỡ khám phá quan trọng nào?
- b) Tại sao bạn nghĩ rằng sự liên kết lại với các hạt 100S và 130S không hoạt động?

Câu 11:

Protein kháng khuẩn colicin E3 là chất ức chế hiệu quả quá trình tổng hợp protein ở sinh vật nhân sơ. Protein này là một nuclease, đặc biệt tấn công liên kết phosphodiester ở gần cuối 16S RNA. Đề xuất cơ chế tác dụng của colicin E3 đối với quá trình dịch mã.

Câu 12:

Gene *JLB-1*, được biểu hiện trong một số cơ quan của người, chứa bảy exon (1 đến 7) và sáu intron (A đến F). Ba mẫu dò oligonucleotide (I đến III), lai bổ sung tương ứng với exon 2, 4 và 7, được biểu thị bằng dấu hoa thị bên dưới bản đồ gene:

mRNA trưởng thành được phân lập từ ba mô biểu hiện gene *JLB-1* và được nghiên cứu bằng phương pháp lai Northern blot sử dụng ba mẫu dò oligonucleotide được chỉ ra ở trên. Các mẫu dò liên kết với các trình tự



bổ sung trong mRNA. Các mẫu lai Northern blot giữa mỗi mẫu dò và mRNA được phân lập từ tế bào máu, gan và thận được hiển thị. Đối với mỗi mẫu lai Northern blot:

- Giải thích ý nghĩa của kết quả phép lai.
- Xác định quá trình sinh học hoặc các quá trình giải thích cho các mẫu lai quan sát.

Câu 13:

Gene *ADAM12* của người mã hóa một protein liên kết màng có chức năng trong quá trình phát triển tế bào xương và cơ. Trình tự đầu N của protein được mã hóa bởi mRNA *ADAM12* không giống với trình tự đầu N của polypeptide được tìm thấy trong màng tế bào: polypeptide được tìm thấy trong màng tế bào thiếu 28 amino acid đầu tiên của polypeptide được mã hóa bởi mRNA. Sự sắp xếp sau đây thu được khi hai trình tự được so sánh bằng cách sử dụng mã đơn của amino acid:

Trình tự amino acid mã hóa bởi mRNA:

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSAS . . .

Polypeptide trong màng:

-----RGVSLWNQGRADEVVSAS . . .

- Giải thích tại sao trình tự đầu N của polypeptide có trong màng tế bào không giống với polypeptide được mã hóa bởi mRNA của nó.
- Giả sử một sự mất đoạn nhỏ xảy ra trong gene và một mRNA được tổng hợp, dẫn đến việc loại bỏ các codon cho các amino acid PLPVSPARALLLALAGALL từ đầu 5' của mRNA. Bạn mong đợi đột biến này có ảnh hưởng gì đến sự phân bố trong tế bào của protein ADAM12?

Câu 14:

Các nhà nghiên cứu đang nghiên cứu việc sử dụng vaccine RNA để bảo vệ con người chống lại một số bệnh. Để phát triển vaccine, các tế bào cụ thể trước tiên được loại bỏ khỏi một cá thể. Sau đó, các mRNA mã hóa cho các protein cụ thể từ mầm bệnh được đưa vào tế bào. Các tế bào đã thay đổi được tiêm trở lại vào cá thể, nơi các tế bào tạo ra các protein được mã hóa bởi các mRNA được đưa vào. Sau đó, cá thể tạo ra phản ứng miễn dịch đối với các protein sẽ giúp bảo vệ cá nhân khỏi phát triển bệnh nếu tiếp xúc với mầm bệnh trong tương lai.

Khi được đưa vào tế bào, các mRNA được sử dụng cho vaccine phải ổn định để chúng không bị phân hủy trước khi các protein mã hóa được tạo ra. Các nhà nghiên cứu đã phát triển một số mũ đầu 5' sửa đổi mà họ đưa ra giả thuyết có thể làm cho các mRNA được đưa vào tế bào ổn định hơn so với các mRNA có mũ GTP bình thường. Để kiểm tra tác động của các mũ sửa đổi, các nhà nghiên cứu đã tạo ra các mRNA chỉ khác nhau về cấu trúc mũ của chúng (không có mũ, mũ bình thường hoặc mũ được biến đổi I, II hoặc III). Họ đã đưa cùng một lượng của mỗi mRNA vào các nhóm tế bào khác nhau và đo lượng thời gian cần thiết để một nửa số mRNA phân hủy (thời gian bán rã của mRNA) và tổng lượng protein được dịch mã từ các mRNA (**Bảng C14**).

Bảng C14

Cấu trúc mũ 5'	Thời gian bán rã mRNA $\pm 2SE_{\bar{x}}$ (giờ sau khi đưa vào tế bào)	Tổng lượng protein được dịch mã mRNA $\pm 2SE_{\bar{x}}$ (so với lượng tạo ra ở mRNA có mũ bình thường)
Không có mũ	1.41 ± 0.02	0.011 ± 0.000
Mũ GTP bình thường	16.10 ± 1.83	1.000 ± 0.007
Mũ được biến đổi I	15.50 ± 1.57	4.777 ± 0.042
Mũ được biến đổi II	27.00 ± 2.85	13.094 ± 0.307
Mũ được biến đổi III	18.09 ± 0.81	6.570 ± 0.075

- Dựa trên dữ liệu, hãy **xác định** cấu trúc mũ nào có nhiều khả năng bảo vệ phần đầu của các mRNA khỏi sự phân hủy.
- Dựa trên dữ liệu về các mRNA có mũ biến đổi, hãy **mô tả** mối quan hệ giữa thời gian bán rã của mRNA và tổng lượng protein được tạo ra.
- Sau khi kiểm tra dữ liệu về thời gian bán rã của mRNA và lượng protein được tạo ra, các nhà nghiên cứu đã đưa ra giả thuyết rằng mỗi phân tử mRNA có nắp được biến đổi I được dịch mã thường xuyên hơn so với từng phân tử mRNA có nắp GTP bình thường. **Đánh giá** giả thuyết của họ bằng cách so sánh dữ liệu trong Bảng 1.
- Việc đưa các mRNA vào tế bào cho phép các tế bào tạo ra các protein lạ mà bình thường chúng không thể tạo ra. Giải thích tại sao việc sản xuất một loại protein ngoại lai có thể có nhiều khả năng hơn từ việc đưa mRNA vào tế bào hơn là DNA.

Câu 15:

Nhìn chung, nếu chúng ta biết trình tự DNA bộ gene của một gene, chúng ta có thể dự đoán một cách đáng tin cậy trình tự nucleotide của RNA được mã hóa bởi gene đó. Phát biểu này có đúng với tRNA ở sinh vật nhân sơ không? Còn tRNA ở sinh vật nhân thực thì sao?

Câu 16:

Rifampicin là một hợp chất bán tổng hợp được làm từ rifamycin B, một loại kháng sinh được phân lập từ *Streptomyces mediterranei*. Rifampicin là một loại thuốc chống mycobacteria đã được phê duyệt, là một thành phần tiêu chuẩn của chế độ điều trị kết hợp để điều trị bệnh lao và nhiễm trùng tụ cầu kháng penicillin.

Các nghiên cứu gần đây cho thấy bệnh lao kháng rifampicin đang trở nên phổ biến hơn. Ví dụ, 2% mẫu từ một cuộc khảo sát ở Botswana được phát hiện kháng thuốc. **Bảng** đưa ra

Bảng C16

Chủng <i>E. coli</i>	Rifampicin* ($\mu\text{g/mL}$)
Kiểu đại	< 5
Asp516Tyr ở tiểu đơn vị β	> 50

* Nồng độ rifampicin tại điểm ngừng tăng trưởng của *E. coli*

một số kết quả từ *E. coli* kiểu đại và *E. coli* với một sự thay đổi amino acid duy nhất trong tiểu đơn vị của β của RNA polymerase (Asp thành Tyr ở vị trí amino acid 516) và phản ứng tăng trưởng của chúng với môi trường chứa rifampicin.

- Bạn giải thích dữ liệu như thế nào?
- Tiểu đơn vị có vai trò gì trong RNA polymerase?

c) Mô tả một cơ chế vi khuẩn kháng rifampicin.

Câu 17:

Trong một số trường hợp hiếm hoi, bộ máy dịch mã gặp phải một codon không thể giải mã nhanh chóng do thiếu tRNA hoặc yếu tố giải phóng cụ thể. Trong những trường hợp này, ribosome có thể tạm dừng và sau đó dịch chuyển một nucleotide đơn lẻ và bắt đầu dịch một khung đọc khác. Sự cố như vậy được gọi là **dịch mã dịch khung (translational frameshifting)**. Yếu tố giải phóng RF-2 ở *E. coli*, được dịch mã từ mRNA chứa codon kết thúc UGA bên trong, được tạo ra bằng cách dịch khung. Giải thích làm thế nào hiện tượng này có thể điều hòa sự tạo thành RF-2.

Câu 18:

U nguyên bào võng mạc là một loại ung thư hiếm gặp, trong đó các khối u phát triển ở võng mạc mắt. Các khối u phát sinh do mất gene *Rb* mã hóa chất ức chế khối u. U nguyên bào võng mạc di truyền thường xuất hiện ở trẻ em chỉ được thừa hưởng một bản sao chức năng của *Rb*. Giải thích tại sao dạng u nguyên bào võng mạc không di truyền thường xuất hiện muộn hơn trong cuộc đời.

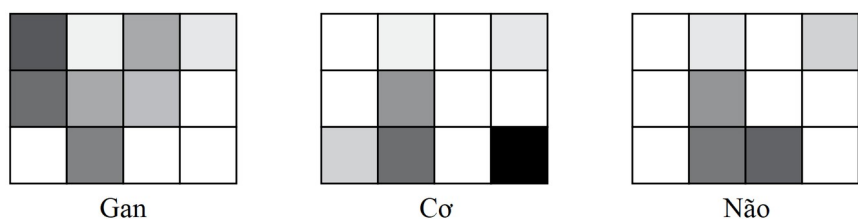
Câu 19:

Enzyme polynucleotide phosphorylase (PNPase) của vi khuẩn là một exoribonuclease 3' → 5' có tác dụng phân hủy mRNA.

- Enzyme xúc tác phản ứng giải phosphoryl hóa (phosphorolysis), như glycogen phosphorylase, chứ không phải phản ứng thủy phân. Viết phương trình phản phân giải phosphoryl hóa mRNA.
- Trong ống nghiệm, PNPase cũng xúc tác quá trình đảo ngược phản ứng giải phosphoryl hóa. Phản ứng này đạt được điều gì và nó khác với phản ứng do RNA polymerase thực hiện như thế nào?
- PNPase bao gồm một vị trí gắn kết cho các ribonucleotide dài, có thể thúc đẩy quá trình xử lý của enzyme (khả năng của enzyme xúc tác cho các phản ứng liên tiếp mà không giải phóng cơ chất của nó). Tại sao đây lại là một lợi thế cho hoạt động chính của PNPase *in vivo*?

Câu 20:

Các nhân tế bào được phân lập từ não, gan và cơ. Sau đó, các nhân này được ủ với α - $[^{32}\text{P}]\text{UTP}$ trong các điều kiện cho phép tổng hợp



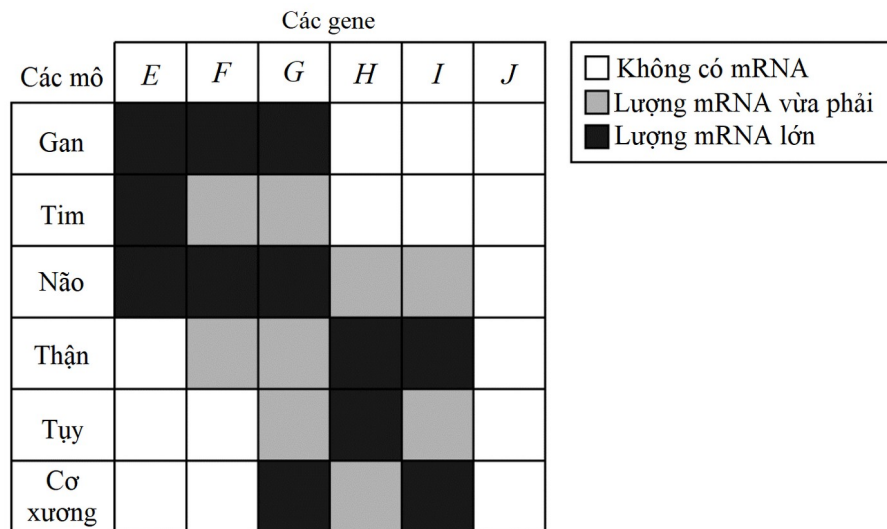
RNA, ngoại trừ việc có mặt chất ức chế khởi tạo RNA. RNA phóng xạ được phân lập và ủ với các trình tự DNA khác nhau đã được gắn vào chip gene (gene chip). Trong các biểu đồ liên kết, độ đậm nhạt cho biết số lượng mRNA tương đối được gắn vào mỗi chuỗi DNA.

- Tại sao cường độ lai giữa các gene lại khác nhau?

- b) Ý nghĩa của việc một số phân tử RNA thể hiện các kiểu lai hóa khác nhau trong các mô khác nhau là gì?
- c) Một số gene được biểu hiện ở cả ba mô. Bạn đoán bản chất của những gen này là gì?
- d) Nêu lý do tại sao lại có chất ức chế khởi mào trong hỗn hợp phản ứng.

Câu 21:

Một nhà nghiên cứu đang nghiên cứu các kiểu biểu hiện gen ở chuột. Nhà nghiên cứu đã thu thập các mẫu từ sáu mô khác nhau ở một con chuột khỏe mạnh và đo lượng mRNA từ sáu gene. Dữ liệu được hiển thị trong Hình C21.



Hình C21

- a) Dựa trên dữ liệu được cung cấp, hãy xác định gene có khả năng mã hóa protein là thành phần thiết yếu của quá trình đường phân. Giải thích.
- b) Nhà nghiên cứu quan sát thấy rằng các mô có hàm lượng mRNA của gene H cao không phải lúc nào cũng có protein gene H. Cung cấp lý do để giải thích làm thế nào các mô có hàm lượng mRNA của gene H cao có thể không có protein gene H.

Câu 22:

Em đang nghiên cứu locus mã hoá một protein nhỏ có trình tự được biểu diễn dưới đây. Trình tự được biểu diễn từ vị trí khởi đầu phiên mã ở đầu 5' đến vị trí kết thúc phiên mã ở đầu 3'. Trình tự này mã hóa hai chuỗi polypeptide khác nhau. Có một trình tự intron ngắn được hiển thị với các chữ cái **in đậm**.

5. 10. 15. 20. 25. 30. 35. 40. 45. 50. 55.

5' –CTACGTACTATGTATTCC**GATCT**ACTCGATCTAGTCGCATTCCGATAAGATCGTAC– 3'

3' –GATGCATGATACATAAGG**CTAGAT**ATGAGCTAGATCAGCGTAAGGCTATTCTAGCATG– 5'

- a) mRNA cuối cùng sau cải biến được tạo ra từ gene này có bao nhiêu nucleotide (không bao gồm mũ 5' và đuôi polyA đầu 3')?
- b) Hãy chỉ ra mạch được sử dụng làm khuôn trong phiên mã cho chuỗi polypeptide ngắn hơn, là mạch 5'-3' hay mạch 3'-5' (đọc theo hướng từ trái sang phải)?
- c) Tỷ lệ số lượng amino acid trong chuỗi polypeptide dài hơn / ngắn hơn (bao gồm cả amino acid được mã hoá bởi codon mở đầu) là bao nhiêu?
- d) Amino acid thứ ba của chuỗi polypeptide ngắn hơn là gì, nếu Met được tính là amino acid đầu tiên? Ghi kí hiệu một chữ cái của amino acid đó dựa vào bảng mã di truyền.

Câu 23:

Nhiều loại kháng sinh, hầu hết được phân lập từ nấm, hoạt động bằng cách ức chế sự khởi đầu, kéo dài hoặc kết thúc các peptide trong quá trình tổng hợp protein. Các hợp chất này rất hữu ích trong nghiên cứu tổng hợp protein, chủ yếu là do một loại kháng sinh riêng lẻ thường can thiệp vào quá trình tổng hợp protein ở một bước được xác định rõ trong quá trình dịch mã phức tạp. Giả sử rằng bạn đã phân lập được một loại kháng sinh mới từ nấm *Pilobolus*. Loại kháng sinh này ức chế sự phát triển của vi khuẩn và quá trình tổng hợp protein của vi khuẩn nhưng không ảnh hưởng đến sự phát triển hoặc quá trình tổng hợp protein trong tế bào nhân chuẩn. Làm thế nào bạn có thể xác định thành phần của bộ máy dịch mã bị ảnh hưởng bởi loại kháng sinh mới này?

Câu 24:

Bệnh cơ tim ở trẻ sơ sinh là một chứng rối loạn nghiêm trọng có thể gây tử vong trong năm đầu đời do khiếm khuyết chức năng của cơ tim do rối loạn chức năng ty thể. Một nghiên cứu, được thực hiện bởi Götz và cộng sự [(2011). *Am. J. Hum. Genet.* 88:635–642), đã xác định được hai đột biến nguyên nhân khác nhau trong gene mã hoá cho alanyl-tRNA synthetase của ty thể (mtAlaRS). Một đột biến làm thay đổi gốc leucine ở vị trí amino acid 155 thành arginine (p.Leu155Arg). Đột biến còn lại biến arginine ở vị trí 592 thành tryptophan (p.Arg592Trp). Enzyme mtAlaRS có miền đầu tận cùng N (amino acid 36–481) xúc tác quá trình aminoacyl hóa tRNA và miền chỉnh sửa bên trong (amino acid 484–782) xúc tác quá trình khử acyl hoá (deacylation) trong trường hợp tRNA được nạp nhầm amino acid.

- Xem xét vị trí của bệnh gây ra đột biến nhằm nghĩa trong gen mtAlaRS trong phạm vi các miền protein đã biết của enzyme này. Bạn có thể dự đoán gì về cách những đột biến này làm suy yếu quá trình tổng hợp protein trong ty thể theo những cách khác nhau?
- Bạn dự đoán đột biến nào sẽ làm suy giảm khả năng dịch mã ở ty thể nghiêm trọng hơn, và tại sao?

Câu 25:

Các quan sát gần đây chỉ ra rằng cắt nối thay thế (hay cắt nối có lựa chọn) là một cách phổ biến để sinh vật nhân thực mở rộng các chức năng gene của chúng. Các nghiên cứu chỉ ra rằng khoảng 50% gene của con người thể hiện sự cắt nối thay thế và khoảng 15% các đột biến gây bệnh liên quan đến sự cắt nối thay thế bất thường. Các mô khác nhau cho thấy các tần số cắt nối thay thế khác nhau đáng kể, trong đó não chiếm khoảng 18% các sự kiện như vậy [Xu et al. (2002). *Nucl. Acids Res.* 30:3754–3766].

- Định nghĩa về cắt nối thay thế và suy đoán về chiến lược tiến hóa mà cách cắt nối thay thế mang lại cho các sinh vật.
- Tại sao một số mô có thể tham gia vào quá trình cắt nối thay thế nhiều hơn những mô khác?

Câu 26:

Bảng sau đây liệt kê trình tự 12 nucleotide trước codon khởi đầu và nucleotide đầu tiên sau codon khởi đầu của 8 gene α -globin và 8 gene β -globin của động vật có xương

sống được báo cáo bởi Kozak (1987). Các trình tự được viết từ -12 đến +4 với trình tự codon mở đầu (MĐ) bằng chữ in hoa.

Bảng C26

	Trình tự của gene
	-12 MĐ+4
Họ gene α-globin	
Người trưởng thành	agagaacccaccATGg
Phôi người	caccctgccgccATGt
Khỉ đầu chó	ccagcgcgggacATGg
Chuột trưởng thành	caggaagaaaccATGg
Thỏ trưởng thành	gaaggaaccaccATGg
Phôi dê	tcagctgccaccATGt
Vịt trưởng thành	ggagctgcaaccATGg
Phôi gà	ctctcctgcacaATGg
Họ gene β-globin	
Thai người	agtccagacgccATGg
Phôi người	aggcctggcatcATGg
Thỏ trưởng thành	aaacagacagaATGg
Phôi thỏ	agaccagacatcATGg
Gà trưởng thành	ccaaccgccgccATGg
Phôi gà	cccgcctgccaccATGg
<i>Xenopus</i> (chi cá Cóc có móng) trưởng thành	tcaactttggccATGg
Ấu trùng <i>Xenopus</i>	tctacagccaccATGg

Sử dụng dữ liệu trong bảng này để:

- Xác định trình tự liên ứng của 16 gene α -globin và β -globin đã chọn.
- So sánh trình tự liên ứng của các gene globin này với trình tự liên ứng thu được từ nghiên cứu lớn hơn về 699 gene của động vật có xương sống trong **Câu 8 phần Câu hỏi và bài tập có hướng dẫn trả lời**.

PHẦN C: KẾT LUẬN

1. Kết luận

Sau khi thực hiện, chuyên đề đã thực hiện được các vấn đề sau:

- Nghiên cứu các vấn đề lý thuyết của hai cơ chế di truyền phân tử chính trong quá trình biểu hiện gene đó là phiên mã và dịch mã. Từ đó, chuyên đề đã khái quát, hệ thống các kiến thức cơ bản và chuyên sâu về cơ chế phiên mã và dịch mã, giúp giáo viên và học sinh chủ động trong nghiên cứu, giảng dạy, học tập.

- Chuyên đề cũng đã đưa ra được một số các câu hỏi vận dụng kiến thức về các cơ chế này để giúp học sinh có cơ hội luyện tập, vận dụng các kiến thức trong quá trình học tập và ôn tập cho các kì thi học sinh giỏi.

- Thông qua nghiên cứu và xây dựng chuyên đề, bản thân tác giả cũng trau dồi thêm nhiều kiến thức, phát triển chuyên môn của bản thân trong công tác giảng dạy tại trường THPT chuyên và bồi dưỡng các đội tuyển học sinh giỏi.

2. Đề xuất

Do những hạn chế về thời gian nghiên cứu và nội dung chuyên đề được xây dựng theo hiểu biết và cách tiếp cận của tác giả, những kết quả nội dung chuyên đề trên đây mới chỉ là bước đầu, tác giả hy vọng sẽ có thêm những chuyên đề tiếp tục nghiên cứu, phát triển các nội dung liên quan đến cơ chế biểu hiện gene ở ở các nhóm sinh vật khác nhau cũng như việc ứng dụng những hiểu biết này vào thực tế đời sống nhằm tạo nguồn tài liệu phục vụ cho công tác dạy và học môn Sinh học ở trường THPT chuyên và đáp ứng nhu cầu tài liệu cho công tác bồi dưỡng học sinh giỏi các cấp.

Chuyên đề của tôi chắc hẳn không tránh khỏi những thiếu sót, nên rất mong nhận được sự góp ý tận tình của quý thầy, cô để chuyên đề hoàn thiện hơn!

TÀI LIỆU THAM KHẢO

TÀI LIỆU TIẾNG VIỆT

- [1]. **Đinh Đoàn Long (chủ biên), Đỗ Lê Thăng (2013).** *Cơ sở di truyền học phân tử và tế bào (Tái bản lần thứ nhất)*. Nhà xuất bản Đại học Quốc gia Hà Nội.
- [2]. **Harvey Lodish, Rnaold Berk, Chris A. Kaiser và cs (2014).** *Sinh học phân tử của tế bào – Tập 2: Di truyền học và sinh học phân tử (Nguyễn Xuân Hưng và cs dịch theo bản tái bản lần thứ 7)*. Nhà xuất bản Trẻ.
- [3]. **Harvey Lodish, Rnaold Berk, Chris A. Kaiser và cs (2015).** *Sinh học phân tử của tế bào – Tập 3: Cấu trúc và chức năng của tế bào 1 (Nguyễn Xuân Hưng và cs dịch theo bản tái bản lần thứ 7)*. Nhà xuất bản Trẻ.
- [4]. **Nguyễn Tấn Thành, Nguyễn Thành Phi Long, Phùng Thuý Linh, Đỗ Lê Phúc Thịnh (2022).** *Di truyền học (Bài tập chuyên đề)*. Nhà xuất bản Đại học Quốc gia Hà Nội.
- [5]. **Nguyễn Thị Hồng Vân, Bùi Thị Việt Hà (2011).** *Giáo trình di truyền học sinh vật nhân sơ*. Nhà xuất bản Giáo dục Việt Nam.
- [6]. **Bộ Giáo dục và Đào tạo.** *Đề thi chọn học sinh giỏi Quốc gia THPT và Đề thi chọn đội tuyển Olympic*.

TÀI LIỆU TIẾNG ANH

- [1]. **Robert J. Brooker (2018).** *Genetics: Analysis & Principles (6e)*. New York: McGraw-Hill Education.
- [2]. **John L. Bowman, Mark L. Sanders (2019).** *Genetic Analysis: An Integrated Approach (3e)*. New York: Pearson Education, Inc.
- [3]. **David P. Clark, Nanette J. Pazdernik, Michelle R. McGehee (2019).** *Molecular Biology (3e)*. London, United Kingdom: Academic Press, an imprint of Elsevier.
- [4]. **Michael M. Cox, Michael O'Donnell, Jennifer A. Doudna (2012).** *Molecular Biology: Principles and Practice*. New York: W. H. Freeman and Company.
- [5]. **William S. Klug, Michael R. Cummings, Charlotte A. Spencer, Michael A. Palladino, Darrell J. Killian (2019).** *Concepts of genetics (12e)*. Hoboken, New Jersey: Pearson Education, Inc.
- [6]. **Jocelyn E. Krebs, Elliott S. Goldstein, Stephen T. Kilpatrick (2014).** *Lewin's genes XI (11e)*. United States of America: Jones & Bartlett Learning
- [7]. **Harvey Lodish, Rnaold Berk, Chris A. Kaiser et al. (2016).** *Molecular Cell Biology (8e)*. New York: W. H. Freeman and Company.
- [8]. **Laurence A. Moran, H. Robert Horton, K. Gray Scrimgeour, Marc D. Perry (2012).** *Principles of biochemistry (5e)*. United States of America: Pearson Education, Inc.
- [9]. **Peter J. Russell (2010).** *iGenetics: a molecular approach (3e)*. San Francisco: Pearson Education, Inc.

- [10]. **Lisa Urry, Michael Cain, Steven Wasserman, Peter Minorsky, Jane Reece (2017).** *Campbell biology (11e)*. New York: Pearson Education, Inc.
- [11]. **Donald Voet, Judith G. Voet, Charlotte W. Pratt (2013).** *Fundamentals of Biochemistry: Life at the Molecular Level (4e)*. United States of America: John Wiley & Sons, Inc.
- [12]. **John Wilson, Tim Hunt (2013).** *Molecular Biology of the Cell: The Problems Book (6e)*. United States of America: Garland Science, Taylor & Francis Group, LLC.