

LÊ TRẦN BÌNH (Chủ biên)
QUYỀN ĐÌNH THI

Cơ sở công nghệ SINH HỌC

TẬP MỘT - CÔNG NGHỆ GEN



NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM

LÊ TRẦN BÌNH (Chủ biên) – QUYỀN ĐỊNH THI

CƠ SỞ CÔNG NGHỆ

SINH HỌC

TẬP MỘT – CÔNG NGHỆ GEN

NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM

**Công ty cổ phần Dịch vụ xuất bản Giáo dục Hà Nội -
Nhà xuất bản Giáo dục Việt Nam giữ quyền công bố tác phẩm**

1

VẬT CHẤT DI TRUYỀN

Trong tế bào tồn tại nhiều loại phân tử có kích thước và khối lượng lớn từ hàng chục đến hàng trăm kDa, thường do nhiều phân tử tiểu phân và phân tử cơ bản hợp thành, ví dụ như tinh bột và cellulose do các phân tử đường glucose, mỡ do các acid béo và glycerin, protein do các amino acid và nucleic acid do các nucleotide liên kết với nhau mà thành. Các loại đại phân tử trong thế giới vật chất sống có cấu trúc và chức năng khác nhau, trong đó nucleic acid, chủ yếu là deoxyribonucleic acid (DNA) có vai trò quan trọng nhất là vật chất di truyền.

1.1. CẤU TRÚC VÀ CHỨC NĂNG NUCLEIC ACID

Nucleic acid là nhóm đại phân tử quan trọng nhất, có cấu trúc cơ bản thống nhất và có chức năng vật chất di truyền trong mọi sinh vật, bao gồm khả năng tự nhân đôi trong chu trình phân chia tế bào và khả năng di truyền các tính trạng từ thế hệ này qua thế hệ khác trong quá trình sinh sản.

Thông tin được chuyển từ thế hệ này sang thế hệ tiếp theo dưới hai hình thức : (1) Hình thức cấu trúc, cụ thể là trong sinh sản hữu tính thì tế bào trung thụ tinh hay trong sinh sản vô tính thì tế bào con đều nhận được một bộ những cấu trúc có tổ chức chặt chẽ tồn tại trước đó phản ánh thực trạng và tính chất cấu trúc tế bào đang tồn tại. (2) Hình thức gene, cụ thể là một genome cần thiết để tiếp tục hình thành các cấu trúc tiếp theo trong quá trình phát triển của cơ thể.

Thông tin di truyền hoạt động thông qua khả năng tạo ra rất nhiều loại phân tử protein khác nhau. Mỗi một gene có chức năng biểu hiện thành một chuỗi polypeptide. Nhận thức rằng mỗi protein được tạo bởi một chuỗi nhất định các amino acid và phát minh ra bản chất gene có bản chất là DNA đã đặt ra một câu hỏi quan trọng là làm sao một trình tự nucleotide trong DNA lại có thể đại diện cho một trình tự amino acid trong protein.

1.1.1. Thành phần nucleic acid

Mỗi phân tử DNA có một tính chất rất quan trọng là không phụ thuộc vào thành phần và trật tự sắp xếp của các nucleotide. Nói theo một cách khác là trình tự nucleotide trong DNA có vai trò quan trọng thực chất không phải vì cấu trúc của phân tử DNA, mà vì mật mã di truyền cho trình tự các amino acid tạo nên chuỗi polypeptide tương ứng. Mỗi quan hệ giữa trình tự các nucleotide trong phân tử DNA và trình tự amino acid trong phân tử protein được gọi là mật mã di truyền.

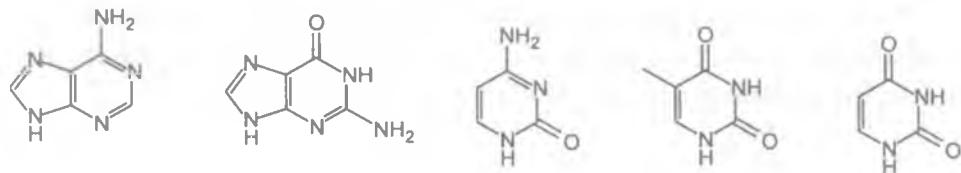
Cấu trúc hoặc hoạt tính xúc tác của mỗi loại protein phụ thuộc vào trình tự sơ cấp của các amino acid được sắp xếp trong phân tử protein đó. Toàn bộ thông tin cần thiết mã hoá trong trình tự nucleotide gene quyết định trình tự amino acid trong mỗi phân tử protein. Bằng cách thức như vậy, chỉ cần một kiểu cấu trúc cơ bản, gene có khả năng đại diện cho vô số kiểu protein khác nhau. Chức năng xúc tác và chức năng cấu trúc của các sản phẩm protein khác nhau trong một tế bào là hai chức năng quyết định sự hình thành kiểu hình của từng loại tế bào và của từng cá thể sinh vật. Đương nhiên, bên cạnh hai chức năng chính vừa nêu, một số gene còn có chức năng được nhận biết bởi các phân tử điều khiển chủ yếu là protein. Trong trường hợp này, chức năng của DNA được quyết định trực tiếp bởi chính trình tự, không thông qua mật mã trung gian nào. Hai loại vùng DNA với chức năng mã hoá protein và chức năng điều khiển tạo nên toàn bộ hệ thống thông tin di truyền.

Dột biến là những biến đổi di truyền được của thông tin di truyền. Vì DNA là vật chất di truyền nên dột biến có bản chất tất yếu là những thay đổi trong trình tự nucleotide. Thông qua mật mã di truyền, mỗi thay đổi trong trình tự nucleotide sẽ dẫn đến những thay đổi trình tự amino acid trong protein. Thông qua dột biến, chúng ta có thể so sánh gene lành với gene dột biến, xác định sự thay đổi protein do gene dột biến gây ra, xác định tính trạng, kiểu hình thay đổi do protein mã hoá trong gene dột biến để từ đó xác định được chức năng của gene.

Thiết kế cấu trúc phân tử DNA là một thiết kế đặc biệt tạo ra khả năng thực hiện được việc bảo tồn và duy trì mãi mãi trình tự của chúng. Được tạo bởi hai sợi tương đồng nhau theo một nguyên tắc biệt trước, mỗi một phân tử DNA trong thực tế mang tới hai lần lượng thông tin cần thiết. Nhưng chính bằng cách thức đó, mỗi một sợi được kiểm soát bởi sợi kia và mỗi thay đổi sẽ được sửa chữa để bảo đảm tính ổn định. Vì hai sợi liên

kết với nhau không qua liên kết hoá trị nên không cần tốn kém nhiều năng lượng, chúng có thể tách rời nhau trong quá trình nhân đôi ở những điều kiện sinh lý thường.

Nucleic acid là một chuỗi các tiểu đơn vị liên kết hoá học với nhau. Một tiểu đơn vị bao gồm một base hữu cơ (cấu trúc mạch vòng hõn (dị) nguyên gồm nguyên tố carbon và nitơ), một phân tử đường pentose (đường 5 carbon mạch vòng) và một gốc phosphate.

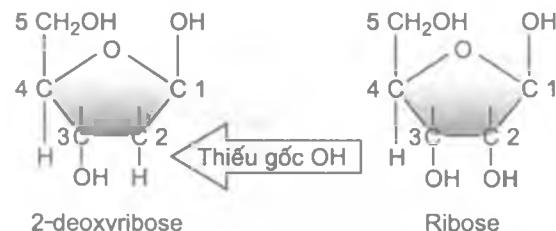


Hình 1.1. Cấu trúc cơ bản của các base purine (A, G) và pyrimidine (C, T, U).

Base chứa nitơ được chia thành hai nhóm là *pyrimidine* và *purine* (hình 1.1), trong đó pyrimidine chỉ có một mạch vòng 6, còn purine có một vòng kép gồm một vòng 6 và một vòng 5 (hình 1.1).

Mỗi phân tử nucleic acid đều chứa 4 loại base, trong đó 2 base purine là adenine và guanine luôn cùng có mặt ở DNA và RNA, trong khi đó 2 base pyrimidine trong DNA là cytosine và thymine, nhưng trong RNA lại thấy uracil thay vì thymine. Sự sai khác duy nhất giữa thymine và uracil là thymine có hơn một gốc methyl ở vị trí C₃. Thông thường, các base được ký hiệu theo chữ cái đầu tiên, như vậy : DNA chứa A, G, C, T ; còn RNA chứa A, G, C, U.

Hai loại đường pentose có trong 2 loại nucleic acid khác biệt nhau ở chỗ : Trong DNA là đường 2-deoxyribose, trong khi đó ở RNA là đường ribose. Từ đó hai loại nucleic acid được gọi theo tên của đường : deoxyribonucleic acid (DNA) và ribonucleic acid (RNA). Hai loại đường này chỉ khác nhau một

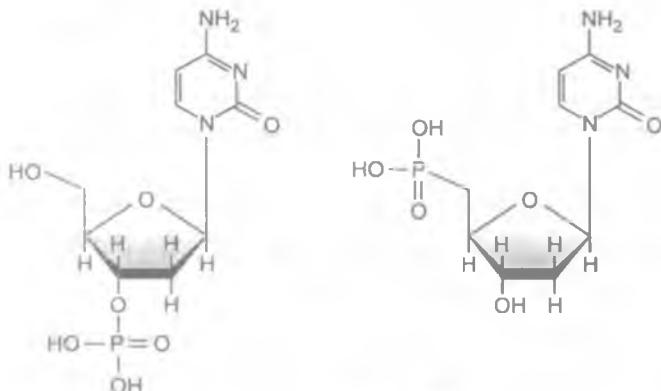


Hình 1.2. Hai loại đường pentose có trong nucleic acid : 2-deoxyribose trong DNA và ribose trong RNA. Các số đánh dấu vị trí nguyên tử trong mạch vòng của phân tử đường. Base nitơ gắn vào vị trí C₁.

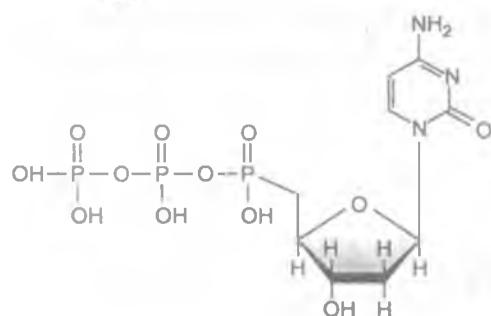
điểm là ở 2-deoxyribose, nhóm OH không tồn tại ở vị trí C₂ của vòng pentose (hình 1.2).

Khung phân tử của base liên kết với vòng phân tử đường pentose ở vị trí C₁ thông qua liên kết glucoside với nguyên tử N₁ của vòng purine hay N₃ của vòng pyrimidine. Để tránh nhầm lẫn với việc đánh dấu vị trí các nguyên tử trong khung phân tử base, vị trí của các nguyên tử trong khung pentose được thêm một dấu phẩy mũ (').

Phân tử liên kết giữa một base với một phân tử đường pentose được gọi là nucleoside (hình 1.3). Khi liên kết thêm với một gốc phosphate thì phân tử base-đường-phosphate được gọi là nucleotide (hình 1.4).



Hình 1.3. Nucleoside có thể liên kết với gốc phosphate tại vị trí 5' hoặc 3'.



Hình 1.4. Một phân tử nucleoside gắn ở vị trí 5' với 3 gốc phosphate giàu năng lượng.

Danh pháp hoá học của các tiểu phần này được nêu trong bảng 1.1.

Bảng 1.1. Danh pháp hoá học và cách viết tắt của các tiểu phần tạo nên nucleic acid.

Base	Nucleoside	Nucleotide	Viết tắt	
			RNA	DNA
Adenine	Adenosine	Acid adenylic	AMP	dAMP
Guanine	Guanosine	Acid guanylic	GMP	dGMP
Cytosine	Cytosine	Acid cytidylic	CMP	dCMP
Thymine	Thymidine	Acid thymidylic		dTMP
Uracil	Uridine	Acid uridylic	UMP	

Nucleotide là vật liệu xây dựng để làm nên nucleic acid. Các nucleotide liên kết với nhau thành một chuỗi polynucleotide có khung xương gồm một phân tử đường và một gốc phosphate thay nhau liên kết thành mạch dài. Trong mạch đó, vị trí 5' của một phân tử đường pentose này nối với vị trí 3' của phân tử đường pentose tiếp theo thông qua một gốc phosphate (hình 1.5).

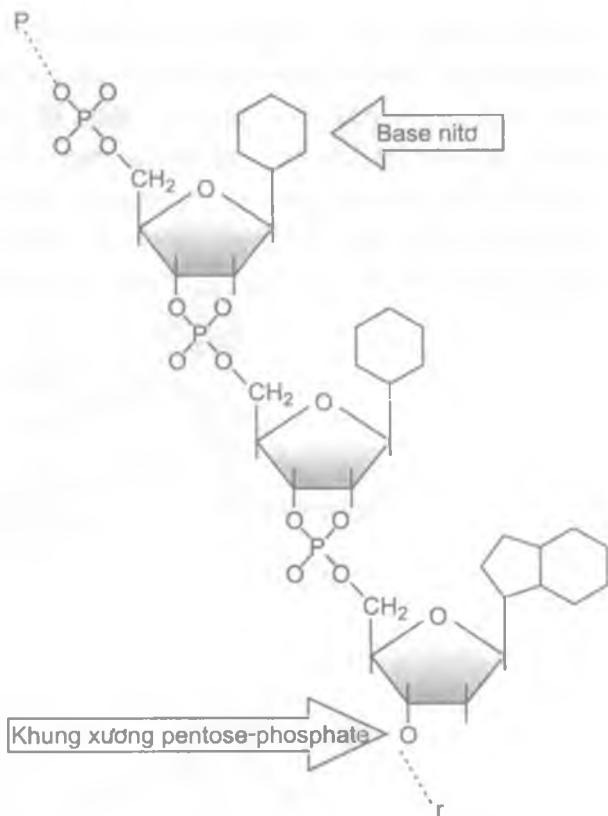
1.1.2. Cấu trúc deoxyribonucleic acid

Kết quả phân tích thành phần các base có trong DNA của các loài khác nhau có sự khác nhau rất đặc trưng cho phép rút ra kết luận rằng *trình tự các base trong DNA là hình thức mà thông tin di truyền được lưu giữ.*

Mô hình cấu trúc phân tử DNA dạng xoắn kép được Watson và Crick xây dựng năm 1953 là dựa trên những kết quả nghiên cứu sau đây :

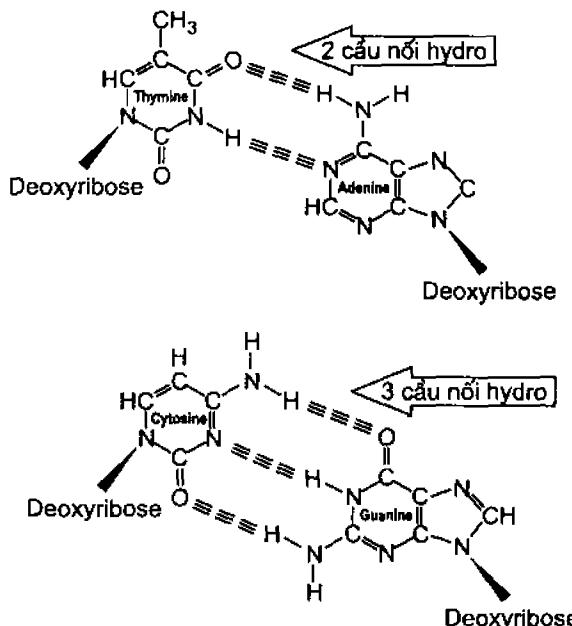
Ánh chụp X-quang phân tử DNA cho thấy DNA có hình dạng xoắn đều đặn, mỗi vòng xoắn dài 34 Å (3,4 nm) và đường kính là 20 Å (2 nm). Vì khoảng cách giữa 2 nucleotide là 3,4 Å, cho nên mỗi vòng xoắn phải có 10 nucleotide.

Mật độ của DNA cho thấy vòng xoắn của DNA gồm hai chuỗi polynucleotide. Sợi xoắn DNA có đường kính khá đều, điều này chỉ có



Hình 1.5. Một chuỗi polypeptide hình thành nhờ liên kết 5'-3' giữa các phân tử đường pentose và gốc phosphate tạo nên một mạch xương sống cho các base nitơ treo nhô ra từ các vị trí C₁ của gốc đường pentose.

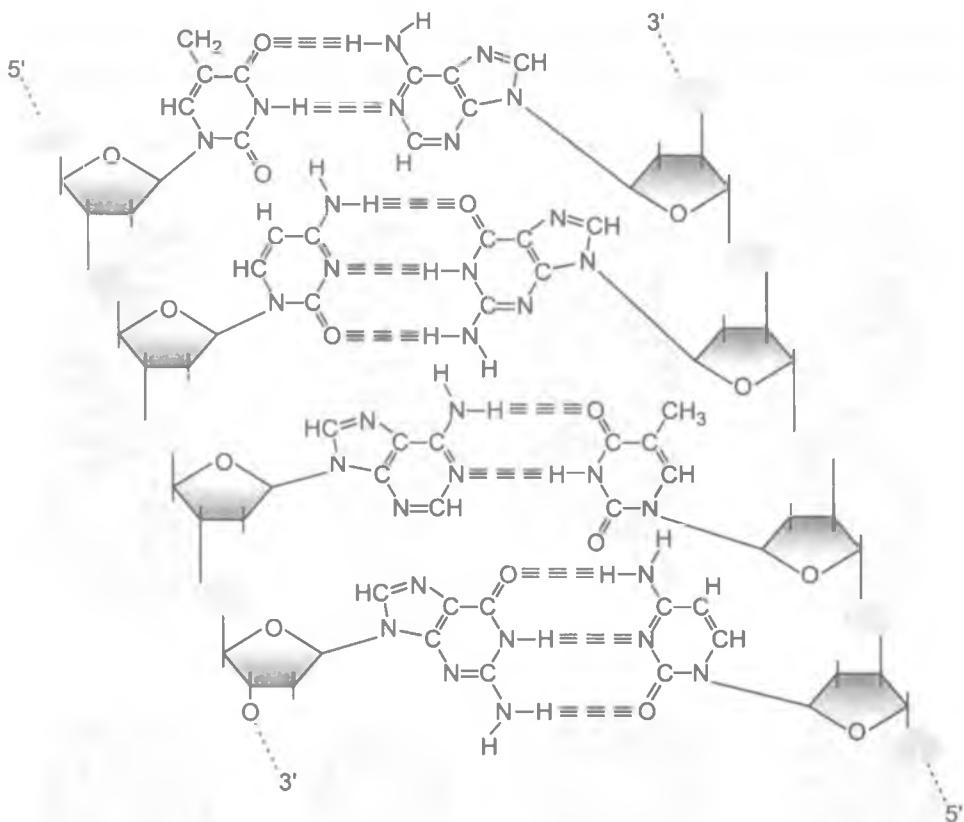
thể giải thích được là do các base nitơ kết thành cặp với nhau về phía trong đường xoắn và mỗi cặp chỉ có thể bao gồm một phân tử purine và một phân tử pyrimidine. Nếu hai phân tử purine tạo một cặp thì kích thước cặp này sẽ lớn hơn rất nhiều cặp của hai phân tử pyrimidine và đường kính của sợi xoắn sẽ có đoạn to, đoạn nhỏ khác nhau. So sánh thành phần các base có trong DNA thì thấy tỷ lệ của G và của C luôn bằng nhau và tỷ lệ của A và của T cũng luôn bằng nhau.



Hình 1.6. Cặp base bổ sung giữa A và T thông qua hai cầu nối hydro và giữa G và C thông qua ba cầu nối hydro tồn tại trong phân tử DNA xoắn kép.

Watson và Crick đã nêu ra kết luận rằng hai chuỗi polynucleotide trong sợi DNA xoắn kép liên kết với nhau thông qua các cầu nối hydro giữa các base nitơ. Cụ thể là : G chỉ có thể liên kết với C qua ba cầu nối hydro và A chỉ có thể liên kết với T thông qua hai cầu nối hydro (hình 1.6). Hiện tượng các base liên kết với nhau thông qua cầu nối hydro được gọi là hiện tượng cặp đôi base và hai base trong một cặp được gọi là các **base bổ sung (complementary base)**.

Mô hình sợi DNA xoắn kép đòi hỏi hai chuỗi polynucleotide chạy theo chiều ngược nhau. Quan sát dọc theo phân tử thì có một sợi hướng theo chiều $5' \rightarrow 3'$ và sợi kia theo chiều $3' \rightarrow 5'$ (hình 1.7).



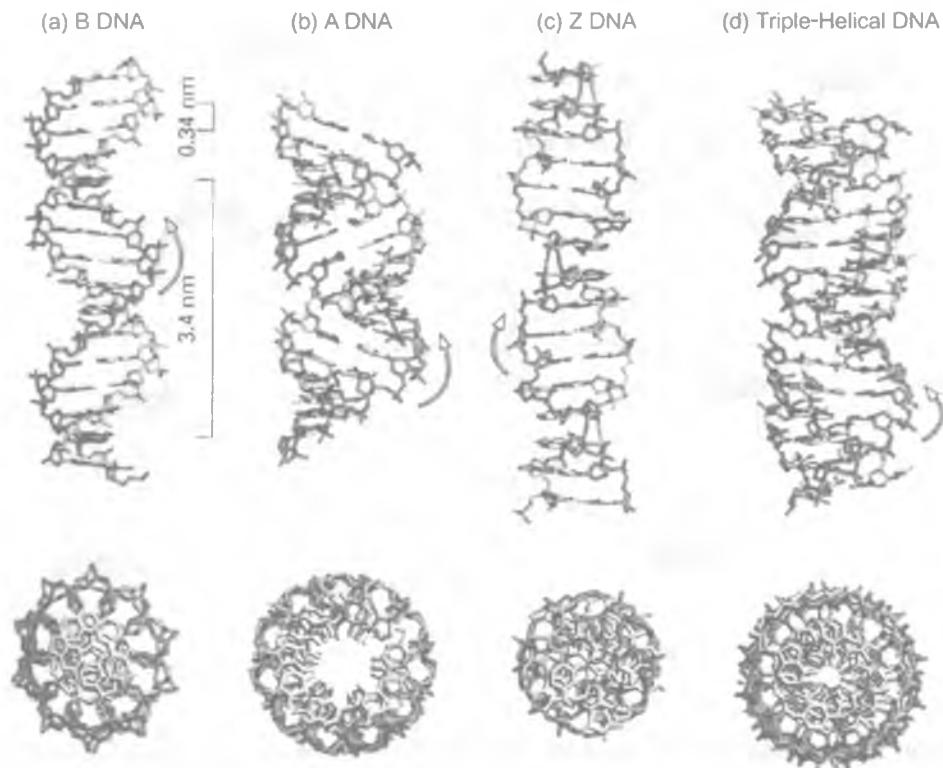
Hình 1.7. Sợi xoắn kép DNA giữ được kích thước ổn định nhờ các cặp base bổ sung luôn được hình thành giữa một purine và một pyrimidine. Sự khác nhau chính là giữa C và G có ba cầu nối hydro và giữa A và T có hai cầu nối hydro.

Bên ngoài là khung đường-phosphate mang điện tích âm do gốc phosphate. Khi hòa tan DNA trong dung dịch, điện tích của chúng trở nên trung hoà do liên kết với các ion Na^+ có trong dung dịch. Trong điều kiện *in vivo*, các loại protein tích điện dương thường liên kết với DNA tạo điều kiện để sắp xếp DNA trong tế bào.

Mỗi cặp base vặn quanh trục của mạch xoắn kép một góc 36° , như vậy cứ 10 cặp base sẽ hoàn thành một vòng xoắn. Hình xoắn có kích thước phần gờ rộng là 22 \AA và phần rãnh hẹp là 12 \AA (hình 1.8). Sợi xoắn quay theo chiều kim đồng hồ quanh trục.

Hình 1.8 minh họa các mô hình cấu trúc DNA khác nhau đã được tìm thấy. Hai mạch polynucleotide tạo thành một chuỗi xoắn kép với những đặc điểm chung và riêng như sau : Xương sống của chuỗi xoắn kép là

mạch đường phosphate nằm bên ngoài, các base nitơ hướng vào bên trong. Trên hình 1.8 mô tả hình nhìn từ bên và hình nhìn thẳng theo trục đứng.



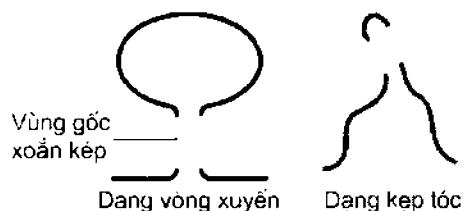
Hình 1.8. Mô hình các cấu trúc DNA khác nhau đã được tìm thấy.

(a) Cấu trúc B là cấu trúc chặt hơn của phân tử DNA có vòng xoắn dài 3,4 nm và 10 cặp đôi nucleotide, khoảng cách giữa các cặp nucleotide là 0,34 nm. Rãnh chính và rãnh phụ trên cấu trúc khá rõ. (b) Cấu trúc A có 11 cặp nucleotide trong mỗi vòng xoắn và các cặp base thường có độ nghiêng. Ngoài ra, trong lòng phân tử DNA cấu trúc A có lỗ hổng. Dạng cấu trúc xoắn này cũng thường tìm thấy trong cấu trúc xoắn DNA-RNA hoặc RNA-RNA. (c) Cấu trúc Z xoắn theo chiều phải (cùng chiều kim đồng hồ) và có hình dạng zig-zag. Mỗi vòng xoắn có tới 12 cặp nucleotide và được tạo từ cấu trúc B dưới tác động của tác nhân Za. (d) Cấu trúc xoắn sợi ba (triple-helical) xảy ra khi toàn bộ các purine (A, G) của một sợi phù hợp với toàn bộ các pyrimidine (T, C) của sợi kia, tạo điều kiện cho một sợi thứ ba tham gia vào cấu trúc xoắn.

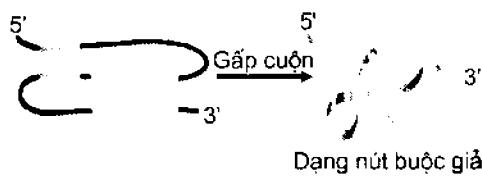
1.1.3. Cấu trúc ribonucleic acid

Về nguyên tắc, cấu trúc hoá học của RNA tương tự như DNA, chỉ có hai điểm khác biệt là : các phân tử đường ribose trong RNA mang nguyên gốc OH tại vị trí 2' (hình 1.2) và thymine của DNA được thay bằng uracil trong RNA (hình 1.1). Nhóm OH tại vị trí C₂ của phân tử đường ribose làm cho phân tử RNA kém bền vững về mặt hoá học so với DNA, vì thế, RNA thường bị phân giải thành các nucleotide trong dung dịch kiềm, trong khi đó thì DNA không bị như vậy. Cũng như DNA, các phân tử RNA là sợi polynucleotide dài có khả năng tạo thành sợi đôi, dạng thẳng hay dạng vòng. RNA cũng có thể tham gia tạo sợi kép cùng một sợi DNA. Vì thế, phân tử lai này có cấu trúc phân tử tương tự như dạng cấu trúc B của DNA.

(a) Cấu trúc bậc 2



(b) Cấu trúc bậc 3



Hình 1.9. Cấu trúc bậc hai và bậc ba của RNA dạng vòng xuyên, dạng kẹp tóc hay dạng nút buộc giả.

cách xa nhau khoảng 50 đến một vài trăm nucleotide. Các cấu trúc này góp phần tạo nên những dạng cấu trúc bậc ba phức tạp hơn dạng nút buộc giả (pseudoknot) của phân tử RNA (hình 1.9).

Riêng các phân tử tRNA có những hình dáng thích hợp cho chức năng vận chuyển amino acid trong quá trình sinh tổng hợp protein. Các phân tử rRNA có kích thước to lớn hơn và thường có cấu trúc không gian ba chiều khá ổn định ở vùng chúng tồn tại trong tế bào. Cấu trúc bậc hai

Không giống như DNA có cấu trúc không gian ba chiều thống nhất là sợi xoắn kép, các loại phân tử RNA có các cấu trúc phân tử khác nhau. Sự khác nhau về kích thước và cấu trúc phân tử cho phép các phân tử RNA khác nhau thực hiện các chức năng đặc thù khác nhau trong tế bào. Cấu trúc bậc hai đơn giản nhất của phân tử RNA sợi đơn là hiện tượng tạo các cấu trúc dạng kẹp tóc (hairpin) do 5-10 nucleotide cặp đôi hoặc cấu trúc dạng vòng xuyên (stem-loop) cũng do các nucleotide cặp đôi, nhưng chúng

và bậc ba cũng thấy xuất hiện ở những phân tử mRNA. Rõ ràng là các phân tử RNA cũng có những đoạn phân tử có cấu trúc ổn định liên kết với những đoạn năng động và ít bị cấu trúc hoá.

Đoạn cuối gấp của phân tử RNA không những tương đương về cấu trúc với các đoạn xoắn α và các chuỗi β của phân tử protein mà còn có cả chức năng xúc tác. Các RNA xúc tác được gọi là các ribozyme có khả năng cắt chuỗi RNA. Có những chuỗi RNA có khả năng cắt bỏ đoạn RNA hình thành từ intron và nối hai đoạn RNA hình thành từ exon với nhau. Đó là quá trình *cắt gọt RNA* (RNA splicing). Đây là hiện tượng xảy ra trong quá trình thành thực hóa các phân tử mRNA trong tế bào nhân chuẩn (hoàn chỉnh, eucaryote) và cả tế bào vi khuẩn với nhân sơ (chưa hoàn chỉnh, procaryote). Đáng chú ý hơn nữa là có những phân tử RNA được cắt gọt bằng ngay chính đoạn RNA intron của nó trong quá trình tự thành thực hóa.

1.1.4. Chức năng của deoxyribonucleic acid

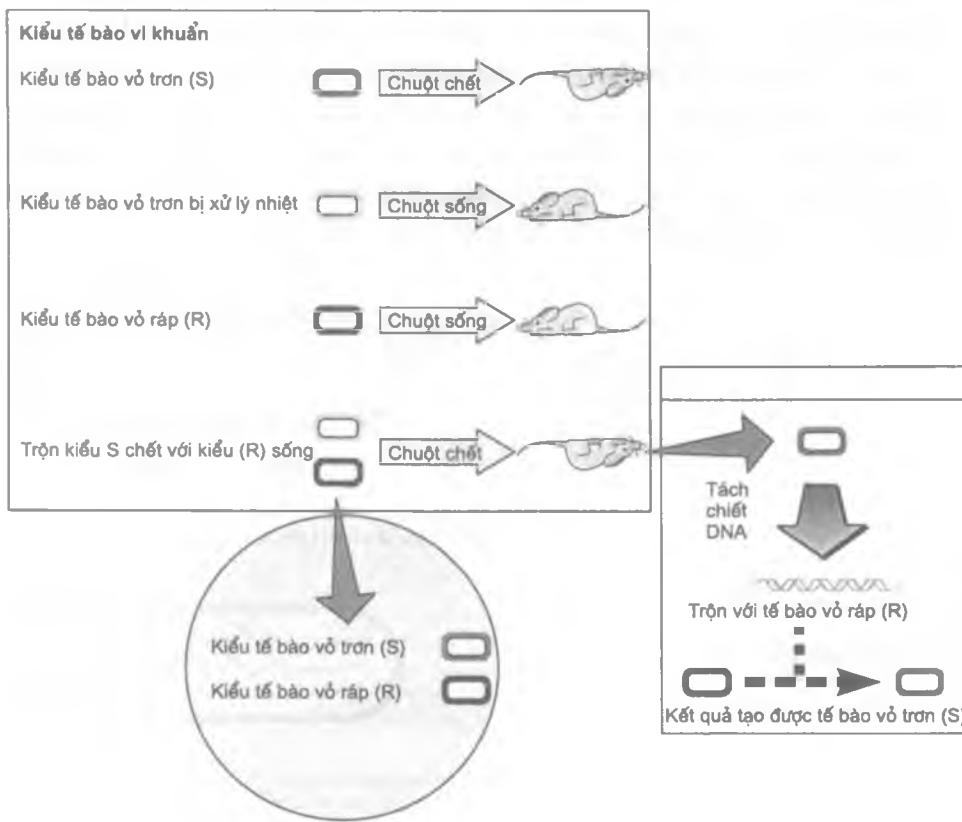
ĐNA có hai chức năng chính là lưu giữ thông tin di truyền của mọi sinh vật và chuyển tải thông tin di truyền từ thế hệ này sang thế hệ khác.

a) Chức năng lưu giữ thông tin

DNA là vật liệu di truyền ở vi khuẩn : Ý tưởng cho rằng vật liệu di truyền có thể là nucleic acid này sinh từ thí nghiệm biến nạp của Griffith năm 1928. Vi khuẩn *Pneumococcus* gây bệnh pneumonia làm cho chuột bị chết. Yếu tố gây độc của vi khuẩn được xác định là thành phần polysaccharide của thành tế bào bảo đảm cho vi khuẩn không bị vật chủ tiêu huỷ. Thông thường tế bào của *Pneumococcus* tạo vỏ polysaccharide làm cho bể ngoài tế bào có dạng tròn (S). Bên cạnh kiểu tế bào tròn còn có những tế bào biến đổi, mất khả năng tạo vỏ polysaccharide, làm cho bể mặt tế bào hết tròn mà là ráp (R) và kiểu tế bào này mất luôn khả năng gây bệnh.

Khi làm chết bằng xử lý nhiệt, tế bào tròn (S) không thể gây bệnh, nhưng khi trộn chúng với tế bào ráp (R) thì chuột thí nghiệm lại bị chết vì mắc bệnh pneumonia và kiểu tế bào gây bệnh (S) của *Pneumococcus* tái xuất hiện trong chuột chết. Phân tích kỹ chủng vi khuẩn gây bệnh mới xuất hiện cho thấy tính chất của vỏ polysaccharide của chúng giống như của kiểu tế bào (S) đã bị xử lý nhiệt. Như vậy, tế bào chết (S) đã biến nạp tính chất vỏ tròn sang tế bào sống kiểu ráp (R). Năm 1944, Avery đã tinh chế

thành phần hoá học của kiểu tế bào trơn (S) cần thiết để biến nạp thành công trong điều kiện ống nghiệm. Đó là deoxyribonucleic acid (DNA) (hình 1.10). Điều đáng ngạc nhiên là trước đó DNA đã được biết đến như một thành phần hoá học quan trọng nhất của nhiễm sắc thể trong các tế bào nhân chuẩn, nhưng lại không biết trong thành phần tế bào *Pneumococcus* cũng có DNA. Như vậy, vật chất di truyền của thế giới sinh vật tế bào nhân sơ cũng là DNA. Kết quả này đã nhất thể hoá quan niệm về tính thống nhất của nền tảng di truyền ở vi khuẩn và sinh vật bậc cao.



Hình 1.10. Yếu tố biến nạp là DNA.

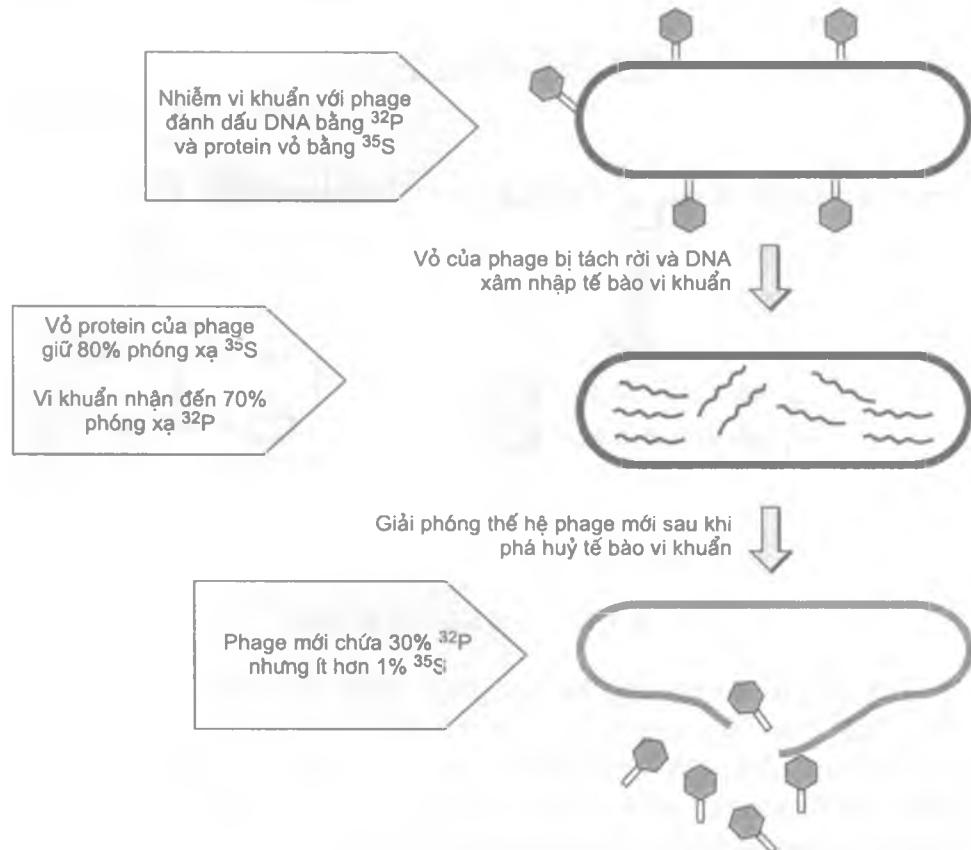
Tuy nhiên, trong công bố gốc phát minh này được thể hiện như sau : Chất cảm ứng, căn cứ vào tính chất hoá học và lý học của nó có vẻ như là một loại DNA cao phân tử và có độ nhớt cao. Mặt khác, sản phẩm polysaccharide ở vỏ tế bào được tổng hợp có sự tham gia của tác nhân biến nạp lại là một polysaccharide không chứa nitơ. Như vậy, rõ ràng là chất gây nên cảm ứng tổng hợp và sản phẩm được tổng hợp

khác nhau hoàn toàn về bản chất hoá học và chức năng sinh học và cả hai đều được coi là thành phần quyết định đặc điểm của tế bào mà chúng tạo nên.

Qua đó, chúng ta thấy rõ sự khác biệt giữa vật chất di truyền và sản phẩm biểu hiện gene, một nguyên lý cơ bản có tính tuyệt đối thống nhất trong mọi sinh vật.

b) Chức năng chuyển tải thông tin – DNA vật chất di truyền phổ biến

DNA là vật liệu di truyền ở virus : Sau khi chứng minh chức năng di truyền của DNA, chúng ta thảo luận đến tính phổ biến của DNA là vật chất di truyền của các đối tượng hệ thống sinh vật khác nhau, ví dụ như thực khuẩn T2. Đây là một loại virus có khả năng nhiễm vào tế bào vi khuẩn *E. coli*. Một hạt T2 khi gặp tế bào vi khuẩn, nó sẽ bám vào thành tế bào, một ít vật chất được truyền từ virus vào trong tế bào, sau 20 phút, tế bào chủ sẽ bị phân huỷ và hàng nghìn thế virus mới được giải phóng (hình 1.11).



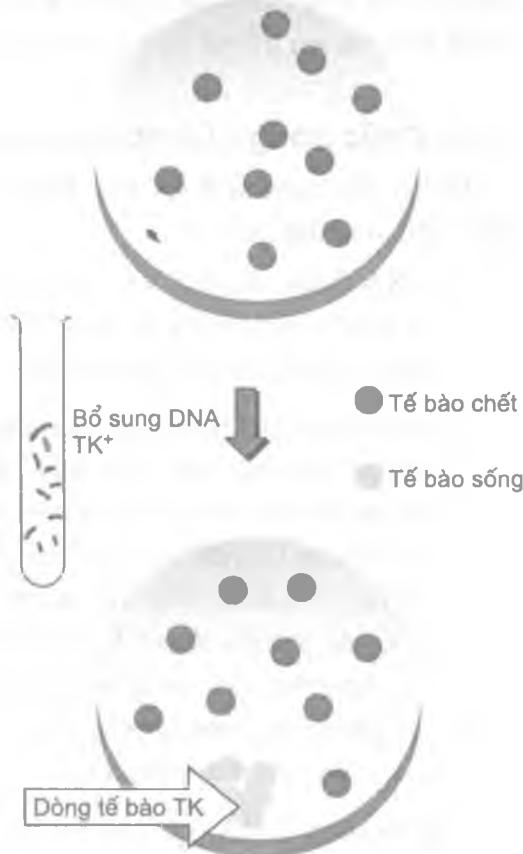
Hình 1.11. Vật chất di truyền của thực khuẩn thể T2 là DNA.

Thí nghiệm đánh dấu vỏ protein của T2 bằng đồng vị phóng xạ ^{35}S và DNA bằng ^{32}P thì nhận thấy chỉ có DNA của virus T2 xâm nhập vào tế bào vi khuẩn chủ, protein vỏ ở ngoài hoàn toàn. Những hạt T2 mới hình thành sau khi phân huỷ tế bào mang đến 30% lượng phóng xạ ^{32}P ban đầu, trong khi đó lượng ^{35}S của chúng chưa đến 1%.

Như thế có thể khẳng định DNA của T2 xâm nhập vào tế bào vi khuẩn, sử dụng toàn bộ bộ máy di truyền của tế bào chủ để tái bản DNA của virus T2 và tạo ra protein vỏ của chúng để đóng gói thành thê virus hoàn chỉnh trước khi thoát ra ngoài thông qua quá trình tiêu huỷ toàn bộ tế bào chủ mà chúng ta gọi là hiện tượng thực bào.

DNA là vật liệu di truyền ở sinh vật nhân hoàn chỉnh: Khi bổ sung DNA vào một quần thể tế bào động vật nuôi cấy *in vitro*, phân tử DNA xâm nhập vào tế bào và kết quả thu được là quần thể tế bào này có thể tổng hợp được protein mới. Cụ thể là dòng tế bào mất khả năng tạo ra thymidine kinase (TK) không thể sống trên môi trường thiếu thymidine nhưng khi nhận được DNA có đoạn gene TK thì một số tế bào có khả năng

Tế bào thiếu gene Thymidine kinase (TK)
chết trên môi trường thiếu thymidine



Một vài tế bào đã tiếp nhận gene TK sống sót và phát triển thành dòng tế bào mới

Hình 1.12. Tế bào nhân hoàn chỉnh có khả năng tiếp nhận DNA ngoại lai để có những tính trạng mới.

tạo ra TK và có thể sinh trưởng và phát triển thành một dòng tế bào biến nạp (hình 1.12).

Qua đó cho thấy DNA có thể được chuyển giữa các tế bào khác nhau mà vẫn duy trì được chức năng của nó.

Có thể khái quát vật liệu di truyền của hầu hết các sinh vật đã biết và nhiều loại virus đều là DNA. Đương nhiên, cũng có một số loại virus đã dùng ribonucleic acid (RNA) làm vật liệu di truyền. Mặc dù cấu trúc phân tử của RNA có một số điểm khác với DNA nhưng về nguyên tắc, chúng thực thi vai trò vật liệu di truyền đúng như DNA.

1.1.5. Chức năng của ribonucleic acid

Có ba loại phân tử RNA thực hiện các chức năng phối hợp trong quá trình sinh tổng hợp protein :

1. RNA thông tin (mRNA) mang thông tin di truyền được sao chép từ DNA dưới dạng các bộ mã di truyền gồm 3 nucleotide nối tiếp nhau, mỗi mã tương ứng với một amino acid nhất định.
2. RNA vận chuyển (tRNA) làm chức năng chìa khoá hoá giải mã di truyền trên mRNA. Mỗi loại amino acid có một số loại tRNA tương ứng làm nhiệm vụ kết gắn với amino acid đó và vận chuyển nó đến vị trí tiếp theo ở đầu cuối đang mọc dài ra của chuỗi polypeptide tương ứng với mã di truyền trên phân tử mRNA. Cách thức các phân tử tRNA kết gắn với một amino acid được chọn lọc phù hợp cho bước tổng hợp polypeptide tiếp theo là thông qua bộ mã gồm 3 nucleotide có trình tự bổ sung với 3 nucleotide của mã di truyền trên phân tử mRNA.
3. Ribosome RNA (rRNA) kết hợp với một nhóm các loại protein để tạo nên thể ribosome. Phức hệ cấu trúc ribosome này chuyển động dọc theo phân tử mRNA và xúc tác quá trình lắp ghép các amino acid thành chuỗi protein. Ribosome liên kết các tRNA và các phân tử phụ kiện khác cần thiết cho quá trình sinh tổng hợp protein. Mỗi thể ribosome được cấu thành bởi hai tiểu đơn vị – nửa lớn và nửa bé, mỗi tiểu đơn vị đều chứa riêng một hay nhiều phân tử rRNA.

a) RNA thông tin (mRNA) có chức năng mang thông tin từ DNA dưới dạng mã di truyền bộ ba nucleotide

RNA chứa các loại ribonucleotide là adenine, cytosine, guanine và uracil, trong khi đó DNA chứa các deoxyribonucleotide là adenine, cytosine, guanine và thymine. Nếu mỗi một nucleotide là mã di truyền cho một amino acid thì chỉ có 4 trong số 20 amino acid có mã di truyền trong trình tự xếp mạch thẳng của protein. Mã di truyền từ 4 loại nucleotide phải được cấu trúc sao cho ít nhất cả 20 amino acid đều có mã riêng rẽ.

Nếu 2 nucleotide tham gia cấu thành một mã cho một amino acid thì chỉ có $4^2 = 16$ mã từ được tạo thành. Số này không đủ cho 20 amino acid. Vì thế khi dùng 3 nucleotide cho một mã di truyền thì thu được $4^3 = 64$ mã từ được tạo ra. Toán học cho phép tính toán bộ mã từ với ít nhất từ 3 nucleotide trở lên mới đảm bảo cho tất cả 20 amino acid có được mã riêng. Trong thực tế, mã di truyền tồn tại dưới dạng bộ ba nucleotide. Một bộ ba nucleotide được gọi là một mã di truyền. Trong số 64 tổ hợp bộ ba được tạo nên bởi 4 nucleotide khác nhau có 61 bộ ba là mã di truyền của các amino acid khác nhau, còn 3 bộ là mã dừng (mã kết thúc). Bảng 1.2 cho thấy phần lớn các amino acid đều có nhiều hơn một mã di truyền. Riêng methionine và tryptophan là chỉ có một mã di truyền mà thôi, ngược lại leucine, serine và arginine, mỗi amino acid này có đến 6 mã di truyền khác nhau, chúng được gọi là mã di truyền "đồng nghĩa".

Sinh tổng hợp protein trong tế bào nhân sơ và nhân chuẩn thường bắt đầu bằng amino acid methionine. Trong hầu hết các phân tử mRNA, mã khởi đầu là AUG nhưng đôi khi, có thể là GUG ở một vài loại mRNA của vi khuẩn và CUG ở một vài loại mRNA của tế bào nhân chuẩn. Ba bộ mã UAA, UGA và UAG không mã hoá cho amino acid nào, mà là tín hiệu dừng, tạo nên điểm kết thúc ở đầu carboxyl của chuỗi protein.

Bảng 1.2. Mã di truyền từ RNA thành amino acid.

Vị trí thứ nhất (đầu 5')	Vị trí thứ hai				Vị trí thứ ba (đầu 3')
	U	C	A	G	
U	UUU : Phe	UCU : Ser	UAU : Tyr	UCU : Ser	U
	UUC : Phe	UCC : Ser	UAC : Tyr	UCC : Ser	C
	UUA : Leu	UCA : Ser	UAA : (kết thúc)	UCA : Ser	A
	UUG : Leu	UCG : Ser	UAG : (kết thúc)	UCG : Ser	G
C	CUU : Leu	CCU : Pro	CAU : His	CCU : Pro	U
	CUC : Leu	CCC : Pro	CAC : His	CCC : Pro	C
	CUA : Leu	CCA : Pro	CAAGln	CCA : Pro	A
	CUG : Leu (Met)	CCG : Pro	CAG : Gln	CCG : Pro	G
A	AUG : Ile	ACU : Thr	ACU : Thr	AGU : Ser	U
	AUC : Ile	ACC : Thr	ACC : Thr	AGC : Ser	C
	AUA : Ile	ACA : Thr	ACA : Thr	AGA : Arg	A
	AUG : Met	ACG : Thr	ACG : Thr	AGG : Arg	G
G	GUU : Val	GCU : Ala	GCU : Ala	CAU : His	U
	GUC : Val	GCC : Ala	GCC : Ala	CAC : His	C
	GUA : Val	GCA : Ala	GCA : Ala	CAAGln	A
	GUG : Val (Met)	GCG : Ala	GCG : Ala	CAG : Gln	G

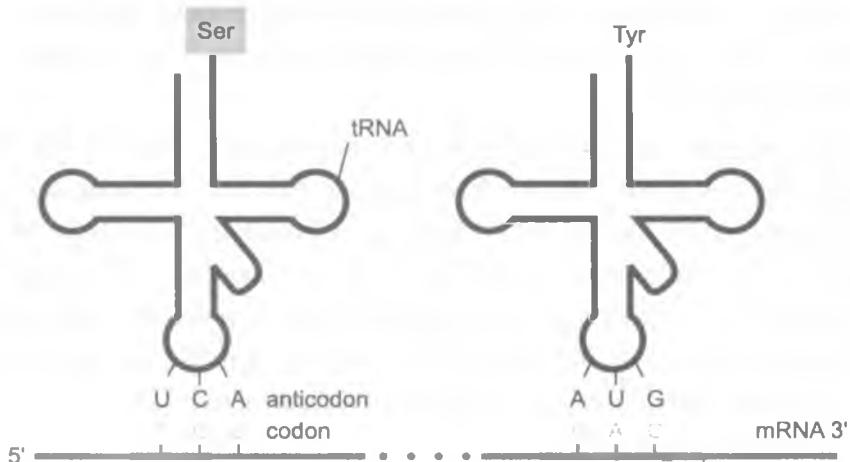
Trình tự nucleotide bắt đầu bằng mã khởi đầu chạy tới điểm có mã kết thúc được gọi là một khung đọc. Trình tự các bộ ba nucleotide trên phân tử mRNA quyết định trình tự amino acid trong protein, đồng thời trên phân tử mRNA đó cũng có tín hiệu bắt đầu và tín hiệu kết thúc quá trình tổng hợp chuỗi protein. Trong một khung đọc, trên phân tử mRNA, các mã di truyền được sắp xếp theo trình tự nối tiếp nhau. Cũng có những trường hợp trên cùng một phân tử mRNA có tới hai khung đọc đè chồng lên nhau, kết quả là có tới hai chuỗi polypeptide được tạo ra với trình tự hoàn toàn khác nhau.

Ngoài ra, giữa nhân và cơ quan tử như ty thể hay lục lạp của cùng một tế bào cũng có sự sai khác về mã di truyền của bộ máy sinh tổng hợp protein. Hiện tượng sai khác nhau chỉ là những hiện tượng riêng lẻ, còn lại nhìn chung trong thế giới sinh vật hầu hết mã di truyền có tính thống nhất.

b) RNA vận chuyển (tRNA) có cấu trúc gấp xoắn làm chức năng giải mã

Bước tiếp theo để hiểu dòng chảy thông tin từ DNA sang protein được quyết định như thế nào để trình tự nucleotide trên phân tử mRNA được chuyển đổi sang trình tự amino acid trên phân tử protein.

Quá trình giải mã cần phải có hai loại phân tử môi giới : đó là tRNA và enzyme với tên gọi là aminoacyl-tRNA synthetase.



		Base thứ hai trong codon					
		U	C	A	G		
Base thứ nhất trong codon	U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp		
	C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg		
	A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg		
	G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly		

Hình 1.13. Chức năng của phân tử tRNA trong quá trình giải mã DNA.

Mỗi phân tử tRNA có hai chức năng : (1) Liên kết với amino acid nhất định và (2) vận chuyển amino acid đó đến đúng chỗ thông qua "cặp đôi base bổ sung" vào đúng vị trí của mã di truyền tạo cho chuỗi protein được chấp thêm một amino acid vào đúng vị trí phía đầu phân tử đang dài dần ra. Mỗi phân tử tRNA chỉ được nhận biết bởi một trong số 20 loại aminoacyl-tRNA-synthetase mà thôi. Mỗi enzyme này cũng chỉ xúc tác việc kết gắn một trong 20 amino acid với phân tử tRNA nhất định mà thôi. Khi mối liên kết aminoacyl-tRNA hoàn tất, phức hợp này được chuyển đến đúng vị trí của mã di truyền tương thích trên phân tử mRNA tạo cho phân tử protein được nối dài thêm một amino acid nữa ở phía đầu carbocyl (hình 1.13).

Kết quả nghiên cứu về tRNA cho thấy trong tế bào vi khuẩn có khoang 30-40 loại phân tử tRNA và trong tế bào thực vật hoặc động vật có khoảng 50-100 loại tRNA khác nhau. Con số này cho thấy số loại phân tử tRNA lớn hơn con số 20 amino acid có trong protein và nhỏ hơn con số mã di truyền (64) có trên mRNA. Điều này phù hợp với thực tế mỗi amino acid có một hay nhiều mã di truyền và một phân tử tRNA có thể giải mã một hay nhiều mã di truyền trên phân tử mRNA.

Các phân tử tRNA có kích thước dài khoảng 70-80 nucleotide, ở đầu 3' luôn kết thúc bằng trình tự ba nucleotide CCA được gắn thêm vào khi quá trình tổng hợp phân tử tRNA đã kết thúc. Chức năng của tRNA hoàn toàn phụ thuộc vào cấu trúc không gian ba chiều của chúng. Trong dung dịch, tRNA thường có cấu trúc xoắn vành khuyên, nếu vẽ trên mặt phẳng thì có hình lá cỏ tam giác giống lá cây chua me đất của ta (hình 1.14).



Hình 1.14. Cấu trúc không gian 3D và phẳng của phân tử tRNA.

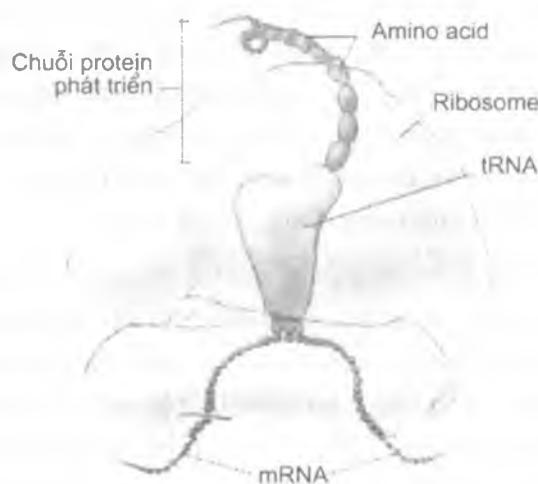
Cuống lá cấu trúc xoắn kép theo nguyên lý cặp đôi base của Watson và Crick, ở cả ba lá đều có cấu trúc xoắn đôi nhưng trên đầu có thêm một vòng khuyên với 7 hoặc 8 nucleotide, trong đó có 3 nucleotide ở giữa tạo thành bộ đôi mā. Điều đáng lưu ý là ngoài các nucleotide thông thường còn thấy xuất hiện những nucleotide hiếm như : dihydrouridine (D), inosine (I), thymine (T, như DNA), pseudouridine (ψ) và gốc methyl hoá (m).

Các enzyme aminoacyl-tRNA synthetase nhận diện được phân tử tRNA nhất định và xúc tác phản ứng gắn amino acid tương ứng vào cuống – đầu không có vòng khuyên của phân tử tRNA.

c) RNA ribosome (rRNA) là thành phần cấu trúc của bộ máy sinh tổng hợp protein

Trong thực tế, nếu các thành phần tham gia quá trình dịch mã mRNA hoạt động dưới dạng hòa tan trong dung dịch thì tốc độ dịch mã tạo protein xảy ra rất chậm. Hiệu quả tổng hợp protein tăng lên rất mạnh khi gắn các phân tử mRNA và các phân tử aminoacyl-tRNA vào một cấu trúc phức hợp RNA–protein dưới dạng bào quan là các thể ribosome. Bộ máy hai bộ phận này điều khiển quá trình kéo dài phân tử polypeptide đạt tốc độ 3-5 amino acid/giây. Một phân tử protein kích thước nhỏ với 100-200 amino acid được tổng hợp trong vòng 1 phút hoặc nhanh hơn, ngược lại cần 2-3 giờ để làm ra phân tử actin có trong bắp cơ với chiều dài phân tử là 30000 amino acid. Như vậy bộ máy này cần rất chính xác và có độ bền cao.

Thể ribosome được tạo bởi nhiều loại phân tử rRNA và trên 50 loại protein, sắp xếp thành hai nửa (tiểu phần) – nửa lớn và nửa nhỏ. Protein và rRNA của hai nửa hoàn toàn khác nhau. Nửa nhỏ chứa các phân tử rRNA riêng rẽ, gọi chung là *RNA nhỏ*, trong khi



Hình 1.15. Cấu trúc hai nửa của phân tử ribosome.

Trong chương này, chúng ta tập trung nghiên cứu về đặc điểm cấu trúc của các phân tử protein như thế nào mà chúng có thể thực hiện được các chức năng sinh học do chúng đam nhận.

1.2.1. Thành phần của protein

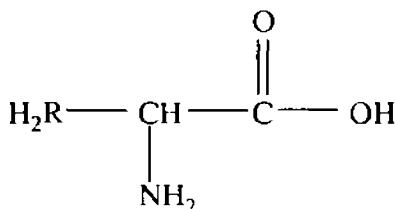
Protein được thiết kế có thể liên kết với bất kỳ loại hình phân tử nào, từ những ion kim loại đơn giản đến phân tử chất béo, chất đường, nucleic acid và các protein khác. Chúng có khả năng xúc tác hàng loạt loại phản ứng hóa học, tạo nên những cấu trúc tế bào đa dạng, giám sát việc vận chuyển các chất vào trong và ra ngoài tế bào, kiểm soát nồng độ các chất trao đổi, hoạt động như những bộ cảm ứng và công tắc, tạo nên chuyển động và kiểm soát chức năng của gene. Cấu trúc không gian ba chiều của phân tử protein góp phần quyết định, cho phép protein thực hiện được những chức năng nêu trên một cách hiệu quả và với tính chính xác cao. Sự sắp xếp không gian, hình dạng trong không gian 3 chiều là những thông tin trọng tâm để hiểu được các protein hoạt động ra sao.

Một trong những lĩnh vực chính của nghiên cứu sinh vật học hiện đại là vì sao các protein được cấu tạo bởi chỉ có 20 amino acid khác nhau mà chúng lại có thể thực hiện được một mang lưới các chức năng khổng lồ như thế. Khác với cấu trúc phân tử chia nhánh phức tạp của các chất carbohydrate, các phân tử protein chỉ có cấu trúc chuỗi các amino acid không phân nhánh. Cấu trúc ít gấp của các phân tử protein hình thành nhờ các liên kết không hoá trị của các đoạn trình tự amino acid.

Amino acid, peptide và protein

Protein là một loại đại phân tử có khối lượng phân tử từ 10000 đến một vài triệu được hợp thành bởi các amino acid.

Công thức hoá học phổ biến của các amino acid là :



Trong đó, gốc amine gắn ở vị trí α của phân tử acid hữu cơ.

Mặc dù trong thế giới thực vật đã phát hiện thấy có tới 100 loại amino acid khác nhau, nhưng trong protein của tất cả các loài sinh vật chỉ có 20 amino acid (hình 1.16). Thông thường, gốc amine được gắn vào vị trí α và tất cả các amino acid trừ glycine trung tính đều quang hoạt theo chiều L. Gốc R có thể rất khác nhau, qua đó người ta phân biệt các amino acid theo các nhóm như :

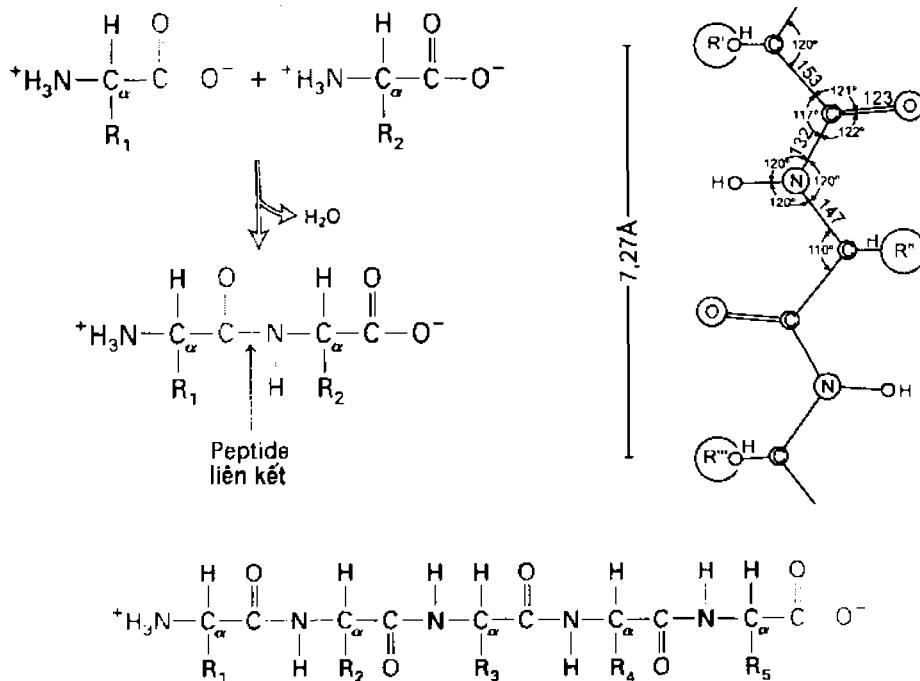
– Hydrophylic (ưa nước) : gồm những amino acid tan trong nước, có chuỗi R ion hoá dạng base (dương tính) như lysine, arginine, histidine ; dạng acid (ion âm tính) như aspartic, glutamic hay phân cực nhưng không mang diện tích như serine, threonine, asparagine và glutamine. R cũng có thể chỉ là một nguyên tử hydro như ở glycine.

– Hydrophobic (ky nước) : Mang chuỗi R có cấu trúc mạch thẳng hay mạch vòng của mạch carbon hữu cơ không hoặc tan rất ít trong nước. Gốc R của các amino acid như alanine, valine, leucine, isoleucine và methionine chỉ là những mạch carbohydrate thông thường, trừ nguyên tử lưu huỳnh trong methionine và tất cả đều không phân cực. Phenylalanine, tyrosine và tryptophan có gốc R gồm những carbohydrate mạch vòng.

– Có ba amino acid đặc biệt là cysteine, glycine và proline với những tính chất riêng rẽ do gốc R của chúng tạo nên. Cysteine có một nhóm sulfite (SH), có khả năng bị oxy hoá tạo thành cầu nối disulfite –S–S– với một phân tử cysteine khác. Cầu nối này tạo nên các loại cấu trúc gấp xoắn ba chiều ở nhiều loại protein ngoại bào. Phân tử glycine là amino acid nhỏ nhất vì gốc R của nó chỉ là một nguyên tử H, nó có thể lọt vào những vị trí nhỏ nhất trong cấu trúc protein. Còn proline lại có một liên kết giữa nhóm NH_3 với gốc CH_2 tạo nên những nút xoắn cố định trong chuỗi protein.

Amino acid hydrophobic (ưa nước)								
Amino acid kiềm			Amino acid			Amino acid phân cực		
Asparagine (Asn hoặc N)	Asparagine (Asn hoặc N)	Asparagine (Asn hoặc N)	Asparagine (Asn hoặc N)	Asparagine (Asn hoặc N)	Asparagine (Asn hoặc N)	Asparagine (Asn hoặc N)	Asparagine (Asn hoặc N)	Glutamine (Gln hoặc Q)
Amino acid hydrophobic (kỵ nước)						Amino acid đặc biệt		
Methionine (Met hoặc M) 1.59	Leucine (Leu hoặc L) 1.61	Isoleucine (Ile hoặc I) 1.72	Leucine (Leu hoặc L) 1.61	Methionine (Met hoặc M) 1.59	Phenylalanine (Phe hoặc F) 1.69	Tyrosine (Tyr hoặc Y) 1.56	Tryptophan (Trp hoặc W) 1.72	Proline (Pro hoặc P) Proline (Pro hoặc P) Proline (Pro hoặc P)

Hình 1.16. Cấu trúc phân tử của 20 loại amino acid thông thường phân chia thành 3 nhóm chính gồm (1) amino acid kiềm, (2) amino acid kị nước (acid) và (3) amino acid đặc biệt.



Hình 1.17. Liên kết peptide được tạo bởi phản ứng giữa một gốc amine của một amino acid này với một gốc carboxyl của amino acid kia (hình trái). Liên kết peptide là liên kết hoá bền chất tạo nên trình tự mạch đơn của phân tử protein cơ bản (hình phải).
 Mỗi trình tự protein đều có đầu amine (còn gốc NH₃ tự do và đầu carboxyl có gốc COOH tự do (hình dưới).

Thông thường, gốc carboxyl (COOH) của phân tử amino acid này có khả năng phản ứng với gốc amine của phân tử amino acid kia giải phóng một phân tử nước và tạo ra một mối liên kết hóa học gọi là liên kết peptide (hình 1.17). Cứ theo cách thức đó, nhiều phân tử amino acid liên kết với nhau thành một chuỗi dài. Các peptide bao gồm 2 gốc amino acid được gọi là dipeptide, 3 amino acid là tripeptide, dưới 10 amino acid gọi chung là oligopeptide, chúng được gọi chung là polypeptide. Các đại phân tử peptide có tới hàng trăm amino acid.

1.2.2. Cấu trúc protein

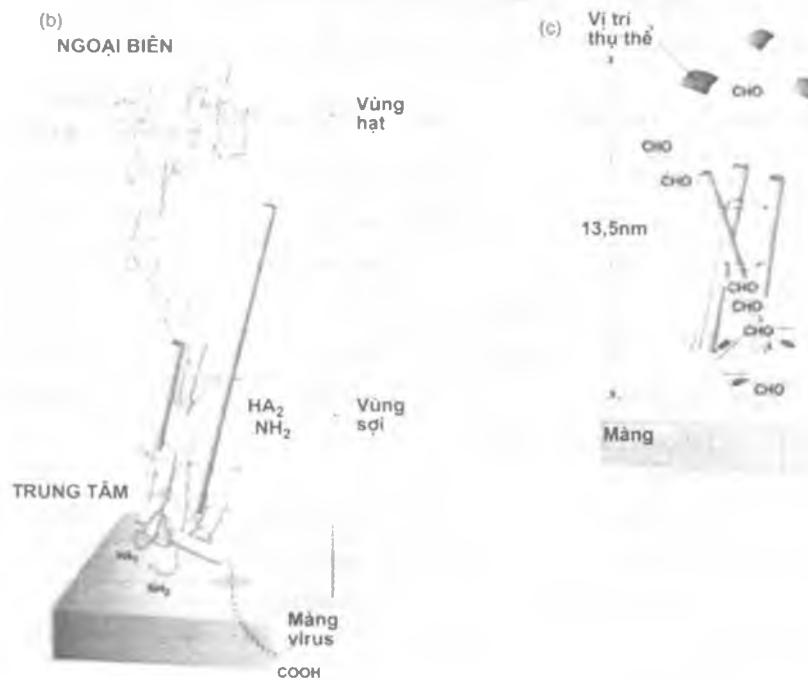
Cấu trúc thông thường của protein được mô tả theo trình tự bốn bậc. Hình 1.18 trình bày khái quát bốn bậc cấu trúc của phân tử hemagglutinin, một loại protein bề mặt của virus cúm. Protein này gắn vào bề mặt tế bào, cả tế bào người và gây nên quá trình nhiễm virus cúm.

(I) Cấu trúc bậc một (= cấu trúc sơ cấp)

Trong phân tử protein, các amino acid liên kết xếp hàng với nhau theo một trình tự nhất định tạo nên một mạch thẳng thông qua mối liên kết peptide được hình thành giữa gốc COOH của amino acid đứng trước với gốc NH₂ của amino acid đứng tiếp theo sau. Mỗi loại protein có một trình tự hoàn toàn riêng biệt, khác với trình tự trong phân tử protein loại khác. Các phân tử của một loại protein có trình tự amino acid hoàn toàn giống nhau. Trình tự amino acid trong phân tử protein được gọi là cấu trúc bậc một (hay còn gọi là cấu trúc sơ cấp) của protein. Tính toán lý thuyết cho thấy khả năng hình thành các loại cấu trúc bậc một của protein có số lượng vô cùng lớn. Nếu mỗi protein có 100 amino acid thì với 20 amino acid khác nhau có thể tạo được 20¹⁰⁰ loại protein có cấu trúc bậc một khác nhau. Trong thực tế, số lượng amino acid của các đại phân tử protein có khi lên đến vài trăm và cả nghìn.

(a)
DALLGDPHICDVFOINETWOLFVERSIAFSNCYPVDVFDYASLRLVASSGTLFIFTGFTWTGV

TONGGSNACKKPGSOPPSALNWLTISGOSTYPIVILVTFMPNNNDIKLYANGIMHPSTNOEOTS



Hình 1.18. Bốn bậc cấu trúc của phân tử hemagglutinin.

Hình 1.18 minh họa bốn bậc cấu trúc của phân tử hemagglutinin, có cấu trúc nhiều tiểu phần với ba tiểu đơn vị giống nhau, trong đó mỗi tiểu đơn vị có hai chuỗi HA1 và HA2. Hình (a) trình bày cấu trúc bậc một là một trình tự amino acid từ vị trí 69 đến vị trí 195 của chuỗi HA. Chuỗi này virus dùng để gắn với màng tế bào. Ký hiệu các amino acid viết bằng loại một chữ cái. Bên dưới trình tự amino acid là cấu trúc bậc hai mô tả các đoạn polypeptide tạo thành các chuỗi xoắn α , sợi β và búi rối. Hình (b) mô tả cấu trúc bậc ba là kết quả gấp xếp các chuỗi, sợi và búi thành cấu trúc không gian hoàn chỉnh có chiều cao là 13,5 nm, chia thành hai vùng : Vùng xa màng virus gấp xếp thành cấu trúc dạng hạt ; Vùng gần màng thì tạo thành cấu trúc dạng sợi bởi các chuỗi xoắn và sợi. Hình (c) mô tả cấu trúc bậc bốn gồm ba tiểu đơn vị HA như mô tả ở hình (b). Trên đỉnh mỗi tiểu đơn vị có một vị trí thụ thể giúp thể virus gắn vào màng của tế bào mục tiêu. Cũng như nhiều loại protein màng, HA cũng có nhiều vị trí liên kết với các chuỗi carbohydrate (CHO).

Bảng 1.3 trình bày số lượng amino acid trong một số loại peptide và protein thông dụng.

Bảng 1.3. Số lượng amino acid liên kết với nhau trong một phân tử của một số loại peptide, protein và phức hợp enzyme protein.

Tên peptide	aa	Tên protein	aa	Tên phức hợp	Số lượng amino acid (aa)
Phaloidin	7	RNase	124	Protamine	4000-12000
Oxytocin	7	Lysozyme	129	Histone	5000-30000
Vasopessin	9	Hemoglobin	7	Pronein ribosome	12000-31000
Eleoison	11	Chuỗi α	141	RNase	12700
Gramicidin	12	Chuỗi β	141	Trypsin	12700
Glukagon	29	Myoglobin	153	Elastase	25900
Proinsulin	84	Protein virus khâm thuốc lá	158	Pepsin	33000
Insulin	7	Prolactin	198	Hexokinase	45000
Chuỗi A	29	Trypsin	223	Hemoglobin	68000
Chuỗi B	30	Chymotrypsin A	230	Albumin huyết thanh	68000
Calcitonin	84	Elastase	153	Lactate dehydrogenase	33000

Corticotropin	39	Chymo-trypsinogen A	245	Catalase	250000
Salmin	58	Tryptophan synthetase A	267	Fibronogene	330000
Cytochrome C		Pepsin	338	Urease	480000
Tim ngựa	104			Myosin	620000
<i>Naurospora</i>	107			Thyreoglobulin	630000
Nấm men	108			Glutamate dehydrogenase	1000000

Trình tự amino acid sắp xếp trong phân tử peptide và protein hoàn toàn do trình tự nucleotide trong phân tử DNA cũng như trong phân tử RNA thông tin quyết định thông qua mã di truyền theo nguyên tắc bộ 3 nucleotide là một mã cho 1 amino acid. Vì bộ mã di truyền mang tính phổ cập cho mọi sinh vật nên chương trình vi tính có thể tính toán được trình tự của mọi protein từ thông tin về trình tự gene và ngược lại. Đến nay đã có hàng triệu trình tự protein được ngân hàng dữ liệu Swiss Protein Database lưu giữ và xử lý.

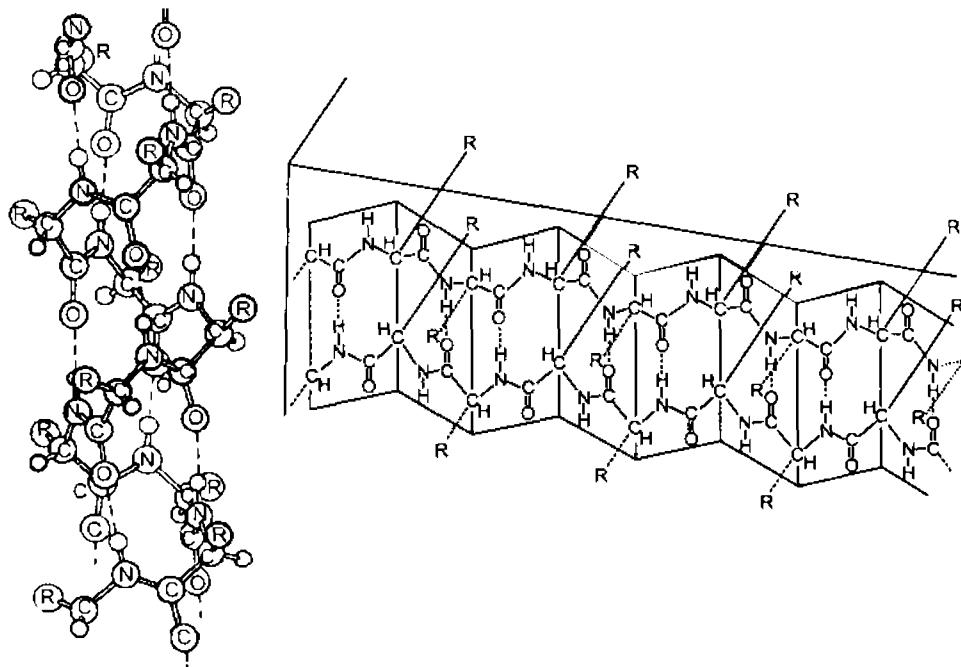
Điều đáng lưu ý là trong quá trình tiến hoá, nhiều loại phân tử protein cùng chức năng sinh học vẫn duy trì được những vùng bảo thủ, tức là ở những đoạn phân tử protein này không có sự thay đổi lớn, mà thường chỉ xảy ra những thay thế trong trình tự amino acid. Chính vì vậy, có thể sử dụng thông tin về protein của loài này để tìm kiếm và phát hiện ra phân tử cùng loại ở loài khác.

(II) Cấu trúc bậc hai (thứ cấp) của phân tử protein

Các phân tử amino acid trong chuỗi protein có nhiều nhóm chức có khả năng tạo ra các liên kết giữa chúng với nhau trong một chuỗi peptide hay giữa các chuỗi với nhau. Các mối liên kết đó là :

- Liên kết đồng hoá trị bao gồm (1) liên kết giữa gốc amine và gốc acid ; (2) liên kết disulfide.
- Liên kết ion giữa gốc COOH của một amino acid có hai gốc COOH (ví dụ Glu) với gốc NH_3^+ của amino acid khác có hai gốc NH_3^+ (ví dụ : Lys).
- Liên kết đối cực do tính phân cực của một số amino acid.
- Liên kết thông qua cầu nối hydro giữa nguyên tử O có trong liên kết C=O và H trong bất kỳ gốc chứa H nào (ví dụ gốc NH).

Như vậy, cấu trúc thứ cấp được định nghĩa là cấu trúc không gian bậc một của phân tử protein. Cần cứ vào hình thái của trật tự cấu trúc không gian người ta phân biệt các loại cấu trúc thứ cấp. Đó là (I) Xoắn lò xo α bao gồm kiểu phân tử những protein thể hạt hoặc protein thể sợi và (II) Bản kép β bao gồm kiểu protein dạng gấp nếp và cấu trúc dạng cuộn chỉ. Cả hai kiểu cấu trúc này đều cùng tồn tại trong một phân tử protein, ví dụ như trong phân tử lysozyme có nhiều vùng thuộc cấu trúc xoắn α , đồng thời cũng có những vùng thuộc cấu trúc bản β (hình 1.19).

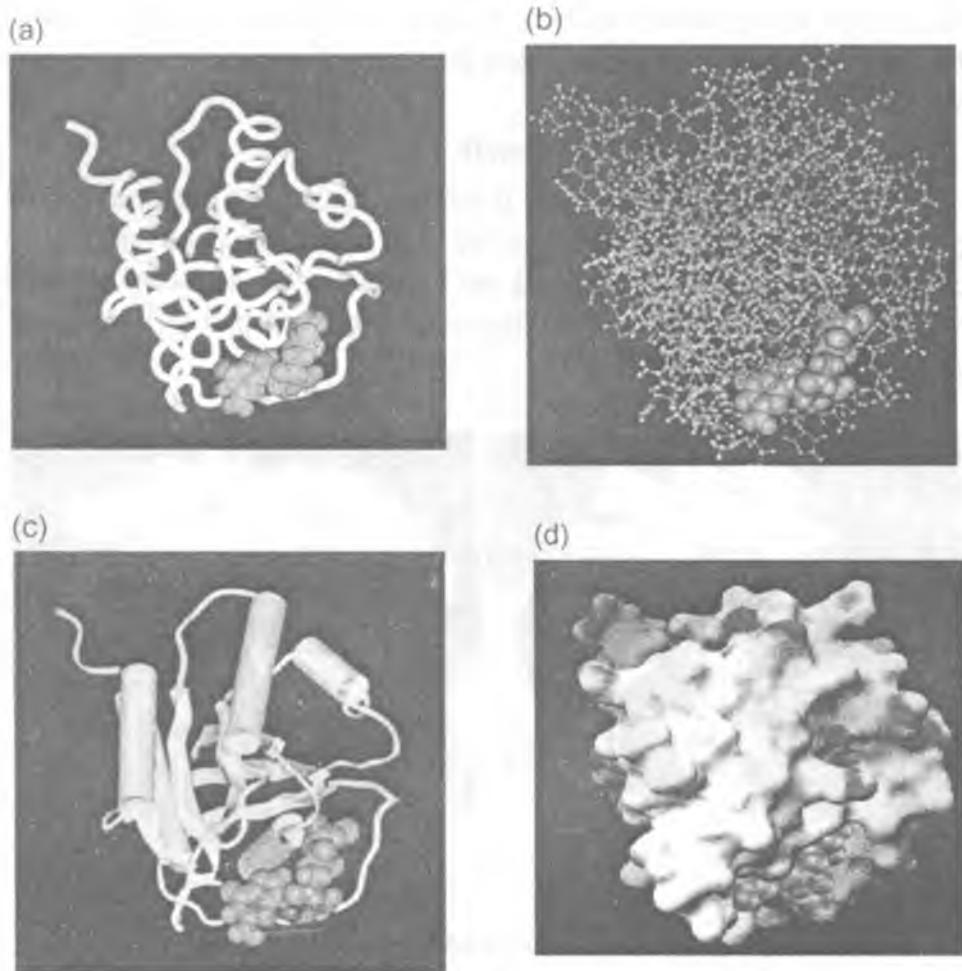


Hình 1.19. Hai kiểu cấu trúc điển hình của chuỗi peptide : kiểu chuỗi xoắn α và kiểu bản kép β .

(III) Cấu trúc bậc ba của protein

Đó là cấu trúc không gian ba chiều của phân tử protein khi sợi peptide dạng xoắn thông qua các liên kết xếp nếp và gấp xoắn thành một cấu trúc không gian cố định. Hình 1.20 trình bày cấu trúc bậc ba của phân tử protein liên kết guanine theo nhiều cách trình bày khác nhau. Ngược với cấu trúc bậc hai, trong đó các cấu trúc được thiết lập nhờ các cầu nối hydro, cấu trúc bậc ba được duy trì ở trạng thái bền vững nhờ các tương tác kị nước của các gốc R của các amino acid và nhờ các cầu

nối disulfide. Nhờ các lực này mà các chuỗi xoắn α , các chuỗi β , các cuộn được liên kết với nhau thành một khung dàn giáo vững chắc. Bởi thế, hình dạng và kích thước của một protein không những chỉ phụ thuộc vào trình tự amino acid của nó và còn phụ thuộc vào số lượng và cách sắp xếp của các chuỗi xoắn, chuỗi đơn và các cuộn trong cấu trúc bậc hai.



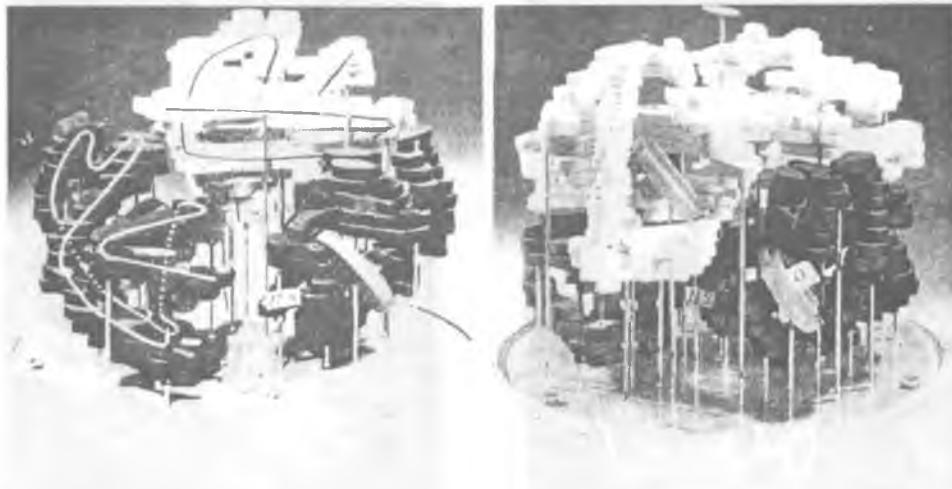
Hình 1.20. Các hình thức biểu diễn mô hình cấu trúc phân tử Ras, một loại protein liên kết.

Hình 1.20 trình bày các hình thức biểu diễn mô hình cấu trúc phân tử Ras, một loại protein liên kết. Phân tử được biểu thị dạng cầu trên các hình là phân tử cơ chất guanosine diphosphate. (a) Khung xương của phân tử Ras được gấp nén để có một cấu trúc gọn với thể tích nhỏ nhất.

(b) Mô hình bóng-que cho thấy vị trí của mọi nguyên tử trong phân tử protein này. (c) Sơ đồ cấu trúc loại này cho thấy các cuộn xoắn α và chuỗi β được sắp xếp như thế nào trong phân tử Ras này. Giữa các cuộn xoắn và các chuỗi thường là các vòng khuyên. (d) Bề mặt ưa nước của phân tử Ras, có những vùng tích điện dương và điện âm. Rõ ràng bề mặt của protein không nhẵn mà có các mảng lồi, lõ hổng và khe hở. Sự phân bố không đều của các vùng tích điện âm dương là cơ sở cho việc kết gân với các loại phân tử cơ chất.

(IV) Cấu trúc bậc bốn của protein

Cấu trúc bậc bốn của protein là kết quả của sự liên kết nhiều chuỗi peptide có cấu trúc bậc ba với nhau để hình thành một phức hợp protein có hoạt tính sinh học. Hình 1.21 là mô hình của phân tử protein được mô hình hóa thông qua phần mềm silicongraphic hiện rất phổ biến trong các chương trình sinh tin học.



Hình 1.21. Mô hình cấu trúc bậc bốn của hemoglobin.

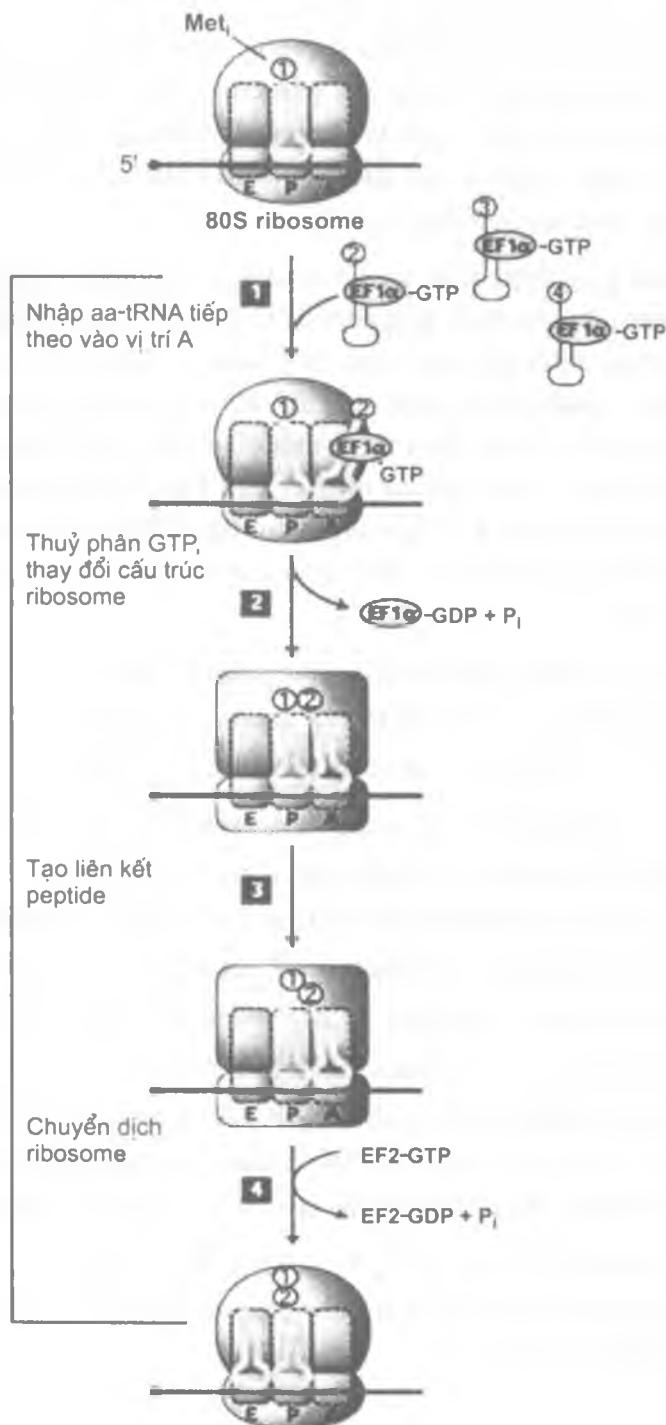
1.2.3. Sinh tổng hợp protein

Quá trình sinh tổng hợp phân tử protein xảy ra nhờ các thể ribosome. Phân tử mRNA sau khi được tổng hợp đã liên kết với ribosome tạo thành những chuỗi cùm polysome. Tuy nhiên, quá trình sinh tổng hợp một phân tử polypeptide lại xảy ra trong mỗi một hạt ribosome. Mỗi ribosome khi trượt qua một phân tử mRNA thì để lại

sau nó một chuỗi polypeptide. Hình 1.22 cho thấy trong mỗi hạt ribosome có hai khoang, mỗi khoang rộng bằng đúng kích thước của 3 nucleotide. Khoang thứ nhất là nơi gắn của phân tử tRNA kèm chuỗi polypeptide, cho nên được gọi là khoang P, khoang liên kề cũng rộng đúng bằng 3 nucleotide, là nơi gắn của phân tử tRNA mang một amino acid, cho nên có tên gọi là khoang A.

Quá trình gắn thêm một amino acid vào chuỗi polypeptide xảy ra thông qua việc chuyển chuỗi polypeptide từ khoang P sang khoang A, cầu nối peptide được thiết lập giữa chuỗi polypeptide với phân tử amino acid làm cho chuỗi polypeptide được kéo dài thêm một gốc amino acid nữa. Lúc này ribosome chuyển dịch thêm một đoạn dài bằng đúng một bộ ba nucleotide về đầu 3' của phân tử mRNA làm cho chuỗi polypeptide mới kéo dài nằm vào khoang P và khoang A trở nên trống để cho một phân tử tRNA với một amino acid mới xâm nhập vào và quá trình tổng hợp lại lặp lại từ bước một.

Chu trình sinh tổng hợp protein thông qua kéo dài chuỗi polypeptide xảy ra ở trong từng hạt ribosome của chuỗi polysome của tế bào nhân chuẩn (hình 1.23). Một hạt ribosome 80S mang Met-tRNA^{Met} tại vùng P được lắp vào, một phức bộ tứ mang một aa thứ hai (aa₂) tương ứng với mã di truyền trên mRNA được gắn vào vị trí A (bước một). Kế đến, thể ribosome thực hiện việc biến đổi cấu trúc dẫn đến GTP trong phức EF2-GTP bị thuỷ phân (bước hai), phân tử rRNA lớn xúc tác quá trình hình thành liên kết peptide giữa Met và aa₂ (bước ba). Thuỷ phân GTP tạo nên thay đổi cấu trúc của ribosome khiến cho thể ribosome trượt về trước một bộ mã (codon) trên phân tử mRNA và đẩy phân tử tRNA chưa adenyl hoá về vị trí E và phân tử tRNA mang đoạn peptide vừa tạo được sang vị trí P (bước bốn). Chu trình lắp lại từ đầu bằng bước gắn một phức bộ tứ mang aa₃ tại vị trí A hiện đang trống. Phân tử tRNA tự do được đẩy ra ngoài ở bước hai khi thể ribosome thay đổi cấu trúc do quá trình thuỷ phân GTP gây ra.



Hình 1.22. Quá trình sinh tổng hợp protein bên trong thể ribosome.

Protein hình thành từng bước trên ribosome

Khái quát lại ta thấy để sinh tổng hợp protein cần có bốn yếu tố : (1) Thể ribosome liên kết với (2) mRNA thành từng chuỗi polysome, (3) Các phân tử tRNA và (4) các amino acid. Nhiệm vụ của bộ máy sinh tổng hợp protein là chuyển trình tự mã di truyền trên mRNA thành trình tự amino acid trên chuỗi protein. Quá trình này được gọi là quá trình dịch mã.

Mã khởi đầu AUG được nhận biết bởi methionyl-tRNA^{Met}

Hầu hết các phân tử mRNA đều có mã AUG với chức năng khởi đầu quá trình sinh tổng hợp các chuỗi protein. Mỗi chuỗi protein đều được bắt đầu bằng methionine vì thế trong tế bào nhân chuẩn và tế bào nhân chưa chuẩn đều có hai loại tRNA chuyên vận chuyển methionine : một tRNA chuyên cho việc vận chuyển methionine vào vị trí khởi đầu quá trình sinh tổng hợp protein tại điểm P trên tiểu phần nhỏ của ribosome và một tRNA chuyên vận chuyển methionine vào những vị trí nằm trong chuỗi protein tại điểm A trên tiểu phần lớn của ribosome. Tuy nhiên, chỉ có một aminocyl-tRNA synthetase làm nhiệm vụ xúc tác việc gắn kết methionine vào cả hai loại tRNA này. Thường thì phức hợp methionyl-tRNA khởi đầu được ký hiệu thêm một chỉ số "i" ($\text{Met-tRNA}_{\text{i}}^{\text{Met}}$) để phân biệt với phức hợp methionyl-tRNA thông thường. Sự sai khác giữa methionyl-tRNA khởi đầu và methionyl-tRNA trong phân tử lại có liên quan đến sự hoạt hoá methionine, chỉ có methionine đã hoạt hoá mới có khả năng liên kết với tRNA đặc hiệu để làm chức năng khởi đầu chuỗi protein. Đối với các vi sinh vật thì hiện tượng gốc amine của methionine bị formyl hoá (+CHO, ký hiệu là fMet-tRNA $_{\text{i}}^{\text{Met}}$) được thay thế cho quá trình hoạt hoá.

Sinh tổng hợp protein ở vị khuẩn bắt đầu ngay ở đầu 5' của mRNA

Để khởi đầu quá trình sinh tổng hợp protein trên phức hệ ribosome vi khuẩn, một tương tác liên kết xảy ra giữa protein có tên gọi là yếu tố khởi đầu (initiation factor, IF) với tiểu phần nhỏ (30S) của ribosome. Sau đó, phức hệ này cùng với fMet-tRNA $_{\text{i}}^{\text{Met}}$ gắn kết với mRNA vào gần vị trí

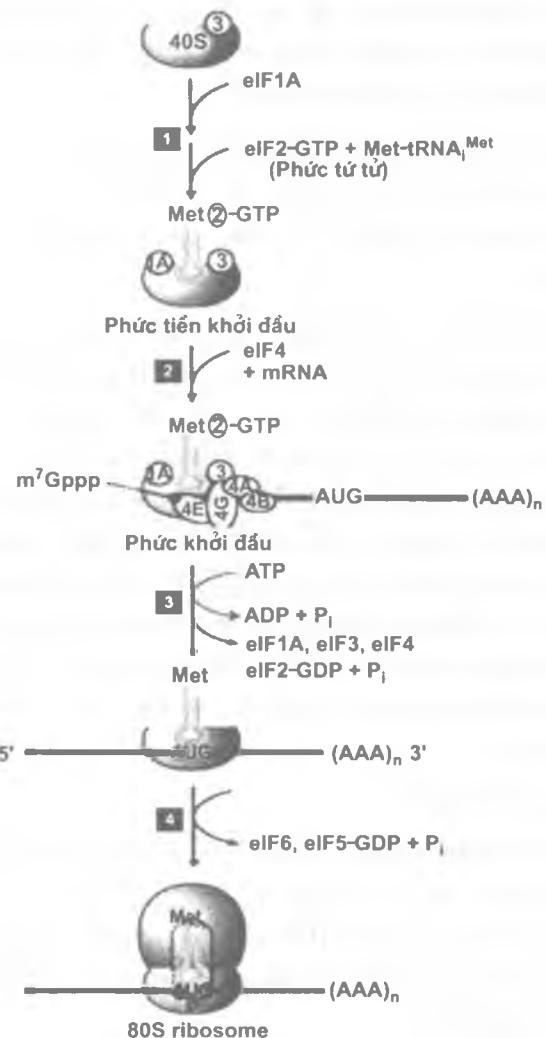
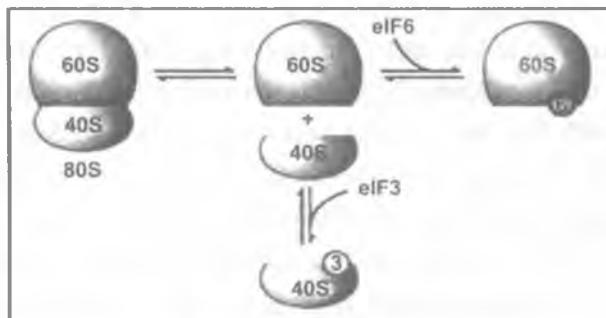
của mã di truyền khởi đầu AUG thông qua tác động tương hỗ của IF1 và IF3, hình thành phức hệ khởi đầu 30S.

Để nhận biết vị trí khởi đầu, phân tử rRNA nhỏ (16S) có trong tiểu phân nhỏ của ribosome tương tác với một đoạn gồm 8 nucleotide trên phân tử mRNA có tên gọi theo tác giả phát minh ra nó là *Shine-Dalgarno*. Đoạn này thường nằm gần mã khởi đầu AUG, tạo cặp đôi base với các nucleotide ở vùng đầu 3' của phân tử 16S rRNA. Như thế rRNA vi khuẩn làm luôn cả nhiệm vụ khởi động quá trình dịch mã phân tử mRNA. Thường những phân tử mRNA dài có nhiều đoạn Shine-Dalgarno bên trong nằm gần vị trí mã khởi đầu của các protein được mã hoá. Vì tiểu phân nhỏ của thể ribosome có thể gắn độc lập với các đoạn Shine-Dalgarno nên trên một phân tử mRNA dài có thể xảy ra nhiều quá trình tổng hợp protein khác nhau vào cùng một thời điểm.

Việc lắp ghép phức hệ khởi đầu 70S hoàn thành khi tiểu phân lớn dùng năng lượng của phản ứng thuỷ phân GTP để liên kết với IF2. Lúc này, phân tử tRNA khởi đầu được đưa vào vị trí P của thể ribosome hoàn chỉnh.

Sinh tổng hợp protein ở tế bào nhân (chuẩn) bắt đầu ở đầu 5' hoặc giữa phân tử mRNA

Cũng như ở bộ máy sinh tổng hợp protein của tế bào vi khuẩn, quá trình sinh tổng hợp protein ở tế bào nhân chuẩn cũng bắt đầu bằng việc tạo ra phức hệ protein khởi đầu (eIF) với tiểu phân nhỏ của ribosome trước khi kết gắn với mRNA. Tuy nhiên, ở tế bào nhân chuẩn có những điểm khác cơ bản : Có hai phân tử protein khởi đầu là eIF3 (kích thước lớn với 8 protein tiểu phân) và eIF6 chịu trách nhiệm giữ cho hai tiểu phân của thể ribosome tách rời nhau sau khi ribosome đó kết thúc việc tham gia vào quá trình sinh tổng hợp một phân tử protein trước đó để tiếp tục tham gia vào sinh tổng hợp phân tử protein tiếp theo. Trước tiên, một phức hệ bộ ba gồm eIF2 đã gắn với GDP và Met-tRNAP^{Met} kết gắn bền vững với tiểu phân nhỏ 40S của ribosome thông qua hai yếu tố là eIF3 và eIF1A (hình 1.23).



Hình 1.23. Bắt đầu quá trình sinh tổng hợp protein xảy ra ở chuỗi polysome của tế bào nhân chuẩn.

Hình 1.23 minh họa quá trình sinh tổng hợp protein bắt đầu xảy ra ở chuỗi polysome của tế bào nhân chuẩn (trong phân đóng khung). Khi một thể ribosome tách rời khỏi điểm kết của quá trình dịch mã thì các tiểu phân 40S và 60S liên kết với các yếu tố khởi đầu eIF3 và eIF6 tạo nên phức mới để bắt đầu một chu trình sinh tổng hợp protein khác. Bước một và hai : Bổ sung phức tiểu phân 40S-eIF3 tạo nên phức khởi đầu. Bước ba : Sàng lọc mRNA thông qua liên kết phức khởi đầu với tiểu phân nhỏ và Met-tRNA^{Met} tại codon khởi đầu. Bước bốn : Liên kết với tiểu phân 60S tạo ra thể ribosome hoàn chỉnh 80S sẵn sàng phiên mã mRNA. Hai yếu tố khởi đầu eIF2 và eIF5 là những protein liên kết GTP thực hiện việc thuỷ phân GTP trong quá trình dịch mã.

Tế bào có thể điều tiết quá trình sinh tổng hợp protein thông qua việc phosphoryl hoá một gốc serine trên eIF2 liên kết với GDP làm cho phức này không gắn với Met-tRNAl^{Met} được, qua đó, quá trình sinh tổng hợp protein bị ức chế.

Để nhận biết điểm khởi đầu sinh tổng hợp protein trên phân tử mRNA tiểu phân nhỏ của ribosome tế bào nhân chuẩn cần nhiều yếu tố khác nhau. Hầu hết các mRNA của tế bào nhân chuẩn chỉ có điểm khởi đầu ở ngay đầu 5' của nó. Hệ protein thuộc eIF4 làm nhiệm vụ tương tác giữa tiểu phân nhỏ và đầu 5' của mRNA với cấu trúc methyl hoá. Cấu trúc methyl hoá này được nhận biết bởi eIF4F và phần biến đổi bị loại bỏ thông qua hoạt tính helicase có trong eIF4F. Lúc này, phức hợp khởi đầu đã hoàn thiện với nhóm protein eIF3 trượt dọc theo phân tử mRNA và dừng lại tại vị trí mã khởi đầu AUG. Tuy nhiên, việc chọn lọc vị trí dừng AUG được thực hiện nhờ một đoạn nucleotide có tên gọi là đoạn Kozak (tác giả phát minh) :

mRNA 5'-ACCAUUGG-3'

Nucleotide A đứng trước AUG có vẻ là nucleotide quan trọng nhất quyết định điểm khởi đầu của quá trình dịch mã. Ở đây đã chứng minh được mối tương tác giữa phức eIF3, tiểu phân 40S của ribosome và phân eFO4G của phức eIF4, nó quyết định sự chính xác của điểm khởi đầu quá trình sinh tổng hợp protein.

Sau khi vào được vị trí chính xác của mã di truyền khởi đầu một tiểu phân nhỏ có gắn với Met-tRNAl^{Met} sẽ liên hợp với tiểu phân lớn 60S

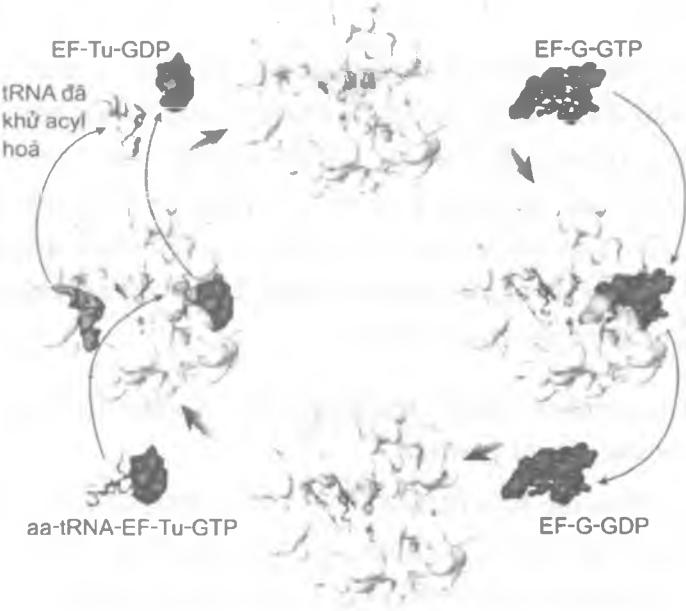
của ribosome tạo nên phức ribosome-mRNA khởi đầu kích thước 80S (hình 1.23).

Thông thường thì quá trình dịch mã các phân tử mRNA của tế bào nhân chuẩn được bắt đầu ở đầu 5' nhưng trong trường hợp tế bào này bị nhiễm virus thì việc dịch mã các mRNA của thể virus xâm nhập không có "mũ" methyl hoá, nên được khởi đầu tại những vị trí ở giữa phân tử. Vị trí đó được gọi là *diểm nội nhập của ribosome* (internal ribosome entry site, IRES). Đến nay, cũng đã phát hiện được có những mRNA của tế bào có thể được dịch mã thông qua IRES.

Mỗi aminoacyl-tRNA trượt qua ba điểm ribosome trong quá trình kéo dài chuỗi protein

Phức ribosome-Met-tRNAM^{Met} có kích thước 70S của tế bào vi khuẩn hay 80S của tế bào nhân chuẩn sau khi gắn đúng vị trí khởi đầu trên mRNA sẵn sàng làm nhiệm vụ kết gắn các amino acid với nhau theo nguyên tắc phiên dịch khung đọc mã di truyền. Cũng không khác với giai đoạn khởi đầu, ở giai đoạn kéo dài cũng cần có những yếu tố kéo dài (elongation factor, EF) hỗ trợ. Bước quan trọng của giai đoạn kéo dài này là (1) sự xâm nhập của các phân tử aminoacyl-tRNA kế tiếp vào ribosome, (2) sự hình thành liên kết peptide và (3) sự chuyển động của ribosome sang mã di truyền tiếp theo đọc trên phân tử mRNA. Cụ thể là phức chất aminoacyl-tRNA thứ hai (sau methionine) được mang đến chỗ phức hệ bộ ba kết gắn sẵn với EF1 γ -GTP trong tế bào nhân chuẩn (và là EF-Tu-GTP trong tế bào vi khuẩn) và được gắn vào vị trí A của ribosome (hình 1.24).

Ta đã biết rằng phức chất khởi đầu Met-tRNAM^{Met} đang gắn với ribosome ở vị trí P với bộ giải mã phù hợp cho mã khởi đầu AUG. Nếu bộ giải mã của phân tử tRNA của phức chất aminoacyl-tRNA thứ hai phù hợp với mã thứ hai của mRNA thì một liên kết chắc chắn giữa phức chất này và ribosome được tạo ra tại vị trí A của ribosome, năng lượng cho việc đó được tạo ra nhờ thuỷ phân phức chất EF1 γ -GTP (hoặc EF-Tu-GTP). Nếu bộ giải mã không phù hợp thì phức chất aminoacyl-tRNA đó sẽ khuếch tán đi nơi khác, không có phản ứng kéo dài phân tử protein xảy ra.



Hình 1.24. Bắt đầu quá trình sinh tổng hợp protein xảy ra ở chuỗi polysome của tế bào nhân chuẩn.

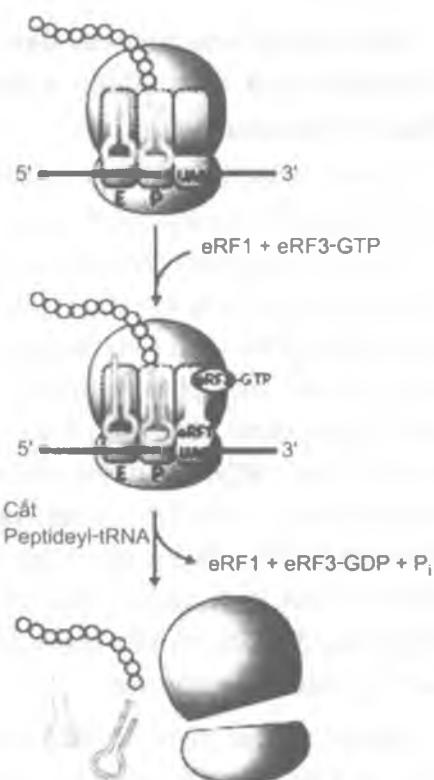
Với phức chất khởi đầu Met-tRNA^{Met} tại vị trí P và phức chất aminoacyl-tRNA tại vị trí A gốc α -amino của amino acid thứ hai sẽ phản ứng với methionine đã hoạt hoá đang gắn trên phân tử tRNA khởi đầu và tạo nên cầu nối liên kết peptide. Nghiên cứu hoá sinh tìm được bằng chứng rằng *phản ứng peptidyltransferase* đóng vai trò hoạt hoá amino acid của phức chất aminoacyl-tRNA đang đến lượt vào cuộc tại vị trí P. Ở tế bào vi khuẩn, phân tử 23S rRNA trong tiểu phân lớn của ribosome đảm nhận chức năng peptidyltransferase. Thí nghiệm loại bỏ hết protein của tiểu phân lớn, chỉ để lại 23S rRNA, ngay cả chỉ còn lại 23S rRNA tinh sạch thì phản ứng peptidyltransferase vẫn xảy ra.

Trong quá trình sinh tổng hợp peptide và vị trí kết gắn của tRNA thay đổi thì ribosome chuyển động dọc theo phân tử mRNA với một quãng dài bằng đúng chiều dài của một mã di truyền và qua đó thêm một amino acid được bổ sung vào chuỗi peptide. Bước chuyển dịch này được xúc tác nhờ EF2-GTP ở tế bào nhân chuẩn và nhờ EF-G-GTP ở tế bào vi khuẩn. Tại hình 1.24 ta thấy rõ, sau khi cầu nối peptide được tạo thành phân tử tRNA^{Met} không còn mang methionine hoạt

hoá nữa và đã thoát ra ngoài ribosome tại vị trí E. Lúc này có một phức bộ ba mới với amino acid tiếp theo xâm nhập vào ribosome để tiến hành việc nối dài chuỗi peptide bằng amino acid tiếp theo này. Cứ như vậy, chu trình kéo dài chuỗi được thực hiện.

Sinh tổng hợp protein kết thúc khi đạt đến mã dừng

Cũng như giai đoạn khởi đầu và giai đoạn kéo dài thì giai đoạn kết thúc cũng đòi hỏi có những phân tử protein đặc hiệu cao làm tín hiệu quyết định sự dừng lại việc duy trì phức hệ mRNA-ribosome-tRNA-peptidyl. Các nghiên cứu vừa qua phát hiện thấy có hai yếu tố kết thúc (termination factor) tồn tại trong mỗi loại tế bào nhân chưa chuẩn và tế bào nhân chuẩn. Chúng được đặt tên là RF1 và RF2 đối với vi khuẩn và eRF1 và eRF2 ở tế bào nhân chuẩn. Yếu tố RF1 cũng như eRF1 có cấu trúc phân tử tương tự tRNA, có khả năng nhận biết mã dừng UAG trên mRNA, RF2 và eRF2 nhận biết UGA và cả hai RF và eRF đều nhận biết UAA. Ngoài ra, còn tìm thấy RF3 và eRF3 liên kết với GTP giàu năng lượng. RF3 tương hỗ với các RF trong việc cắt peptidyl-tRNA khỏi phức hệ mRNA-ribosome-tRNA-peptidyl giải phóng phân tử protein vừa tổng hợp xong (hình 1.25).



Hình 1.25. Kết thúc quá trình dịch mã ở tế bào nhân chuẩn.

Hình 1.25 minh họa kết thúc quá trình dịch mã ở tế bào nhân chuẩn. Khi thế ribosome mang chuỗi polypeptide mới tổng hợp tiến đến một mã

ngừng (UAA, UGA, UAG) yếu tố giải phóng eRF1 xâm nhập vào phức ribosome, gắn vị trí A cùng với eRF3-GTP. Thuỷ phân GTP dẫn đến cắt tách chuỗi polypeptide khỏi tRNA ở vị trí P giải phóng tRNA và hai tiểu phân của thể ribosome.

Việc đóng gói (gấp xoắn) phân tử protein vừa tổng hợp thành sản phẩm tự nhiên với cấu trúc không gian 3 chiều được thực hiện bởi một loại protein khác có tên là *chaperone*. Còn thêm những yếu tố giải phóng khác làm nhiệm vụ tháo dỡ ribosome, giải phóng các tiểu đơn vị, mRNA và tRNA kết thúc để chúng có thể tham gia vào chu trình sinh tổng hợp protein mới.

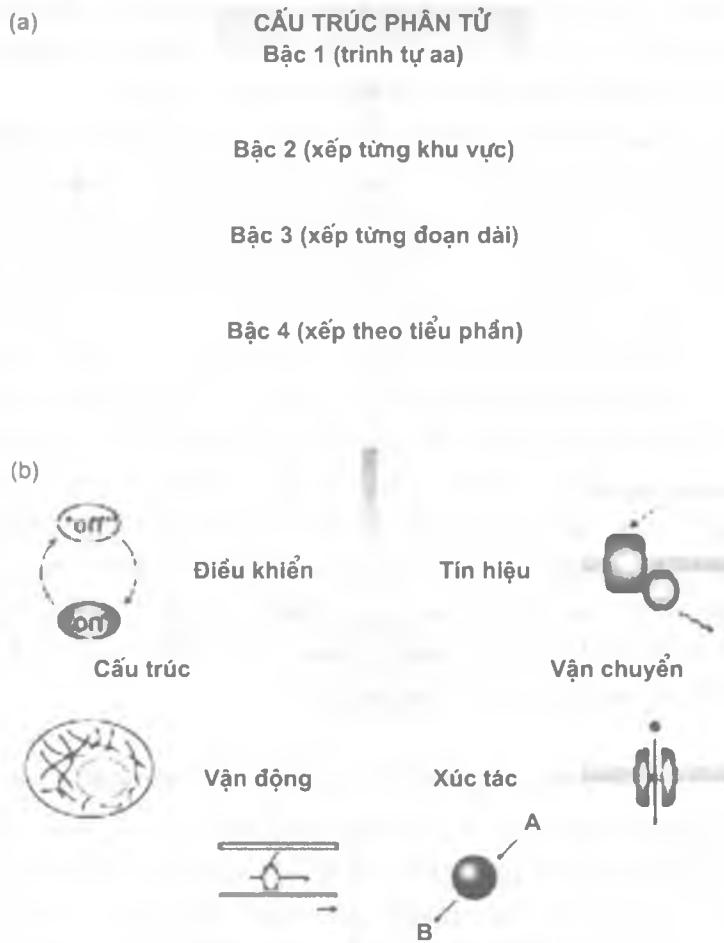
Sinh tổng hợp protein đạt hiệu quả cao nhờ hiện tượng dịch mã đồng thời trên nhiều ribosome và quay vòng tái sử dụng nhanh ribosome

Như đã trình bày, một protein có kích thước trung bình được tổng hợp trong quãng thời gian 30-60 giây. Hai hiện tượng góp phần tăng nhanh tốc độ sinh tổng hợp protein là (1) Một phân tử mRNA đồng thời được nhiều ribosome dịch mã và (2) Các thể ribosome được tái sử dụng nhanh khi kết thúc chu trình tổng hợp một protein. Hiện tượng nhiều ribosome cùng gắn với một phân tử mRNA, mang những chuỗi peptide dài dần được chứng minh bằng kính hiển vi điện tử khi nghiên cứu chất kết lỏng ly tâm. Một mRNA gắn với nhiều ribosome được gọi là *polyribosome* hay *polysome*. Còn hiện tượng tăng cường tái sử dụng các tiểu phân ribosome thì khó chứng minh hơn bằng thực nghiệm. Thông thường, khi quan sát dưới kính hiển vi điện tử các chuỗi polysome có cấu trúc hình vòng tròn, điều đó có liên quan đến việc gia tăng khả năng tái sử dụng các tiểu phân của ribosome.

Ngoài ra, các phân tử mRNA của tế bào nhân chuẩn có một chuỗi poly(A). Cấu tạo đuôi poly(A) này có chức năng gia tăng quá trình dịch mã. Trong tế bào nhân chuẩn, một protein kết gắn với đoạn poly(A) đã được phát hiện và đặt tên là *poly(A) binding protein I* (PABI). Khi bị đột biến mất đoạn poly(A), sinh tổng hợp protein ở tế bào nấm men xảy ra rất yếu. Vai trò của PABI và hai thành phần nữa của nấm men là eIF4 (4E và 4G) có khả năng tương tác với mRNA để cho phân tử kép thành vòng tròn. Thông qua việc tạo vòng phân tử mRNA bằng cách nối được poly(A) với

đầu 5' các tiểu phân ribosome có nhiều khả năng nhanh chóng trở lại tham gia vào quá trình sinh tổng hợp phân tử protein mới.

1.2.4. Xếp gấp, biến đổi và phân huỷ protein



Hình 1.26. Tổng quát cấu trúc và chức năng của protein.

Hình 1.26 minh họa tổng quát cấu trúc và chức năng của protein. (a) Trình tự thẳng của chuỗi aa (bậc một), xếp gấp thành cấu trúc xoắn lò xo hay cấu trúc tấm bản (bậc hai), sau đó chúng đóng gói thành thể hạt hay thể sợi (bậc ba). Các protein riêng rẽ có thể liên kết với nhau như những tiểu phân của một phức hệ (bậc bốn) với các tiểu đơn vị giống hay

khác nhau. (b) Protein thực hiện chức năng xúc tác các phản ứng sinh hoá học trong tế bào (enzyme), mang các phân tử nhỏ hay ion vào ra qua màng tế bào (vận chuyển), tiếp nhận và truyền các tín hiệu của ngoại cảnh vào các hệ thống cảm nhận của tế bào (cảm ứng tín hiệu), điều hoà các hoạt động (điều khiển), hình thành các cấu trúc như nhiễm sắc thể, màng tế bào, nguyên sinh chất (cấu trúc), tạo ra động lực cho các chuyển động (vận động). Các chức năng này hình thành thông qua mối tương tác và biến đổi cấu trúc của phân tử protein đã được xếp gấp.

Như đã trình bày ở các phần trên, protein được sinh tổng hợp trong bộ máy tế bào là ribosome tuân thủ theo nguyên tắc dịch mã di truyền của phân tử mRNA. Tiếp theo, sản phẩm protein này được "chế tác" thông qua xếp gấp phân tử và biến đổi – kể cả biến đổi hoá học để có được trạng thái thành thực cuối cùng. Về nguyên tắc mỗi protein có n gốc amino acid có thể được xếp gấp thành 8^n kiểu cấu trúc khác nhau. Lý do là mỗi amino acid có tới 8 trạng thái hoá lập thể trong liên kết tạo thành khung xương peptide. Tuy nhiên, trong thực tế mỗi phân tử chỉ tồn tại ở một trạng thái, được gọi là trạng thái tự nhiên (*native state*) có độ bền vững cao nhất. Có hai cơ chế hạn chế việc xếp gấp bị lỗi đối với phân tử protein. Đó là ở mức phân tử việc xếp gấp chỉ xảy ra thông qua cơ chế đơn giản không qua nhiều bước và ở mức tế bào cũng có những hệ thống thuộc tế bào ngăn cản việc xếp gấp sai. Sau khi phân tử protein hoàn thành chức năng sinh học của chúng thì chúng trở thành mục tiêu phân huỷ.

Trình tự amino acid quyết định cách thức xếp gấp của protein

Thực nghiệm trong điều kiện ống nghiệm cho phép rút ra kết luận rằng trình tự amino acid quyết định cách thức xếp gấp của phân tử protein. Các yếu tố hoá lý của môi trường như nhiệt năng khi đun nóng, độ pH cực đoan làm thay đổi điện tích của các chuỗi phụ của amino acid, hoá chất như urea hoặc guanidium hydrochloride ở nồng độ 6-8 M làm yếu các liên kết không hoá trị vốn đang là lực đảm bảo cấu trúc tự nhiên của protein. Việc thay đổi cấu trúc không gian ba chiều của phân tử protein được gọi là biến tính không những làm thay đổi cấu trúc mà còn làm mất hoạt tính của protein. Hầu hết các protein bị biến tính đều kết tủa do các liên kết kị nước của phân tử tương tác với các liên kết khác hình thành một trạng thái không tan của protein trong dung dịch.

Rất nhiều protein bị biến tính trong dung dịch 8M urea và β-mercaptoethanol (làm giảm các cầu nối disulfide) nhưng có khả năng hồi tính chất cũ khi thông qua thẩm tách các chất gây biến tính bị loại khỏi dung dịch. Trong quá trình hồi tính chất cũ, các cầu nối disulfide, cầu hydro và các liên kết kị nước được tái thiết giúp cho phân tử tái tạo và duy trì được cấu trúc tự nhiên ban đầu. Qua đó, phân tử protein lại có được khả năng hoạt động bình thường.

Thông qua các thí nghiệm biến tính và tái tính trên phân tử ribonuclease các nhà nghiên cứu đã chứng minh được tính chất xếp gấp trong cấu trúc tự nhiên của protein là do trình tự amino acid của nó quyết định.

Chaperone hỗ trợ việc xếp gấp protein trong cơ thể sinh vật

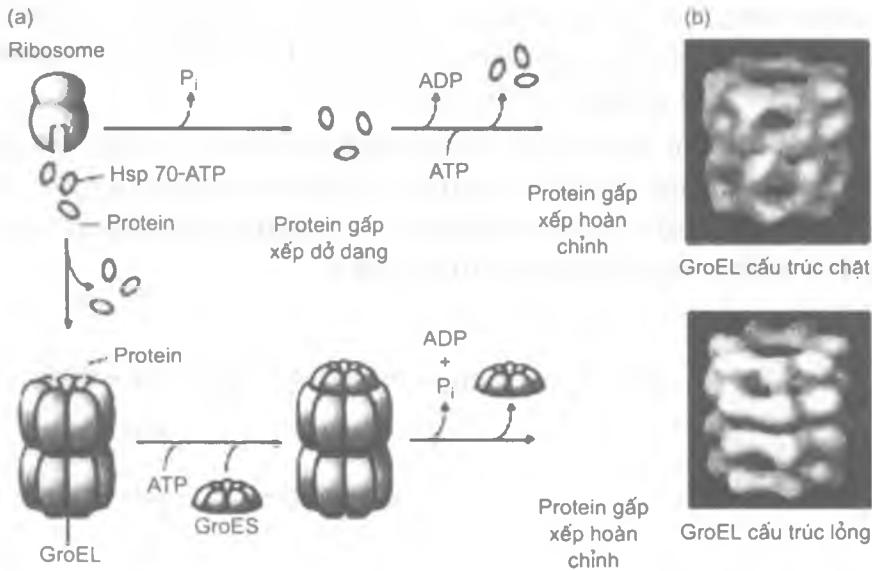
Trong tự nhiên, tế bào cần thiết lập quá trình xếp gấp protein một cách nhanh chóng nhằm giảm thiểu sự lãng phí năng lượng dùng cho sinh tổng hợp các protein và kể cả trong phân huỷ chúng nếu chúng không hoạt động được vì bị lỗi trong xếp gấp. Trên 95% protein trong tế bào đều tồn tại ở trạng thái cấu trúc tự nhiên mặc dù trong tế bào nồng độ protein được xem là rất cao (100 mg/ml), nếu trong điều kiện ống nghiệm thì nồng độ này thường gây nên hiện tượng protein bị kết tủa vì biến tính. Nguyên nhân giúp cho tế bào sinh vật có thể duy trì được tình trạng tự nhiên của protein một cách hiệu quả như vậy là vì tế bào có loại protein có chức năng xếp gấp protein, tên gọi là chaperone. Người ta đã thấy ở mọi loại sinh vật từ vi khuẩn đến con người, chaperone có mặt trong mọi cấu trúc và khoang của tế bào, liên kết với một vài loại protein khác và có thể đóng vai trò quyết định trong cơ chế xếp gấp phân tử protein. Có hai họ chaperone : (1) *chaperone phân tử* liên kết với protein đã hoặc chưa xếp gấp nhằm làm cho chúng bền vững và không bị phân huỷ và (2) *chaperonin* làm nhiệm vụ xếp gấp protein. Chaperone có hoạt tính ATPase và hoạt tính bảo vệ đối với các protein khác chủ yếu phụ thuộc vào năng lượng thuỷ phân ATP.

Nhóm chaperone phân tử có họ Hsp70, trong đó có Hsp70 trong lưới màng của tế bào và của ty thể, có Bip trong lưới nội màng và DnaK trong vi khuẩn. Hsp70 xuất hiện rất nhanh khi tế bào bị gây sốc nóng và tìm thấy ở mọi loại sinh vật. Khi liên kết với ATP, phân tử Hsp70 tồn tại ở

trạng thái mở tạo ra một khoang kín nước cho phân tử protein chưa xếp gấp gắn vào. Khi thuỷ phân ATP, phân tử Hsp70 chuyển về trạng thái đóng và phân tử protein đích được giải phóng (hình 1.27). Các chaperone phân tử thường liên kết với mọi phân tử protein vừa mới được tổng hợp trong ribosome. Khi được giải phóng khỏi chaperone thì có đến 85% protein vi sinh hình thành được xếp gấp tự nhiên. Ở tế bào nhân chuẩn, tỷ lệ này còn cao hơn.

Để xếp gấp được các protein làm nhiệm vụ khung xương cho tế bào như là actin hay tubulin cần phải có sự tham gia của chaperonin. Chaperonin tế bào nhân chuẩn, được gọi là TCiP, là một phân tử lớn hình trụ tròn với 8 tiểu phân Hsp60. Loại chaperonin tương tự của tế bào vi khuẩn có tới 14 tiểu đơn vị giống nhau, được gọi là GroEL. Cho đến nay cơ chế hoạt động của GroEL được nghiên cứu kỹ hơn là của TCiP. Thông thường, các protein cần xếp gấp được đưa vào trong lòng ống trụ, liên kết với thành bên trong của phân tử chaperonin và xếp gấp thành cấu trúc tự nhiên cần thiết. Bước tiếp theo cần có ATP, nhờ năng lượng này ống GroEL phình rộng ra và phân tử protein mới xếp gấp thoát ra ngoài. Bước này cần thêm sự hỗ trợ của một protein co-chaperonin tên là GroES làm nhiệm vụ như chiếc nắp đậy ống GroEL. Vì trên phân tử chaperonin tế bào nhân chuẩn TCiP không có nắp đậy GroES nên chắc chắn là có sự khác biệt ở bước này trong tế bào nhân chuẩn. Hơn thế nữa, cho đến lúc này mới thấy TCiP của tế bào nhân chuẩn có lòng ống chỉ đủ cho phân tử có kích thước nhỏ hơn 55 KDa vào được. Rõ ràng là cần nghiên cứu sâu hơn về quá trình xếp gấp protein ở tế bào nhân chuẩn.

Hình 1.27 mô tả xếp gấp protein thông qua chaperone và chaperonin. (a) Nhiều protein xếp gấp cấu trúc 3D của chúng thông qua sự hỗ trợ của phân tử protein tương tự Hsp70. Loại protein tạo khuôn phân tử này liên kết tạm thời với chuỗi polypeptide mới hình thành vừa tách ra từ một thể ribosome. (b) Cách thức xếp gấp thích hợp đối với một số protein khác ở tế bào nhân chưa chuẩn, phụ thuộc vào chaperonin như GroEL, một loại phân tử rỗng hình trụ gồm 14 tiểu phân hình vòng nhẫn, giống nhau, khối lượng 60000 chồng lên nhau.

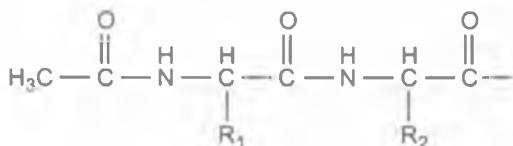


Hình 1.27. Xếp gấp protein thông qua chaperone và chaperonin.

Hoạt tính sinh học của protein thay đổi sau quá trình biến đổi hoá học và chế tác

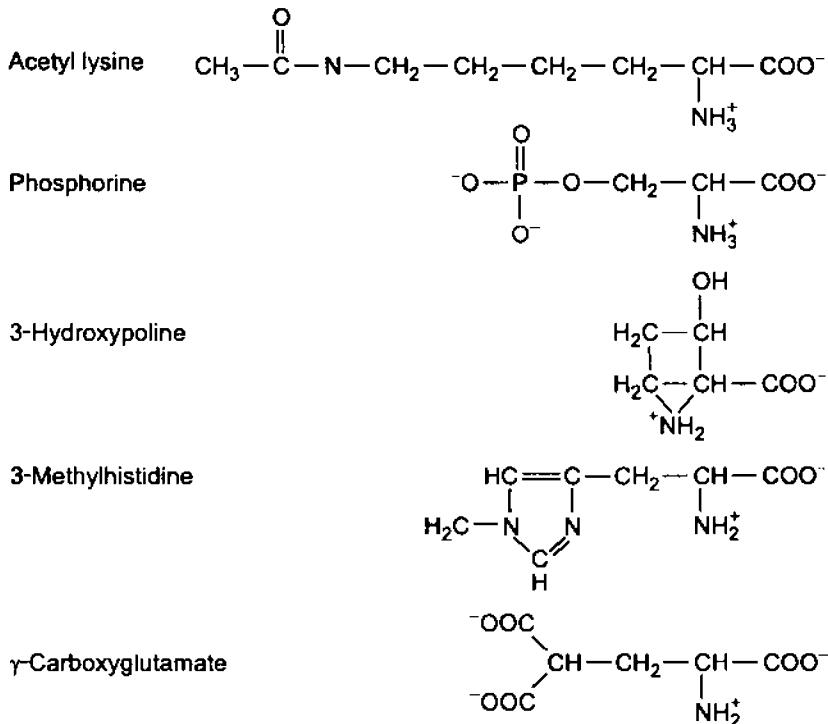
Hầu như tất cả các phân tử protein sau khi được tổng hợp trong ribosome đều trải qua những biến đổi hóa học, qua đó tuỳ theo cách thức biến đổi mà hoàn thiện về hoạt tính sinh học, thay đổi thời gian tồn tại hoặc vị trí lưu giữ trong tế bào. Biến đổi protein được chia thành hai dạng thức : (1) Biến đổi hóa học và (2) Chế tác. Biến đổi hóa học bao gồm liên kết một nhóm chức hóa học vào gốc amine kết thúc hay gốc carboxyl hay tái hoạt hóa một gốc hóa học của các chuỗi nhánh thuộc một gốc amino acid bên trong phân tử protein. Các biến đổi này thường ở trạng thái thuận nghịch. Chế tác bao gồm việc cắt bỏ một phần chuỗi peptide và nói chung là không thuận nghịch.

Acetyl hóa là việc gắn một gốc acetyl ($\text{CH}_3\text{CO}-$) vào nhóm chức amine của gốc amino acid cuối thuộc đầu kết N rất phổ biến đối với gần 80% các chuỗi protein :



Hiện tượng acetyl hoá này đóng vai trò quan trọng trong việc duy trì tuổi thọ của các protein, nếu không được acetyl hoá các phân tử protein rất nhanh chóng bị protease nội bào phân huỷ.

Đối với protein tham gia thiết lập màng tế bào thì các gốc gắn đuôi chuỗi protein thường được kết gắn thêm các gốc dài tương tự lipid. Việc gắn đuôi kỵ nước này có chức năng neo các protein vào lớp lipid, giữ cho tế bào có những protein nhất định trong màng.



Hình 1.28. Chuyển đổi thông thường xảy ra đối với các gốc amino acid trong phân tử protein. Nổi bật nhất là việc bổ sung những nhóm chức (in đậm) vào các gốc amino acid trong chuỗi polypeptide sau khi tổng hợp.

Các gốc bên trong chuỗi protein có thể bị biến đổi thông qua gắn với các nhóm chức hóa học rất khác nhau vào các nhánh phụ của chúng. Một biến đổi rất quan trọng là quá trình phosphoryl hoá các gốc serine, threonine và tyrosine. Có nhiều ví dụ về hoạt tính của protein thay đổi thế nào khi chúng bị phosphoryl hoá hay bị khử phospho (dephosphoryl hoá). Mạch nhánh của asparagine, serine và threonine lại là vị trí thường xảy ra quá trình glycosyl hoá với những chuỗi carbohydrate mạch thẳng hay

mạch nhánh. Thường những protein bị glycosyl hoá có mặt trong các dịch tiết của tế bào và mô. Hình 1.28 trình bày thêm những loại biến đổi hoá học khác đối với protein đã được tìm thấy.

Không giống như quá trình biến đổi hoá học xảy ra theo nguyên tắc thuận nghịch, quá trình chế tác protein xảy ra theo nguyên tắc không thuận nghịch. Thường là có hiện tượng bẻ một cầu nối peptide ở đầu C hay đuôi N thông qua hoạt động của một protease. Cắt thông qua hoạt tính phân giải protein kiểu này có thể liên quan đến hoạt hoá hoặc bắt hoạt phân tử protein, đặc biệt là các enzyme hoạt động trong quá trình đông và giải đông máu. Quá trình cắt phân tử cũng có thể liên quan đến việc hình thành các hormone dạng peptide, như EGF hoặc insulin tạo ra từ một polypeptide tiền chất có kích thước lớn hơn.

Một kiểu chế tác bất thường có tên gọi là *tự cắt gọt* (self splicing). Hiện tượng này quan sát thấy ở tế bào vi khuẩn và tế bào nhân chuẩn dạng sơ khai. Cách thức tự cắt gọt có thể so sánh với các nhà biên tập phim, cắt bỏ một đoạn giữa chuỗi polypeptide và dính nối hai đầu với nhau để tạo ra phân tử mới. Khác với hiện tượng chế tác, quá trình tự cắt gọt này xảy ra theo nguyên lý tự xúc tác, nghĩa là không có enzyme khác tham gia. Đoạn peptide thừa tự loại khỏi chuỗi polypeptide giống như quá trình chế tác phân tử RNA. Hiện tượng này cũng được quan sát thấy ở một số loài động vật không xương sống khi chúng cần những peptide cho một số bước trong quá trình phát triển.

Tế bào phân huỷ protein bằng nhiều phương cách

Tế bào có cả phương cách nội bào và ngoại bào để phân huỷ protein. Cách thức ngoại bào chủ yếu dựa trên hoạt động của protease phân huỷ, chúng cắt protein thành các đoạn polypeptide trong dịch đường tiêu hoá. Tham gia vào đây có các loại endoprotease như trypsin và chymotrypsin, chúng bẻ khung protein thành các gốc kiềm và gốc thơm, kế đến là các exopeptidase với hoạt tính cắt bỏ các gốc amino acid ở đầu N (aminopeptidase) và gốc ở đầu C (carboxypeptidase) và cuối cùng là các peptidase hoạt động cắt các oligopeptide thành các dipeptide và tripeptide và các amino acid riêng rẽ. Các phân tử amino acid nhỏ này được hấp thụ qua màng ruột và đi vào máu.

Tuổi thọ của protein nội bào có thể ngắn như trường hợp của cyclin trong phân bào nguyên nhiễm là một vài phút nhưng cũng dài bằng cả

tuổi thọ của sinh vật như trường hợp protein của thuỷ tinh thể mắt. Tế bào cũng có một số loại protease nội bào để phân huỷ các protein xếp gấp bị lỗi, protein biến tính, protein cần giảm nồng độ hay protein do tế bào tiếp nhận. Phương cách phân huỷ nội bào thường do các enzyme trong lysosome hoặc cơ quan tử có màng bao bọc với nội chất mang tính acid. Một cách thức tiêu huỷ protein đặc biệt thông qua quá trình ubiquitin hoá : trước hết phân tử protein được gắn vào chuỗi bên của gốc lysine, sau đó phân tử protein đã bị ubiquitin hoá bị tiêu huỷ nhờ một thể proteasome có cấu trúc đa tiểu phân hình trụ to, cần có ATP cung cấp năng lượng và kết quả thu được là các peptide và phân tử ubiquitin nguyên vẹn.

Chỉ có những protein được nhận biết bởi enzyme ubiquitin hoá mới có thể bị phân huỷ theo cách thức này. Tuy nhiên, có nhiều loại enzyme khác có khả năng gắn protein đích với các chất phụ trợ khác, ví dụ như trình tự Arg-X-X-Leu-Gly-X-Ileu-Gly-Asx của cyclin phân bào nguyên nhiễm được enzyme ubiquitin hoá E1 nhận biết. Những chuỗi protein giàu proline, glutamic acid, serine và threonine của PEST lại được nhận biết bởi một loại enzyme ubiquitin hoá khác. Thông thường tuổi thọ của protein nội bào thường bị quyết định bởi đặc tính của gốc nằm ở đầu N với khả năng bị ubiquitin hoá nhanh, chẳng hạn các protein chỉ tồn tại trong 3 phút nếu ở đầu N chứa Arg, Lys, Phe, Leu hoặc Trp. Ngược lại những protein chứa Cys, Ala, Ser, Thr, Gly, Val hoặc Met ở đầu N có thể tồn tại lâu tới 30 giờ. Như đã trình bày ở phần trên, hầu hết các protein đều có Met ở đầu nhưng trong quá trình thành thực phân tử protein sau tổng hợp có những biến đổi như gắn thêm một amino acid làm giảm độ bền của đầu N khiến cho protein này sớm trở thành mục tiêu phân huỷ.

Protein xếp gấp bất thường là nguyên nhân phát sinh bệnh tật

Như đã trình bày ở trên, mỗi loại protein thường được xếp gấp thành một cấu trúc tối ưu về năng lượng dựa theo trình tự amino acid của nó. Bằng chứng mới đây cho thấy protein cũng có thể được xếp gấp thành một cấu trúc không gian ba chiều khác mà chưa rõ vì lý do gì. Hiện tượng xếp gấp lỗi như thế không những làm cho protein mất hoạt tính, mà còn là mục tiêu phân huỷ. Kết quả tích tụ những phân đoạn peptide chưa phân huỷ hết dẫn đến phát sinh các loại bệnh thoái hoá nhất định với đặc trưng

là xuất hiện những vết protein không tan trong các cơ quan khác nhau như gan và não.

Bệnh Alzheimer là ví dụ điển hình của sự xuất hiện các mảng và tóm sợi rối protein trong vùng não hư hỏng. Các sợi protein xuất hiện từ những sản phẩm phân huỷ protein tự nhiên vốn là tiền chất của amyloid – một protein có khả năng vận chuyển qua màng và một loại protein liên kết các bô vi ống (Tau). Mảnh protein tìm thấy trong cơ quan khác là các phân đoạn thuỷ phân protein tự nhiên như gelsolin, một loại protein liên kết, actin hay albumin huyết thanh. Các phân đoạn protein được giải phóng ra trùng hợp với nhau thành một búi sợi rất bền. Quá trình thoái hoá não như trong căn bệnh Alzheimer gây ra bởi các prion, một loại protein có khả năng truyền nhiễm do protein não phân huỷ tạo ra.

1.2.5. Chức năng xúc tác sinh học của enzyme

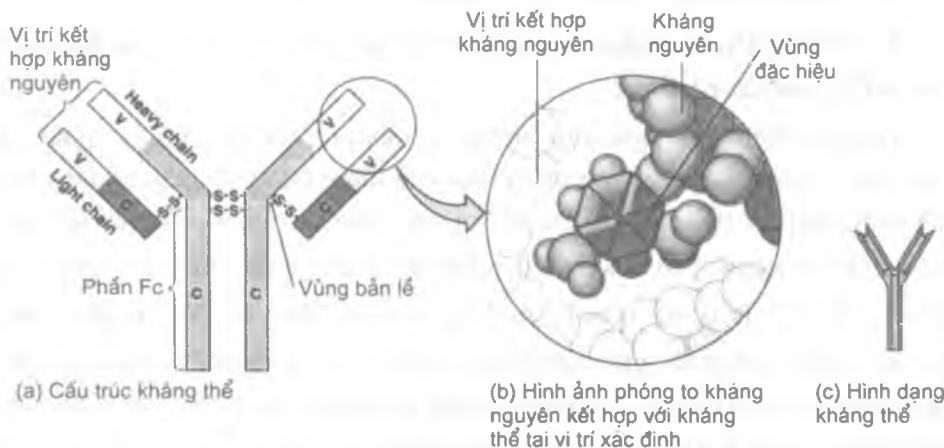
Sự tương thích phân tử quyết định tính đặc hiệu và ái lực của liên kết protein-phổi tử

Hai đặc điểm đánh giá mối tương tác của protein với phổi tử (phân tử liên kết) : *Tính đặc hiệu* liên quan đến khả năng của một protein liên kết với một phân tử này nhiều hơn so với một phân tử khác. Còn *ái lực* thể hiện mức độ mạnh yếu của mối liên kết đó. Hằng số K_d của một phức hệ protein-phổi tử là tỷ số nghịch của hằng số cân bằng K_{eq} trong phản ứng liên kết. Mỗi tương tác giữa protein và phổi tử càng mạnh thì hằng số K_d càng nhỏ. Cả hai đặc tính – tính đặc hiệu và ái lực của một protein đối với một phổi tử hoàn toàn phụ thuộc vào cấu trúc của vị trí liên kết trong phân tử protein dành cho phổi tử được thiết kế như một hốc trong khuôn đúc. Đối với những tương tác có tính đặc hiệu cao và ái lực cao thì hình dáng và tính chất hoá học của bề mặt hốc liên kết thường rất tương thích với phổi tử. Đặc điểm này được gọi là *tính tương thích phân tử*.

Khả năng protein có thể phân biệt các loại phân tử khác nhau được phát triển rất cao ở trong nhóm protein có tên gọi là *kháng thể* (antibody), là những protein được sản sinh ra khi phản ứng lại các tác nhân gây nhiễm – *kháng nguyên* (antigen) như vi khuẩn hay virus hoặc một vài chất lạ như protein hay polysaccharide của phấn hoa, lông thú, bụi thải trại nhà. Sự xuất hiện của kháng nguyên gây nên hiện tượng cơ thể tiếp nhận tạo ra một lượng lớn protein kháng thể có khả năng kết gắn với

kháng nguyên vào những vùng khác nhau, gọi là các epitope, *vùng đặc hiệu*. Có thể mô tả kháng thể là đầu dò đối với kháng nguyên tạo nên phức chất kháng thể – kháng nguyên khởi đầu phản ứng bảo vệ của hệ thống miễn dịch trong tế bào.

Tất cả các phân tử kháng thể đều có hình chữ Y ghép bởi hai chuỗi nặng giống nhau và hai chuỗi nhẹ cũng giống hệt nhau (hình 1.29a). Mỗi cánh tay của phân tử kháng thể có một chuỗi nhẹ liên kết với chuỗi nặng thông qua cầu nối disulfide. Ở đầu cuối mỗi cánh tay có 6 vòng khuyên có mức độ biến đổi cao các amino acid. Chính các vòng khuyên này làm chức năng của vị trí liên kết với *vùng xác định tính tương thích* (complementary-determining region, CDR). Sự biến đổi trình tự các amino acid trong các vòng khuyên là cơ sở để các phân tử kháng thể có thể nhận biết một cách đặc hiệu các kháng nguyên khác nhau.

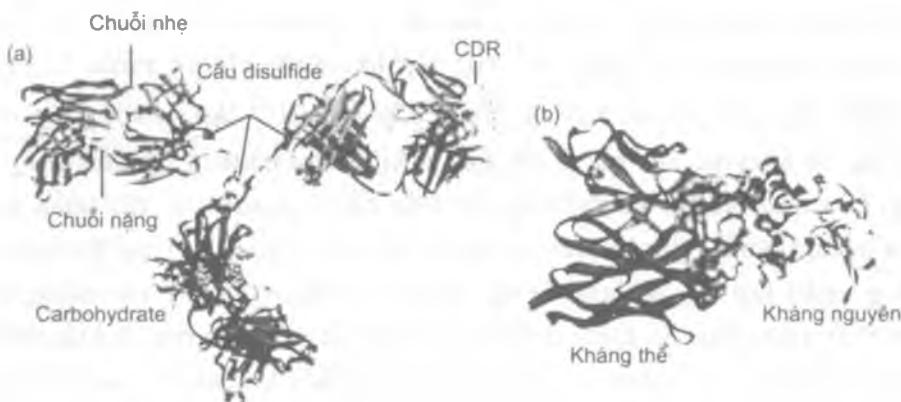


Hình 1.29. Cấu trúc của kháng thể và vùng xác định tính tương thích.

Mỗi tương tác giữa kháng thể và vùng epitope của kháng nguyên dựa trên tính tương thích phân tử. Bề mặt của vị trí liên kết kháng nguyên của kháng thể phù hợp hoàn toàn với vùng epitope tương đương như chiếc găng hợp với bàn tay (hình 1.29b). Khi giữa hai bề mặt tiếp xúc với nhau, có hàng loạt liên kết không hoá trị được thiết lập làm cơ sở cho mối liên kết đặc thù giữa kháng thể và kháng nguyên.

Tính đặc hiệu của kháng thể chuẩn xác tới mức chúng có thể phân biệt được tế bào của từng cá thể trong một loài và trong một vài trường hợp có thể phân biệt được protein khác biệt nhau chỉ một amino acid.

Dựa vào đặc tính chuyên biệt cao mà kháng thể được sử dụng làm nguyên liệu cho rất nhiều nghiên cứu quan trọng, ví dụ như để tinh chế những loại protein nhất định.



Hình 1.30. Cấu trúc kháng thể và tương tác kháng thể – kháng nguyên.

Hình 1.30 mô tả cấu trúc kháng thể và tương tác kháng thể – kháng nguyên. (a) Mô hình rubăng của một phân tử kháng thể : Mỗi phân tử kháng thể bao gồm 2 chuỗi nặng giống hệt nhau liên kết với nhau thông qua cầu nối disulfide và 2 chuỗi nhẹ cũng giống hệt nhau liên kết với nhau qua cầu nối disulfide. (b) Cách tích hợp giữa kháng thể và đoạn epitope của kháng nguyên (ví dụ lysozyme lòng trắng trứng gà). Hai phân tử tiếp xúc nhau được thông qua vùng quyết định bổ sung (CDR).

Enzyme : chất xúc tác đặc hiệu và hiệu quả cao

Khác với kháng thể chỉ đơn giản gắn với phân tử tương tác và trình diện chúng trong hệ thống miễn dịch, các enzyme kích thích phản ứng hóa học của các phân tử tương tác – lúc này được gọi là *các cơ chất*. Hầu như mọi phản ứng hóa học xảy ra trong tế bào đều được xúc tác bởi một loại enzyme chuyên biệt. Cũng như mọi chất xúc tác, enzyme không quyết định quy mô phản ứng và không tham gia vào sản phẩm. Enzyme làm gia tăng tốc độ phản ứng thông qua làm giảm năng lượng hoạt hóa. Trong ống nghiệm các chất xúc tác hóa học như than hoạt tính hay bạch kim chỉ kích hoạt phản ứng dưới điều kiện nhiệt độ và áp suất đặc biệt cao hoặc độ pH thấp hoặc trong dung môi hữu cơ. Trong khi đó các

enzyme xúc tác phản ứng trong tế bào ở môi trường nước, nhiệt độ 37°C, áp suất 1 atm và pH từ 6,5 đến 7,5.

Hai tính chất rất riêng khiến cho enzyme có khả năng xúc tác ở điều kiện bình thường trong tế bào là : *Lực xúc tác* cực lớn và *tính chuyên biệt* rất cao. Năng lực xúc tác lớn được chứng minh là phản ứng do enzyme xúc tác xảy ra nhanh gấp 10^6 - 10^{12} lần so với phản ứng không có enzyme xúc tác ở cùng điều kiện như nhau. Còn tính chuyên biệt của enzyme thì có thể thấy ở ví dụ các enzyme xúc tác phản ứng của amino acid. Như ta biết amino acid có thể tồn tại ở hai trạng thái đồng phân lập thể là dạng D và dạng L nhưng trong tế bào sống chỉ tìm thấy trạng thái L. Điều đó được giải thích bởi enzyme xúc tác phản ứng của amino acid dạng L nhanh hơn rất nhiều lần so với dạng D mặc dù cả hai dạng đồng phân lập thể của một amino acid đều có kích thước và có cùng gốc R như nhau.

Đến nay, trong ngân hàng dữ liệu enzyme đã ghi nhận, sắp xếp và phân loại được 3700 loại enzyme khác nhau, mỗi loại xúc tác một phản ứng hóa học duy nhất hoặc một nhóm phản ứng giống nhau. Có những enzyme nhất định được tìm thấy ở hầu hết các loại tế bào vì chúng xúc tác phản ứng của các sản phẩm thông dụng của tế bào (như protein, nucleic acid và phospholipid) hoặc tham gia vào việc sản sinh ra năng lượng thông qua biến đổi đường glucose và oxy thành carbon dioxide và nước. Những enzyme khác lại chỉ thấy ở những loại tế bào nhất định bởi chúng làm nhiệm vụ xúc tác phản ứng hóa học chỉ xảy ra ở riêng loại tế bào này (như enzyme chuyển hoá tyrosine thành dopamine làm nhiệm vụ chuyển tải tín hiệu trong tế bào thần kinh). Mặc dù hầu hết enzyme tồn tại trong tế bào, cũng có một vài enzyme được tiết ra ngoài và hoạt động ở những vùng ngoại bào như máu, khoang tiêu hoá, đôi khi kể cả bên ngoài cơ thể.

Hoạt tính xúc tác của một số enzyme lại không thuộc phạm vi sinh tổng hợp hoặc phân giải các phân tử. Chẳng hạn như những protein điều hòa hoặc enzyme tín hiệu nội bào xúc tác phản ứng phosphoryl hoá các protein khác hoặc một số protein vận chuyển xúc tác quá trình phân ATP gắn với việc vận chuyển các phân tử qua màng tế bào.

Trung tâm hoạt động của enzyme thực hiện chức năng liên kết cơ chất và xúc tác phản ứng

Một số gốc amino acid của chuỗi enzyme protein có vai trò rất quan trọng trong việc tạo nên tính đặc hiệu và khả năng xúc tác. Trong cấu trúc tự nhiên, các gốc R này được dồn về một vùng và tạo nên trung tâm hoạt động của phân tử enzyme. Mỗi trung tâm hoạt động có hai vùng chức năng quan trọng : Một vùng nhận biết cơ chất, liên kết với cơ chất và một vùng thực hiện việc xúc tác phản ứng biến đổi cơ chất. Ở một số enzyme thì vùng xúc tác có thể là một phần của vùng liên kết, ở những enzyme khác thì hai vùng có cấu trúc hoàn toàn phân biệt nhau.

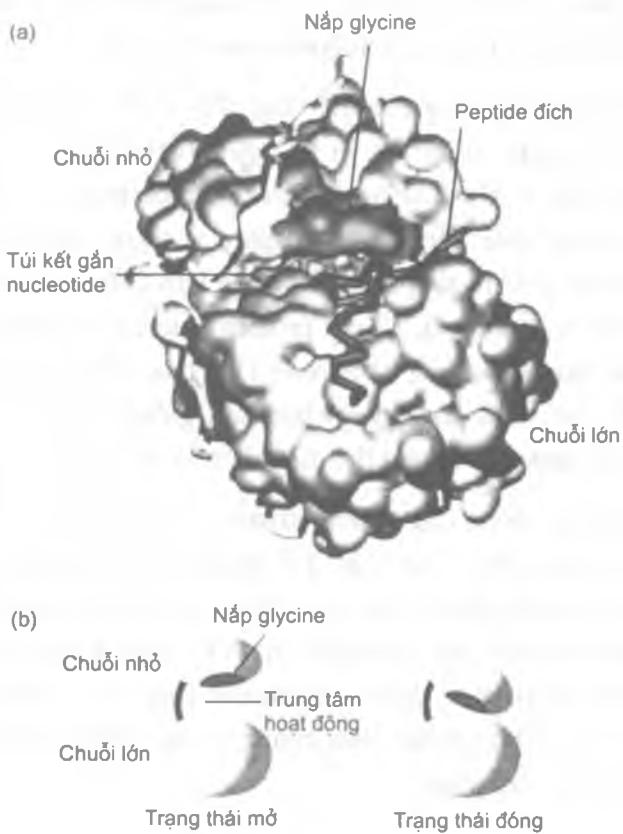
Để minh họa chức năng của hai vùng thuộc trung tâm hoạt động của một enzyme chúng ta có thể lấy ví dụ của protein kinase A một enzyme hoạt động phụ thuộc vào ATP vòng (cATP). Protein kinase này cùng với các protein kinase khác gắn thêm phosphor vào gốc serine, threonine và tyrosine có trong nhiều loại protein nội bào làm chức năng phản ứng đối với các tín hiệu từ bên ngoài. Vì các protein kinase của tế bào nhân chuẩn thuộc một họ lớn các kinase nên trung tâm hoạt động của chúng và cơ chế thực hiện quá trình phosphoryl hoá khá giống nhau. Vì vậy protein kinase A có thể được coi là một đại diện điển hình của họ enzyme này.

Trung tâm hoạt động của protein kinase A nằm ở vùng 240 gốc thuộc lõi kinase của tiểu đơn vị xúc tác. Lõi kinase khá bảo thủ ở tất cả các kinase, có chức năng liên kết với cơ chất (gồm ATP và một đoạn peptide đích) và sau đó chuyển gốc phosphate từ ATP sang serine, threonine hay tyrosine của chuỗi peptide đích. Lõi kinase gồm một chuỗi lớn và một chuỗi nhỏ, với một khe nứt sâu giữa chúng khiến cho trung tâm hoạt động có các gốc nằm cả ở hai vùng.

Hoạt động liên kết cơ chất của protein kinase

Cấu trúc của điểm kết gắn ATP của lõi xúc tác kinase tương thích với cấu trúc của cơ chất dạng nucleotide như ATP. Vòng adenine của ATP gắn vừa vặn vào khe nứt giữa chuỗi lớn và chuỗi nhỏ. Có một trình tự rất bảo thủ gồm Gly-X-Gly-X-Gly-X-Val (X là một amino acid bất kỳ) giữ cho vòng adenine ở vị trí ổn định (hình 1.31).

Ở trạng thái mở, hai chuỗi lớn và chuỗi nhỏ của vùng lõi mở đủ rộng cho phân tử cơ chất ATP lọt vào và kết gắn. Khi trung tâm hoạt động bị cơ chất chiếm xong thì hai chuỗi cùng chuyển động về vị trí đóng. Kiểu thay đổi cấu trúc bậc ba – một ví dụ điển hình của sự vừa vặn do cảm ứng (*induced fit*) giúp cho đoạn peptide đích tiếp cận đủ gần để tiếp nhận gốc phosphate từ ATP. Sau khi phản ứng phosphoryl hoá hoàn thành, sản phẩm tác động lên hai chuỗi lớn và nhỏ để chúng quay về vị trí mở, khiến cho sản phẩm được giải phóng ra.



Hình 1.31. Protein kinase A và biến đổi cấu trúc phân tử khi liên kết với cơ chất.

Hình 1.31 mô tả protein kinase A và biến đổi cấu trúc phân tử khi liên kết với cơ chất. (a) Mô hình tiểu phân xúc tác của protein kinase A gắn với cơ chất ; Vùng bảo thủ của lõi kinase được thể hiện trên bề mặt

phân tử. Một trình tự giàu glycine bẫy ATP vào một khe sâu giữa tiểu phân lớn và tiểu phân nhỏ. Chuỗi lớn liên kết với peptide đích. Cấu trúc của protein kinase khá bảo thủ trong các loại kinase của sinh vật nhân chuẩn. (b) Sơ đồ trạng thái mở và đóng của cấu trúc kinase. Khi không có cơ chất, cấu trúc kinase ở trạng thái mở. Sự gắn kết với cơ chất khiến cho tiểu phân lớn và nhỏ xoay để cho vị trí liên kết peptide gần với vị trí gắn kết ATP khiến nắp glycine trùm lên vị trí adenine của ATP, đưa ATP vào khe giữa hai tiểu phân.

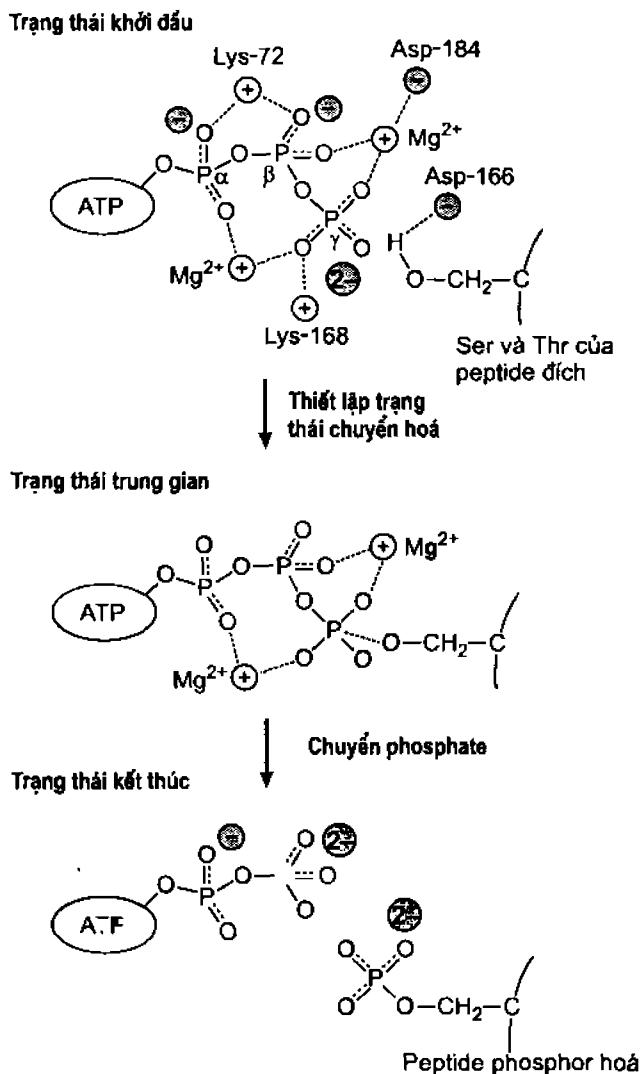
Hiện tượng xoay từ vị trí mở sang vị trí đóng đồng thời gây nên chuyển động của nắp Gly lên khe gắn ATP. Nắp Gly kiểm soát sự xâm nhập của ATP và sự thoát ra ngoài của ADP tại trung tâm hoạt động. ATP có thể xâm nhập và liên kết với rãnh của trung tâm hoạt động, tại vị trí đóng nắp Gly ngăn cản không cho ATP thoát ra khỏi rãnh trung tâm hoạt động. Sau khi phosphate được chuyển từ ATP liên kết sang peptide liên kết, nắp Gly phải xoay về vị trí mở trước khi ADP có thể thoát ra. Độ đặc động học tốc độ thoát ra của ADP chậm hơn tốc độ phosphoryl hoá 20 lần, cho thấy ảnh hưởng của nắp Gly lên tốc độ của phản ứng kinase. Đột biến tại nắp Gly cho thấy tính linh hoạt của nắp Gly bị ức chế và tốc độ xúc tác của protein kinase A giảm đáng kể.

Hoạt động chuyển gốc phosphor thông qua protein kinase

Sau khi cơ chất được kết gắn, phân tử protein kinase A giữ cấu trúc ở trạng thái đóng để cho quá trình phosphoryl hoá gốc serine hoặc gốc threonine hoặc gốc tyrosine của đoạn peptide đích (hình 1.32). Như mọi phản ứng hóa học nhóm phosphate cần chuyển và nhóm hydroxyl tiếp nhận được dồn về một vùng. Quá trình gắn kết, bền vững hoá của sản phẩm trung gian được hoạt hoá bởi protein kinase A (hình 1.31). Sự hình thành sản phẩm gây nên cảm ứng để phân tử enzym trở lại cấu trúc mở, cho phép ADP và phân tử peptide đã phosphoryl hoá thoát ra.

Hình 1.32 trình bày cơ chế phosphoryl hoá thông qua protein kinase. Khởi đầu ATP và protein đích được gắn vào trung tâm hoạt động của protein kinase. Điện tử của gốc phosphate chuyển sang Lys và Mg^{2+} . Vòng trong in màu chỉ những gốc aa của lõi phân tử kinase có chức năng gắn cơ chất và chuyển gốc phosphor. Ở trạng thái trung gian liên kết mới được tạo ra giữa serine hoặc threonine tại vị trí nguyên tử oxy và gốc phosphate γ , tạo ra sản phẩm trung gian hoá trị năm. Ở trạng thái cuối,

liên kết ester giữa phosphate α và γ bị bẻ gãy, hình thành ADP và phân tử peptide được phosphoryl hoá tại gốc serine hay threonine. Cơ chế phosphoryl hoá của các loại protein kinase khác cũng xảy ra tương tự.

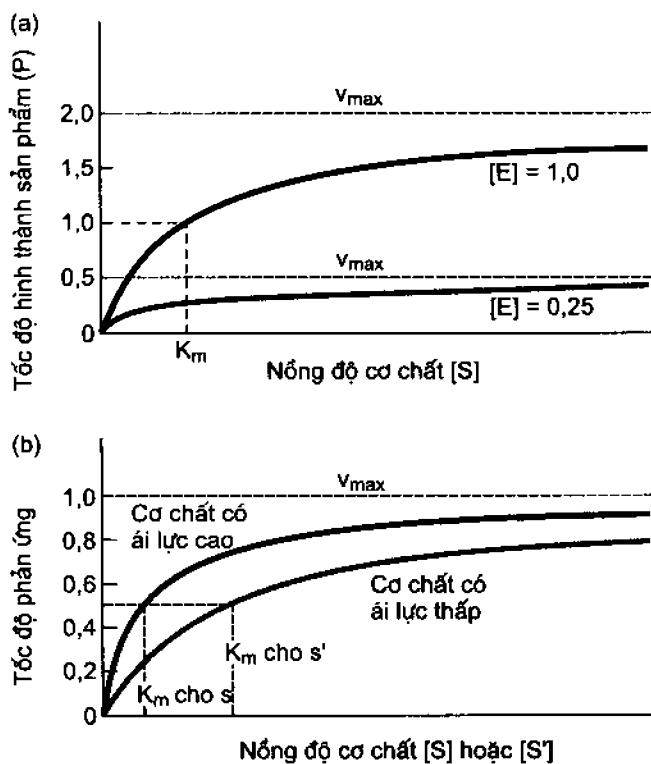


Hình 1.32. Cơ chế phosphoryl hoá thông qua protein kinase.

V_{max} và K_m biểu thị đặc điểm của phản ứng enzyme

Tác động xúc tác của một enzyme lên một cơ chất nhất định được mô tả thông qua hai chỉ số : V_{max} là tốc độ tối đa của phản ứng khi cơ chất

đạt nồng độ bão hòa và K_m (hằng số Michaelis) biểu thị ái lực của một enzyme đối với cơ chất của nó (hình 1.33). K_m được định nghĩa là nồng độ của cơ chất mà ở đó tốc độ phản ứng đạt được bằng đúng 1/2 giá trị cực đại ($1/2 V_{max}$). Hằng số K_m càng nhỏ thì enzyme càng có khả năng liên kết với cơ chất từ trong dung dịch và nồng độ cơ chất càng thấp mà phản ứng vẫn đạt được tốc độ bán cực đại.



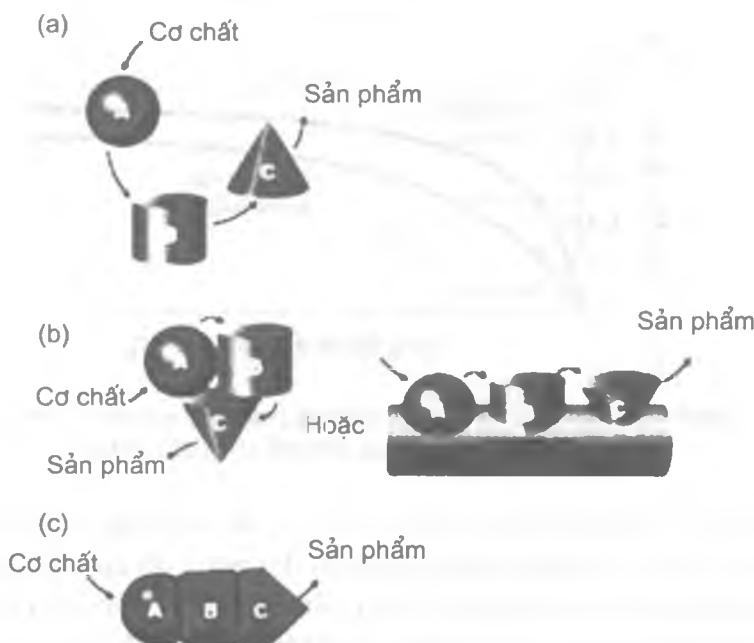
Hình 1.33. Hằng số K_m và V_{max} của một phản ứng xúc tác bởi enzyme được xác định thông qua biến đổi nồng độ cơ chất.

Hình 1.33 mô tả hằng số K_m và V_{max} của một phản ứng xúc tác bởi enzyme được xác định thông qua biến đổi nồng độ cơ chất. Hình dạng đồ thị động học biểu thị đặc điểm của phản ứng xúc tác enzyme đơn giản chuyển hóa cơ chất S thành sản phẩm P. Tốc độ phản ứng được đo trực tiếp sau khi đưa enzyme vào dung dịch cơ chất trước khi nồng độ cơ chất thay đổi một cách rõ rệt. (a) Biểu đồ tốc độ ban đầu của hai nồng độ enzyme $[E]$ khác nhau như một hàm số của nồng độ cơ chất $[S]$.

Nồng độ [S] cho thu được tốc độ phản ứng bằng $1/2$ tốc độ cực đại được gọi là hằng số Michaelis (K_m), một đại lượng nói lên ái lực giữa enzyme E và cơ chất S. Tăng gấp đôi nồng độ enzyme làm cho tốc độ phản ứng cũng tăng tương ứng, tức là tốc độ tối đa cũng tăng gấp đôi, nhưng hằng số K_m vẫn giữ không đổi. (b) Biểu đồ của tốc độ ban đầu ở phản ứng với một cơ chất có ái lực cao S với enzyme và một cơ chất với ái lực thấp S' với enzyme. Ở đây tốc độ tối đa giữ không đổi, nhưng hằng số K_m của S cao hơn của S'.

Nồng độ của những phân tử nhỏ trong một tế bào có sự khác nhau rất lớn, vì thế hằng số K_m của các enzyme tác động lên chúng cũng rất khác nhau. Nhìn chung, nồng độ nội bào của cơ chất thường lớn hơn trị giá K_m của enzyme tác động lên chúng.

Các enzyme của cùng một quá trình trao đổi chất thường liên kết vật lý với nhau



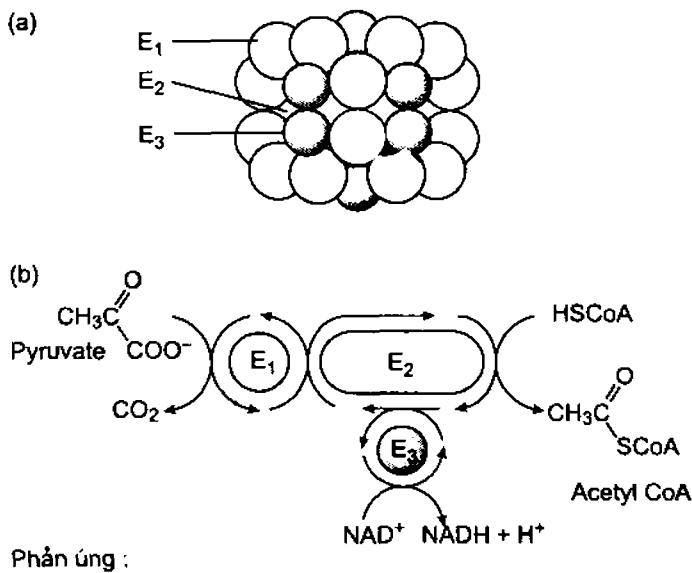
Hình 1.34. Sự tiến hóa của enzyme đa chức năng.

Hình 1.34 mô tả sự tiến hóa của enzyme đa chức năng. Sơ đồ trên trình bày một cơ chất được chuyển hoá thành sản phẩm thông qua ba enzyme. (a) Khi các enzyme có mặt tự do tại nội bào hay trong một khoang nào đó thì các sản phẩm trung gian phải khuếch tán từ enzyme này sang enzyme khác, quá trình sẽ bị chậm lại. (b) Quá trình khuếch tán giảm đi hay bị loại hoàn toàn nếu ba enzyme trên liên hợp với nhau thành một phức hệ đa tiểu phân. (c) Hiện tượng tích hợp chặt chẽ nhất xảy ra nếu nhiều chức năng xúc tác của các enzyme được dung hợp với nhau ở mức gene di truyền thành một protein duy nhất.

Các enzyme của một chu trình trao đổi chất (như phản ứng glucose thành pyruvate) thường cùng có mặt trong một khoang nội bào, ví dụ như trong túi cytosol hoặc trong màng của cơ quan tử. Ở cùng trong một khoang thì sản phẩm của phản ứng này có thể khuếch tán sang enzyme tiếp theo của cùng một dãy phản ứng. Tuy nhiên, khuếch tán vẫn là chuyển động ngẫu nhiên và xảy ra chậm (hình 1.33). Để thăng tình trạng này tế bào thiết lập những cơ chế gồm các enzyme của một chu trình trao đổi chất về cùng một khu vực.

Đơn giản nhất là đưa các enzyme của cùng một chu trình trao đổi chất vào một cấu trúc enzyme nhiều tiểu phân dạng giàn giáo (hình 1.34b). Với cách thức này, sản phẩm của một phản ứng được băng chuyển trực tiếp sang enzyme tiếp theo trong chu trình trao đổi chất. Ví dụ như trong chu trình tạo khử hydro của pyruvate phải có một phức hệ ba enzyme tham gia vào việc chuyển hoá pyruvate thành acetyl CoA trong ty thể (hình 1.35).

Trong một vài trường hợp, các protein riêng rẽ được dung hợp với nhau ở mức di truyền tạo nên một đa chuỗi với chức năng đa enzyme (hình 1.34c). Ví dụ việc đồng phân hoá citrate thành isocitrate trong chu trình acid citric được xúc tác bởi aconitase, một polypeptide duy nhất, nhưng xúc tác hai phản ứng riêng biệt (1) khử hydro của citrate thành *cis*-aconitate và sau đó (2) hydrat hoá *cis*-aconitate thành isocitrate.



Hình 1.35. Cấu trúc và chức năng của pyruvate dehydrogenase, một enzyme phức hệ đa tiểu phần làm chức năng chuyển hoá pyruvate thành acetyl CoA.

Hình 1.35 mô tả cấu trúc và chức năng của pyruvate dehydrogenase, một enzyme phức hệ đa tiểu phần làm chức năng chuyển hoá pyruvate thành acetyl CoA. (a) Phức hệ bao gồm 24 bản giống nhau của pyruvate decarboxylase (E_1), 24 bản của lipoamide transacetylase (E_2) và 12 bản của dihydrolipoyl dehydrogenase (E_3). Các tiểu phần E_1 và E_3 gắn bên ngoài lõi do các tiểu phần E_2 tạo nên. (b) Phản ứng xúc tác bởi enzyme phức hệ này bao gồm nhiều chất trung gian gắn trực tiếp với enzyme. Việc tích hợp cấu trúc một cách chặt chẽ 3 enzyme làm tăng tốc độ phản ứng và giảm thiểu hiệu ứng phụ.

1.2.6. Chức năng vận động cơ học của protein

Motor phân tử và hoạt động cơ học của tế bào

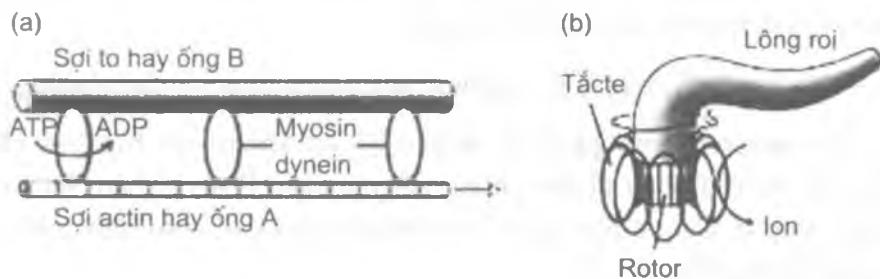
Một đặc điểm chung của tất cả tế bào là có khả năng chuyển động theo một hướng nhất định. Nhiều quá trình tế bào biểu hiện dưới dạng các loại chuyển động hoặc ở mức tế bào hoặc ở mức phân tử : Điểm chung của mọi chuyển động đều là hậu quả tác động của một lực nào đó.

Hình thức chuyển động ngẫu nhiên, không định hướng theo kiểu chuyển động Brown chỉ đạt được với những cự ly rất ngắn. Còn vật liệu tế bào thường được vận chuyển qua một khoảng cách lớn và theo một hướng nhất định. Loại vận chuyển này được thực hiện nhờ có các protein đóng vai trò các motor. Trước hết, sẽ trình bày các kiểu motor khác nhau, rồi sau đó là cơ chế tạo ra lực của các motor này.

Motor phân tử biến đổi năng lượng thành chuyển động

Ở mức độ nano của tế bào và phân tử thì chuyển động chịu tác động của nhiều lực rất khác so với thế giới bình thường. Chẳng hạn, nồng độ protein cao (200 – 300 mg/ml) trong tế bào chất ngăn cản không cho cơ quan tử và các túi nhỏ khuếch tán nhanh hơn 100 $\mu\text{m}/3\text{h}$. Kể cả một con vi khuẩn có kích thước cỡ μm cũng không thể chuyển động nổi một nm nếu nó ngừng bơi chủ động. Để tạo ra lực cần thiết cho nhiều loại vận động của tế bào thì tế bào cần có những enzyme chuyên dụng, thường được gọi là enzyme vận động làm nhiệm vụ chuyển hóa năng lượng thải ra khi thủy phân ATP thành lực cơ học.

Hình 1.36 minh họa so sánh hoạt động của motor chuyển động thẳng và chuyển động xoay tròn. (a) Trong cơ hay bó sợi của sinh vật nhân chuẩn vùng đầu của protein vận động gắn kết với sợi actin mỏng hoặc vi ống A. Thuỷ phân ATP ở vùng đầu gây ra chuyển động thẳng của sợi so với sợi to bên trên. (b) Trong motor xoay tròn của màng vi khuẩn, bộ phận tắc te được cố định vào màng. Ion chảy qua tắc te tạo ra lực xoay khiến cho motor bên trong và lông roi xoay tròn.



Hình 1.36. So sánh hoạt động của motor chuyển động thẳng và chuyển động xoay tròn.

Protein motor tạo ra được chuyển động thẳng hay chuyển động xoay tròn (bảng 1.4). Một vài loại protein motor là thành phần của cụm thiết bị phân tử nhưng không tham gia vào chuyển động dọc theo bộ khung sợi của tế bào như là các chất myosin, kinesin và dynein làm nhiệm vụ vận tải dọc theo các bó sợi microfilament hay microtubule (hình 1.36a). DNA và RNA polymerase là những protein dạng motor chuyển động thẳng vì chúng vận động dọc theo phân tử DNA khi tiến hành tái bản DNA hay phiên mã DNA. Ngược lại, motor xoay được gặp ở mao động của lông roi vi khuẩn, hay khi DNA được đóng gói vào vỏ protein của virus hay trong quá trình sinh tổng hợp ATP. Lực đẩy khi vi khuẩn bơi thường được tạo ra bởi phức hệ protein motor có mặt trong màng vi khuẩn. Các ion chảy ngược dòng với gradient hoá điện qua một vòng cố định các protein, có tên gọi là stator, nằm ở màng. Lực quay tạo ra bởi stator làm xoay các protein vòng trong và dính vào lông roi (hình 1.36b). Tương tự vậy, trong trường hợp của các phân tử ATP synthetase ty thể hoặc phức F_0F_1 một dòng chảy vào của ion qua màng trong của ty thể gây nên hiệu ứng xoay đối với phân F_0 của tiểu phân γ , nó tác động lên vòng các tiểu phân α và β của tiểu đơn vị F_1 . Tương tác giữa tiểu phân γ và tiểu phân β dẫn đến sự tổng hợp ATP.

Từ quan sát hoạt động của các protein motor có thể rút ra 3 tính chất chung của chúng :

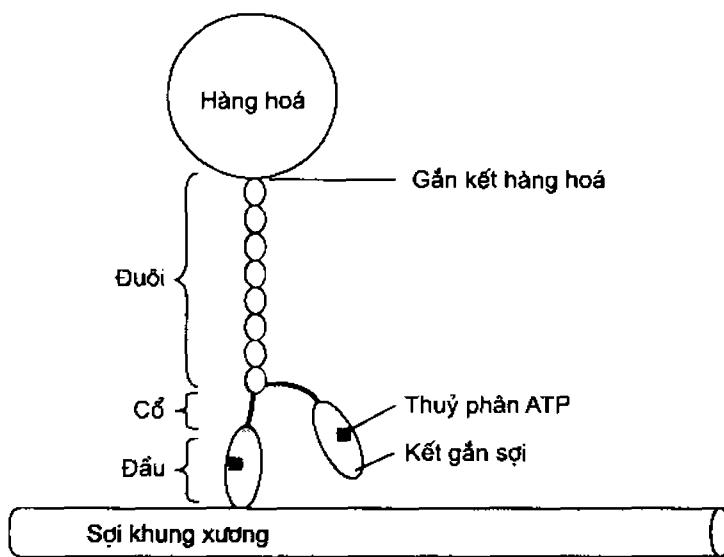
- Khả năng chuyển hoá nguồn năng lượng ATP hay thang ion thành chuyển động thẳng hay xoay vòng.
- Khả năng liên kết và chuyển dịch dọc theo các bó sợi khung nội bào, sợi acid nucleic hay phức hệ protein.
- Kết quả cuối cùng là tạo ra được sự chuyển động có định hướng.

Loại motor protein gắn kết với các bó sợi khung nội bào làm chức năng vận tải các phân tử khi chúng chuyển chỗ. Phân tử hàng hoá của chúng trong tế bào cơ hay roi tế bào nhân chuẩn bao gồm những bó sợi hay các vi ống B (hình 1.37).

Loại motor protein này cũng có khả năng vận chuyển cả nhiễm sắc thể hay các túi màng dọc theo các vi ống hay vi sợi (hình 1.37).

Bảng 1.4. Các motor phân tử chọn lọc (NLN : nguồn năng lượng).

Motor	NLN	Cấu trúc, thành phần	Vị trí nội bào	Chuyển động tạo được
MOTOR CHUYỂN ĐỘNG THẲNG				
DNA polymerase	ATP	Polymerase đa tiểu phần trong replisome	Nhân	Chuyển dọc theo phân tử DNA khi tái bản
RNA polymerase	ATP	Polymerase đa tiểu phần trong phức phiên mã kéo dài	Nhân	Chuyển dọc theo phân tử khi phiên mã
Ribosome	GTP	Yếu tố kéo dài 2 (EF2) liên kết với ribosome	Tế bào chất/màng reticulum tế bào chất	Chuyển dịch theo mRNA khi dịch mã
Myosin	ATP	Chuỗi năng và nhẹ, vùng đầu có hoạt tính ATPase và điểm kết gắn với vi sợi	Tế bào chất	Vận chuyển vi túi và co thắt
Myosin	ATP	Chuỗi năng và nhẹ, vùng đầu có hoạt tính ATPase và điểm kết gắn với vi sợi	Tế bào chất	Vận chuyển vi túi và nhiễm sắc thể khi phân bào
Ribosome	GTP	Đa chuỗi năng, trung và nhẹ, vùng đầu có hoạt tính ATPase và điểm kết gắn vi ống	Tế bào chất	Vận chuyển vi túi và cử động roi và lồng mao tế bào nhân chuẩn
MOTOR CHUYỂN ĐỘNG XOAY				
Motor lông roi vi khuẩn	Thang $H^+ Na^+$	Stator và motor protein, roi	Màng tế bào	Xoay lông roi
ATP synthase F_0F_1	Thang H^+	Hạt F_0 và F_1 do tiểu đơn vị tạo nên	Màng trong của ty thể, màng thylakoid lục lạp và màng vi khuẩn	Xoay tiểu phần γ dẫn đến tổng hợp ATP
Motor vỏ virus	ATP	Connector, RNA và ATPase	Nhân	Xoay connector để đóng gói DNA



Hình 1.37. Hoạt động của các loại protein motor.

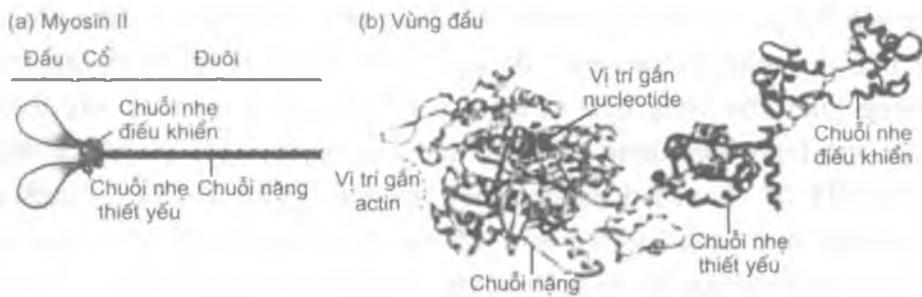
Hình 1.37 trình bày về hoạt động của các loại protein motor. Vùng đầu của các protein motor như myosin, dynein và kinesin gắn vào sợi khung xương nội bào (vi sợi hay vi ống) và vùng đuôi gắn với nhiều loại "hang" ví dụ như các bóng nhỏ dạng túi màng. Thuỷ phân ATP giúp cho vùng đầu chuyển động dọc theo bó sợi (da cam). Chu trình lặp lại giúp cho hàng được chuyển về phía bên phải.

Mọi phân tử myosin đều có vùng đầu, cổ và đuôi với chức năng khác nhau

Để tiếp tục giới thiệu về các đặc tính của protein motor, chúng ta thảo luận đến myosin II thường thấy chuyển động dọc theo sợi actin trong tế bào cơ khi co cơ. Một loại myosin khác làm nhiệm vụ vận chuyển các vi túi dọc theo sợi actin trong khung nội bào. Myosin II và các thành viên khác của họ myosin đều có cấu tạo bởi một hay hai chuỗi nặng và nhiều chuỗi nhẹ. Chuỗi nặng được tổ chức thành 3 vùng có cấu trúc và chức năng riêng biệt (hình 1.38).

Hai vùng đầu hình hạt được chuyên hoá cho hoạt tính ATPase làm nhiệm vụ thuỷ phân ATP kèm với chuyển động. Một đặc điểm hạn chế của myosin ATPase là nó phải được actin hoạt hoá. Nếu thiếu actin thì myosin chuyển hoá ATP thành ADP và phosphate rất chậm chạp. Nhưng nếu khi

myosin tạo phức với actin thì tốc độ ATPase của myosin nhanh gấp 4 đến 5 lần so với khi thiếu actin. Quả là chỉ khi đầu ATPase của myosin được gắn kết với actin thì hoạt động ATPase của nó mới đạt cực đại. Kế sau vùng đầu là chuỗi xoắn α của vùng cổ thuộc chuỗi nhẹ. Chức năng của vùng cổ là thay đổi cấu trúc của vùng đầu nhằm điều chỉnh hoạt tính của vùng đầu. Vùng đuôi có cấu tạo giống một cái cần câu chứa điểm liên kết xác định hoạt tính chuyên biệt của một vài loại myosin nhất định.



Hình 1.38. Cấu trúc của phân tử myosin II.

Hình 1.38 mô tả cấu trúc của phân tử myosin II. (a) Myosin II là một phân tử protein 2 tiểu phân với 2 chuỗi nặng giống nhau (trắng) và 4 chuỗi nhẹ (lam và lục). Mỗi chuỗi nặng chuyển hóa năng lượng từ thuỷ phân ATP thành chuyển động. Hai chuỗi nhẹ liên kết với vùng cổ của chuỗi nặng. Trình tự aa của vùng đuôi có cấu trúc xoắn cục hai lần tạo ra cấu trúc dimer. (b) Mô hình 3D của vùng đầu cho thấy đầu myosin có hình cong kéo dài và lõm vào thành một khe. Vị trí gắn nucleotide nằm ở một đầu và vị trí gắn actin nằm ở đầu kia của khe này. Bao bọc quanh vùng cổ là hai chuỗi nhẹ làm cho cổ thêm độ cứng. Hình vẽ mô tả trạng thái cấu trúc gắn ADP.

Nghiên cứu các phân đoạn myosin thông qua thuỷ phân bằng enzyme phân huỷ protein và nhiều xạ tia X cho ra hình ảnh về vùng đầu, vùng cổ và vùng đuôi loại protein này, trong đó chỉ rõ vùng liên kết ATP và liên kết actin. Vùng đầu kéo dài của đầu myosin gắn với đầu kết thúc của một chuỗi xoắn α của vùng cổ (hình 1.38b).

Hai phân tử của chuỗi nhẹ nằm ở cuối của vùng đầu cho thấy chúng ôm vùng cổ tạo thành một cái kẹp hình chữ C. Ở vị trí này, vùng cổ được

cùng cố thêm bởi các chuỗi nhẹ và qua đó có thể điều hoà được hoạt tính của vùng đầu.

Thay đổi cấu trúc vùng đầu chuyển hoá thuỷ phân ATP thành chuyển động

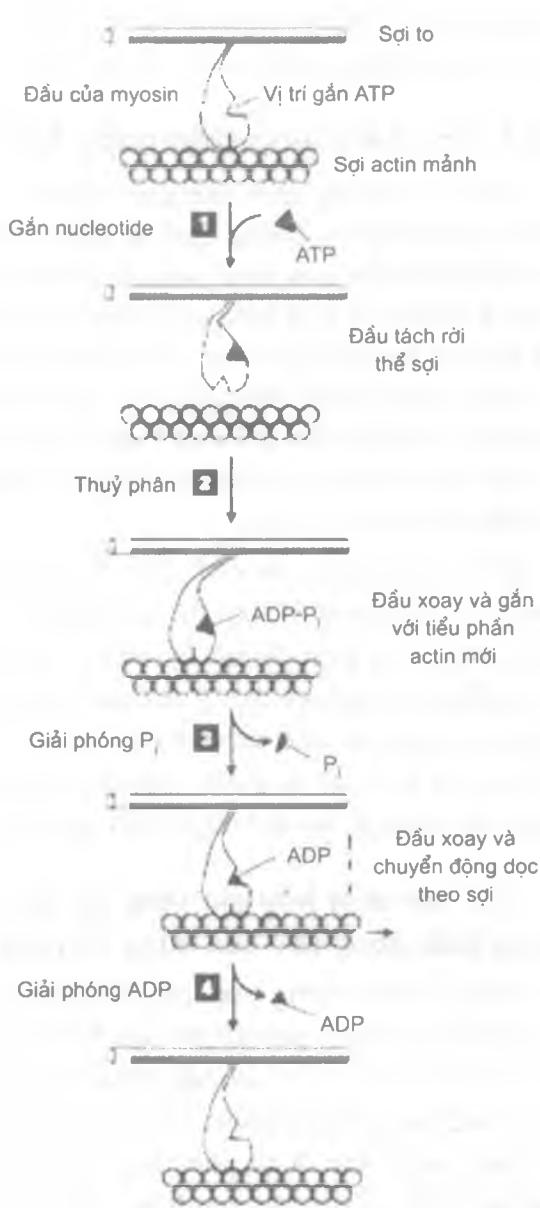
Kết quả nghiên cứu quá trình co thắt của cơ bắp cung cấp những bằng chứng đầu tiên về sự vận động của đầu myosin dọc theo sợi actin. Cơ chế co thắt của cơ được làm sáng tỏ thông qua phương pháp đo chuyển động *in vitro* và phép đo lực từng phân tử riêng rẽ. Nhờ những số liệu thu được từ các phép đo đặc trên, cùng với mô hình cấu trúc không gian của vùng đầu myosin, các nhà nghiên cứu đã xây dựng được mô hình hoạt động sử dụng năng lượng thu giải phóng từ thuỷ phân ATP để chuyển động dọc theo sợi actin. Cho đến nay, người ta cho rằng tất cả các loại myosin đều tạo ra chuyển động theo một cơ chế. Thông thường thì vùng đuôi của myosin đều liên kết với một túi nhỏ hay một búi sợi trong bó cơ. Điều cơ bản mà mô hình này trình bày là sự thuỷ phân một phân tử ATP luôn đi kèm với từng bước chuyển động dọc theo sợi actin của phân tử myosin.

Như trình bày ở hình 1.39, phân tử myosin trải qua một loạt các bước chuyển hoá trong mỗi bước của quá trình chuyển động. Trong mỗi chu kỳ, phân tử myosin phải biến đổi ít nhất 3 lần trạng thái cấu trúc : Một trạng thái ATP chưa liên kết với actin, một trạng thái ADP-Pi liên kết với actin và trạng thái sau khi hoàn tất bước sinh lực. Câu hỏi chính là túi liên kết nucleotide và vùng gắn actin cách đó được hoạt hoá đồng thời như thế nào ? Và những biến đổi xảy ra tại các vùng đó đã được chuyển hoá thành lực chuyển động như thế nào ? Kết quả nghiên cứu cấu trúc của myosin khi có nucleotide hay đồng phân nucleotide được thay thế vào trong nhiều bước của một chu kỳ cho thấy sự liên kết và thuỷ phân một nucleotide gây nên một biến đổi nhỏ ở vùng đầu, biến động nhỏ đó đã được khuếch đại thành một chuyển động mạnh hơn ở vùng cổ. Sự biến đổi nhỏ ở vùng đầu xảy ra tại một khu vực bật tắt (switch) tạo bởi điểm gắn nucleotide và điểm gắn actin. Một khu vực "chuyển đổi" ở cuối vùng đầu hoạt động như một điểm tựa cho vùng cổ uốn xuống và bật lên như một chiếc đòn bẩy.

Cấu trúc tương tự cấu trúc bật-tắt, chuyển đổi, cánh tay đòn bẩy trong kinesin cũng có chức năng chính trong quá trình chuyển động của protein motor kinesin dọc theo vi ống. Cấu trúc cơ bản trong chuyển động của dynein còn chưa được biết vì đến nay chưa xác định được cấu trúc không gian 3 chiều của phân tử protein này.

Hình 1.39 biểu thị mô hình hoạt động của myosin chuyển năng lượng thuỷ phân ATP thành chuyển động dọc theo sợi actin. Mô hình này mô tả chu trình hoạt động của đầu myosin II thường có trong bó cơ. Các myosin khác kết gắn với màng cùng hoạt động theo chu trình như vậy.

Bước một : Gắn ATP là mở khe trên đầu myosin, làm yếu vị trí gắn actin. Bước hai : Giải phóng actin, đầu myosin thuỷ phân ATP khiến cấu trúc thay đổi, đầu chuyển về gần vị trí (+) của sợi actin. Bước ba : Tháo bỏ Pi từ



Hình 1.39. Mô hình hoạt động của myosin chuyển năng lượng thuỷ phân ATP thành chuyển động dọc theo sợi actin.

túi liên kết ATP, đầu myosin lại thay đổi hình thể, giải phóng làm cho actin chuyển động. Bước bốn : Giải phóng ADP. Hoàn tất chu trình.

1.2.7. Cơ chế chung điều hoà chức năng protein

Hầu hết các quá trình trong tế bào không xảy ra một cách độc lập với một tốc độ cố định. Thông thường, hoạt tính xúc tác của các enzyme hay tốc độ hình thành một phức hợp đại phân tử được điều hoà sao cho lượng sản phẩm hoặc sự xuất hiện các phức đại phân tử vừa đủ nhu cầu cần thiết của tế bào. Kết quả là nồng độ cụ thể của cơ chất và sản phẩm thường biến động phụ thuộc theo điều kiện của tế bào. Dòng chảy của vật liệu vào một chu trình trao đổi chất enzyme được điều hoà bằng nhiều cơ chế, có một vài cơ chế còn điều hoà cả các chức năng của các protein không có chức năng làm xúc tác.

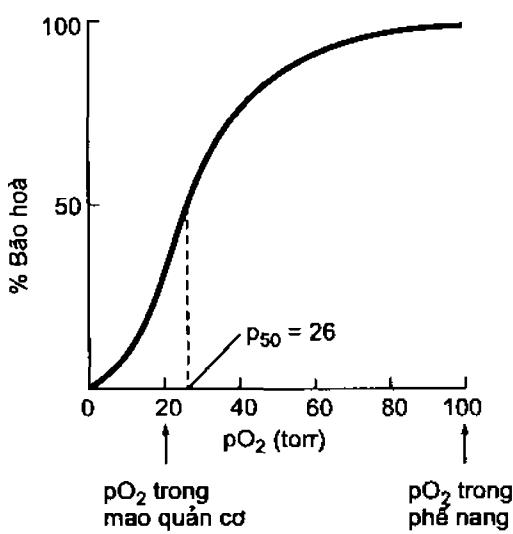
Một trong những cơ chế quan trọng nhất trong điều hoà chức năng protein là trạng thái dị lập thể (allostery). Dị lập thể được nói tới theo nghĩa rộng là sự biến đổi bất kỳ trong cấu trúc bậc ba và bậc bốn hoặc cả hai của phân tử protein thông qua quá trình liên kết với phối tử dưới dạng một chất hoạt hoá, chất ức chế, cơ chất hay cả ba. Cơ chế điều hoà dị lập thể đặc biệt hay xảy ra ở các enzyme hoặc protein đa tiểu phân. Sau đây cơ chế điều hoà dị lập thể được trình bày cụ thể.

Liên kết phối hợp gia tăng độ nhạy của một protein đối với những biến động nhỏ của nồng độ phối tử

Trong nhiều trường hợp, đặc biệt là khi protein liên kết với một vài loại phân tử của một phối tử thì xuất hiện hiện tượng thay đổi mức độ liên kết : Cụ thể là liên kết với một phân tử của phối tử tác động ảnh hưởng lên liên kết với phân tử phối tử tiếp theo. Dạng thức này của cơ chế dị lập thể được gọi là tác động phối hợp (cooperativity), cho phép nhiều loại protein đa tiểu phân đáp ứng hiệu quả hơn lên những thay đổi nhỏ của nồng độ các phối tử, nếu không như vậy thì sẽ khó có thể đạt được. Trong trường hợp của tác động phối hợp dương tính thì sự liên kết tiếp theo sẽ gia tăng và ngược lại trong trường hợp của tác động phối hợp âm tính thì liên kết tiếp theo bị ức chế.

Hemoglobin là một ví dụ kinh điển về tác động phối hợp dương tính : Mỗi tiểu phân của bốn tiểu đơn vị trong hemoglobin chứa một phân tử

heme với một ion Fe nằm trong vòng porphyrin. Gốc heme là thành phần liên kết oxygen của phân tử hemoglobin. Khi có một nguyên tử oxygen liên kết với một trong bốn tiểu đơn vị của hemoglobin thì có hiệu ứng thay đổi cấu trúc khu vực xảy ra đổi với tất cả các tiểu đơn vị còn lại làm cho hằng số K_m đổi với oxygen giảm đi giúp cho các tiểu đơn vị này có thể liên kết thêm được với các phân tử oxygen với nồng độ thấp hơn, tạo ra biểu đồ phản ứng liên kết hình sigma (hình 1.40). Kết quả là hemoglobin có thể tải được nhiều oxygen hơn từ mô xung quanh so với điều kiện thông thường.



Hình 1.40. Sự phối hợp dương tính trong liên kết oxygen với hemoglobin.

Hình 1.40 mô tả sự phối hợp dương tính trong liên kết oxygen với hemoglobin. Mỗi phân tử hemoglobin có 4 vị trí liên kết oxygen ; ở trạng thái bão hòa cả bốn vị trí này đều chứa đủ oxygen. Nồng độ oxygen được đo bằng áp suất từng phần (pO_2). P_{50} là trạng thái ở đó 50% vị trí liên kết oxygen có oxygen liên kết, tương đương với hằng số K_m của phản ứng enzyme. Đồ thị hyperbol thể hiện mối tương tác phối hợp dương tính giữa các vị trí liên kết oxygen của hemoglobin.

Liên kết phôi tử có thể cảm ứng giải phóng dị lập thế đổi với tiểu đơn vị xúc tác hoặc chuyển hóa sang trạng thái có hoạt tính khác

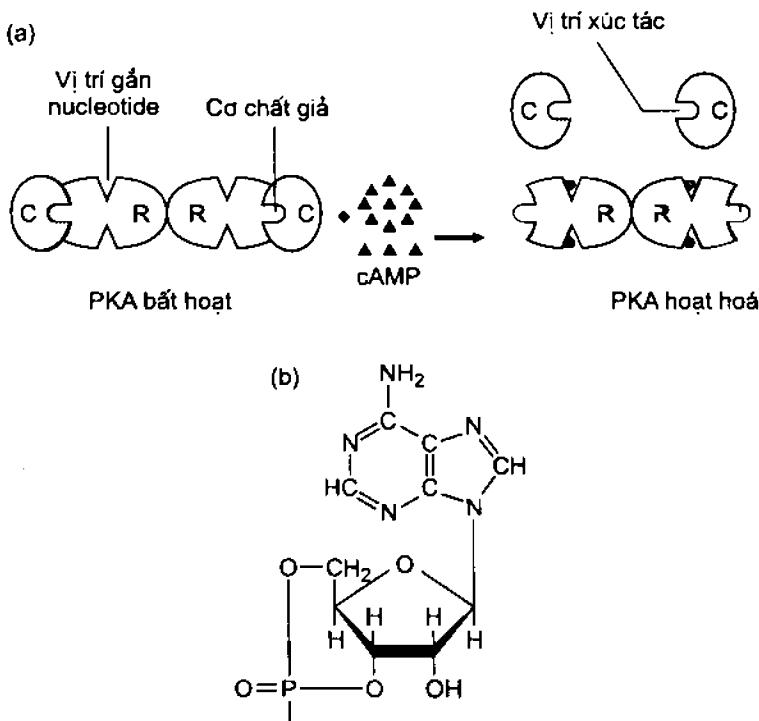
Ở phần trước, protein kinase A được minh họa cho khả năng liên kết và xúc tác của trung tâm hoạt động của một enzyme. Enzyme này có thể tồn tại trong trạng thái bất hoạt gồm bốn tiểu đơn vị, trong đó có hai tiểu đơn vị xúc tác và hai tiểu đơn vị điều hoà. Mỗi tiểu đơn vị điều hoà chứa một trình tự cơ chất giả (pseudosubstrate) gắn với trung tâm hoạt động

của tiểu đơn vị xúc tác. Thông qua ức chế khả năng liên kết cơ chất tiểu đơn vị điều hoà ức chế hoạt tính của tiểu đơn vị xúc tác.

Protein kinase A bất hoạt được hoạt hoá bởi AMP vòng (cAMP), một loại phân tử thông tin thứ cấp. Việc kết gắn cAMP vào tiểu đơn vị điều hoà gây nên cảm ứng thay đổi cấu trúc của vùng cơ chất giả khiến cho chúng mất khả năng liên kết với tiểu đơn vị xúc tác. Như vậy, sự có mặt của cAMP làm cho phức hệ 4 tiểu đơn vị đang bất hoạt phân ly thành hai tiểu đơn vị đơn xúc tác được hoạt hoá và một phức hệ gồm hai tiểu đơn vị điều hoà (hình 1.41). Tương tự như thế, việc kết gắn các phân tử hormone khác nhau vào thụ thể bề mặt tế bào cũng gây nên cảm ứng làm tăng nồng độ cAMP nội bào, dẫn đến trạng thái kích hoạt protein kinase A. Khi tín hiệu biến mất và nồng độ cAMP hạ thấp thì hoạt tính của protein kinase A cũng ngừng thông qua biểu hiện tái lập ráp phức hệ bộ bốn tiểu đơn vị về trạng thái bất hoạt. Việc kết gắn cAMP vào tiểu đơn vị điều hoà thể hiện tác động phối hợp dương, cụ thể là một biến đổi nhỏ trong nồng độ của phân tử dị lập thể tạo ra biến đổi lớn trong hoạt tính của protein kinase A.

Nhiều enzyme đa tiểu đơn vị chịu tác động của phép chuyển hoá dị lập thể khiến cho mối tương tác giữa tiểu đơn vị này với tiểu đơn vị kia bị thay đổi nhưng không làm cho cả phức hệ phân ly thành các tiểu phần như trường hợp của protein kinase A. Theo dạng thức dị lập thể này thì hoạt tính của một protein ở trạng thái liên kết với phối tử có khác với trạng thái không liên kết. Chẳng hạn như GroEL chaperonin thì bộ máy xếp gấp protein thành ống trun này được tạo nên bởi hai tiểu đơn vị hình vòng nhẫn liền kề nhau, cho phép chúng có thể tồn tại ở trạng thái liên kết peptide chật chẽ hay trạng thái giải phóng peptide lỏng lẻo (hình 1.27). Quá trình kết gắn ATP và phân tử co-chaperonin GroES với một trong hai vòng nhẫn ở trạng thái liên kết chật chẽ gây nên hiện tượng kéo dài gấp 2 lần khe GroEL, làm cho phân tử chuyển sang trạng thái xếp gấp peptide lỏng lẻo.

Hình 1.41 mô tả hoạt hoá protein kinase A thông qua kết tử. Khi nồng độ cAMP thấp thì PKA là một phức tử từ bất hoạt. Khi gắn kết với cAMP vào tiểu đơn vị điều khiển (R) thì hình thể của tiểu phần này thay đổi giải phóng tiểu phần xúc tác hoạt hoá (C). (b) AMP vòng (cAMP) là một dẫn xuất của adenosine monophosphate đóng vai trò tín hiệu ngoại bào thông qua gia tăng nồng độ và có khả năng hoạt hoá nhiều loại protein.



Hình 1.41. Hoạt hoá protein kinase A thông qua kết tử.

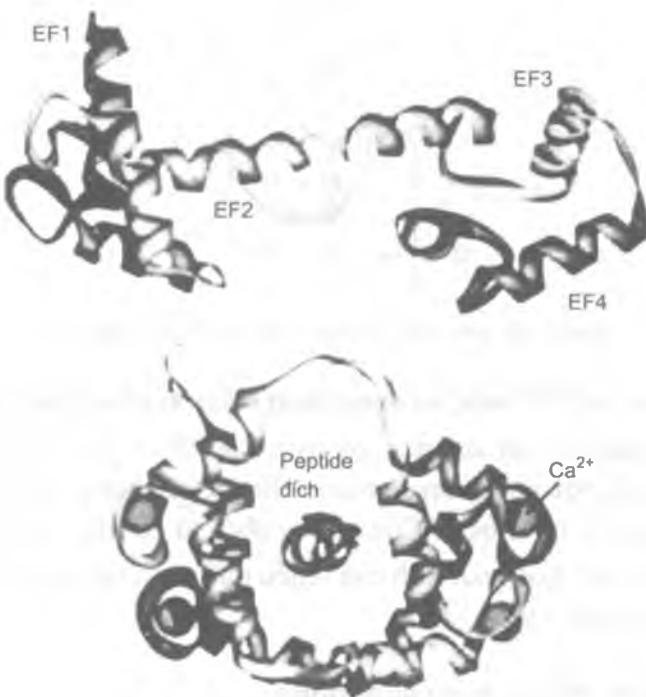
Calcium và GTP được sử dụng rộng rãi trong hoạt hoá protein

Trong những ví dụ vừa nêu, oxygen, cAMP và ATP đã gây nên thay đổi dị lập thể đối với những protein đích của chúng như hemoglobin, protein kinase A và GroEL. Còn có hai phối tử có hiệu ứng dị lập thể là Ca^{2+} và GTP điều hoà hoạt tính của nhiều quá trình nội bào thông qua các protein rất phổ biến khác.

Kích hoạt thông qua calmodulin

Nồng độ Ca^{2+} tự do trong tế bào chất thường được giữ rất thấp ($\approx 10^{-7} \text{ M}$) thông qua protein vận chuyển của màng có chức năng bơm Ca^{2+} ra khỏi tế bào hoặc vào vùng lưới nội bào (ER). Nồng độ Ca^{2+} trong nội bào có thể tăng từ 10 đến 100 lần thông qua việc giải phóng Ca^{2+} từ ER hoặc thu nhận từ vùng ngoại bào. Việc gia tăng nồng độ Ca^{2+} được điều hoà bởi các protein liên kết Ca^{2+} , cụ thể là các protein thuộc họ EF (EF hand family). Tất cả các protein này đều có dạng xoắn helix – vòng khuyên loop – xoắn helix.

Dạng tiền chất của calmodulin được tìm thấy ở tất cả các loại tế bào nhân chuẩn, có thể tồn tại dưới dạng một protein đơn tiểu phân hoặc một tiểu phân của một protein đa tiểu phân. Một phân tử có hình quả tạ, phân tử calmodulin chứa 4 điểm liên kết Ca^{2+} với hằng số $K_D \approx 10^{-6} \text{ M}$. Sự liên kết Ca^{2+} vào calmodulin gây nên biến đổi trong cấu trúc của chúng khiến chúng có thể gắn kết với nhiều loại protein đích khác, qua đó làm cho những protein này có thể "bật" hoặc "tắt" hoạt tính của chúng (hình 1.42). Calmodulin và các protein EH tương tự với vai trò là các protein bật-tắt cùng với Ca^{2+} kích hoạt hoạt tính của các protein khác.



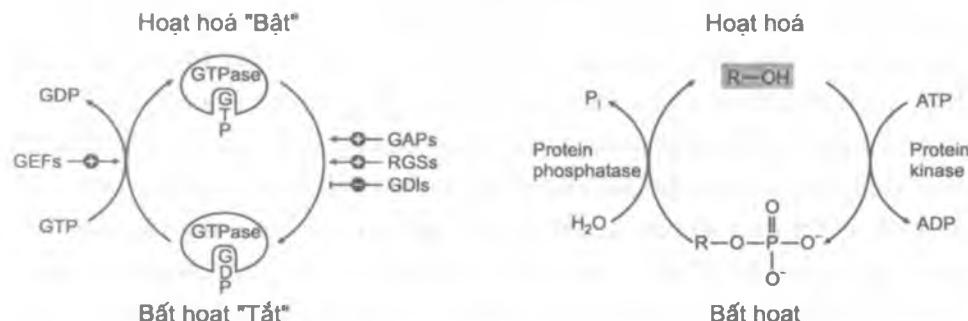
Hình 1.42. Cơ chế bật tắt do phân tử Ca^{2+} /calmodulin.

Hình 1.42 mô tả cơ chế bật tắt do phân tử Ca^{2+} /calmodulin. Calmodulin là một protein nội bào khá phổ biến có 4 vị trí gắn Ca^{2+} tại các chuỗi EF. Mỗi chuỗi EF có một mẫu hình thể xoắn-khuyên-xoắn. Khi nồng độ Ca^{2+} nội bào đạt $5 \times 10^{-7} \text{ M}$, sự kết gắn với Ca^{2+} , khi đó hình thể của Ca^{2+} calmodulin thay đổi khiến cho phân tử calmodulin bọc quanh peptide đích làm chúng thay đổi hoạt tính.

Bật tắt thông qua protein liên kết nucleotide

Một nhóm protein bật-tắt nội bào khác được biết đến thuộc siêu họ GTPase. Siêu họ này bao gồm protein đơn tiểu phân Ras và tiểu đơn vị G_α của nhóm protein G tam tiểu phân. Cả hai Ras và G_α đều kết gắn với màng nguyên sinh chất, hoạt động chuyển tín hiệu trong tế bào và đóng vai trò chủ chốt trong quá trình nhân và phân hoá tế bào. Thành viên khác của siêu họ GTPase có chức năng trong sinh tổng hợp protein, trong vận chuyển protein giữa nhân và nguyên sinh chất, trong hình thành các túi màng và trong quá trình dung hợp chúng với màng đích và trong sắp xếp khung actin của tế bào.

Tất cả các protein bật tắt thuộc siêu họ GTPase đều tồn tại dưới hai dạng (hình 1.43, trái) : (1) Dạng hoạt hoá "bật" với GTP (guanosine triphosphate) liên kết có khả năng kích hoạt tính của những protein đích nhất định và (2) dạng "tắt" gắn với GDP (guanosine diphosphate). Hoạt tính GTPase của dạng protein "bật" thuỷ phân GTP liên kết một cách chậm chạp thành GDP để tạo ra dạng bất hoạt. Sự trao đổi tiếp theo của GDP với GTP để tái tạo nên dạng hoạt hoá xảy ra cũng khá chậm. Sự hoạt hoá là biểu hiện tạm thời, được gia tăng hay kìm hãm thông qua những protein khác có khả năng hoạt động như những chất điều hoà di lập thể đổi với protein "bật-tắt".



Hình 1.43. Vòng công tắc bật tắt của GTPase thông qua trạng thái hoạt hoá và bất hoạt (trái). Điều khiển hoạt tính protein thông qua công tắc kinase/phosphatase (phải).

Hình 1.43 mô tả vòng công tắc bật tắt của GTPase thông qua trạng thái hoạt hoá và bất hoạt (trái). Biến đổi từ dạng hoạt hoá sang dạng bất hoạt thông qua thuỷ phân GTP nhờ GTPase và protein tín hiệu điều hoà (RGS) và bất hoạt thông qua guanine nucleotide disociation inhibitor (GDI).

Hình 1.43 cũng mô tả điều khiển hoạt tính protein thông qua công tắc kinase/phosphatase (phải). Vòng phosphoryl hoá và dephosphoryl hoá một protein thường là cơ chế tế bào thông thường để điều khiển hoạt tính một protein. Trong ví dụ này, protein đích R bị bất hoạt khi phosphoryl hoá và hoạt hoá khi bị dephosphoryl hoá. Có những protein lại chịu ảnh hưởng ngược lại.

Phosphoryl hoá protein vòng và dephosphoryl hoá điều hoà nhiều chức năng tế bào

Như đã đề cập đến, một trong những cơ chế phổ biến nhất để điều hoà chức năng protein là quá trình phosphoryl hoá – sự bổ sung hoặc làm mất đi nhóm phosphate từ các gốc serine, threonine hoặc tyrosine. Protein kinase xúc tác quá trình phosphoryl hoá, còn phosphatase xúc tác quá trình dephosphoryl hoá. Mặc dù cả hai phản ứng nhất thiết xảy ra không thuận nghịch nhưng tương tác hoạt tính giữa kinase và phosphatase cung cấp cho tế bào một cơ chế "đóng-mở" để bật hoặc tắt hoạt động của nhiều loại protein (hình 1.43, phải). Phosphoryl hoá làm thay đổi diện tích và thường dẫn đến thay đổi cấu trúc. Hệ quả này làm thay đổi khả năng liên kết phối tử của protein khiến cho hoạt tính của nó hoặc tăng hoặc giảm đi.

Gần 3% toàn bộ protein nấm men là protein kinase hoặc phosphatase, chứng tỏ sự quan trọng của quá trình phosphoryl và dephosphoryl hoá ngay trong những tế bào còn khá đơn giản. Tất cả các họ protein, bao gồm protein cấu trúc, enzyme, kênh màng và các phân tử tín hiệu đều được điều hoà thông qua cơ chế "bật-tắt" của kinase-phosphatase. Các protein kinase và phosphatase khác nhau thường đặc thù cho các protein đích khác nhau và có thể điều hoà nhiều chu trình tế bào. Một số enzyme hoạt động lên một số hay nhiều protein đích. Loại sau cùng rất cần thiết trong phối hợp hoạt tính của nhiều protein mà chỉ cần thông qua một công tắc kinase/phosphatase duy nhất là có thể điều hoà được. Loại kinase hoặc phosphatase có một protein đích là phổ biến, vì thế chúng tạo nên cả một mạng lưới điều khiển.

Một số protein được hoạt hoá hoặc bất hoạt thông qua chia tách phân tử

Cơ chế điều khiển vừa đề cập đến ở trên bật tắt thuận nghịch hoạt tính một protein. Sự điều khiển một vài protein lại xảy ra theo cơ chế

hoàn toàn khác : Đó là sự hoạt hoá và bất hoạt không thuận nghịch thông qua chia tách phân tử. Cơ chế này thường xảy ra ở những protein có chức năng hormone (ví dụ insulin) và protease tiêu hoá. Ví dụ điển hình cho những enzyme loại đó là trypsin và chymotrypsin được tổng hợp ở tuy và tiết vào ruột non dưới dạng phân tử trypsinogen và chymotrypsinogen bất hoạt. Enterokinase và aminopeptidase tiết ra từ tế bào thành ruột non chuyển hoá trypsinogene thành trypsin và đến lượt trypsin lại tách chymotrypsinogene thành chemotrypsin. Việc hoạt hoá chậm trễ các loại protease này trước khi chúng đến được ruột non tránh được hiện tượng chúng sẽ tiêu hoá các mô của tuyến tuy là nơi sản sinh ra chúng.

Kiểm soát vị trí và nồng độ protein là cách thức điều hòa bậc cao

Hoạt tính của protein được điều khiển một cách liên tục nhằm giúp cho nhiều protein trong một tế bào có thể hoạt động một cách hài hoà. Ví dụ, nhiều chu trình trao đổi chất được kiểm soát chặt chẽ ở mọi thời điểm. Phản ứng tổng hợp xảy ra khi cần có sản phẩm của những phản ứng này, phản ứng phân huỷ xảy ra khi phân tử cần bỏ đi. Tất cả các cơ chế mô tả trước đây tác động lên một protein ngay tại nơi nó hoạt động, bật hay tắt hoạt tính của nó.

Tuy nhiên, hoạt động bình thường trong tế bào đòi hỏi phải phân chia protein thành những khoang nhất định như ty thể, nhân, lysosome. Liên quan đến hoạt động của các enzyme sự phân khoang không những chỉ tạo ra cơ hội để điều khiển việc cung cấp cơ chất hoặc giải phóng sản phẩm mà còn cho phép các phản ứng có thể xảy ra cùng một thời điểm tại nhiều phân khác nhau của tế bào.

Ngoài sự phân khoang, các quá trình của tế bào còn được điều tiết thông qua sinh tổng hợp và phân huỷ protein. Ví dụ protein thường được tổng hợp với tốc độ thấp khi tế bào chỉ cần ít hoặc không cần hoạt tính xúc tác của chúng. Khi tế bào cần nhiều hơn (chẳng hạn cơ chất cảm ứng enzyme, kháng nguyên kích thích lymphocyte B tạo kháng thể), tế bào phản ứng lại bằng quá trình sinh tổng hợp mới các phân tử protein. Sau đó, nguồn protein được sinh tổng hợp giảm đi khi cơ chất giảm hoặc tế bào trở nên hết hoạt hoá. Tín hiệu ngoại bào thường đóng vai trò công cụ cảm ứng biến đổi tốc độ sinh tổng hợp và phân huỷ protein. Những sự thay đổi

có điều khiển như thế đóng vai trò chủ đạo trong chu trình phân chia và trong phân hoá tế bào.

Trong dị lập thể, sự liên kết của một phôi tử (cơ chất, chất hoạt hoá hoặc chất kìm hãm) gây nên thay đổi cấu trúc hoặc chuyển đổi dị lập thể làm cho hoạt tính hoặc ái lực đối với phôi tử khác của protein thay đổi.

Trong các protein đa tiểu phân như hemoglobin, khi liên kết với nhiều phôi tử việc liên kết với phôi tử này có thể làm tăng ái lực liên kết đối với phôi tử tiếp theo. Các enzyme gắn kết hợp lực với cơ chất có biểu đồ động học hình sigma tương tự như đồ thị liên kết oxygen của hemoglobin.

Nhiều cơ chế dị lập thể hoạt động theo nguyên tắc công tắc bật tắt hoạt tính protein một cách thuận nghịch.

Sự liên kết dị lập thể với các phôi tử có thể làm chuyển đổi một protein từ trạng thái hoạt hoá sang trạng thái khác hoặc giải phóng các tiểu đơn vị đã hoạt hoá.

Hai họ protein công tắc nội bào điều khiển hàng loạt quy trình của tế bào : (1) Calmodulin và protein liên kết Ca^{2+} thuộc họ EF và (2) thành viên của siêu họ GTPase (như Ras và G_α) có thể xoay vòng từ trạng thái hoạt hoá khi gắn với GTP sang trạng thái bất hoạt khi gắn với GDP.

Phosphoryl và dephosphoryl hoá các gốc amino acid thông qua kinase và phosphatase tạo nên cơ chế điều hoà bật-tắt hoạt tính nhiều loại protein.

Cơ chế không dị lập thể trong điều khiển hoạt tính protein còn có sự chia tách phân tử, qua đó chuyển hoá không thuận nghịch các tiền enzyme thành các enzyme hoạt hoá ; phân khoang protein và kích thích nhờ tín hiệu ngoại bào đối với sinh tổng hợp và phân giải protein.

2

SAO CHÉP, PHIÊN MÃ VÀ DỊCH MÃ

2.1. THÀNH PHẦN CẤU TẠO VÀ LIÊN KẾT HÓA HỌC

Sinh học phân tử quan tâm đầu tiên tới mối quan hệ tương tác giữa các đại phân tử thông tin DNA (deoxyribonucleic acid), RNA (ribonucleic acid) và các phân tử này được sử dụng để tổng hợp các polypeptide như thế nào, thành phần cơ bản của tất cả các protein. Trong một số virus, RNA là vật liệu di truyền, nhưng trong tất cả các tế bào khác, thông tin di truyền được lưu trữ ở các phân tử DNA. Các vùng chọn lọc của các phân tử DNA tế bào làm khuôn để tổng hợp các phân tử RNA. Phần lớn các phân tử RNA đến lượt được sử dụng để chỉ định tổng hợp các polypeptide, hoặc trực tiếp hoặc bằng cách trợ giúp ở các bước khác nhau trong biểu hiện gene. Do phần lớn biểu hiện gene được dành cho tổng hợp polypeptide, các protein là điểm cuối chức năng chính của khuôn DNA và chiếm phần lớn khối lượng khô của tế bào. Khái niệm protein được dẫn từ tiếng Hy Lạp *proteios*, nghĩa là "thuộc hàng đầu tiên" và phản ánh vai trò quan trọng của các protein trong các chức năng tế bào đa dạng, như enzyme, thụ thể, protein dự trữ, protein vận chuyển, yếu tố phiên mã, phân tử tín hiệu, hormone...

2.1.1. Cấu trúc polymer mạch thẳng

DNA, RNA và polypeptide là các polymer được xác định bằng một trình tự mạch thẳng của các đơn vị lặp lại đơn giản.

Trong eukaryote, các phân tử DNA đơn lẻ được tìm thấy trong các nhiễm sắc thể của nhân tế bào và trong ty thể, cũng như trong diệp lục của tế bào thực vật. Chúng là những polymer lớn, với một khung mạch thẳng và các gốc phosphate. Đường trong các phân tử DNA là deoxyribose, một đường 5 carbon và các gốc đường tiếp theo được liên kết bởi các liên kết phosphodiester đồng hoá trị. Gắn đồng hoá trị vào nguyên tử carbon số 1' (1 phẩy) của mỗi gốc đường là một base nitơ.

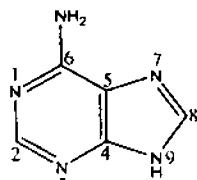
Bốn dạng base được tìm thấy : adenine (A), cytosine (C), guanine (G) và thymine (T) và chúng bao gồm các đị vòng của các nguyên tử carbon và nitrogen. Chúng có thể được chia thành hai lớp : các purine (A và G) có hai đị vòng liên kết với nhau ; các pyrimidine (C và T) chỉ có một vòng đơn. Một đường với một base gắn vào được gọi là một nucleoside. Một nucleoside với một nhóm phosphate gắn vào nguyên tử carbon 5' hoặc 3' tạo thành một nucleotide, đơn vị lặp lại cơ bản của sợi DNA (bảng 2.1 và hình 2.1). Thành phần của các phân tử RNA tương tự như thành phần của các phân tử DNA nhưng khác biệt ở chỗ chúng chứa các gốc đường ribose trong chỗ của deoxyribose và uracil (U) thay cho thymine (bảng 2.1 và hình 2.1).

Bảng 2.1. Các base nói chung được tìm thấy trong các nucleic acid với các nucleoside và nucleotide tương ứng.

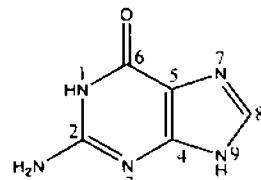
BASE	NUCLEOSIDE (base + đường)	NUCLEOTIDE (nucleoside + phosphate)		
		mono phosphate	tri phosphate	tri phosphate
Purine				
Adenine	Adenosine Deoxyadenosine	AMP (Adenylyate) dAMP	ADP dADP	ATP dATP
Guanine	Guanosine Deoxyguanosine	GMP (Guanylyate) dGMP	GDP dGDP	GTP dGTP
Purine				
Guanine	Cytidine Deoxycytidine	CMP (Cytidylate) dCMP	CTP dTCTP	CTP dTCTP
Guanine	Thymidine Deoxythymidine	[TMP] Thymidylate) dTTP	CTP dTCTP	[TTP] dTTP
Guanine	Uridine Deoxyuridine	UMP (Uridylate) dUMP	UDP dUDP	UTP dTUTP

Trong bảng trên, ngoặc đơn biểu thị các nucleotide ít được tìm thấy. Các nucleotide với một monophosphate đơn thường được mô tả với cái

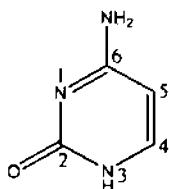
tên có hậu tố -ine của base được thế chỗ bằng hậu tố -ylate, như đã chỉ ra trong các ví dụ về ribonucleotide.



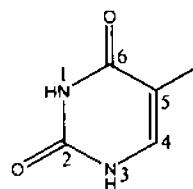
Adenine (A)



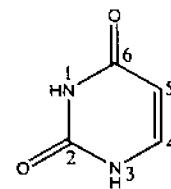
Guanine (G)



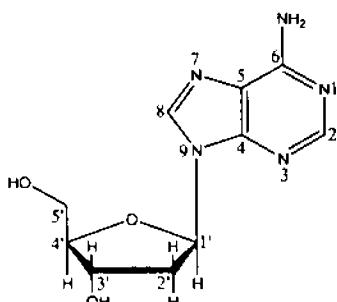
Cytosine (C)



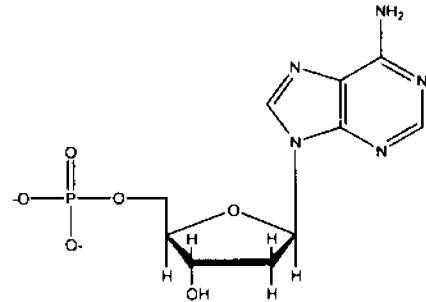
Thymine (T)



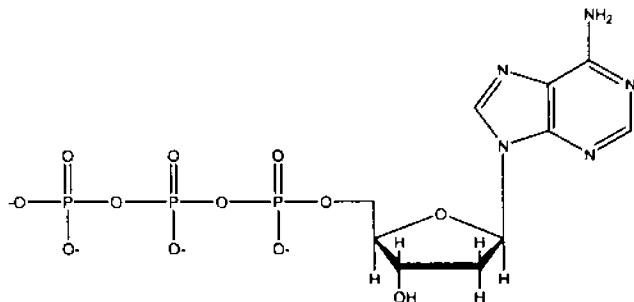
Uracil (U)



Adenosine



Adenosine 5'-monophosphate



2'-Deoxyadenosine 5'-triphosphate (dATP)

Hình 2.1. Cấu trúc của các base, nucleoside và nucleotide.

Nét vẽ đậm ở đáy các vòng đường có nghĩa cho thấy rằng mặt phẳng của vòng được đặt ở góc 90° so với mặt phẳng của base tương ứng (có nghĩa là nếu mặt phẳng của base được trình bày nằm trên mặt phẳng của trang giấy, thì các nguyên tố carbon 2' và 3' của đường có thể được nhìn theo hướng đi ra từ trang giấy và nguyên tố oxygen như hướng xuống mặt trang giấy). Cách đánh số trong đường deoxyribose và ribose được chốt lại với 5 nguyên tố carbon và được đánh số $1' \rightarrow 5'$ nhưng việc đánh số của các base bao gồm cả các nguyên tử carbon lẫn nitrogen tồn tại nội trong các dị vòng. Các nguyên tố hydroxyl và hydrogen được nối với carbon 2' cho thấy sự khác biệt cơ bản giữa các gốc đường ribose và deoxyribose. Nhóm phosphate được biểu thị tiếp theo là α , β , γ ... theo trạng thái gắn với vòng đường.

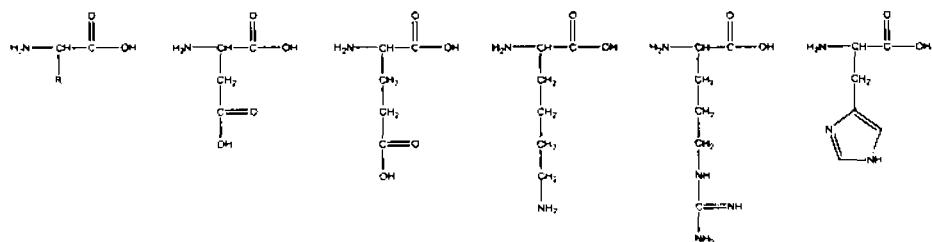
Protein có thành phần là một hoặc nhiều phân tử polypeptide mà có thể được cải biến bằng cách bổ sung các chuỗi phụ carbohydrate khác nhau hoặc các nhóm hóa học. Giống như DNA và RNA, các phân tử polypeptide là những polymer bao gồm một trình tự mạch thẳng các đơn vị lặp lại, trong trường hợp này là các amino acid. Chúng bao gồm một nhóm amino tích điện dương và một nhóm carboxylic acid (carboxyl) tích điện âm nối với nhau bằng nguyên tố carbon trung tâm được gắn với các chuỗi phụ nhận dạng. Hai mươi amino acid khác nhau có thể được chia nhóm thành các lớp khác nhau, phụ thuộc vào bản chất của các chuỗi phụ (hình 2.2).

Việc chia lớp dựa theo như sau :

- Các amino acid kiềm mang một chuỗi phụ với tổng điện tích dương ; một nhóm amino (NH_2) hoặc vòng histidine trong một chuỗi phụ cần một H^+ ở pH sinh lý.
- Các amino acid tính *acid* mang một chuỗi phụ với tổng điện tích âm ; một nhóm carboxyl (COOH) trong chuỗi phụ mất đi một H^+ ở pH sinh lý để tạo ra COO^- .
- Các amino acid có cực không tích điện là điện tích trung hoà nhưng mang các chuỗi phụ với các nhóm hóa học có cực, chúng phân biệt nhau bởi có điện tích từng phần, ví dụ các nhóm hydroxyl ($\text{O}^{\delta-} - \text{H}^{\delta+}$), các nhóm sulfhydryl ($\text{S}^{\delta-} - \text{H}^{\delta+}$)... Giống như các amino acid kiềm và tính acid, các amino acid có cực có thể phản ứng với các nhóm khác mang điện tích.

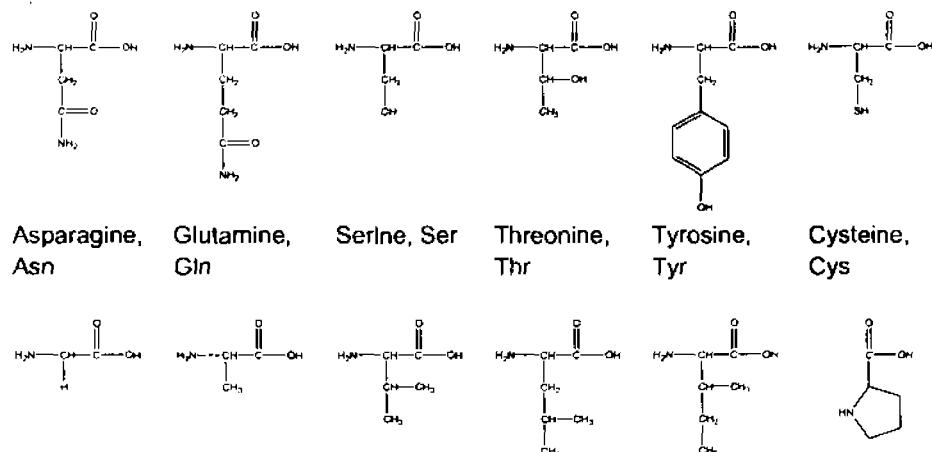
– Các amino acid *trung tính không phân cực* là hydrophobic (ky nước). Chúng thường tương tác với một và các nhóm ky nước khác.

Phân cực, tích điện

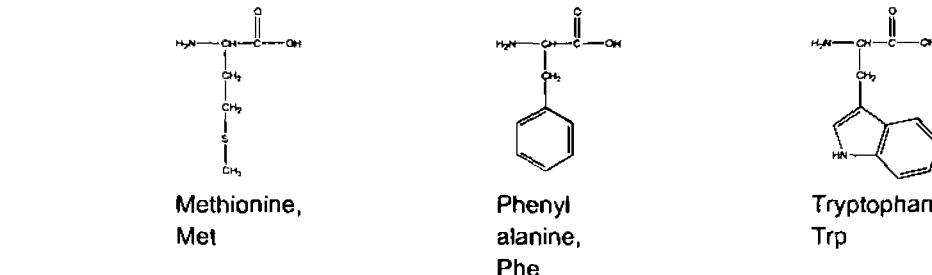


Công thức chung	Acid, aspartic aspartate, Asp	Glutamic acid, glutamate, Glu	Lysine, Lys	Arginine, Arg	Histidine, His
-----------------	-------------------------------	-------------------------------	-------------	---------------	----------------

Phân cực không tích điện



Asparagine, Asn	Glutamine, Gln	Serine, Ser	Threonine, Thr	Tyrosine, Tyr	Cysteine, Cys
-----------------	----------------	-------------	----------------	---------------	---------------



Hình 2.2. Công thức của các amino acid.

Các amino acid trong một nhóm phụ (ví dụ các amino acid acidic hoặc basic amino acid) rất tương tự về hoá học. Các nhóm đã nêu là các nhóm hoá học có cực. Quy ước đánh số các nguyên tử carbon để đặt tên nguyên tử carbon trung tâm là α và các carbon tiếp theo của chuỗi phụ mạch thẳng là β , γ , δ ... (xem hình phân chuỗi phụ lysine). Nói chung, các amino acid có cực là ưa nước và các amino acid không cực là kỵ nước, glycine (có một chuỗi phụ rất nhỏ) và cysteine (nhóm sulphydryl của nó không có cực như nhóm hydroxyl) chiếm các vị trí trung gian ở phạm vi ưa nước – kỵ nước. Proline thường trong chuỗi phụ liên kết với nguyên tử nitrogen của nhóm NH_2 cũng như nguyên tử carbon trung tâm.

Các polypeptide được hình thành bằng phản ứng mất nước giữa nhóm amino của một amino acid và nhóm carboxyl của amino acid tiếp theo để hình thành một khung lặp lại (-NH-CHR-CO-), ở đó các chuỗi phụ R phân biệt một amino acid này với một amino acid khác (hình 2.2).

2.1.2. Liên kết trong các đại phân tử

Các liên kết đồng hoá trị tạo ra độ bền trên các phân tử DNA, RNA và polypeptide, nhưng các liên kết không đồng hoá trị yếu hơn rất quan trọng trong việc cho phép các mối quan hệ giữa các phân tử trong cấu trúc bền vững.

Độ bền của các polymer nucleic acid và protein trước tiên phụ thuộc vào các mối liên kết đồng hoá trị mạnh nối các nguyên tố thành phần của các khung xương mạch thẳng. Thêm vào các mối liên kết đồng hoá trị, một số các mối liên kết không đồng hoá trị yếu (bảng 2.2) rất quan trọng trong các mối tương tác giữa những phân tử này và giữa các nhóm bên trong một phân tử nucleic acid hoặc protein đơn. Diễn hình, các mối liên kết không đồng hoá trị như vậy yếu hơn các mối liên kết đồng hoá trị bởi một hệ số hơn 10 lần. Các mối liên kết đồng hoá trị không giống như vậy, cường độ của chúng được xác định chỉ bởi các nguyên tố đặc biệt liên quan, tuy nhiên, cường độ của các mối liên kết không đồng hoá trị cũng chủ yếu phụ thuộc vào môi trường nước của chúng. Cấu trúc của nước là một phức hợp đặc biệt, với một mạng lưới thay đổi nhanh của liên kết không đồng hoá trị xảy ra giữa các phân tử H_2O đơn lẻ. Lực trội hơn hẳn trong cấu trúc này là liên kết hydrogen, một liên kết điện tử

giữa một nguyên tố hydrogen tích điện dương một phần và một nguyên tố tích điện âm một phần nó trong trường hợp của các phân tử nước, là một nguyên tố oxygen.

Bảng 2.2. Liên kết không đồng hóa trị yếu.

Dạng liên kết	Bản chất liên kết
Hydrogen	Liên kết hydrogen hình thành khi một nguyên tố hydrogen được đặt xen kẽ giữa hai nguyên tố hút điện tử, thường là oxygen hoặc nitrogen. Xem chương 1 về tầm quan trọng trong cấu trúc và chức năng của nucleic acid và protein.
Ion	Mối tương tác ion xảy ra giữa các nhóm tích điện. Chúng có thể rất mạnh trong tinh thể nhưng trong môi trường nước, các nhóm tích điện được che chở bởi các phân tử H_2O và các ion khác trong dung dịch nên rất yếu. Tuy nhiên, chúng có thể rất quan trọng trong chức năng sinh học, như trong trường hợp nhận biết cơ chất-enzyme.
Van der Waals	Bất kỳ hai nguyên tử nào gần nhau đều cho thấy một tương tác liên kết hút yếu nhờ vào điện tích dao động (lực hút Van der Waals) cho tới khi chúng cực gần thì đẩy nhau rất mạnh (lực đẩy Van der Waals). Mặc dù các lực đẩy Van der Waals riêng lẻ rất yếu nhưng có thể trở thành quan trọng khi điều chỉnh tốt giữa các bề mặt của hai đại phân tử.
Van der Waals	Nước là một phân tử phân cực. Khi các phân tử ky nước hoặc các nhóm hoà học được đặt vào một môi trường nước, chúng buộc phải giảm tối đa các tác dụng phá vỡ của chúng trong mạng lưới phức hợp của liên kết hydrogen giữa các phân tử nước. Các nhóm ky nước buộc phải cùng nhau chống giữ nhau các liên kết ky nước, mà bản chất của lực hút này là do lực đẩy chung của các phân tử nước.

Các phân tử tích điện có khả năng hòa tan cao trong nước. Do các điện tích phosphate trong các nucleotide thành phần của chúng, cả DNA và RNA được tích điện âm (*polyanion*). Phụ thuộc vào thành phần amino acid của chúng, các protein có thể mang một tổng điện tích dương (basic protein) hoặc một tổng điện tích âm (các acidic protein). Mỗi liên kết hydrogen tiềm năng của các phân tử nước nghĩa là các phân tử với các nhóm phân cực (bao gồm DNA, RNA và protein) có thể hình thành các mối tương tác đa chiều với các phân tử nước, dẫn đến hòa tan chúng. Bởi vậy, ngay các protein tích điện trung tính thường sẵn sàng hòa tan nếu chứa một số amino acid phân cực tích điện hoặc trung tính. Ngược lại, các protein liên kết màng thường được xác định tính chất bởi một hàm

lượng amino acid kỵ nước cao, chúng bền về nhiệt động học trong môi trường kỵ nước của màng lipid.

Các liên kết đồng hoá trị không giống chúng cần cung cấp năng lượng đáng kể để bẻ gãy chúng, các liên kết không đồng hoá trị không thay đổi được tạo ra và bị bẻ gãy ở các nhiệt độ sinh lý. Chúng dễ dàng cho phép các mối tương tác phân tử thuận nghịch và tạm thời, chúng cần thiết cho chức năng sinh học. Trong trường hợp của các nucleic acid và protein, chúng đóng một số vai trò chìa khoá, bảo đảm sao chép của DNA trung thực, phiên mã của RNA, nhận biết codon – codon đối... Mặc dù hoạt động đơn lẻ yếu nhưng hoạt động tổ hợp nhiều liên kết không đồng hoá trị có thể góp phần lớn vào độ bền của cấu trúc (cấu dạng) của những phân tử này và cũng có thể là quan trọng để biệt hoá hình dạng các đại phân tử (xem hình 2.6B ví dụ về liên kết hydro nội phân tử tạo hình dạng của phân tử RNA vận chuyển).

2.1.3. Tầm quan trọng của liên kết hydrogen

Liên kết hydrogen giữa các phân tử trong các nucleic acid

Liên kết này rất quan trọng để hình thành các nucleic acid sợi kép như sau :

DNA sợi kép. Độ bền của helix sợi kép được duy trì bằng liên kết hydrogen giữa các cặp base A-T và C-G (mục 2.2.1 và hình 2.4).

Các sợi kép DNA-RNA. Thường chúng hình thành trong khi phiên mã RNA và liên kết hydrogen làm vững chắc thêm các dạng cặp base sau đây : A-U ; C-G và cũng như A-T (liên kết của A trong sợi RNA với T trong sợi DNA ; mục 2.3.3 và hình 2.11).

Bảng 2.3. Bắt cặp codon – đổi codon cho phép bắt cặp đôi lỏng lẻo (wobble) tại vị trí base thứ ba của codon.

Base ở đầu 5' của codon đối tRNA	Base nhận biết ở đầu 3' của mRNA codon
A	chỉ có U
A	A
A	C hoặc U
A	A hoặc G

RNA sợi kép. Các bộ genome của một số virus có sợi kép RNA-RNA nhưng ngoài ra các sợi kép RNA-RNA tạm thời hình thành trong tất cả các tế bào trong khi tổng hợp RNA và biểu hiện gene. Cắt gọt RNA cần nhận biết biên exon-intron tiếp theo sau liên kết hydro gen giữa các thế phiên mã RNA không cắt gọt và các phân tử RNA nhân tế bào nhỏ (xem 2.4.1). Ngoài ra, nhận biết codon – đổi codon liên quan tới liên kết hydro giữa các phân tử RNA : mRNA và tRNA (xem 2.5.1 và hình 2.19). Liên kết hydro giữa các phân tử RNA tham gia vào bắt cặp base không những A-U và C-G mà còn G-U (bảng 2.3).

Liên kết hydrogen nội phân tử trong các nucleic acid

Liên kết này quan trọng để tạo ra cấu trúc bậc hai trong các phân tử DNA và RNA, như trong trường hợp của sự hình thành các kẹp tóc trong DNA và các cánh tay phức hợp của tRNA (mục 2.2.1 và hình 2.6). Bắt cặp base có thể gồm các cặp base A-U, G-C và cả G-U.

Liên kết hydrogen nội phân tử trong các protein

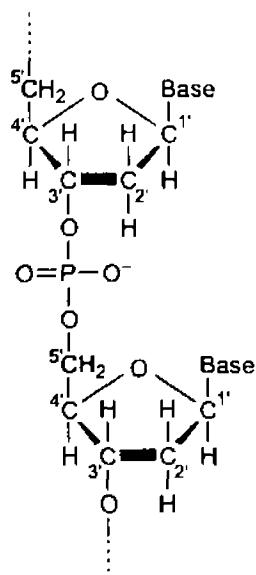
Một số đơn vị cơ bản của cấu trúc bậc hai protein, như α -helix, dải gấp β và vòng xoắn β được hình thành bởi liên kết hydrogen nội bộ chuỗi (mục 2.5.5 và hình 2.23).

2.2. CẤU TRÚC VÀ SAO CHÉP DNA

2.2.1. Cấu trúc của DNA : một helix kép đối song

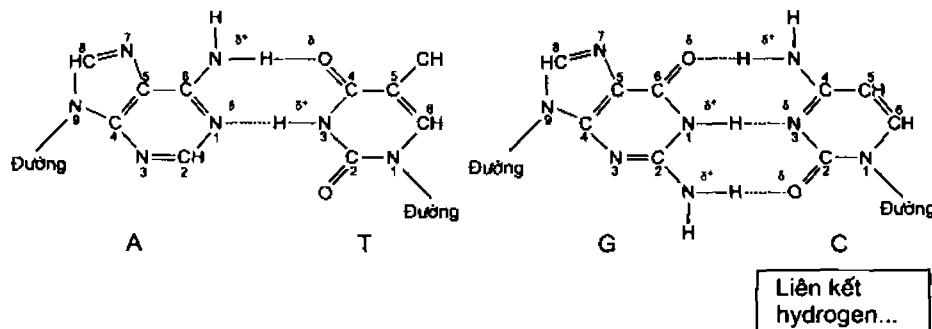
Cột trụ mạch thẳng của phân tử DNA và của phân tử RNA bao gồm các gốc đường và nhóm phosphate xen kẽ. Trong mỗi trường hợp, mối liên kết nối một gốc đường riêng lẻ với các gốc đường bên cạnh là liên kết phosphodiester 3'-5'. Điều này có nghĩa là một nhóm phosphate liên kết nguyên tử carbon 3' của đường với nguyên tử carbon 5' của đường tiếp theo (hình 2.3).

Trong khi các phân tử RNA trong một tế bào thường tồn tại ở dạng các phân tử đơn, cấu



Hình 2.3. Liên kết phosphodiester 3'-5'.

trúc DNA là một helix sợi kép, ở đó hai phân tử DNA (các sợi DNA) được giữ với nhau bằng các liên kết hydrogen yếu để hình thành một DNA sợi kép. Liên kết hydrogen tồn tại giữa các base đối bên, các cặp base của hai sợi của DNA sợi kép theo nguyên lý Watson-Crick : adenine (A) liên kết đặc hiệu với thymine (T) và cytosine (C) liên kết đặc hiệu với guanine (G) (hình 2.4). Kết quả là thành phần base của DNA từ các nguồn tế bào khác nhau không ngẫu nhiên : hàm lượng của adenine tương đương với hàm lượng của thymine và hàm lượng của cytosine tương đương với hàm lượng của guanine. Bởi vậy, thành phần base của DNA có thể được xác định rõ ràng qua định lượng thành phần %GC (= %G + %C). Ví dụ, nếu một nguồn DNA được định lượng là 42%GC, thành phần base có thể được suy ra là : G : 21% ; C : 21% ; A : 29% ; T : 29%.

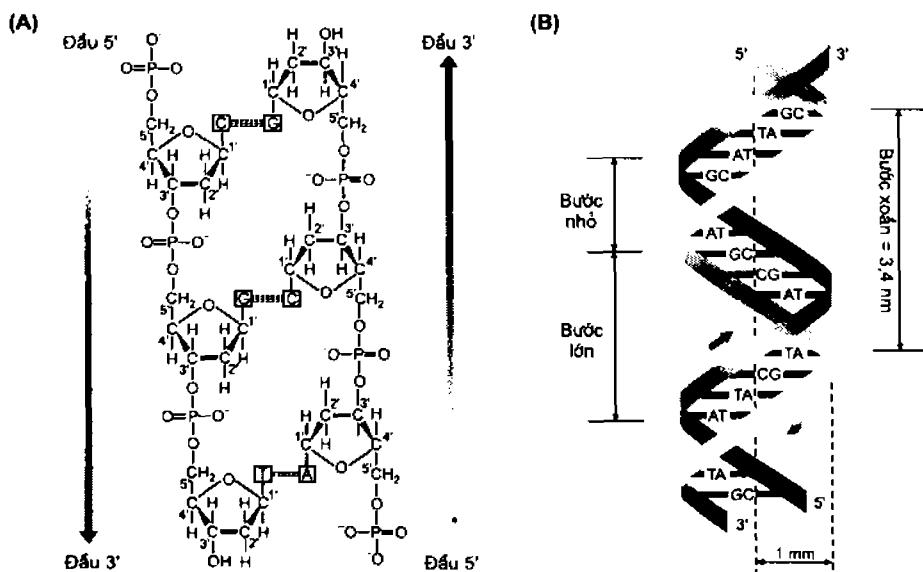


Hình 2.4. Các cặp base A-T và G-C lần lượt có hai và ba liên kết hydrogen.

Điện tích dương phần nhỏ trên các nguyên tử hydrogen và điện tích âm phần nhỏ trên các nguyên tử oxygen và nitrogen được biểu thị là δ^+ và δ^- , tương ứng.

DNA có thể chấp nhận các dạng khác nhau của cấu trúc xoắn. A-DNA và B-DNA là các helix xoáy phải (chúng cuộn xoắn theo chiều kim đồng hồ đối với người quan sát). Chúng lần lượt có 11 và 10 cặp base mỗi vòng xoay. Z-DNA là helix xoáy trái có 12 cặp base mỗi vòng xoay. Dưới các điều kiện sinh lý, hầu hết DNA trong genome vi khuẩn hoặc tế bào nhân chuẩn là dạng B-DNA, ở đó mỗi sợi xoắn có một bước xoắn *pitch* (khoảng cách tính cho một vòng xoắn đơn của helix) là 3,4 nm. Do các liên kết phosphodiester nối các nguyên tử carbon số 3' và số 5' của các gốc đường kế tiếp, một đầu của mỗi sợi DNA, được gọi là đầu 5', sẽ có một gốc đường cuối trong đó nguyên tử carbon số 5' không được nối

với gốc đường bên cạnh (hình 2.5). Đầu kia được định nghĩa là đầu 3' nhờ sự vắng mặt của liên kết phosphodiester ở nguyên tử carbon số 3' của gốc đường cuối cùng. Hai sợi của DNA sợi kép được gọi là đối song bởi vì chúng luôn luôn liên kết với nhau theo một cách mà hướng $5' \rightarrow 3'$ của một sợi DNA này đối lại với hướng của sợi kia (hình 2.5).



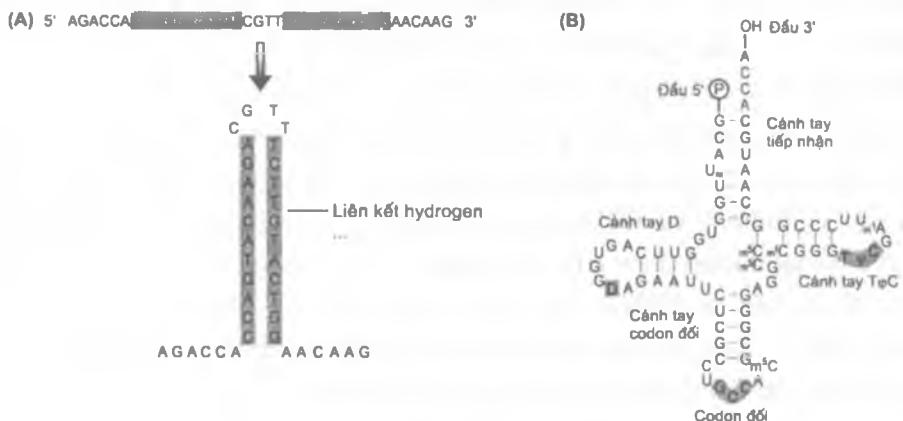
Hình 2.5. Cấu trúc DNA là một helix sợi kép đối song.

Hình 2.5 minh họa cấu trúc DNA là một helix sợi kép đối song. (A) Bản chất đối song của hai sợi DNA. Hai sợi đối song bởi vì chúng có chiều ngược nhau của liên kết giữa nguyên tử carbon số 3' với nguyên tử carbon số 5'. Cấu trúc trên là một trinucleotide sợi kép, trình tự của nó được trình bày : 5' pCpGpT-OH 3' (sợi DNA bên trái) / 5' pApCpG-OH 3' (sợi DNA bên phải) (ở đó p = liên kết phosphodiester và -OH = nhóm OH cuối cùng ở đầu 3'). Điều này được viết tắt thông thường bằng cách cắt các ký hiệu 'p' và 'OH' và cho trình tự chỉ trên một sợi (ví dụ trình tự có thể tương đương được trình bày bởi 5' CGT 3' hoặc 5' ACG 3'). (B) Cấu trúc helix đúp của DNA. Hai sợi được xoắn với nhau để tạo ra xoắn *plectonemic*. **Bước xoắn pitch** của mỗi helix là khoảng cách một vòng xoắn đơn gồm 10 nucleotide trong phân tử B-DNA.

Thông tin di truyền được mã hoá bởi trình tự mạch thẳng của các base trong sợi DNA (cấu trúc đầu tiên). Tiếp theo, hai sợi DNA của một

DNA sợi kép có các trình tự bổ trợ cho nhau (hoặc cho thấy tính bổ trợ base) và trình tự của các base của một sợi DNA có thể sẵn sàng được suy ra nếu biết trình tự DNA của sợi bổ trợ của nó. Bởi vậy, người ta thường mô tả một trình tự DNA bằng cách viết trình tự của các base của chỉ một sợi và theo chiều 5'→3'. Đây là chiều tổng hợp của các phân tử DNA mới trong quá trình sao chép DNA và cũng của sợi RNA mới sinh được tổng hợp trong quá trình phiên mã (xem phần dưới). Tuy nhiên, khi mô tả trình tự của một vùng DNA bao gồm hai base láng giềng (thực sự là một dinucleotide) trên một sợi DNA, thông thường chèn thêm một 'p' để biểu thị một liên kết phosphodiester nối nhau (ví dụ CpG có nghĩa là một cytidine được liên kết đồng hoá trị với một guanosine bên cạnh trên cùng một sợi DNA, trong khi một cặp base CG có nghĩa là một cytosine trên một sợi DNA được liên kết hydrogen với một guanine trên sợi bổ trợ, hình 2.5).

Mỗi liên kết hydrogen nội phân tử cũng tạo ra các phân tử RNA-DNA và RNA sợi kép, rất quan trọng để biểu hiện gene. Ngoài ra, liên kết hydrogen có thể xảy ra giữa các base nội trong một phân tử DNA hoặc phân tử RNA. Các trình tự có đoạn lặp lại ngược bổ trợ vị trí gần nhau có thể tạo ra các cấu trúc kép tóc hoặc các thòng lọng, được bền hoá bằng liên kết hydrogen giữa các base ở cổ chai thòng lọng (hình 2.6A). Những liên kết này cùng với cấu trúc bậc một sẽ tạo ra các cấu trúc bậc hai của phân tử. Các phân tử RNA nhất định, như RNA vận chuyển (tRNA), có mức độ đặc biệt cao về cấu trúc bậc hai (hình 2.6B).



Hình 2.6. Liên kết hydrogen nội phân tử trong DNA và RNA.

Hình 2.6 minh họa liên kết hydrogen nội phân tử trong DNA và RNA.

(A) Sự hình thành một thòng lọng kép tóc sợi kép nội trong một sợi DNA đơn. Các trình tự dược trình bày trong sợi DNA ở trên đầu đại diện cho các trình tự lặp lại ngược chúng có thể liên kết hydrogen để tạo ra cấu trúc kép tóc ở dưới. (B) RNA vận chuyển (tRNA) có cấu trúc bậc hai rộng rãi. Ví dụ chỉ ra là một gene người tRNA^{Glu}. Các nucleotide hiếm gặp như : d,5,6-dihydrouridine ; ψ, pseudouridine (5-ribosyl uracil) ; m⁵C, 5-methylcytidine ; m¹A, 1-methyladenosine. Cấu trúc lá cỏ tam diệp được bền hoá bởi liên kết hydrogen nội phân tử rộng rãi, với các cặp base thông thường G-C, A-U và cả cặp base hiếm G-U. Bốn cánh tay được công nhận : cánh tay tiếp nhận có thể gắn một amino acid vào đầu 3' ; cánh tay TψC được xác định bởi trinucleotide này ; cánh tay D chứa các gốc dihydrouridine và cánh tay codon đối chứa một codon trinucleotide đối ở trung tâm của thòng lọng. Cấu trúc bậc hai của các tRNA thực sự không thay đổi : luôn luôn có bảy cặp base trong thân của cánh tay tiếp nhận, năm cặp trên cánh tay TψC, năm cặp trên cánh tay codon đối và ba hoặc bốn cặp trên cánh tay D.

Trong trường hợp liên kết hydrogen ở sợi kép RNA-RNA và liên kết hydrogen nội phân tử ở phân tử RNA, đôi khi cặp base GU dược tìm thấy (hình 2.6B). Kiểu bắt cặp này đặc biệt không bền nhưng không phá hỏng đáng kể sợi kép RNA-RNA.

2.2.2. Tính bán bảo thủ và bán liên tục

Trong quá trình tổng hợp DNA (sao chép DNA), hai sợi DNA của mỗi nhiễm sắc thể được tháo xoắn bởi một enzyme helicase và mỗi sợi DNA chỉ đạo sự tổng hợp của sợi DNA bổ trợ để tạo ra hai sợi kép DNA con, mỗi sợi đều giống hệt phân tử gốc (hình 2.7). Do mỗi sợi kép DNA con chứa một sợi từ phân tử gốc và một sợi DNA mới được tổng hợp, quá

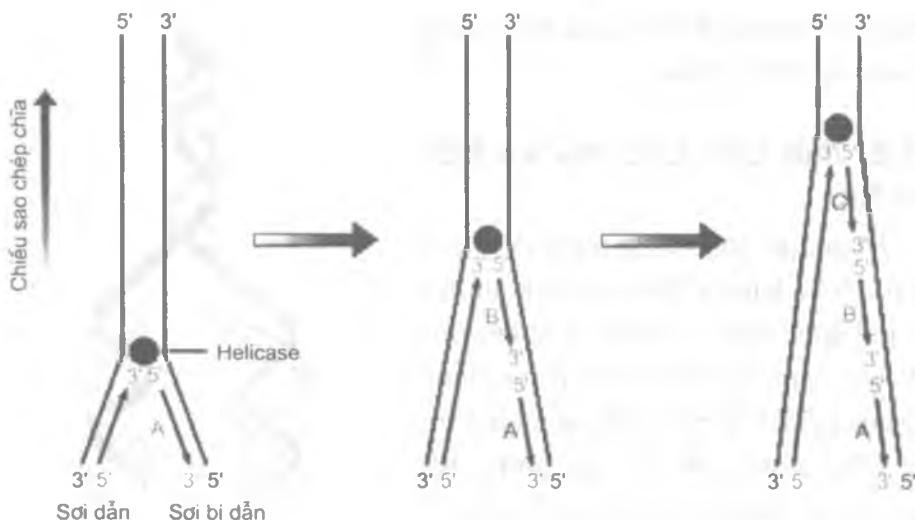


Hình 2.7. Sao chép DNA là bán bảo thủ.

trình sao chép được mô tả là bán bảo thủ. Enzyme DNA polymerase xúc tác phản ứng tổng hợp các sợi DNA mới, sử dụng bốn loại deoxynucleoside triphosphate (dATP, dCTP, dGTP, dTTP) là các tiền thân nucleotide.

Hình 2.7 mô tả quá trình sao chép DNA là bán bảo thủ. Sợi kép DNA gốc bao gồm hai sợi DNA đối song, bổ trợ cho nhau (tô đậm), nó tháo xoắn và sau đó hoạt động riêng làm khuôn cho tổng hợp các sợi DNA đối song và bổ trợ mới (không tô). Mỗi sợi kép DNA con chứa một sợi DNA gốc và một sợi DNA mới, hình thành một sợi kép DNA có cấu trúc giống hệt sợi kép DNA gốc. Hình này chỉ cho thấy kết quả của việc nhân đôi DNA nhưng không cho thấy quá trình xảy ra như thế nào (xem hình 2.8).

Sao chép DNA được khởi động ở những điểm đặc biệt, được gọi là những điểm khởi động sao chép. Bắt đầu từ một điểm khởi động như thế, sự khởi động sao chép DNA cho kết quả là một cài chia sao chép, ở đó, sợi kép DNA gốc rẽ đôi thành hai sợi kép DNA con. Do hai sợi của sợi kép DNA gốc là đối song nhưng từng sợi đơn lẻ hoạt động làm khuôn để tổng hợp sợi con đối song và bổ trợ, nó tuân theo nguyên tắc là hai sợi con phải chạy theo hai chiều đối nhau, nghĩa là chiều phát triển sợi phải là $5' \rightarrow 3'$ đối với một sợi con, sợi dẫn (leading strand) nhưng $3' \rightarrow 5'$ đối với sợi con khác, sợi bị dẫn (lagging strand) (hình 2.8).



Hình 2.8. Bất đối xứng của tổng hợp sợi trong khi sao chép DNA.

Hình 2.8 mô tả tính bất đối xứng của tổng hợp sợi trong khi sao chép DNA. Các sợi gốc nằm ngoài, các sợi mới tổng hợp nằm trong. Helicase tháo xoắn DNA sợi kép để cho phép từng sợi đơn lẻ được sao chép. Khi sao chép tiến triển, chiều 5'→3' tổng hợp của sợi dẫn cùng chiều với chiều của cái chia chuyển động và như vậy, tổng hợp có thể tiếp tục. Tổng hợp 5'→3' của sợi bị dẫn, tuy nhiên, là theo chiều ngược lại với chiều chuyển động của chia sao chép. Nó cần được tổng hợp thành từng đoạn một (phân đoạn Okazaki), trước tiên A, sau đó B, sau đó C tiếp theo chúng được gắn lại bằng DNA ligase để tạo ra sợi DNA liên tục. Tổng hợp những phân đoạn này được khởi động sử dụng một mồi RNA. Trong các tế bào eukaryote, các sợi dẫn và sợi bị dẫn được tổng hợp, tương ứng, bằng các DNA polymerase δ và α (xem bảng 2.4).

Bảng 2.4. Nǎm loại cùa DNA polymerase động vật có vú.

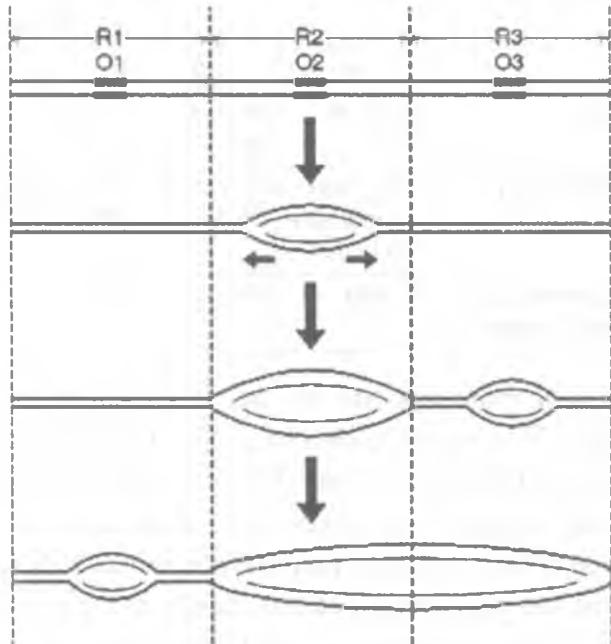
	Loại				
	α	α	ε	α	ε
Vị trí	Nhân tế bào	Nhân tế bào	Ty thể	Nhân tế bào	Nhân tế bào
Chức năng	Tổng hợp và mồi cùa sợi bị dẫn	Sửa chữa DNA	Sao chép DNA ty thể	Tổng hợp sợi dẫn	Sửa chữa DNA
Hoạt tính 3'→5' Exonuclease	Không	Không	Không	Không	Không

Các phản ứng xúc tác bởi DNA polymerase tham gia bổ sung vào nhóm 3' hydroxyl tự do của chuỗi DNA phát triển, một nửa dNMP được cung cấp bởi các tiền thân dNTP (hai gốc phosphate ngoại biên của dNTP – đó là các gốc β và γ (hình 2.1) – được cắt ra và nhóm pyrophosphate thu được (PPi) bị bỏ đi). Yêu cầu này tạo ra một sự bất đối xứng trong quá trình sao chép DNA : chỉ có sợi dẫn sẽ có một nhóm 3' hydroxyl tự do ở điểm rẽ đồi. Điều này sẽ cho phép bổ sung tiếp của các nucleotide và kéo dài tiếp tục theo cùng chiều, trong đó chia sao chép chuyển động. Tuy nhiên, tổng hợp của sợi bị dẫn phải được thực hiện như là một chuỗi luỹ tiến của các phân đoạn nhỏ (thường dài 100-1000 nucleotide), thường được gọi là các phân đoạn Okazaki. Do chí sợi dẫn được tổng hợp liên tục, tổng hợp sợi DNA được gọi là bán không liên tục. Mỗi phân đoạn của sợi bị dẫn được tổng hợp theo chiều 5'→3', chúng sẽ theo chiều ngược lại với

chiều mà chia sao chép chuyển động. Các phân đoạn được tổng hợp thành công nối với nhau đồng hoá trị ở các đầu của chúng sử dụng enzyme DNA ligase và như vậy bảo đảm phát triển chuỗi theo hướng chuyển động của chia sao chép (hình 2.8).

Năm loại DNA polymerase của thú có vú được biết đến, bao gồm một polymerase được dành cho sao chép genome ty thể (bảng 2.4). Trong các nhiễm sắc thể của động vật có vú, sao chép DNA tiến theo hai chiều, tạo ra bâu sao chép từ nhiều điểm khởi động sao chép. Khoảng cách giữa các điểm khởi động sao chép liền kề nhau là khoảng 50-300 kb, một khoảng cách quan trọng trong cấu trúc nhiễm sắc thể. Ở các điểm khởi động khác nhau, sao chép DNA được bắt đầu ở những thời gian khác nhau ở pha S của chu kỳ tế bào nhưng các bâu sao chép liền kề nhau cuối cùng sẽ dung hợp với nhau (hình 2.9). Sao chép DNA là tốn thời gian : các tế bào người trong dịch nuôi cấy cần khoảng 8 giờ để hoàn thiện quá trình.

Hình 2.9 mô tả các nhiễm sắc thể của toàn bộ cơ thể sinh vật có nhiều điểm khởi động sao chép. R1, R2 và R3 biểu thị các đơn vị sao chép kề nhau (replicon) nằm trên cùng một nhiễm sắc thể và có các điểm khởi động sao chép định vị bên trong O1, O2 và O3, tương ứng. Trong ví dụ này, sao chép hai chiều được nhìn thấy tiến triển ban đầu từ O2, sau đó O3 và cuối cùng O1. Sợi ở đáy cho thấy sự dung hợp hai bâu sao chép khởi điểm từ O2 và O3, trước khi R1 replicon kết thúc sao chép.



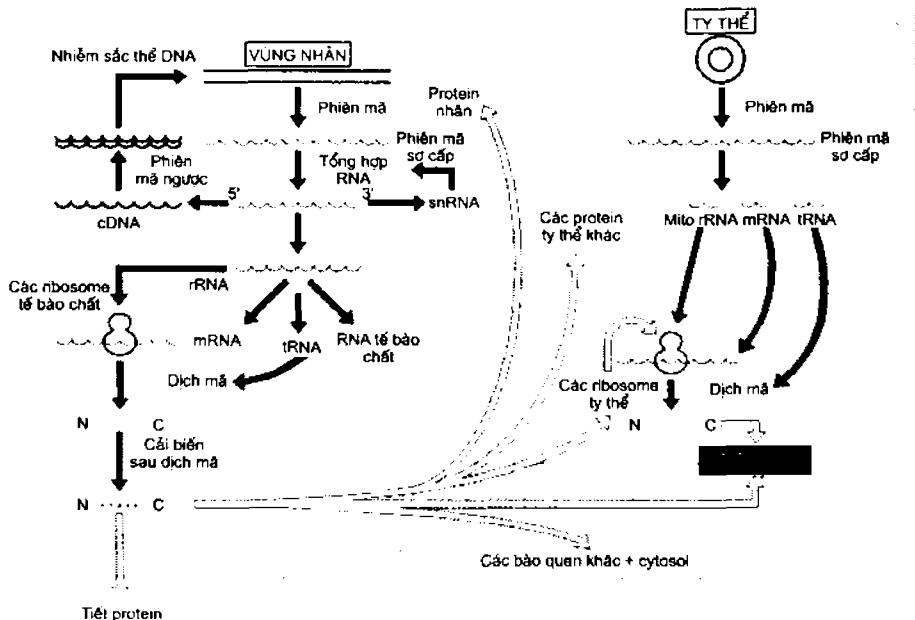
Hình 2.9. Các điểm khởi động sao chép trên nhiễm sắc thể của một sinh vật.

2.3. PHIÊN MÃ RNA VÀ BIỂU HIỆN GENE

2.3.1. Dòng chảy thông tin di truyền

Dòng chảy thông tin di truyền hầu hết theo một con đường : DNA → RNA → protein. Biểu hiện thông tin di truyền trong tất cả các tế bào là một hệ thống một đường rất rộng rãi : DNA định rõ sinh tổng hợp RNA và RNA định rõ tổng hợp polypeptide, sau đó hình thành protein. Do tính phổ biến của nó, dòng chảy DNA → RNA → polypeptide (protein) của thông tin di truyền được mô tả là một thuyết trung tâm (central dogma) của sinh học phân tử. Bước đầu tiên, tổng hợp RNA sử dụng một RNA polymerase phụ thuộc vào DNA, được mô tả là phiên mã và xảy ra trong nhân tế bào của các tế bào eukaryote và tới một phạm vi giới hạn, trong ty thể và lạp lục, chỉ các bào quan khác chúng có một khả năng di truyền bổ sung vào nhân tế bào (xem hình 2.10). Bước thứ hai, tổng hợp polypeptide, được mô tả là dịch mã và xảy ra trong các ribosome, các phức hợp RNA-protein lớn được tìm thấy trong tế bào chất cũng như trong ty thể và lạp lục. Các phân tử RNA mà chúng chỉ định polypeptide được biết đến là RNA thông tin (messenger RNA, mRNA). Biểu hiện thông tin di truyền tuân theo một nguyên lý cộng tuyến : trình tự tuyến tính của nucleotide trong DNA được giải mã để cho một trình tự tuyến tính của các nucleotide trong RNA mà chúng được giải mã đến lượt trong các nhóm của nucleotide (*codon*) để cho một trình tự tuyến tính của amino acid trong sản phẩm polypeptide.

Hình 2.10 mô tả quá trình biểu hiện gene trong một tế bào động vật. Một phần nhỏ của các phân tử RNA nhân tế bào có thể được chuyển thành cDNA tự nhiên bằng các reverse transcriptase được mã hoá bởi virus và tế bào và sau đó tích hợp vào DNA nhiễm sắc thể ở các vị trí khác nhau. Ty thể cũng tự tổng hợp rRNA và tRNA của nó số một số protein chúng tham gia trong hoạt tính phosphoryl hoá oxy (Ox. phosph.). Tuy nhiên, các protein của ribosome ty thể và phân lớn protein trong hệ thống phosphoryl hoá oxy ty thể và các protein ty thể khác được mã hoá bởi các gene trong nhân tế bào và được dịch mã trên các ribosome tế bào chất, trước khi được đưa sang ty thể.



Hình 2.10. Biểu hiện gene trong một tế bào động vật.

Mặc dù DNA là vật liệu di truyền trong tất cả các tế bào ngày nay nhưng trong giai đoạn đầu của quá trình tiến hóa, RNA đảm nhiệm chức năng đó. Các phân tử RNA có thể giống như DNA, trải qua quá trình tự sao chép. Tuy nhiên, nhóm 2' hydroxyl trên các gốc ribose của RNA tạo ra các liên kết đường-phosphate tương đối không bền hóa học. Trong DNA, các gốc deoxyribose chỉ mang các nguyên tử hydrogen ở vị trí 2' và như vậy DNA phù hợp hơn nhiều so với RNA là một thể mang bền vững của thông tin di truyền. Nhiều loại virus khác nhau ngày nay có một genome cấu trúc từ RNA mà không phải là DNA. Các retrovirus như HIV là một loài phụ của các RNA virus, trong đó RNA sao chép nhờ một thể trung gian DNA, sử dụng reverse transcriptase, một DNA polymerase phụ thuộc vào RNA. Ngày nay, rõ ràng rằng các tế bào eukaryote, bao gồm các tế bào động vật có vú, chứa các trình tự DNA nhiễm sắc thể không phải virus, chúng mã hóa các reverse transcriptase tế bào (tham khảo đối với các ví dụ trong genome người). Bởi vì một số trình tự RNA không phải virus hoạt động làm khuôn cho tổng hợp DNA tế bào, nguyên lý của dòng chảy đa chiều của thông tin di truyền không còn hoàn toàn giá trị.

2.3.2. Một DNA biểu hiện một protein hoặc RNA

Chỉ một phần nhỏ của tất cả DNA trong các tế bào được phiên mã. Theo nhu cầu, các tế bào khác nhau phiên mã các vùng khác nhau của DNA (đơn vị phiên mã) chúng là những đơn vị rời rạc, được đặt cách nhau không đều đọc theo trình tự DNA. Tuy nhiên, phần lớn DNA tế bào không bao giờ được phiên mã trong mọi tế bào.

Hơn nữa, chỉ một phần của RNA được tạo ra bằng phiên mã được dịch mã thành polypeptide. Điều này do :

Một số đơn vị phiên mã được biểu hiện để cho một phân tử RNA khác với mRNA và cũng chỉ định trực tiếp các polypeptide. Thay vì các sản phẩm RNA là những phân tử trưởng thành, chúng có thể phục vụ các chức năng khác nhau, như trong trường hợp của ribosomal RNA (rRNA), tRNA và các phân tử RNA nhân tế bào nhỏ (sn) và tế bào chất đa dạng ;

Thể phiên mã sơ cấp (sản phẩm phiên mã đầu tiên) của các đơn vị phiên mã, chúng mã hoá polypeptide là đối tượng cho các sự kiện quá trình RNA. Như là một kết quả, nhiều trình tự RNA ban đầu bị bỏ đi để tạo ra một mRNA nhỏ hơn nhiều (xem 2.4.1) ;

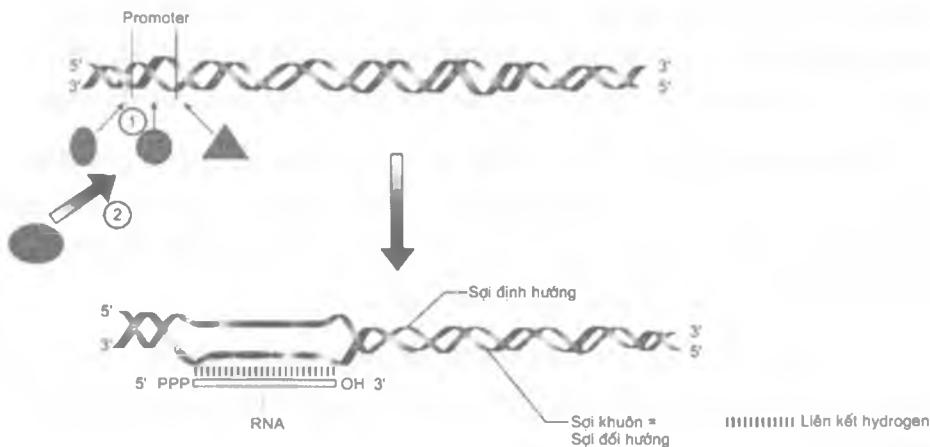
Chỉ một phân trung tâm của mRNA trưởng thành được dịch mã ; các phân của độ dài thay đổi ở mỗi đâu mRNA giữ nguyên không dịch mã (xem 2.5.1).

Trong các tế bào động vật, DNA được tìm thấy trong cả nhân tế bào lẫn ty thể. Tuy nhiên, các ty thể chỉ có một phần nhỏ của DNA tế bào tổng số và một số gene giới hạn ; phần lớn DNA của một tế bào được định vị trên nhiễm sắc thể của nhân tế bào.

Phần của DNA mã hoá trong các genome của các eukaryote phức tạp là tương đối nhỏ. Điều này phản ánh là kết quả của sự không mã hoá của nhiều trình tự nội trong các gene. Một nguyên nhân khác là một phần đáng kể của genome của các eukaryote phức tạp chứa những trình tự lặp lại không chức năng hoặc không được phiên mã thành RNA. Những trình tự này bao gồm các bản sao khuyết của các gene chức năng (các gene giả pseudogene và các phân đoạn gene) và DNA không mã hoá lặp lại cao.

2.3.3. Phiên mã DNA (gene) chỉ định tổng hợp RNA

Phiên mã là quá trình ở đó thông tin di truyền trong một số phân đoạn DNA (gene) chỉ định tổng hợp RNA. Tổng hợp RNA được thực hiện qua việc sử dụng một enzyme RNA polymerase, với DNA làm khuôn và ATP, CTP, GTP và UTP làm tiền thân RNA. RNA được tổng hợp là một sợi đơn, với chiều phiến mã $5' \rightarrow 3'$. Kéo dài chuỗi xảy ra bằng cách bổ sung gốc ribonucleoside monophosphate thích hợp (AMP, CMP, GMP hoặc UMP) vào nhóm 3' hydroxyl tự do và ở đầu 3' của chuỗi RNA phát triển. Các nucleotide như thế được dẫn xuất bằng cách cắt đi gốc pyrophosphate (PPi) từ các tiền thân ribonucleoside triphosphate (rNTP) thích hợp. Điều này nghĩa là nucleotide ở đầu 5' ngoại lệ (*nucleotide khởi động*) sẽ khác với tất cả các nucleotide khác trong chuỗi bằng cách mang một nhóm 5' phosphate.



Hình 2.11. Quá trình phiên mã DNA sang RNA.

Thông thường, chỉ một trong hai sợi DNA hoạt động làm khuôn cho tổng hợp RNA. Trong quá trình phiên mã, DNA sợi kép được tháo xoắn và sợi DNA sẽ hoạt động làm khuôn cho tổng hợp RNA hình thành một thể lai RNA-DNA sợi kép tạm thời với chuỗi RNA phát triển. Do thể phiên mã RNA là hỗ trợ cho sợi khuôn này, thể phiên mã có cùng chiều $5' \rightarrow 3'$ và trình tự base (trừ U thay thế cho T) như sợi đối diện, không làm khuôn của sợi kép. Đối với lý do này sợi không làm khuôn thường được gọi là sợi định hướng (sense strand) và sợi khuôn thường được gọi là sợi đối hướng (antisense strand) (hình 2.11). Trong dãy chứng tài liệu, các trình tự gene, theo thói quen chỉ cho thấy trình tự DNA của sợi định hướng. Sự định hướng của các trình tự tương đối

với một trình tự gene thường được bức chế bởi sợi định hướng và bởi chiều của phiên mã (ví dụ đầu 5' của gene quy cho các trình tự ở đầu 5' của sợi định hướng và các trình tự ngược dòng hoặc xuôi dòng của một gene quy cho các trình tự chúng nằm kề biên gene ở các đầu 5' và 3', tương ứng, của sợi định hướng).

Hình 2.11 minh họa quá trình RNA được phiên mã là một sợi đơn lẻ, nó bổ trợ theo trình tự base với một sợi (sợi khuôn) của một gene.

Các yếu tố phiên mã (transcription factor, TF) khác nhau được cần đến để liên kết với một trình tự promoter ở vùng lân cận ngay của gene 1 và dẫn RNA polymerase 2 để phiên mã gene. Tổng hợp chuỗi được khởi động với một nucleoside triphosphate và kéo dài chuỗi xảy ra bởi bổ sung liên tục của các gốc nucleoside monophosphate được cung cấp bởi rNTP và 3' OH. Điều này có nghĩa rằng đầu 5' sẽ có một nhóm triphosphate, nó có thể chịu cải biến sau đó (ví dụ bằng gắn mũ vào ; xem 2.4.2) và đầu 3' sẽ có một nhóm hydroxyl tự do. Trình tự RNA thường sẽ giống hệt sợi định hướng của gene (trừ U thay cho T) và bổ trợ cho sợi khuôn.

Trong các tế bào eukaryote, ba phân tử RNA polymerase khác nhau được yêu cầu để tổng hợp các loại RNA khác nhau (bảng 2.5). Phân lớn các gene tế bào mã hoá các polypeptide và được phiên mã bởi RNA polymerase II. Tầm quan trọng tăng lên đối với các gene mã hoá RNA như sản phẩm trưởng thành của chúng : Các phân tử RNA chức năng hiện nay được biết đến đóng một vai trò rộng và các chức năng xúc tác đang được gán cho một số trong chúng.

Bảng 2.5. Ba loại RNA polymerase của tế bào nhân chuẩn.

Loại	Gene được phiên mã	Nhận xét
I	28S rRNA ; 18S rRNA ; 5.8S rRNA	Khu trú ở nhân tế bào. Một thể phiên mã sơ cấp đơn lẻ (45S rRNA) được cắt ra thành ba loại rRNA như ghi ở bên
I	Tất cả các gene mã hoá polypeptide ; hầu hết là các gene snRNA	Các thể phiên mã polymerase II là duy nhất được gắn mũ và polyadenyl hoá
III	5S rRNA ; tRNA gene ; U6 snRNA ; 7SL RNA ; 7SK RNA ; 7SM RNA	Promoter cho một số gene được phiên mã bởi RNA polymerase III (ví dụ 5S rRNA, tRNA, 7SL RNA) là nội phân tử đối với gene (xem hình 2.12) và cho các gene khác (ví dụ 7SK RNA) khu trú ngược dòng

2.3.4. Phiên mã các gene eukaryote

Phiên mã của các gene eukaryote cần mối tương tác giữa các yếu tố phiên mã hoạt động cis và các yếu tố phiên mã hoạt động trans. RNA polymerase tế bào nhận chuẩn không thể khởi động phiên mã bởi chính nó. Thay vào đó, các tổ hợp của các yếu tố trình tự ngắn trong các vùng phụ cận ngay bên của một gene hoạt động như là các tín hiệu nhận biết cho các yếu tố phiên mã liên kết với DNA để dẫn và hoạt hoá polymerase. Một nhóm chính của các yếu tố trình tự ngắn như thế thường cụm lại ngược dòng trình tự mã hoá của gene, ở đó chúng cùng nhau thiết lập promoter. Sau một số các yếu tố phiên mã chung liên kết với vùng promoter, một RNA polymerase liên kết với phức hợp yếu tố phiên mã và được hoạt hoá để khởi động tổng hợp RNA từ một vị trí duy nhất. Các yếu tố phiên mã được gọi là hoạt động trans (trans-acting), bởi vì chúng được tổng hợp bởi các gene định vị xa và chúng di chuyển tới các vị trí hoạt động của chúng. Ngược lại, các yếu tố promoter là hoạt động cis (cis-acting); chức năng của chúng là bị giới hạn đối với DNA sợi kép, ở đó, chúng khu trú (bảng 2.6).

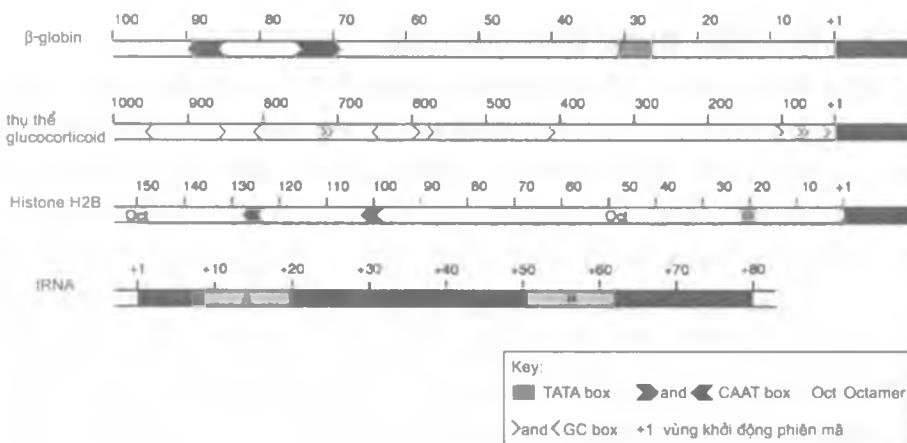
Bảng 2.6. Các ví dụ của các yếu tố hoạt động cis được nhận dạng bởi các yếu tố phiên mã đa dạng.

Yếu tố cis	Trình tự DNA giống hệt hoặc khác	Các yếu tố hoạt động trans liên kết	Nhận xét
Hộp GC	GGGCGG	Spl	Yếu tố Spl có mặt khắp nơi
Hộp TATA	TATAAA	TFIID	TFIIB liên kết với phức hợp TFIID-TATA box để làm bền
Hộp CAAT	CCAAT	Nhiều, ví dụ C/EBP, CTF/NFI	Họ lớn của các yếu tố hoạt hoá trans
CRE (yếu tố đáp ứng cAMP)	GTGACGTA/CAA/G	Họ CREB/ATF, ví dụ ATF-1	Gene hoạt hoá để đáp ứng với cAMP

Trong trường hợp của các gene được phiên mã tích cực bởi RNA polymerase II hoặc ở các bước đặc biệt trong chu kỳ tế bào (ví dụ histone) hoặc trong các dạng tế bào đặc biệt (ví dụ β -globin), các yếu tố promoter luôn bao gồm một hộp TATA, thường là TATAAA hoặc

một biến đổi, ở một vị trí khoảng 25 bp ngược dòng (-25) từ vị trí khởi động phiên mã (hình 2.12). Đột biến ở các yếu tố TATA không ngăn cản sự khởi động phiên mã nhưng gây ra điểm khởi động của phiên mã bị thay thế từ vị trí bình thường.

Hình 2.12 mô tả các promoter tế bào nhân chuẩn bao gồm một tập hợp các yếu tố trình tự ngắn bảo thủ được khu trú ở các khoảng cách tương đối không đổi từ điểm khởi động phiên mã. Các hướng khác nhau cho các yếu tố hộp GC và CAAT được chỉ ra theo hình V : > = hướng bình thường ; < = hướng ngược. Gene thụ thể glucocorticoid bất thường chứa 13 hộp GC ngược dòng (10 theo hướng bình thường ; ba theo hướng ngược lại). Các gene tRNA được phiên mã bởi RNA polymerase III và có một promoter chia đôi bên trong gồm yếu tố A (thường nội trong các nucleotide đánh số +8 đến +19 theo hệ thống đánh số tRNA nucleotide chuẩn) và yếu tố B (thường giữa các nucleotide +52 và +62). Các yếu tố phiên mã đặc biệt liên kết với các yếu tố này và dẫn RNA polymerase III để khởi động phiên mã ở +1.



Hình 2.12. Các promoter tế bào nhân chuẩn.

Các promoter của nhiều gene khác, bao gồm các gene quản gia (housekeeping gene), thiếu các hộp TATA nhưng thường có một hộp GC, chứa các biến số của trình tự bảo thủ GGGCGG. Các yếu tố promoter thường khác bao gồm hộp CAAT, thường ở khoảng -80, nó thường là yếu tố quyết định mạnh nhất của hiệu quả promoter. Tuy nhiên, các hộp GC và CAAT xuất hiện có khả năng tạo chức năng trong hoặc định hướng,

mặc dù các trình tự của chúng là bất đối xứng (hình 2.13). Ngoài các yếu tố phiên mã ngược dòng chung, chúng được nhận biết bởi các yếu tố phiên mã thường gặp, nhiều yếu tố nhận biết đặc biệt khác cũng được biết đến, chúng được công nhận bởi các yếu tố phiên mã giới hạn mô.

Các thể tăng cường (enhancer) bao gồm các nhóm yếu tố trình tự ngăn hoạt động *cis*, chúng có thể tăng cường hoạt tính phiên mã của các gene tế bào nhân chuẩn đặc trưng. Tuy nhiên, không giống như các yếu tố promoter mà các vị trí của chúng tương đối với vị trí khởi động phiên mã, tương đối cố định (hình 2.12), các thể tăng cường được định vị một khoảng thay đổi, thường rất đáng kể từ vị trí khởi động phiên mã và chức năng của chúng là phụ thuộc vào chiêu của chúng. Chúng xuất hiện để liên kết với các protein điều hoà gene và tiếp theo, DNA giữa promoter và thể tăng cường móc lại, cho phép các protein liên kết với thể tăng cường để tương tác với các yếu tố phiên mã liên kết với promoter hoặc với RNA polymerase. Các thể hâm (silencer) là những yếu tố điều hoà tương đương, chúng có thể ức chế hoạt tính phiên mã của các gene đặc trưng.

2.3.5. Biểu hiện gene đặc hiệu mô

Biểu hiện gene đặc hiệu mô tham gia hoạt hoá chọn lọc của các gene đặc trưng và các vùng của chromatin hoạt động phiên mã chấp nhận một cấu trúc mở. Hàm lượng DNA của một dạng đặc trưng của tế bào nhân chuẩn, ví dụ myocyte, thực sự giống hệt của lymphocyte, hepatocyte hoặc bất kỳ dạng tế bào khác của tế bào có nhân từ cùng cơ thể sinh vật. Cái tạo ra sự khác nhau giữa các dạng tế bào là kiểu gene được biểu hiện trong tế bào. Một số tế bào, đặc biệt tế bào não, biểu hiện một số lớn các gene khác nhau. Trong nhiều dạng tế bào khác, một phần lớn các gene là bất hoạt phiên mã. Rõ ràng, các gene được biểu hiện là những gene định rõ chức năng của tế bào. Một số chức năng là những chức năng thông thường chúng cần thiết cho các chức năng tế bào chung và được chỉ định bởi các gene được gọi là các gene quản gia. Biểu hiện của các gene khác có thể bị giới hạn rộng vào một dạng tế bào đặc trưng (biểu hiện gene đặc trưng mô). Tuy nhiên, ngay cả trong trường hợp của các gene cho thấy tính đặc trưng mô quan trọng trong biểu hiện, một số thể phiên mã gene tồn tại ở mức độ rất thấp trong tất cả các dạng tế bào (phiên mã không đúng hoặc lạc vị trí).

Sự khác biệt giữa các vùng hoạt động và bất hoạt phiên mã của DNA trong một tế bào được phản ánh trong cấu trúc của chromatin liên quan. Chromatin bất hoạt phiên mã nói chung chấp nhận một cấu dạng cõ đặc cao và thường liên kết với các vùng của genome, chúng chịu sao chép muộn trong pha S của chu kỳ tế bào. Nó liên quan với liên kết chặt bởi phân tử histone H1. Ngược lại, DNA hoạt hoá phiên mã chấp nhận một cấu dạng mở hơn và thường được sao chép sớm ở pha S. Nó được đánh dấu bởi liên kết tương đối yếu bởi các phân tử histone H1 và acetyl hoá rộng rãi của bốn dạng histone nhân tế bào, ví dụ các histone H2A, H2B, H3 và H4. Ngoài ra, trong chromatin hoạt hoá phiên mã các vùng promoter của các gene động vật có xương sống được xác định tính chất chung bởi sự vắng mặt của các methylated cytosine được methyl hoá (xem phần sau). Các yếu tố phiên mã có thể thay thế các nucleosome và như vậy cấu dạng mở của chromatin hoạt hoá phiên mã có thể được phân biệt bằng thí nghiệm bởi vì nó cũng tạo ra lối vào nuclease : ở nồng độ rất thấp, enzyme DNase I sẽ cắt các vùng dài của DNA tự do không nucleosome. Mặc dù các vùng điều hoà có thể chứa một vài protein liên kết đặc trưng trình tự, cấu trúc chromatin mở được đánh dấu bởi sự có mặt của các vị trí quá nhạy cảm DNase I.

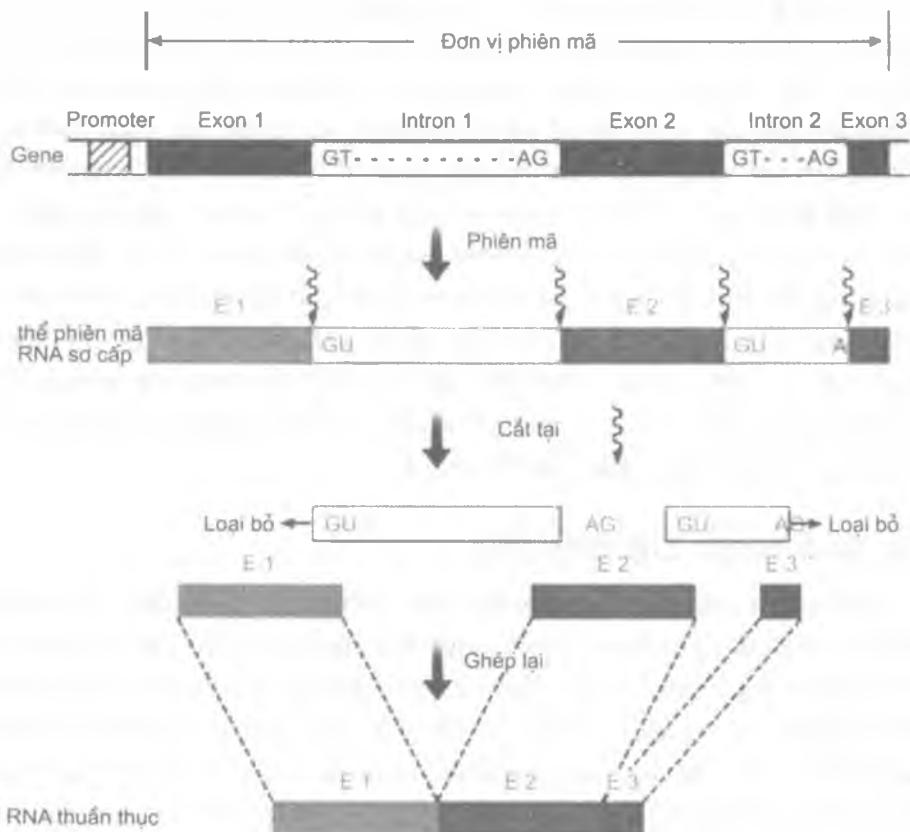
2.4. QUÁ TRÌNH CHẾ BIẾN RNA

Thể phiên mã RNA của hầu hết các gene tế bào nhân chuẩn chịu một loạt các phản ứng chế biến. Thường quá trình này bao gồm cắt các vùng nội bộ không mong muốn và gắn dính các vùng giữ lại với nhau (cắt dán RNA, RNA splicing). Ngoài ra, trong trường hợp của các thể phiên mã RNA polymerase II, một liên kết nucleotide chuyên biệt (7-methylguanosine triphosphate) được bổ sung vào đầu 5' của thể phiên mã sơ cấp (gắn mũ, capping) và adenyl hoá (AMP) các gốc được bổ sung tiếp theo vào đầu 3' của mRNA tạo ra đuôi poly(A) (polyadenyl hoá).

2.4.1. Cắt dán RNA

Cắt dán RNA bao gồm loại bỏ các trình tự RNA intronic từ thể phiên mã sơ cấp và dung hợp các trình tự RNA exonic RNA. Các trình tự mã hoá của hầu hết các gene động vật có xương sống, cả các gene mã hoá polypeptide lẫn các gene mã hoá các phân tử RNA khác mRNA đều được

cắt thành các phân đoạn exon, chúng được phân tách bởi các trình tự xen giữa không mã hoá (intron). Phiên mã bao gồm sản xuất trình tự RNA bổ trợ cho toàn bộ độ dài của gene, bao gồm cả các exon và intron. Tuy nhiên, thể phiến mã RNA chịu sự cắt dán RNA, một loạt các phản ứng xử lý ở đó các phân đoạn intronic RNA được cắt nhỏ và loại bỏ, các phân đoạn exonic RNA được gắn dính đầu này sang đầu kia (ghép nối, spliced) cho một sản phẩm RNA ngắn hơn (hình 2.13).



Hình 2.13. Quá trình cắt dán RNA.

Cơ chế ghép nối RNA phụ thuộc vào sự tương đồng của các trình tự nucleotide ở các mốc giới exon/intron (mỗi nối ghép). Trong trường hợp đặc biệt, nó phụ thuộc vào quy tắc GT-AG : các intron hầu như luôn khởi động với GT (hoặc thực sự GU ở mức độ RNA) và kết thúc với AG (xem hình 2.14).



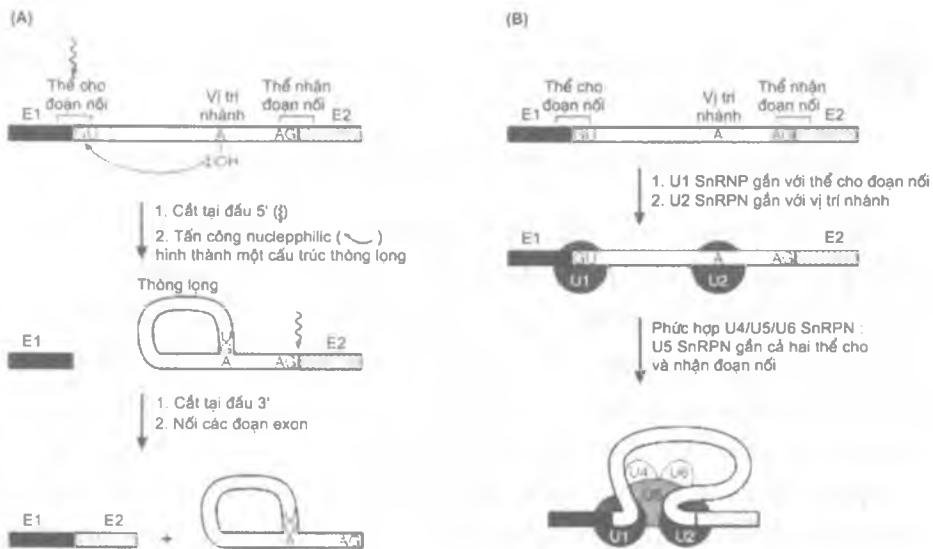
Hình 2.14. Các trình tự bảo thủ ở mức độ DNA đối với thể cho đoạn nối.

Hình 2.14 minh họa các trình tự bảo thủ ở mức độ DNA đối với thể cho đoạn nối, thể nhận đoạn nối và các vị trí nhánh trong các intron các eukaryote phức tạp. Các nucleotide nêu ra là hầu hết bất biến (các intron hiếm gặp cũng tồn tại ở đó thể cho đoạn nối bảo thủ dinucleotide GT được thay thế bởi AT và ở đó thể nhận đoạn nối bảo thủ dinucleotide AG được thay thế bởi AC. Các nucleotide khác đại diện phần lớn nucleotide được tìm thấy ở vị trí đặc biệt này. Trong trường hợp ưu tiên các pyrimidine (C/T hoặc T/C), thì vị trí base không quan trọng. Ví dụ, trình tự bảo thủ của vị trí nhánh được ghi để nêu bật độ tương tự với vị trí nhánh bảo thủ trong các intron nấm men (TACTAAC) nhưng trình tự được cho đổi với vị trí thể nhận đoạn nối không biểu thị sự ưu tiên C đổi với T.

Mặc dù dinucleotide bảo thủ GT (GU) và AG là chủ yếu quan trọng cho việc ghép nối nhưng tự chúng không đủ phát tín hiệu sự có mặt của một intron. So sánh các trình tự dẫn chứng đã hé lộ ra rằng các trình tự gần kề với GT và AG dinucleotide cho thấy một mức độ bảo thủ đáng kể (hình 2.14). Ngoài ra, một trình tự intronic bảo thủ thứ ba được biết đến là chức năng quan trọng trong việc ghép nối được gọi là vị trí nhánh (branch site) nó thường được khu trú rất gần đầu intron, ở khoảng nucleotide trước khi dinucleotide AG cuối cùng (hình 2.14).

Cơ chế ghép nối bao gồm trình tự sau đây :

- Cắt ở đầu 5' mỗi ghép nối ;
- Tán công nucleolytic bởi nucleotide G cuối cùng của vị trí thể cho đoạn nối bất biến A của vị trí nhánh để hình thành một cấu trúc dây thông lọng ;
- Cắt ở đầu 3' mỗi ghép nối, dẫn đến giải phóng intronic RNA như một thông lọng và ghép nối của các phân đoạn exonic RNA (hình 2.15).



Hình 2.15. Cơ chế của ghép nối RNA (GU-AG intron).

Hình 2.15 mô tả cơ chế của ghép nối RNA (GU-AG intron). (A) Cơ chế. Xem hình 2.14 đối với các trình tự bảo thủ của thể cho đoạn nối, thể nhận đoạn nối và vị trí nhánh. Tán công nucleophilic bao gồm nhóm 2' hydroxyl tham gia vào A bảo thủ ở vị trí nhánh và G của GU bảo thủ ở điểm khởi động của intron và kết quả một liên kết đồng hoá trị mới liên kết hai nucleotide này để cho một cấu trúc nhánh (thòng lọng). (B) Vai trò của snRNP. Các hạt ribonucleoprotein nhân tế bào nhỏ (small nuclear ribonucleoprotein particle, snRNP) là một phần của cơ quan ghép nối (spliceosome). U1 snRNA có một trình tự cuôi cùng bổ trợ cho trình tự bảo thủ thể cho đoạn nối và liên kết với nó bởi bắt cặp base RNA-RNA. Sau khi U1 snRNP đã liên kết, U2 snRNA nhận biết vị trí nhánh bằng một phản ứng bắt cặp base tương tự. Mỗi tương tác giữa các mối nối của thể cho đoạn nối và thể nhận đoạn nối được bền hoá bởi liên kết tiếp theo của một hạt đa snRNP được hình thành trước đó, chứa U4, U5 và U6 snRNA, với U5 snRNP có khả năng liên kết đồng thời với cả hai thể cho đoạn nối và thể nhận đoạn nối.

Các phản ứng trên đây được dàn xếp bởi một phức hợp RNA-protein lớn, được gọi là spliceosome, nó bao gồm năm dạng snRNA (RNA nhân tế bào nhỏ, small nuclear RNA) và hơn 50 protein (Stanley and Guthrie,

1998). Mỗi một phân tử snRNA được gắn chặt với các phân đoạn đặc trưng để hình thành các hạt snRNP và tính đặc hiệu của phản ứng ghép nối được thiết lập bởi bắt cặp base RNA-RNA giữa thể phiên mã RNA và các phân tử snRNA. Trong trường hợp phân lớn các intron trong các gene mã hoá các polypeptide, năm loại snRNA là : U1, U2, U4, U5 và U6 snRNA. Đầu 5' của U1 snRNA có một trình tự UACUUAC bắt cặp base với trình tự bảo thủ thể cho đoạn nối (GUAAGUA). Sau khi U1 snRNP đã liên kết, U2 snRNA nhận biết vị trí nhánh bởi một phản ứng bắt cặp base tương tự, tiếp theo, mối tương tác giữa U1 snRNP và U2 snRNP mang lại hai mối ghép nối gần nhau hơn. Sau đó, một hạt đa snRNP, chứa U4, U5 và U6 snRNAs, liên kết với phức hợp U1-U2 snRNP.

Một dạng thứ hai của spliceosome cũng được biết tới, nó chế biến một loại intron rất hiếm, được gọi là AT-AC intron ở đó các GT và AG dinucleotide bảo thủ được thay thế bởi AT và AC tương ứng. Trong trường hợp này U11 và U12 snRNA thay thế các chức năng của U1 và U2 (xem Tam và Steitz, 1997).

Spliceosome được nhìn thấy để hoạt động theo phương thức tiến lên : một vị trí ghép nối 5' được nhận biết, nó quét trình tự RNA cho tới khi nó gặp vị trí ghép nối 3' tiếp (nó được truyền tín hiệu như một đích bằng trình tự bảo thủ vị trí nhánh được định vị ngay trước nó). Tuy nhiên, trật tự trong đó các trình tự intronic được bỏ đi và các trình tự exonic kề bên của chúng được ghép nối không bị chi phối bởi trật tự tuyến tính của chúng trong thể phiên mã RNA ; thực sự, cấu dạng của RNA được cho là có ảnh hưởng tới tính đến gần của các vị trí nhánh 5'.

2.4.2. Hoạt động của thể phiên mã RNA polymerase II

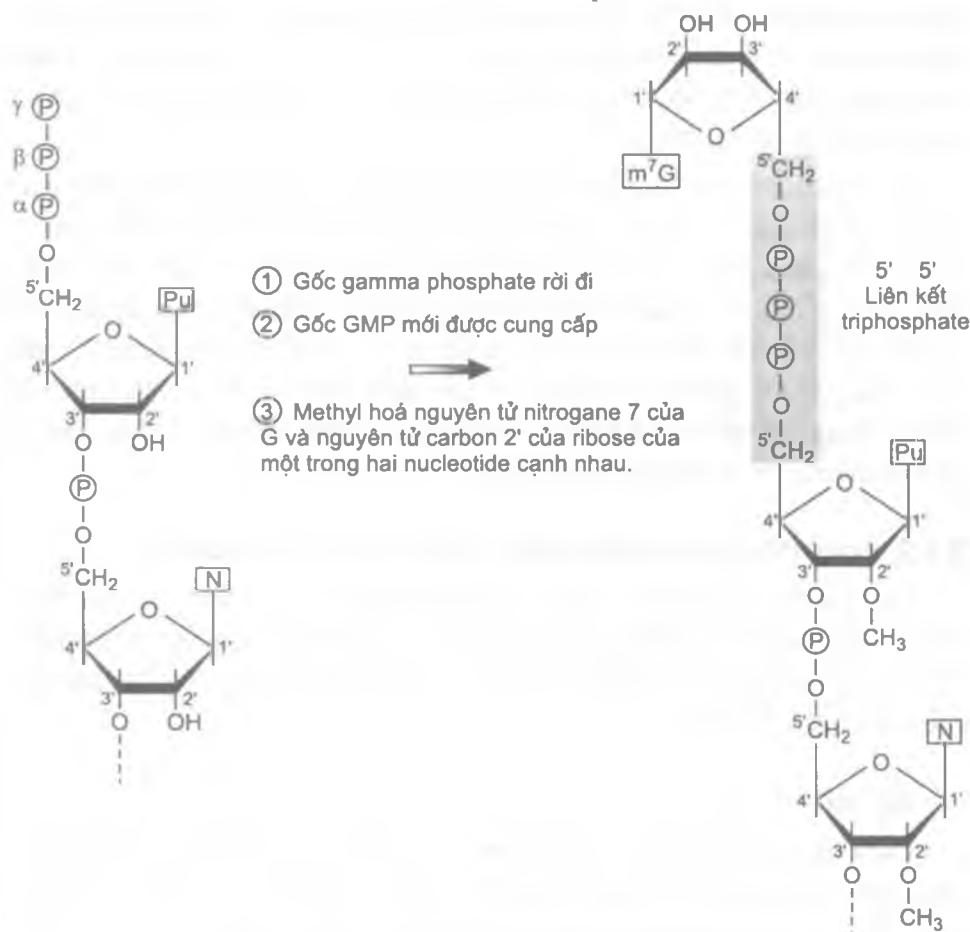
Các nucleotide được chuyên biệt hoá, được bổ sung vào các đầu 5' và đầu 3' của hầu hết các thể phiên mã RNA polymerase II. Ngoài ghép nối RNA, các thể phiên mã RNA polymerase II là đối tượng của hai kết quả chế biến RNA bổ sung.

Bịt mũ

Điều này xảy ra ngay sau khi phiên mã. Trong trường hợp của các thể phiên mã sơ cấp, nó sẽ được xử lý để cho mRNA, một nucleoside bị methyl hoá, 7-methylguanosine (m^7G) được liên kết với nucleotide 5' đầu tiên của thể phiên mã RNA bởi một liên kết phosphodiester 5'-5'. Do liên

kết này bắc cầu có hiệu quả, carbon 5' của gốc m⁷G với carbon 5' của nucleotide đầu tiên, đầu 5' được gọi là bị chặn hay bị mũ (hình 2.16). Các thể phiên mã của các gene snRNA cũng bị mũ nhưng mũ của chúng có thể chịu cải biến bổ sung. Mũ được nhìn thấy trước có một số chức năng có thể :

- Bảo vệ thể phiên mã khỏi 5'→3' exonuclease tấn công (các phân tử mRNA khử mũ bị phân huỷ nhanh chóng) ;
- Tạo điều kiện vận chuyển dễ dàng từ nhân tế bào sang tế bào chất ;
- Tạo điều kiện dễ dàng ghép nối RNA ;
- Đóng một vai trò quan trọng trong việc gắn tiểu đơn vị 40S của các ribosome tế bào chất với mRNA (xem phần sau).

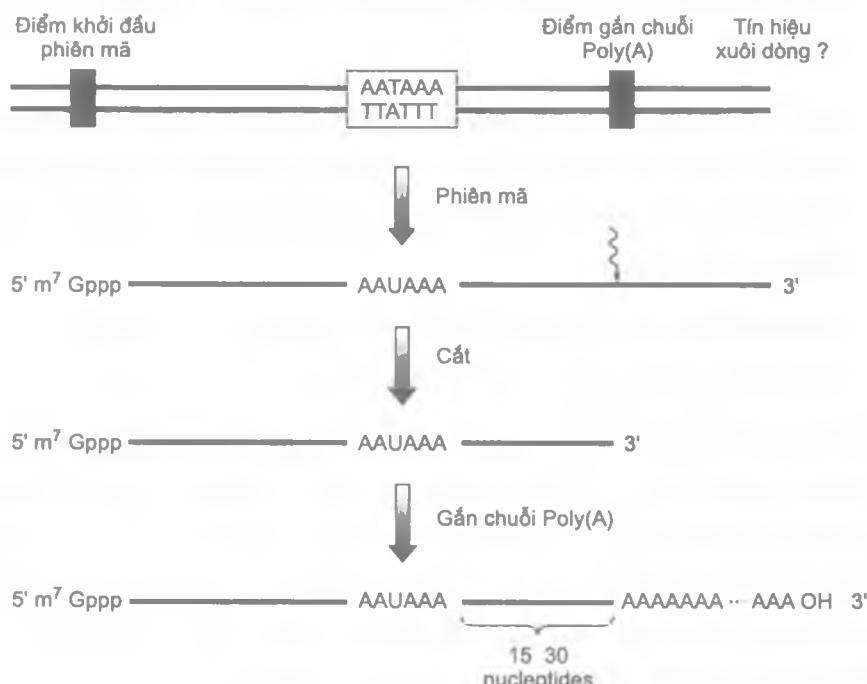


Hình 2.16. Quá trình bịt mũ.

Hình 2.16 minh họa quá trình bịt mũ. Đầu 5' của các phân tử mRNA tế bào nhân chuẩn được bảo vệ bởi một nucleotide chuyên biệt (bịt mũ). Sau khi gốc γ -phosphate của nucleotide đầu 5' rời đi, một gốc GMP mới do tiền thân GTP cung cấp, hình thành một liên kết triphosphate 5'-5' với đầu 5' nucleotide. Các phản ứng tiếp theo dẫn đến methyl hoá nguyên tử nitrogen 7 của G cuối cùng và trong động vật có xương sống, nguyên tử carbon 2' của ribose của một trong hai nucleotide cạnh nhau. N, moi nucleotide ; Pu, purine.

Polyadenyl hoá

Phiên mã bởi cả hai RNA polymerase I và III được biết là ngừng sau khi enzyme nhận biết một vị trí kết thúc phiến mã đặc trưng. Tuy nhiên, việc nhận dạng các vị trí kết thúc có thể cho phiến mã bởi RNA polymerase II là khó khăn bởi vì các đầu 3' của các phân tử RNA được xác định bởi một phản ứng cắt hâu phiến mã. Trình tự AAUAAA là một yếu tố chính mà phát tín hiệu phản ứng cắt 3' cho phần lớn các thể phiến mã polymerase II (các thể phiến mã từ các gene histone và các gene snRNA là những ngoại lệ có tiếng). Phản ứng cắt xảy ra ở một vị trí đặc trưng được khu trú 15-30 nucleotide xuôi dòng của yếu tố AAUAAA (hình 2.17).



Hình 2.17. Quá trình polyadenyl hoá mRNA.

Hình 2.17 : Đầu 3' của hầu hết các phân tử mRNA tế bào nhân chuẩn được polyadenyl hoá. Kết thúc phiên mã của các thể phiên mã RNA polymerase II được phát tín hiệu bởi phản ứng cắt 3' trong RNA được phiên mã. Trong hầu hết các loại mRNA, điều này được tạo ra bởi một tín hiệu ngược dòng AAUAAA trong sự phối hợp, cho đến bây giờ, các tín hiệu xuôi dòng chưa nhận dạng. Phản ứng cắt xảy ra bình thường khoảng 15-30 nucleotide xuôi dòng của yếu tố AAUAAA và các gốc AMP được bổ sung tiếp theo bởi poly(A) polymerase để hình thành đuôi poly(A). Histone mRNA chịu một phản ứng cắt 3' khác.

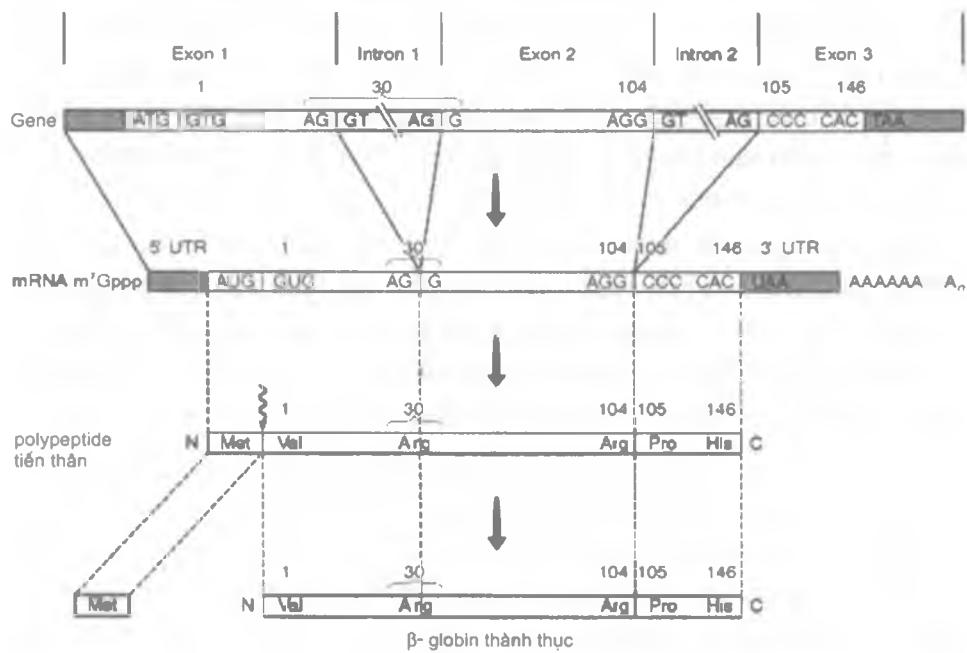
Tiếp theo điểm cắt, phiên mã có thể tiếp tục cho hàng trăm tới hàng nghìn nucleotide cho tới khi kết thúc xảy ra ở một trong các vị trí sau này. Khi phản ứng cắt đã xảy ra xuôi dòng yếu tố AAUAAA, khoảng 200 gốc adenylate (ví dụ AMP) được bổ sung tiếp trong các tế bào động vật có vú bằng enzyme poly(A) polymerase để hình thành đuôi poly(A). Đuôi poly(A) có thể có một số chức năng :

- Tạo điều kiện dễ dàng vận chuyển các phân tử mRNA sang tế bào chất ;
- Làm bền ít nhất một số phân tử mRNA trong tế bào chất (làm ngắn các dài poly(A) được liên kết với phân huỷ mRNA nhưng một số loại mRNA (ví dụ actin mRNA) duy trì bền vững với một ít hoặc không có poly(A)) ;
- Nó có thể tạo điều kiện dễ dàng dịch mã bằng cách cho phép nhận biết được sự tăng cường của mRNA bởi bộ máy ribosome.

Trong trường hợp của các gene histone, chúng là duy nhất trong tổng hợp mRNA mà không trở thành polyadenyl hoá, kết thúc phiên mã cũng bao gồm phản ứng cắt 3' của thể phiên mã sơ cấp. Phản ứng này phụ thuộc vào cấu trúc bậc hai của thể phiên mã RNA, bao gồm một trình tự kẹp tóc ngược dòng bảo thủ và một trình tự xuôi dòng ngắn, chúng bắt cặp base với một trình tự ngắn ở đầu 5' của U7 snRNA.

2.5. DỊCH MÃ

2.5.1. Quá trình tổng hợp polypeptide ở ribosome



Hình 2.18. Biểu hiện gene β-globin người.

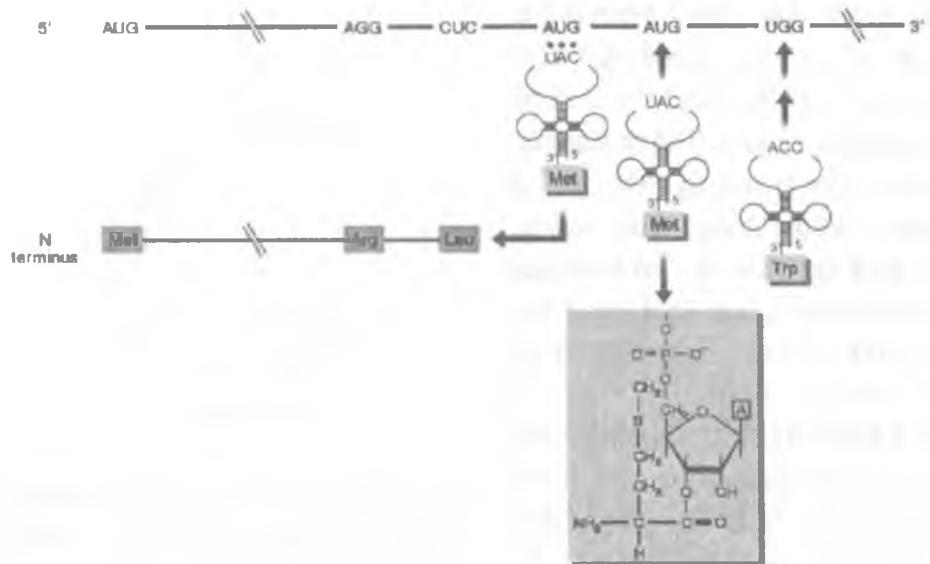
Dịch mã là quá trình ở đó mRNA được giải mã trên các ribosome để chỉ định tổng hợp các polypeptide. Tiếp theo quá trình hậu phiên mã, mRNA được phiên mã từ các gene trên DNA nhân tế bào, di chuyển tới tế bào chất. Ở đây, nó gài khớp với các ribosome và các thành phần khác để định hướng tổng hợp của các polypeptide đặc trưng. Ty thể cũng có các ribosome và một khả năng nhất định cho tổng hợp protein.

Chỉ phân đoạn trung tâm của phân tử mRNA tế bào nhân chuẩn đặc trưng được dịch mã để chỉ định tổng hợp một polypeptide. Các trình tự kẽ biên, 5' và 3', những vùng không dịch mã (untranslated region : 5' UTR ; 3' UTR), được sao bắn gốc từ trình tự dẫn xuất các exon cuối cùng 5' và 3', giống như mũ 5' và đuôi poly(A) 3', trợ giúp trong liên kết và làm bền mRNA trên các ribosome, ở đó dịch mã của phân đoạn trung tâm xảy ra (hình 2.18).

Hình 2.18 mô tả quá trình biểu hiện gene β -globin người. Exon 1 và 3 đều chứa các trình tự không mã hoá (thanh tô đậm) ở các đầu cục của chúng, chúng được phiên mã và có mặt ở các đầu 5' và 3' của β -globin mRNA nhưng không được dịch mã để chỉ định tổng hợp polypeptide. Các vùng không dịch mã 5' và 3' như thế (5' UTR và 3' UTR), được cho là quan trọng trong việc bảo đảm hiệu quả dịch mã cao (xem phần sau). Codon kết thúc UAA đại diện ba nucleotide đầu tiên của vùng không dịch mã 3'. Ghi nhớ rằng sản phẩm dịch mã khởi động có 147 amino acid nhưng methionine đầu N bị loại bỏ bởi xử lý hậu dịch mã để tạo ra một β -globin polypeptide thuần thực. Hai base đầu tiên của codon chỉ định Arg30 được mã hoá bởi exon 1 và base thứ ba được mã hoá bởi exon 2 (có nghĩa là intron 1 phân chia base thứ hai và thứ ba của codon, một ví dụ của một pha 2 intron). Intron thứ hai phân chia codon 104 và 105 và là một ví dụ về pha 0 intron và hình 2.22 cũng là một ví dụ về pha 1 intron.

Các ribosome là các phức hợp RNA-protein lớn được cấu thành từ hai tiểu đơn vị. Trong eukaryote, các ribosome tế bào chất có một tiểu đơn vị lớn 60S và một tiểu đơn vị nhỏ 40S (giá trị S là số đo độ các cấu trúc phân tử lớn sẽ lắng nhanh như thế nào trong máy ly tâm siêu tốc và được chi phối bởi cả hai trọng lượng phân tử và hình dạng). Tiểu đơn vị 60S chứa ba dạng phân tử rRNA : 28S rRNA, 5,8S rRNA và 5S rRNA và khoảng 50 protein ribosome. Tiểu đơn vị 40S chứa một 18S rRNA đơn và khoảng 30 protein ribosome. Các ribosome cung cấp một thể cấu trúc thống nhất cho tổng hợp polypeptide, trong đó, các thành phần RNA chịu trách nhiệm chủ yếu cho chức năng xúc tác của ribosome ; các thành phần protein được cho là tăng cường chức năng của các phân tử rRNA và chúng không xuất hiện là cần thiết cho chức năng ribosome.

Việc lắp ráp một polypeptide mới từ các amino acid thành phần của nó được chi phối bởi một code di truyền bộ tam. Các nhóm tiếp theo của ba nucleotide (codon) trên trình tự mRNA mạch thẳng được giải mã tiếp theo theo thứ tự để chỉ định các amino acid đơn lẻ. Quá trình giải mã được trung gian bởi một tập hợp của các phân tử tRNA, đối với mỗi trong chúng một amino acid đặc trưng được liên kết đồng hoá trị (ở nhóm 3' hydroxyl tự do của tRNA ; xem hình 2.19) bởi một amino acyl tRNA synthetase đặc trưng. Các phân tử tRNA khác nhau liên kết các amino acid khác nhau.



Hình 2.19. Code di truyền được giải mã bởi sự nhận biết codon-codon đối.

Code di truyền được giải mã bởi sự nhận biết codon-codon đối (hình 2.19). Trình tự nucleotide trong trình tự mRNA được dịch ra từ một điểm khởi động dịch mã (thường được đánh dấu bởi trình tự AUG) và tiếp tục theo hướng $5' \rightarrow 3'$ cho tới khi một codon kết thúc đạt được trong khung đọc này (xem hình 2.18). Mỗi codon trong mRNA được nhận biết bởi trình tự codon đối bỗ trợ của một phân tử tRNA với nó một amino acid đặc hiệu được liên kết đồng hoá trị với adenosine ở đầu 3' (xem phân đoạn chèn).

Mỗi tRNA có một trình tự trinucleotide đặc trưng, được gọi là codon đối (anticodon), ở một vị trí quan trọng chủ yếu được khu trú ở trung tâm của một cánh tay tRNA (Hình 2.6B). Vị trí cung cấp tính đặc hiệu cơ chất để dịch code di truyền : đối với một amino acid cần được chèn vào chuỗi polypeptide phát triển, codon thích hợp của phân tử mRNA phải được nhận biết nhờ bắt cặp base với một codon đối bỗ trợ thích hợp của phân tử tRNA thích hợp (hình 2.19).

Một mô hình dịch mã tiên đoán rằng tiểu đơn vị ribosome 40S nhận biết mū 5' trước nhờ sự tham gia của các protein liên kết đặc hiệu với mū. Sau đó, nó quét dọc mRNA cho tới khi nó gặp codon khởi động, hầu hết

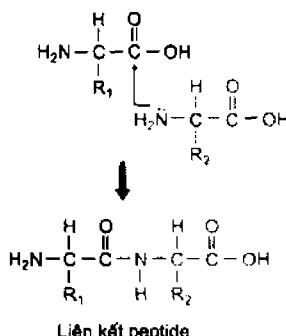
là AUG, chỉ định methionine (một số ít trường hợp được biết ở đó ACG, CUG hoặc GUG được sử dụng thay thế). Thông thường, AUG đầu tiên gặp được sẽ là codon khởi động. Tuy nhiên, AUG được nhận biết có hiệu quả như là một codon khởi động khi nó được ghi vào một trình tự tái tổ hợp, trình tự là : GCCPuCCAUGG. Những thẻ quyết định quan trọng nhất trong trình tự này là G tiếp theo codon AUG và purine, tốt nhất là A, đặt trước nó bởi ba nucleotide

(Kozak, 1996). Tiếp theo, các amino acid liên kề được gắn vào chuỗi polypeptide phát triển bởi một phản ứng mất nước : nhóm amino của amino acid mới tới phản ứng với nhóm carboxyl của amino acid cuối cùng để được gắn vào, cho kết quả một liên kết peptide giữa các gốc liên kề nhau (hình 2.20). Điều này được xúc tác bởi hoạt tính peptidyl transferase, nó tập trung ở thành phần RNA của tiểu đơn vị ribosome lớn.

2.5.2. Code di truyền là suy biến

Code di truyền là code ba chữ cái. Có bốn base có thể để chọn từ đó, ở mỗi một base trong ba vị trí base trong một codon. Bởi vậy có $4^3 = 64$ codon có thể nhưng chỉ có 20 amino acid khác nhau được chỉ định. Kết quả là code di truyền được gọi là suy biến : mỗi amino acid được chỉ định trung bình bởi khoảng ba codon khác nhau. Các amino acid nhất định (như methionine hoặc tryptophan, trong trường hợp của code di truyền nhân tế bào) được chỉ định bởi một codon duy nhất ; các amino acid khác, bao gồm leucine và serine, được chỉ định bởi sáu codon (hình 2.21).

Hình 2.21 biểu thị code di truyền nhân tế bào và ty thể tương tự nhau nhưng không đồng nhất. Các hộp xanh cho thấy bốn codon, chúng được dịch khác nhau trong nhân tế bào và ty thể của tế bào động vật có vú, với dịch ty thể là màu xanh. Bởi vậy, code ty thể có bốn



Hình 2.20. Các polypeptide được tổng hợp bởi sự hình thành liên kết peptide giữa các amino acid liền kề nhau.

codon kết thúc thay vì ba (UAA, UAG, AGA, AGG), hai codon Trp thay vì một (UGA, UGG), bốn codon Arg thay vì sáu (CGA, CGC, CGG, CGU), hai codon Met thay vì một (AUA, AUG) và hai codon Ile thay vì ba (AUC, AUU). Độ suy biến của code di truyền thường phải liên quan tới base thứ ba của codon. Bất kỳ base nào cũng có thể được thay thế (GGN = glycine, CCN = proline... ở đó N là base bất kỳ). Trong các trường hợp khác, bất kỳ purine (Pu) hoặc bất kỳ pyrimidine (Py) sẽ được thay thế (AAPu = lysine, AAPy = asparagine...).

AAA } AAG }	Lys	CAA } CAG }	Gln	GAA } GAG }	Glu	UAA } UAG }	STOP
AAC }		CAC }	His	GAC }	Asp	UAC }	Tyr
AAU }	Asn	CAU }		GAU }		UAU }	
ACA }		CCA }		GCA }		UCA }	
ACG }	Thr	CCG }	Pro	GCG }	Ala	UCG }	
ACC }		CCC }		GCC }		UCC }	Ser
ACU }		CCU }		GCU }		UCU }	
[AGA]	{ Arg	CGA }		GGA }		[UGA]	STOP
[AGG]	{ STOP	CGG }		GGG }		[Trp	
AGC }	Ser	CGC }	Arg	GGC }	Gly	UGG }	Trp
AGU }		CGU }		GGU }		UGC }	Cys
	Ile					UGU }	
[AUA]	{ Met	CUA }		GUA }		UUA }	
AUG }	{ Met	CUG }		GUG }		UUG }	Leu
AUC }	Ile	CUC }	Leu	GUC }	Val	UUC }	
AUU }		CUU }		GUU }		UUU }	Phe

Hình 2.21. Code di truyền nhân tế bào và ty thể tương tự nhau nhưng không đồng nhất.

Mặc dù có 64 codon, số lượng phân tử tRNA tương ứng với các codon đối khác nhau lại ít hơn : chỉ có trên 30 dạng tRNA tế bào chất và chỉ có 22 dạng tRNA ty thể. Dịch mã tất cả 64 codon trong các ribosome tế bào chất và ty thể là có thể bởi vì các nguyên lý bắt cặp base bình thường được nới lỏng khi tới nhận biết codon-anticodon. Giả thuyết quả lắc khẳng định rằng bắt cặp của codon và codon đối tuân theo các nguyên lý A-U và G-C bình thường cho hai vị trí base đầu tiên trong một codon nhưng lắc lư ngoại lệ xảy ra ở vị trí thứ ba và cặp base G-U base cũng được chấp nhận (bảng 2.1).

Dịch mã tiếp tục cho tới khi gặp một codon kết thúc có nghĩa là UAA, UAG hoặc UGA trong trường hợp của mRNA mã hoá nhân tế bào ; UAA,

UAG, AGA hoặc AGG trong trường hợp của mRNA mã hoá ty thể (hình 2.21). Bởi vậy, khung xương của sản phẩm dịch mã sơ cấp sẽ có methionine ở một đầu với nhóm amino tự do (đầu N) và một amino acid ở một đầu khác với nhóm carboxyl tự do (đầu C). Mặc dù các codon được dịch mã trong một khung đọc dịch mã đặc trưng, các gene gối nhau được tìm thấy trong các eukaryote, ở đó các khung đọc dịch mã khác nhau được sử dụng.

Bước ưu thế trong kiểm soát dịch mã là liên kết ribosome. Ngoài ra, mũ 5', 5' UTR (thường < 200 bp) và 3' UTR (thường dài hơn rất nhiều 5' UTR) cả hai đóng các vị trí tối hạn trong sự bổ sung mRNA cho dịch mã. Một vài yếu tố hoạt động cis mà chúng liên quan trong quá trình này đã được xác định tính chất và, một vài yếu tố hoạt động trans mà liên kết với các yếu tố này đã được nhận dạng. Có thể các trình tự 5' và 3' UTR tương tác để tăng cường dịch mã. 3' UTR có vai trò chìa khoá trong điều hoà dịch mã và các tín hiệu cho kiểm soát dịch mã, độ bền vững mRNA và khu trú tất cả được tìm thấy trong vùng này.

Bộ máy tổng hợp protein ty thể là tương tự với bộ máy được tìm thấy trong các ribosome tế bào chát. Ví dụ, các ribosome ty thể chứa hai tiểu đơn vị, một tiểu đơn vị lớn 23S rRNA với một số protein và một tiểu đơn vị nhỏ 16S rRNA với một số protein ribosome. Tuy nhiên, các thành phần được lắp ráp từ các sản phẩm của cả hai gene trong nhân tế bào và ty thể : các thành phần RNA được tổng hợp trong ty thể (2 phân tử rRNA và 22 dạng tRNA và một số thường khoảng 10, dạng mRNA ; tham khảo mô tả của genome ty thể người).

Ở đó, các protein được mã hoá bởi các gene trong nhân tế bào chúng đã được dịch mã trên các ribosome trong nhân tế bào và nhập vào các mitochondrion (RNA polymerase, 22 amino acyl tRNA synthetase và khoảng 80 ribosomal protein).

2.5.3. Cải biến hậu dịch mã

Cải biến hậu dịch mã của các protein thường xuyên xảy ra và có thể bao gồm cả phản ứng cộng các nhóm hoá học đặc trưng vào các amino acid đặc trưng và cắt sản phẩm dịch mã sơ cấp. Các sản phẩm dịch mã sơ cấp thường chịu một loại của các phản ứng cải biến, bao gồm phản ứng

công các nhóm hoá học, chúng được gắn đồng hoá trị vào chuỗi polypeptide ở các mức độ dịch mã và hậu dịch mã. Điều này có thể bao hàm cải biến hoá học (hydroxyl hoá, phosphoryl hoá...) của các chuỗi phụ của các amino acid đơn hoặc bổ sung các dạng khác nhau của các nhóm carbohydrate hoặc lipid (xem bảng 2.7).

Bảng 2.7. Các dạng chính của cải biến polypeptide.

Dạng cải biến (nhóm được bổ sung)	Các amino acid đích	Ghi chú
Phosphoryl hoá (PO_4^-)	Tyrosine, serine, threonine	Đạt được bởi các kinase đặc trưng. Có thể được đảo ngược bởi các phosphatase
Methyl hoá (CH_3)	Lysine	Đạt được bởi các methylase và không đạt được bởi demethylase
Hydroxyl hoá (OH)	Proline, lysine, aspartic acid	Hydroxyproline và hydroxylysine là đặc biệt phổ thông trong collagen
Acetyl hoá (CH_3CO)	Lysine	Đạt được bởi một acetylase và không đạt được bởi deacetylase
Carboxyl hoá (COOH)	Glutamate	Đạt được bởi γ -carboxylase
<i>N</i> -glycosyl hoá (carbohydrate phức hợp)	Asparagine, thường trong trình tự : Asn-X-Ser/Thr	Thường xảy ra trước tiên trong màng lưới nội chất ; X là bất kỳ amino acid khác proline
O-glycosyl hoá (carbohydrate phức hợp)	Serine, threonine, hydroxylysine	Xảy ra trong bộ máy Golgi ; ít thông dụng hơn <i>N</i> -glycosyl hoá
GPI (glycolipid)	Aspartate ở đầu C	Dùng làm protein mỏ neo với lớp ngoài của màng plasma
Myristoyl hoá (C_{14} nhóm acyl béo)	Glycine đầu N	Dùng làm mỏ neo màng
Palmitoyl hoá (C_{16} nhóm acyl béo)	Cysteine để hình thành liên kết S-palmitoyl	Dùng làm mỏ neo màng
Farnesyl hoá (nhóm C_{15} prenyl)	Cysteine ở đầu C	Dùng làm mỏ neo màng
Geranylgeranyl hoá (nhóm C_{20} prenyl)	Cysteine ở đầu C	Dùng làm mỏ neo màng

Cải biến protein bằng cách bổ sung nhóm carbohydrate

Glycoproteins chứa oligosaccharide chúng được gắn đồng hoá trị với các chuỗi bên của các amino acid nhất định. Một số protein trong tế bào chất được glycosyl hoá, có nghĩa là đã gắn carbohydrate và những cái mà mang một gốc đường đơn, *N*-acetylglucosamine, đồng hoá trị liên kết với một gốc serine hoặc threonine. Ngược lại, các protein đó được tiết ra khỏi tế bào hoặc xuất vào lysosome, bộ máy Golgi hoặc màng plasma được glycosyl hoá. Các thành phần oligosaccharide của các glycoprotein được tạo thành trước rộng rãi và bổ sung toàn bộ gộp cả lại vào polypeptide. Hai dạng glycosyl được thừa nhận :

Trong hầu hết các trường hợp, *N*-glycosyl hoá bao gồm vận chuyển bước đầu của một trình tự oligosaccharide chung vào nhóm NH₂ chuỗi bên của một gốc Asn nội trong màng lưới nội chất (ER ; xem bảng 2.7). Cắt gọt tiếp theo các gốc hoặc thay thế với các monosaccharide khác nhau xảy ra trong bộ máy Golgi ;

O-glycosyl hoá – xem bảng 2.7.

Proteoglycans là những protein với *glycosaminoglycan* được gắn vào, thường chứa các đơn vị lặp lại disaccharide chứa glucosamine hoặc galactosamine. Các proteoglycan được xác định tính chất kỹ nhất là các thành phần của chất nền ngoại bào.

Cải biến protein bằng cách bổ sung nhóm lipid

Một số proteins, nhất là các protein màng, được cải biến bởi phản ứng cộng các nhóm acyl béo hoặc prenyl đặc trưng dùng làm các mỏ neo màng. Các ví dụ của các nhóm acyl béo bao gồm nhóm myristoyl, nó là một (C₁₄) lipid được tìm thấy gắn với gốc glycine ở đầu N cực ngoài và tạo khả năng protein cải biến tương tác với thụ quan màng hoặc lớp lipid kép của màng. Một nhóm acyl béo khác dùng làm mỏ neo màng là nhóm (C₁₆) palmitoyl, nó trở nên gắn vào nguyên tử S của gốc cysteine.

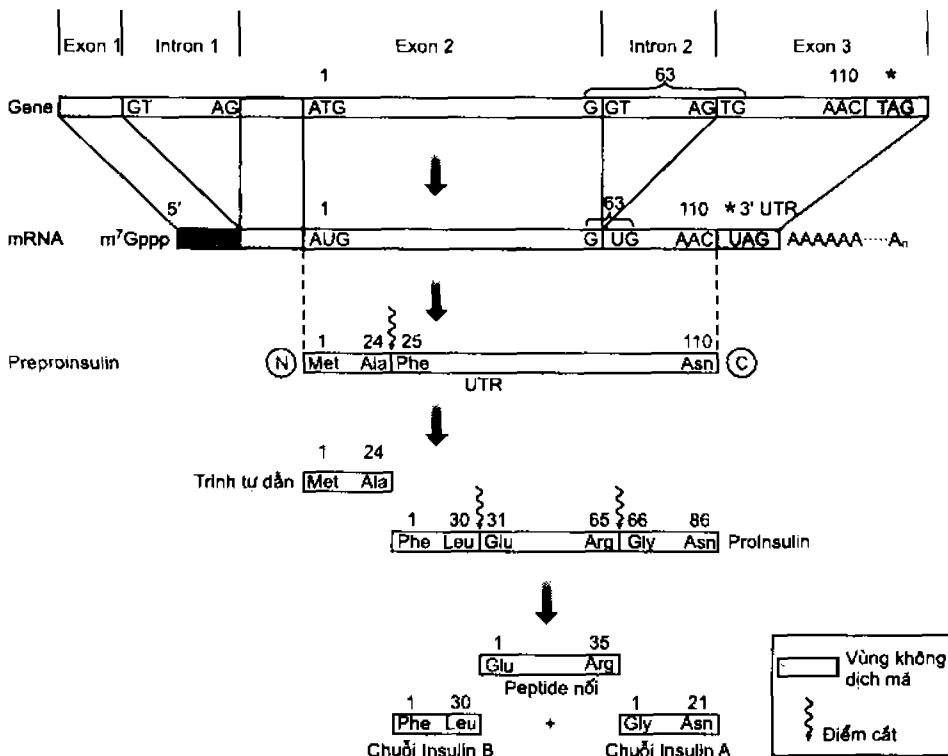
Các nhóm prenyl thường gắn với các gốc cysteine gần đầu C và bao gồm các nhóm farnesyl (C₁₅) và nhóm geranylgeranyl (C₂₀). Nhiều protein tham gia vào việc truyền tín hiệu và nhắm đích protein chứa hoặc một đơn vị farnesyl hoặc geranyl ở đầu C.

Mỏ neo của một protein vào lớp ngoài của màng plasma sử dụng một cơ chế khác : gắn nhóm glycosylphosphatidyl inositol (GPI). Nhóm glycolipid phức hợp chứa một nhóm acyl béo dùng làm mỏ neo màng, nó được liên kết kế tiếp với đơn vị glycerophosphate, một đơn vị oligosaccharide và cuối cùng qua một đơn vị phosphoethanol amine tới đầu C của protein. Toàn bộ protein ngoại trừ mỏ neo GPI được định vị ở không gian ngoại bào.

Cắt hậu dịch mã

Sản phẩm dịch mã sơ cấp cũng có thể chịu phản ứng cắt nội bộ để tạo ra một sản phẩm thuần thực nhỏ hơn. Thỉnh thoảng, methionine khởi động bị cắt từ sản phẩm dịch mã sơ cấp, như trong quá trình tổng hợp β -globin (hình 2.18). Cắt polypeptide đáng kể được quan sát thấy trong trường hợp nhiều protein thuần thực, bao gồm các plasma protein, polypeptide hormone, neuropeptide, yếu tố sinh trưởng... Ví dụ, tất cả các polypeptide được tiết ra và cũng như polypeptide được vận chuyển qua màng nội bào (ví dụ những polypeptide được tổng hợp trong tế bào chất và được vận chuyển tới các ty thể) được tổng hợp trước tiên là các tiền thân, ở đó một trình tự tín hiệu (đôi khi được gọi là một trình tự dẫn) ở đầu N hoạt động như là một tín hiệu nhận biết cho vận chuyển qua màng tế bào (xem phần dưới). Sau đó peptide tín hiệu được cắt ra khỏi polypeptide chính và phân huỷ. Ngoài ra, trong một số trường hợp, một phân tử mRNA đơn chỉ định nhiều hơn một chuỗi polypeptide chức năng như là kết quả của cắt proteolytic của một tiền thân polypeptide lớn (hình 2.22).

Hình 2.22 mô tả tổng hợp insulin bao gồm nhiều phản ứng cắt hậu dịch mã của các tiền thể polypeptide. Intron đầu tiên ngắt vùng không dịch mã 5' ; intron thứ hai ngắt các vị trí base 1 và 2 của codon 63 và được phân loại như một pha I intron (xem hình 2.18 cho các pha intron khác). Sản phẩm dịch mã sơ cấp, preproinsulin, có một trình tự dẫn 24 amino acid, chúng được yêu cầu cho protein đi qua màng tế bào và bởi vậy được loại bỏ sau này. Tiền thể proinsulin chứa một vùng trung tâm, peptide nối, nó được cho là quan trọng trong việc duy trì cấu dạng của các phân đoạn chuỗi A và B, như vậy, chúng có thể hình thành cầu nối disulfide (xem hình 2.24).



Hình 2.22. Quá trình sinh tổng hợp insulin.

2.5.4. Sự tiết protein

Sự tiết protein và xuất vào các khu trú nội bào đặc trưng đòi hỏi các tín hiệu khu trú trong trình tự mã hoá hoặc các dạng đặc trưng của nhóm gắn. Các protein được tổng hợp trong ribosome ty thể có chức năng nội trong ty thể. Tuy nhiên, số protein mà được tổng hợp trong ribosome tế bào chất có chức năng phong phú hơn, chúng có thể được tiết ra từ tế bào ở đó chúng được tổng hợp (như với các hormone và các phân tử tín hiệu giữa các tế bào) hoặc được xuất vào các vùng nội bào đặc trưng, như nhân tế bào (histone, DNA và RNA polymerase, các yếu tố phiên mã, các protein xử lý RNA...), mitochondrion (các protein ribosome ty thể, nhiều thành phần chuỗi hô hấp...), các peroxisome và nhiều cái khác. Để làm được điều này, một tín hiệu khu trú đặc trưng cần được gài vào cấu trúc của polypeptide sao cho nó được gửi tới địa chỉ đúng. Thông thường, tín hiệu định vị có dạng một trình tự peptide ngắn. Nó bao gồm một trình tự tín hiệu (hoặc trình tự dẫn) do một peptidase tín hiệu chuyên biệt loại ra từ protein trong quá trình phân loại.

Các tín hiệu cho xuất sang mạng lưới nội chất và không gian ngoại bào

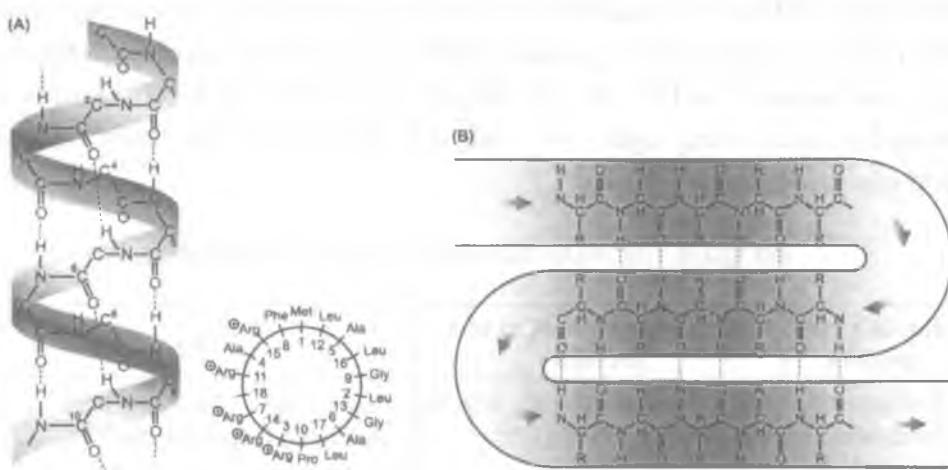
Trong trường hợp các protein tiết, tín hiệu peptide gồm khoảng 20 amino acid đầu tiên ở đầu N và luôn luôn có một số amino acid kỵ nước cần thiết (xem bảng 2.8). Trình tự tín hiệu được dẫn tới màng lưới nội chất ER bởi một hạt nhận biết tín hiệu (signal recognition particle, SRP) là một phức hợp RNA-protein gồm một loại RNA tế bào chất nhỏ, 7SL RNA và protein đặc trưng. Phức hợp SRP liên kết cả chuỗi polypeptide phát triển và ribosome, dẫn chúng tới một protein tiếp nhận SRP trên bề mặt bên tế bào của màng mạng lưới tế bào chất thô (rough endoplasmic reticulum, RER). Sau đó polypeptide có thể đi qua vào lumen của ER, tới nơi để tiết ra từ tế bào. Khi có các phân đoạn kỵ nước, chúng ngăn quá trình vận chuyển lại, như trong trường hợp các protein xuyên màng.

Bảng 2.8. Các ví dụ của các trình tự định vị protein

Nơi đến của protein	Định vị và hình dạng của tín hiệu	Ví dụ
Endoplasmic reticulum và tiết từ tế bào	Peptide đầu N của khoảng 20 amino acid ; rất kỵ nước	Insulin người – 24 amino acid, peptide tín hiệu khác cao : N – Met-Ala-Leu-Trp-Met-Arg-Leu-Leu-Pro-Leu-Leu-Ala-Leu-Leu-Ala-Leu-Trp-Gly-Pro-Asp-Pro-Ala-Ala-Ala
Ty thể	Peptide đầu N ; α -helix với các gốc tích điện dương trên mặt và các vùng kỵ nước khác	Aldehyde dehydrogenase ty thể người – N-terminal 17 amino acid : N-Met-Leu-Arg-Ala-Ala-Ala-Arg-Phe-Gly-Pro-Arg-Leu-Gly-Arg-Arg-Leu-Leu
Nhân tế bào	Trình tự nội bộ của các amino acid ; thường là các amino acid base cộng với các proline ; có thể là sóng đôi	SV40 T kháng thể – tiết : <u>Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val</u>
Lysosome	Bổ sung các gốc mannose 6-phosphate	p53 – bipartite : <u>Lys-Arg-Ala-Leu-Pro-Asn-Asn-Thr-Ser-Ser-Ser-Pro-Gln-Pro-Lys-Lys-Lys</u>

Các tín hiệu khác

Giống như tín hiệu ER, một trình tự tín hiệu đầu N được yêu cầu để đi qua màng thay thế và sau đó được cắt đi. Đặc trưng, một peptide tín hiệu thay thế, ngoài nhiều amino acid kỵ nước ra, một vài amino acid tích điện dương thường được đặt ở các khoảng bốn amino acid. Cấu trúc này là để hình thành một *amphipathic α-helix*, một cấu trúc xoắn với các amino acid tích điện trên bề mặt và các amino acid kỵ nước trên bề mặt khác (xem phần sau và hình 2.23).



Hình 2.23. Cấu trúc bậc hai trong các polypeptide.

Hình 2.23 mô tả các vùng cấu trúc bậc hai trong các polypeptide thường là chiếm ưu thế bởi liên kết hydrogen nội bộ chuỗi. (A) Cấu trúc của một α-helix. *Trái* : chỉ khung xương của polypeptide được trình bày rõ ràng. Carbonyl (CO) oxygen của mỗi liên kết peptide được liên kết hydrogen với hydrogen trên nhóm amide liên kết peptide (NH) của bốn amino acid từ đây, như vậy helix có 3,6 amino acid mỗi vòng xoắn. Để cho rõ ràng, một số liên kết được bỏ đi. Các chuỗi phụ của mỗi amino acid nằm ngoài helix và hầu hết không có không gian tự do nội trong helix. *Phải* : các amino acid tích điện và các amino acid kỵ nước nằm trên các bề mặt khác nhau trong một amphipathic α-helix. Trình tự được trình bày là trình tự peptide tín

hiệu dài 17 amino acid cho aldehyde dehydrogenase ty thể (xem bảng 2.8). (B) Cấu trúc cho một dài gấp β . Liên kết hydrogen tồn tại giữa các nguyên tử CO oxygen và NH hydrogen của các liên kết peptide trên các dài song song gần kề của khung polypeptide. Ví dụ trường hợp liên kết giữa các vùng đối song của khung polypeptide (dài β đối song) và sự thay đổi hướng đột ngột bắt buộc giữa các phần đối song thường đi kèm sử dụng quay β . Các mũi tên đánh dấu hướng từ đầu N sang đầu C. Các dài khán xếp β song song với các phân lân cận chạy theo cùng hướng cũng được tìm thấy.

Các tín hiệu định vị nhân tế bào có thể tìm thấy ở bất kỳ nơi nào trên trình tự polypeptide và thường bao gồm một dài 4-8 amino acid tích điện dương, cùng với các gốc proline liền kề. Tuy nhiên, thường tín hiệu được chia đôi, các amino acid tích điện dương được tìm thấy ở hai block của 2 – 4 gốc, được phân tách bởi khoảng 10 amino acid (bảng 2.7). Các protein trong nhân nào đó thiếu các trình tự định vị nhân nhưng vẫn được vận chuyển vào nhân với sự trợ giúp từ các protein nhân khác mà chúng có các tín hiệu thích hợp.

Các protein lysosome được hướng tới lysosome bằng cách bổ sung một gốc mannose 6-phosphate vào ngăn *cis* của bộ máy Golgi và được nhận dạng bởi một protein thụ quan trong ngăn *trans* của Golgi.

2.5.5. Cấu trúc protein biến đổi và phức tạp

Cấu trúc protein biến đổi, phức tạp và không dễ dàng tiên đoán từ trình tự amino acid của các polypeptide. Protein được cấu tạo từ một hoặc nhiều polypeptide, mỗi một chuỗi có thể là đối tượng cho cải biến hậu dịch mã. Chúng có thể tương tác với các dòng yếu tố đặc trưng (ví dụ, các cation hoá trị 2 như Ca^{2+} , Fe^{2+} , Cu^{2+} , Zn^{2+} hoặc các phân tử nhỏ chúng được yêu cầu cho hoạt tính enzyme chức năng, ví dụ NAD^+) hoặc các cấu tử ligands (mỗi phân tử nó một protein liên kết chuyên). Chúng có thể ảnh hưởng mạnh lên cấu dạng của protein. Ít nhất bốn mức khác nhau của tổ chức cấu trúc đã được phân biệt cho các protein (xem bảng 2.8).

Bảng 2.8. Các mức độ cấu trúc protein

Mức độ	Định nghĩa	Ghi chú
Bậc nhất	Trình tự amino acid tuyển tính trong một polypeptide	Có thể biến đổi nhiều về độ dài từ một peptide nhỏ dài tới hàng nghìn amino acid
Bậc hai	Hướng mà một khung polypeptide theo trong không gian	Có thể biến đổi từng vùng, ví dụ α -helix hoặc dải gấp β
Bậc ba	Cấu trúc không gian ba chiều của một polypeptide	Có thể biến đổi rất nhiều, ví dụ : hình cầu, que, ống, cuộn, dải khẩn ...
Bậc bốn	Cấu trúc tổng thể của protein, bao gồm nhiều chuỗi, với sự kết hợp của các tiểu đơn vị	Cấu trúc được ổn định bằng các cầu disulfide và bởi sự gắn kết với các phối tử ...

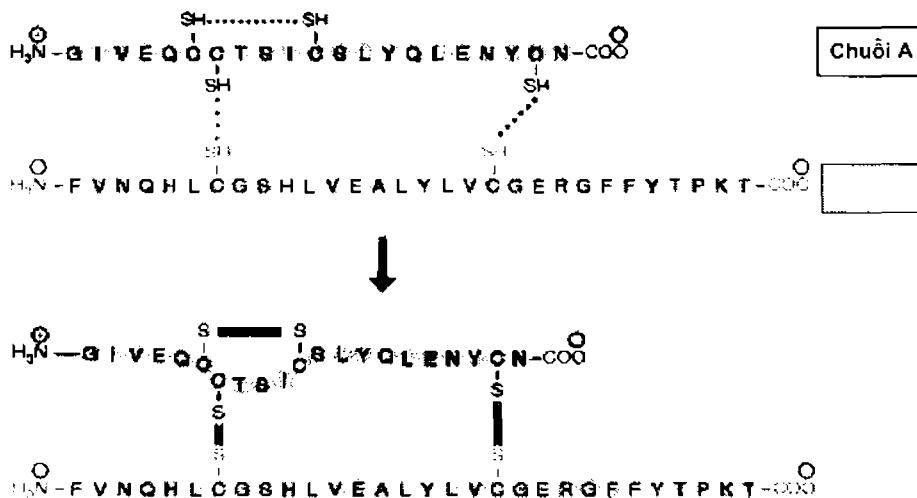
Nội trong một chuỗi polypeptide đơn, có một phạm vi rộng cho liên kết hydrogen giữa các gốc khác nhau ; bất luận là các chuỗi phụ, oxygen của nhóm carboxyl (COOH) của liên kết peptide có thể liên kết hydrogen với hydrogen của nhóm amin NH₂ của một liên kết peptide khác. Các đơn vị cấu trúc cơ bản được định nghĩa bởi liên kết hydrogen giữa các gốc amino acid liền kề của một polypeptide đơn bao gồm :

α -helix. Cái này tham gia hình thành một xylan cứng. Cấu trúc chiếm ưu thế bởi liên kết hydrogen giữa carbonyl oxygen của một liên kết peptide với nguyên tử hydrogen của amino nitrogen của một liên kết peptide khu trú cách bốn amino acid (xem hình 2.24). Các domain liên kết DNA của các yếu tố phiên mã thường là α -helical. Một amphipathic α -helix có các gốc tích điện trên bề mặt và các gốc ky nước trên một bề mặt khác (hình 2.24). Các α -helice giống hệt nhau với một sự sắp xếp lặp lại của các chuỗi phụ không phân cực, chúng xoắn quanh mỗi chuỗi để tạo thành một cấu trúc đặc biệt bền vững được gọi là một siêu xoắn. Xoắn giống như que được tìm thấy trong nhiều protein sợi, như sợi α -keratin của da, tóc và móng tay hoặc fibrinogene của cục máu đông.

Dải gấp β mô tả đặc trưng sự hình thành liên kết hydrogen giữa các liên kết peptide đối ngược nhau trong các khúc song song (thường thực sự đối song) của cùng một chuỗi polypeptide (xem hình 2.24). Các dải gấp β hình thành lõi của hầu hết các protein hình cầu.

Vòng xoắn β : Liên kết hydrogen giữa nhóm CO liên kết peptide của gốc amino acid n của một polypeptide với nhóm NH, liên kết peptide của gốc $n+3$ cho kết quả một vòng xoắn kẹp tóc. Bằng cách cho phép polypeptide bất ngờ quay ngược hướng, các hình dạng cầu nén chất có thể

dược tạo ra. Vòng xoắn β được đặt tên như vậy cũng là do chúng thường nối các sợi đối song trong các dải gấp β (xem hình 2.24).



Hình 2.24. Cầu nối disulfide nội chuỗi và giữa các chuỗi ở insulin người.

Các kiểu cấu trúc phức tạp hơn bao gồm các tổ hợp của các phân tử cấu trúc trên thiết lập các protein domain (các vùng kết đặc của một protein được hình thành từ gấp lại của cấu trúc bậc một, như vậy, các yếu tố của cấu trúc bậc hai có thể được xếp đồng cai này lên cái kia). Các domain như thế thường điện di các đơn vị chức năng tham gia vào liên kết các phân tử khác nhau. Ngoài ra, các cầu disulfide đóng vai trò thường được hình thành giữ các nhóm sulfhydryl (-SH) của các cặp các gốc cysteine tồn tại nội trong cùng chuỗi polypeptide hoặc trên các chuỗi polypeptide khác nhau (xem hình 2.24).

Hình 2.24 mô tả các cầu disulfide nội chuỗi và giữa các chuỗi ở insulin người. Các cầu disulfide ($-S-S-$) được hình thành bởi một phản ứng ngưng giữ các nhóm sulfhydryl (-SH) đối nhau của các gốc cysteine 6 và 11 của chuỗi A và giữa các gốc đã chỉ ra của các chuỗi khác nhau.

Rõ ràng, cấu trúc bậc ba hoặc bậc bốn của các protein được xác định bởi trình tự amino acid sơ cấp. Tuy nhiên, mặc dù các kiểu cấu trúc bậc hai như các α -helices, dải gấp β và vòng β có thể được tiên đoán bằng phân tích trình tự sơ cấp, toàn bộ cấu trúc không gian ba chiều không được, hiện nay, tiên đoán chính xác. Ngoài độ phức tạp cấu trúc của các polypeptide đơn giản, nhiều protein được tổ chức như một khối phức tạp của nhiều tiểu đơn vị polypeptide.

3

ĐỌC TRÌNH TỰ VÀ NGHIÊN CỨU GENOME

3.1. PCR CHUẨN

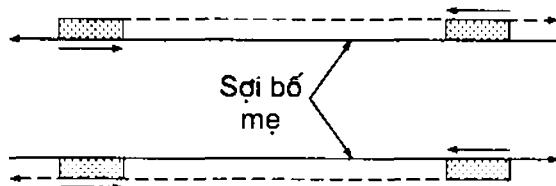
Phản ứng chuỗi sử dụng enzyme polymerase hay còn gọi là PCR (polymerase chain reaction) là một phản ứng tổng hợp sợi DNA đơn dựa vào một sợi DNA đơn khác làm khuôn và một đoạn oligonucleotide làm mồi. Mồi ở đây theo đúng nghĩa như muốn bơm được nước, phải cho một ít nước vào máy bơm để làm mồi, muốn tổng hợp được sợi DNA dài phải cho một đoạn nucleotide ngắn vào. Sợi DNA đơn được tổng hợp nên có trình tự bổ trợ với sợi khuôn.

PCR là một kỹ thuật trong ống nghiệm hay còn gọi là *in vitro* để khuếch đại vùng DNA nằm giữa hai vùng có trình tự được biết. Khuếch đại PCR được thực hiện nhờ sử dụng các mồi oligonucleotide. Chúng thường là các oligo nucleotide sợi đơn, ngắn, bổ trợ cho vùng ngoài của trình tự đã biết.



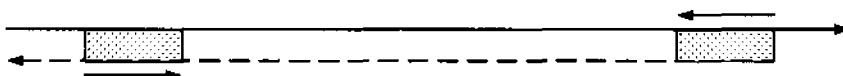
Hình 3.1. Sơ đồ PCR.

Các oligonucleotide dùng làm mồi cho enzyme DNA polymerase và sợi biến tính của phân đoạn DNA lớn được làm sợi khuôn. Kết quả là tổng hợp các sợi DNA mới, bổ trợ cho các sợi khuôn bố mẹ. Những sợi mới này có đầu 5' (chính là đầu 5' của mỗi oligonucleotide), trong khi đó đầu 3' thì chưa xác định được độ dài.



Hình 3.2. Sơ đồ PCR sau chu kỳ thứ nhất.

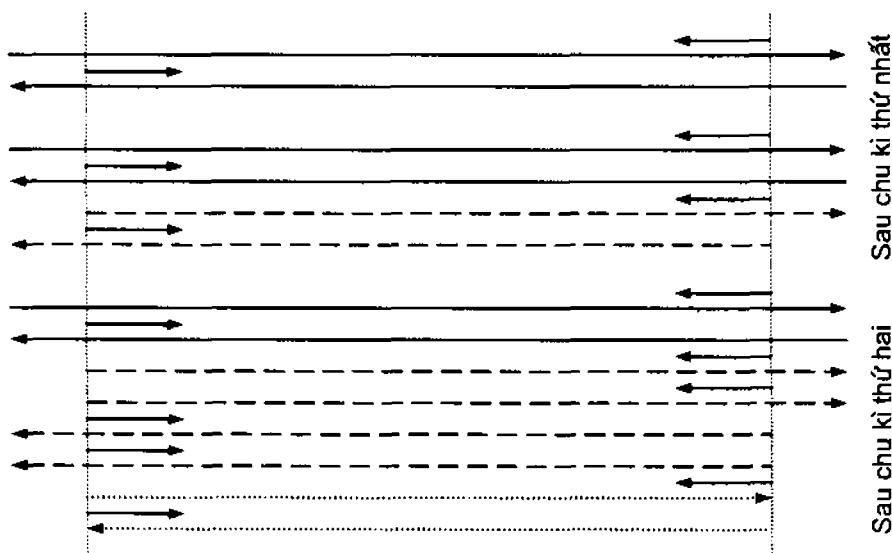
Quá trình tổng hợp định hướng theo oligonucleotide của các sợi DNA con có thể được nhắc lại nếu sợi kép mới được biến tính (bằng nhiệt) và mỗi bổ sung được phép gắn vào sợi khuôn (nhờ hạ xuống nhiệt độ thích hợp). Các bước này bao gồm : a) biến tính, b) gắn mới, c) kéo dài mới tạo ra một chu kỳ trong phương pháp khuếch đại PCR.



Hình 3.3. Sợi DNA mới tổng hợp có thể làm khuôn trong chu kỳ tiếp theo.

Sau mỗi chu kỳ, sợi DNA mới tổng hợp có thể làm khuôn trong chu kỳ tiếp theo.

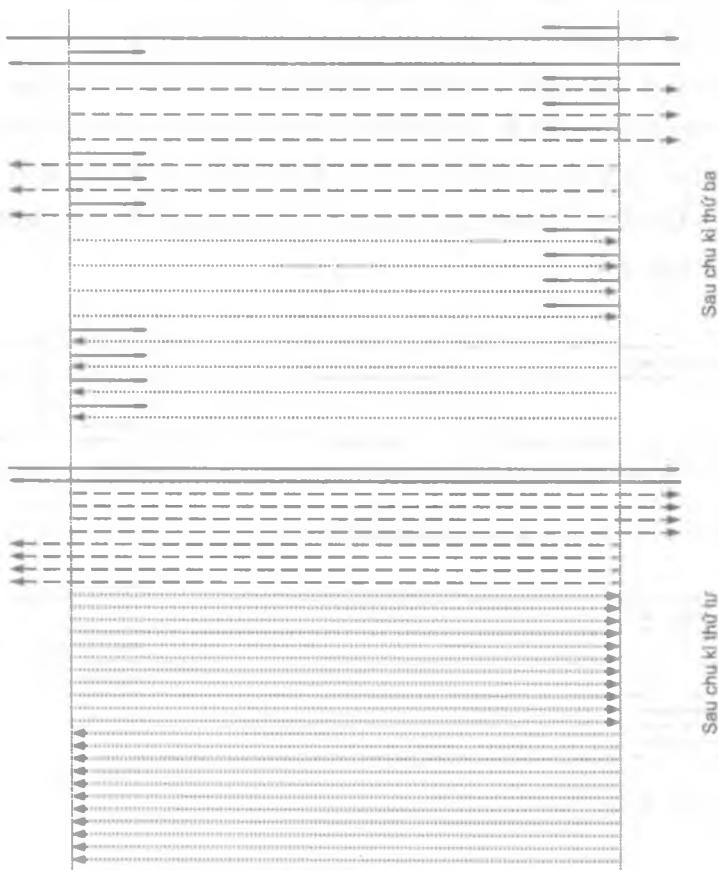
Một sợi trong sợi mới được tổng hợp từ chu kỳ thứ hai trở đi có đầu 5' và 3' được xác định tại vị trí gắn mới oligonucleotide. Sau chu kỳ thứ nhất, tổng số sợi là 4 bao gồm 2 sợi gốc -DNA- và 2 sợi được tổng hợp mới xác định ở một đầu -DNA- và DNA- (hình 3.4). Sau chu kỳ thứ hai, tổng số sợi là 8 bao gồm 2 sợi gốc -DNA-, 4 sợi được tổng hợp mới xác định ở một đầu -DNA- và DNA- và 2 sợi được tổng hợp mới xác định ở cả hai đầu DNA (hình 3.4).



Hình 3.4. Sơ đồ PCR sau chu kỳ thứ nhất và thứ hai.

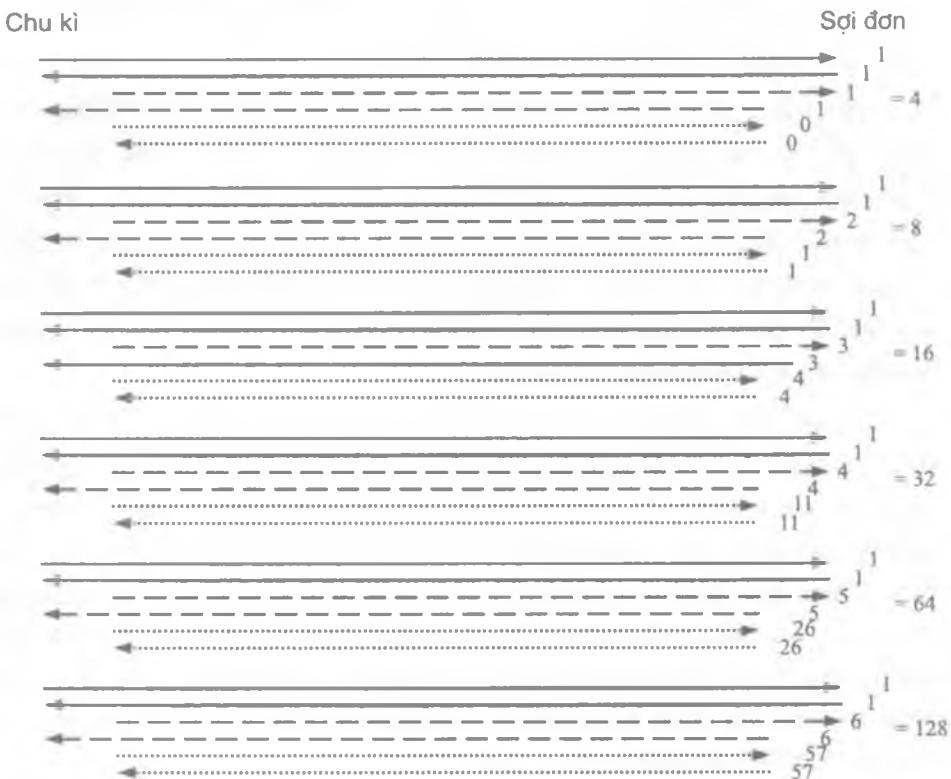
Sau chu kỳ thứ ba, tổng số sợi là 16 bao gồm 2 sợi gốc –DNA–, 6 sợi được tổng hợp mới xác định ở một đầu –DNA và DNA–, 8 sợi được tổng hợp mới xác định ở cả hai đầu DNA (hình 3.5). Sau chu kỳ thứ tư, tổng số sợi là 32 bao gồm 2 sợi gốc –DNA–, 8 sợi được tổng hợp mới xác định ở một đầu –DNA và DNA–, 22 sợi được tổng hợp mới xác định ở cả hai đầu DNA (hình 3.5).

Sau n chu kỳ, sự khuếch đại các phân đoạn tuân theo quy luật sau :
1) Chỉ có đúng 1x bản sợi khuôn ban đầu (sợi kép 2 –DNA–), PCR không bao giờ tái tạo được sợi dài như sợi khuôn, trừ trường hợp hai mồi nằm ở tận cùng hai đầu sợi khuôn. 2) Có n sợi đơn mỗi loại (–DNA và DNA–) với độ dài không xác định, trong đó n là số chu kỳ. Những phân đoạn có độ dài không xác định này có một đầu được giới hạn bởi mồi PCR còn đầu kia không xác định. 3) Có $[2n - (n + 1)]$ bản sao sợi đơn mỗi loại (DNA xuôi và ngược) có độ dài xác định giữa hai mồi PCR.



Hình 3.5. Sơ đồ PCR sau chu kỳ thứ ba và thứ tư.

Ví dụ sau 6 chu kỳ, chúng ta sẽ có 1x bản sao mỗi sợi khuôn đơn ($-$ DNA) ; 6x bản sao mỗi sợi đơn ($-$ DNA và DNA) có độ dài không xác định ; và $[64 - (6 + 1)] = 57$ x bản sao mỗi sợi đơn (DNA xuôi và ngược) có độ dài xác định.



Hình 3.6. Sơ đồ PCR sau chu kỳ thứ năm và thứ sáu.

Trong trường hợp trên, sản phẩm PCR mong muốn sẽ là một sợi kép có độ dài xác định. Câu hỏi đặt ra là thực sự sẽ có bao nhiêu sợi như thế ? Trong thực tế thì : a) Những sợi khuôn ban đầu không nhất thiết phải bắt cặp với nhau ; b) Tương tự như vậy những sợi khuôn có độ dài không xác định cũng không nhất thiết phải bắt cặp với nhau ; c) Hơn nữa, theo số chu kỳ tăng lên, số phân đoạn có độ dài xác định sẽ tăng nhanh hơn các phân đoạn còn lại. Do đó, sợi khuôn ban đầu và những phân đoạn có độ dài không xác định sẽ bắt cặp với những phân đoạn có độ dài xác định. Sau 6 chu kỳ PCR, số phân đoạn có độ

dài xác định có thể là $57 - (6 + 1) = 50$ lần sao phân đoạn có độ dài xác định (50 lần phân đoạn sợi kép có độ dài xác định).

Số sản phẩm phân đoạn sợi kép có độ dài xác định với nồng độ sợi khuôn gốc là 'a' sẽ được tính theo công thức sau đây : $\{(2n - (n + 1)) - (n + 1)\} \times a$ hay $\{(2n - 2(n + 1)) \times a\}$ và thường được đơn giản hóa thành công thức cho dễ nhớ là : $\{(2n - 2n) \times a\}$.

Điển giải công thức này như sau : a) Đối với số chu kỳ 'n' chúng ta có tổng số $(2n \times a)$ sợi kép ; b) Đối với số chu kỳ 'n' sẽ có $[2(n+1) \times a]$ (hay làm chẵn $2na$) sợi kép được tạo thành từ một sợi đơn khuôn hoặc sợi đơn không xác định với một sợi đơn có độ dài xác định (là những sản phẩm không mong muốn). Do đó, tổng nồng độ của sản phẩm mong muốn (sợi kép với độ dài xác định theo mỗi PCR) sẽ là $\{[2n - 2(n+1)] \times a\}$ (ở đó, a là nồng độ sợi khuôn kép ban đầu).

Trong thực tế giá trị khuếch đại lý thuyết không bao giờ đạt được. Có nhiều yếu tố ảnh hưởng tới giá trị này bao gồm : a) Việc bắt cặp hoàn toàn của các sợi con với sợi mồi (ví dụ như hai sợi con bắt cặp với nhau sẽ không cho sản phẩm khuếch đại) ; b) Mất hoạt tính của enzyme vì biến tính nhiệt, đặc biệt trong những chu kỳ về sau ; c) Ngay cả khi không biến tính vì nhiệt, số enzyme sẽ ngày càng ít đi do số phân tử đích càng tăng lên trong những chu kỳ sau (ví dụ sau 25-30 chu kỳ sẽ có rất nhiều phân tử mồi cần được khuếch đại) ; d) Mỗi gắn vào vị trí không đặc hiệu và mồi cho các sản phẩm phụ.

PCR được Kary Mullis phát triển vào năm 1985 do làm việc cho công ty gần Berkeley, California. Phương pháp PCR gốc sử dụng phân đoạn Klenow của *E. coli* DNA polymerase I. Tuy nhiên, enzyme này biến tính ở nhiệt độ thấp hơn nhiệt độ của các sợi khuôn kép. Chính vì thế, sau mỗi chu kỳ, người ta lại bổ sung enzyme mới vào hỗn hợp phản ứng. Đó là một việc nhảm chán. Hơn nữa, mẫu phải di chuyển từ bồn nhiệt độ này sang bồn nhiệt độ khác để thực hiện từng bước biến tính, bắt cặp và tổng hợp do chúng cần các nhiệt độ khác nhau. Điều này lại càng phức tạp hơn.

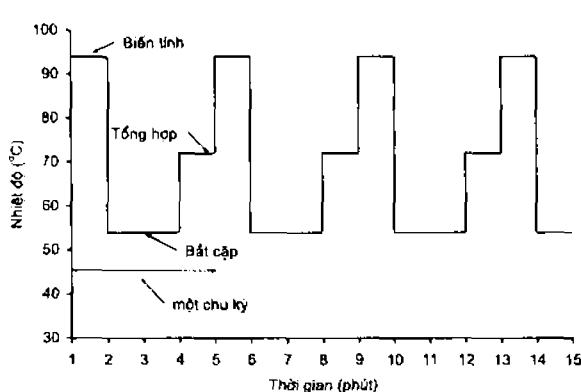
Do đó, hai ưu điểm cho phép quá trình được tự động hóa là : a) Sử dụng các DNA polymerase chịu nhiệt, có khả năng chống biến tính (bất hoạt) ở nhiệt độ cao. Vì thế, một lượng polymerase ban đầu có thể bền qua một số chu kỳ của quá trình. DNA polymerase chịu nhiệt đầu tiên

sử dụng được tách từ vi khuẩn chịu nhiệt *Thermus aquaticus*. Nó được phân lập từ một suối nước nóng ở Yellowstone National Park, ở đó chúng phát triển rất tốt, có thể sao chép DNA ở nhiệt độ hơn 85°C. b) Sự phát triển các bón nhiệt có thể tăng hay giảm nhiệt độ nhanh chóng theo một chương trình lập sẵn. Đó chính là máy chu trình nhiệt hay còn gọi là máy PCR.

Hai phát hiện này đã dẫn tới tự động hóa PCR. Quy trình PCR do hãng Hoffmann-La Roche Inc. đăng ký bản quyền, một hãng hóa chất rất đáng tin cậy.

Các thông số chu trình nhiệt là các yếu tố quyết định PCR thành công. Các bước quan trọng trong mỗi chu kỳ PCR bao gồm : a) biến tính sợi khuôn ; b) bắt cặp mồi ; c) kéo dài mồi.

Bước biến tính sợi khuôn : Biến tính ban đầu sợi khuôn được thực hiện ở nhiệt độ 95-100°C. Plasmid siêu xoắn bền hơn để nóng chảy và có thể cần đun sôi vài phút hoặc có thể được biến tính ban đầu bằng base (NaOH, sử dụng trung hòa pH). Bước biến tính trong lúc PCR (ví dụ từ chu kỳ thứ hai) thường được thực hiện ở nhiệt độ 92-95°C theo kinh nghiệm.



Hình 3.7. Sơ đồ chế độ nhiệt độ chung cho mỗi chu kỳ.

Bước bắt cặp mồi (gắn mồi sợi khuôn) : Nhiệt độ gắn mồi là một thông số quan trọng để thí nghiệm PCR thành công. Nhiệt độ gắn mồi là đặc trưng cho mỗi oligonucleotide : nó là một hàm số của độ dài và thành phần mồi cũng như cường độ ion của đậm phản ứng. Việc tính toán

nhiệt độ bắt cặp có thể được tính theo nhiều cách khác nhau. Nhiệt độ gắn mồi tính toán này là điểm khởi động cho thí nghiệm PCR nhưng nhiệt độ nóng chảy lý tưởng nhất là theo kinh nghiệm.

Bước kéo dài mồi (hay còn gọi là bước tổng hợp) : Kéo dài mồi thường được thực hiện ở 72°C, hay nhiệt độ tối ưu của DNA polymerase. Thời gian của bước kéo dài mồi có thể được tăng lên nếu vùng DNA cần được khuếch đại rất dài. Tuy nhiên, đối với phân lõi các thí nghiệm PCR, bước kéo dài 2 phút là đủ để nhận được phản ứng kéo dài hoàn toàn.

Số chu kỳ thường giữa 25 và 35. Nhiều chu kỳ có nghĩa là số sản phẩm lớn hơn. Tuy nhiên, với số chu kỳ tăng thì khả năng tạo ra các sản phẩm phụ cũng tăng theo (mỗi sai vị trí). Các quy trình có số chu kỳ trên 40 rất hiếm.

Một trong các ưu điểm cho phép phát triển phương pháp PCR là khả năng có các polymerase chịu nhiệt. Nó cho phép enzyme bổ sung ban đầu vượt qua chu kỳ nhiệt độ gần 100°C.

Chọn polymerase cho PCR :

Bảng 3.1. DNA polymerase chịu nhiệt và nguồn gốc của chúng.

DNA Polymerase	Tự nhiên/ tái tổ hợp	Nguồn
Taq	Tự nhiên	<i>Thermus aquaticus</i>
AmpliTaq®	Tái tổ hợp	<i>T. aquaticus</i>
AmpliTaq (Stoffel fragment)®	Tái tổ hợp	<i>T. aquaticus</i>
Hot Tub™	Tự nhiên	<i>Thermus flavis</i>
Pyrostase™	Tự nhiên	<i>T. flavis</i>
Vent™	Tái tổ hợp	<i>Thermococcus litoralis</i>
Deep Vent™	Tái tổ hợp	<i>Pyrococcus GB-D</i>
Tth	Tái tổ hợp	<i>Thermus thermophilus</i>
Pfu	Tự nhiên	<i>Pyrococcus furiosus</i>
ULTma™	Tái tổ hợp	<i>Thermotoga maritima</i>

Đêm và dung dịch MgCl₂ trong các PCR.

Một đêm phản ứng đặc trưng thường có thành phần như sau :

10 mM Tris, pH 8,3 ; 50 mM KCl ; 1,5 mM MgCl₂ ; 0,01% gelatin. Nồng độ MgCl₂ trong hỗn hợp phản ứng cuối cùng thường giữa 0,5 đến

5,0 mM và nồng độ tối đa được xác định theo kinh nghiệm (thường giữa 1,0 – 1,5 mM). Các ion Mg²⁺ : a) Tạo phức hoà tan với dNTP cần để gắn dNTP ; b) Kích thích hoạt tính polymerase ; tăng nồng độ nóng chảy Tm (melting temperature) của mồi/sợi khuôn (ví dụ để làm bền tương tác). Nói chung nồng độ Mg²⁺ thấp dẫn đến hiệu suất thu sản phẩm thấp (hoặc không có sản phẩm) và nồng độ Mg²⁺ cao dẫn tới tích tụ các sản phẩm phụ không đặc hiệu (mỗi sai vị trí).

Bảng 3.2. Các tính chất của DNA polymerase sử dụng trong PCR.
đb : đầu bằng.

	Taq AmpliTaq®	Phân đoạn Stoffel	ULTma™	Deep Vent™	Deep Vent™	ULTma™	ULTma™
Thời gian bán phân huỷ ở 95°C (phút)	>120	80	>120	>120	>120	>120	>120
5' → 3' exo							
3' → 5' exo							
Tốc độ kéo dài (nt/sec)	?	>50	>50	?	?	?	?
Hoạt tính RT	Yếu	Yếu	Có	Có	Có	Có	Có
Tạo ra đầu	?	?	>95% đb	>95% đb	?	?	>5%
Thay thế sợi	Có	Có	Có		Có	Có	Có
Khối lượng (kDa)	Cá	Cá	Yếu	Cá	92	Cá	Cá

Nói chung, mỗi được sử dụng có độ dài 20-30 nucleotide. Tạo ra nhiệt độ bắt cặp thực tế trong vùng nhiệt độ cao mà hầu hết các polymerase bền nhiệt hoạt động.

Tránh dùng mỗi có các vùng trình tự polybase như polydG hay các kiểu lặp lại. Chúng có thể lai với các vùng không thích hợp trên khuôn.

Nên tránh trình tự lặp lại ngược để ngăn mỗi tạo ra cấu trúc bậc hai mà không thể lai được với sợi khuôn.

Tránh dùng các trình tự mỗi bổ trợ cho nhau trong PCR để ngăn chúng lai với nhau. Đặc biệt quan trọng đối với đầu 3' của mỗi.

Nên thiết kế đầu 3' của mồi có nhiều base G và C để nâng cao độ gắn đính đầu, nó sẽ kéo dài.

Khoảng cách giữa các mồi cần phải ngắn hơn 10 kb. Thông thường, hiệu suất sẽ giảm đáng kể khi kéo dài mồi trên 3 kb.

Nhiệt độ nóng chảy (T_m) của mồi : Nhiệt độ nóng chảy T_m của thế lai mồi có thể được tính theo một số công thức khác nhau. Tuỳ theo kinh nghiệm mà tính theo công thức nào.

Công thức 1 : $T_m = [4(G + C) + 2(A + T)]^{\circ}C$. Đây là công thức chung nhất hay được sử dụng, xác định gốc từ phân tử lai oligonucleotide, được thực hiện trong 1 M NaCl. Ở nồng độ muối thấp, công thức chỉ chính xác đối với mồi dài không quá 20 base

Công thức 2 : Công thức này được dùng cho các mồi có độ dài 14 – 70 base : $T_{ms} = \{81,5 + 16,6(\log_{10}[J^+] + 0,41(\%G + C) - (600/l)\}^{\circ}C$. Kinh nghiệm chung là T_m cao thấp hơn 3-5°C so với nhiệt độ được tính toán theo công thức. Ở đó $[J^+]$ = nồng độ phân tử của cation hoá trị đơn (ví dụ Na^+ từ NaCl) và l = độ dài của oligonucleotide. (%G+C) là giá trị phần trăm thực sự và không phải là giá trị phân số (ví dụ giá trị để chèn vào mồi 90% hàm lượng G+C có nghĩa là "90" chứ không phải "0,90").

Công thức 3 : $T_{mp} = \{22 + 1,46([2(G + C) + (A + T)])\}^{\circ}C$. Công thức này được dùng cho mồi có độ dài 20-35 base.

Nhiệt độ bắt cặp tính theo lý thuyết là nhiệt độ tham khảo dùng để khởi động thí nghiệm : a) Nhiệt độ bắt cặp thực sự có thể cao hơn 3-12°C so với nhiệt độ nóng chảy tính toán T_m ; b) Điều kiện nhiệt độ bắt cặp thực sự được xác định theo kinh nghiệm; c) Nhiệt độ bắt cặp cao nhất mà cho sản phẩm PCR tốt nhất được sử dụng.

Bảng 3.3. Ví dụ về cách tính T_m , T_{ms} và T_{mp} (50 mM K⁺).

Độ dài mồi (nt)	G	G	G	G	T_m	T_{ms}	T_{mp}
15 mer	G	G	G	G	46	46	46
15 mer	G	G	G	G	46	52	46
30 mer	G	G	G	8	46	46	46

Tính nồng độ mồi :

Nồng độ phân tử của mồi được tính theo độ hấp phụ của mồi ở bước sóng 260 nm (A260) và hệ số tiêu hao phân tử của mồi tại bước sóng này. Hệ số tiêu hao phân tử của mồi có thể được tính theo trình tự mồi và sau đó cộng thêm giá trị tiêu hao phân tử cho mỗi base tạo nên mồi. Từng base có hệ số tiêu hao phân tử tại bước sóng 260 nm như sau :

Bảng 3.4. Hệ số tiêu hao phân tử ở bước sóng 260 nm.

Dung dịch	Phân tử lượng	Giá trị A260
8.400	8.400	8.400
8.400	1.0	8.400
8.400	1.0	12.010
8.400	12.010	12.010
5'-TAGC-3'	12.010	12.010
5'-TAGC-3'	8.427	0.427

Ví dụ mồi có hệ số tiêu hao phân tử 42660 ở bước sóng 260 nm. Tương tự như vậy, một dung dịch mồi 10 μM sẽ cho một giá trị hấp phụ 0,427 ở bước sóng 260 nm.

3.2. Tối ưu phản ứng PCR

Các yếu tố quyết định sự thành công của PCR có thể chia thành bốn nhóm chính, đó là : khuôn (độ tinh sạch và nồng độ) ; mồi thiết kế (tính đặc hiệu, cấu trúc và nồng độ) ; các thành phần khác (dệm, MgCl_2 , phụ gia, polymerase, dNTP : nồng độ) ; các chế độ phản ứng (nhiệt độ và thời gian). PCR liên quan tới việc chuẩn bị mẫu, hỗn hợp phản ứng và mồi, tiếp theo là phát hiện và phân tích sản phẩm phản ứng. Những yếu tố này được thảo luận kỹ dưới đây.

3.2.1. Khuôn DNA

PCR là một công cụ cực kỳ mạnh để khuếch đại DNA, bởi vậy cần rất ít khuôn DNA. Nhưng để giảm khả năng tạo lỗi của Taq DNA polymerase, có thể sử dụng một nồng độ DNA cao hơn, mặc dù quá nhiều khuôn có thể tăng hàm lượng thể lẩn tạp và giảm hiệu suất.

Thường hàm lượng khuôn DNA trong khoảng 0,01 – 1 ng đối với plasmid hoặc DNA phage và 0,1 – 1 µg đối với DNA nhân tế bào, cho một phản ứng tổng thể tích 50 µl. Hàm lượng khuôn DNA cao hơn thường tăng hiệu suất thu hồi sản phẩm PCR không đặc hiệu nhưng nếu độ tin cậy tổng hợp là tối hạn, nên dùng hàm lượng khuôn DNA tối đa cho phép cùng với số chu kỳ PCR giới hạn để tăng tỷ lệ sản phẩm PCR đặc hiệu.

PCR rất đa dụng. Nhiều loại mẫu nucleic acid đều có thể phân tích được. Hầu hết PCR sử dụng DNA làm đích hơn là RNA, do độ bền và dễ tách chiết của phân tử DNA. Tuân theo một số quy tắc cơ bản, có thể tránh được một số vấn đề trong bước chuẩn bị DNA cho PCR. Các yêu cầu nhất thiết cho mọi mẫu DNA là chứa ít nhất một sợi DNA nguyên vẹn mang vùng cần nhân lên và mọi tạp chất cần được pha loãng đủ sao cho không ức chế bước polymer hoá trong PCR.

Hầu hết các phương pháp thường ngày đều thích hợp để tinh sạch DNA khuôn. Tuy lượng hoá chất (phenol, EDTA, proteinase K...) được sử dụng trong các quy trình tinh sạch DNA ít nhưng chúng ức chế mạnh Taq pol. Tùa DNA với ethanol và rửa tủa DNA nhiều lần với 70% ethanol có tác dụng loại bỏ tạp chất từ mẫu DNA.

Tuy mọi quy trình tinh sạch DNA đều có thể chấp nhận được cho mục đích PCR nhưng cách tốt nhất là sử dụng vài thủ thuật có thể trong khi tách chiết DNA để ngăn lẩn tạp xảy ra với DNA không mong muốn. Thường tỷ lệ pha loãng mẫu 1 : 5 với nước là đủ để pha loãng mọi tạp chất từ quá trình tinh sạch DNA. Phương pháp đơn giản nhất để tách chiết DNA từ tế bào như sau :

1. Dùng một que tăm cạo tế bào dưới móng tay, nạo trong môm hoặc cạo chân lông rồi hòa tế bào vào 20 µl nước.
2. Đối với dịch tế bào, ly tâm 5 phút với 1.200 – 1.500x g. Hoà tủa tế bào vào 1 ml muối đệm phosphate (PBS) và ly tâm thêm 5 phút với 1.200 – 1.500x g. Làm lại lần nữa. Các bước rửa PBS này loại môi trường và các yếu tố ức chế trên bề mặt tế bào. Sau bước rửa cuối cùng, hòa tủa tế bào vào 20 µl nước cất. Quá nhiều mảnh vỡ tế bào sẽ ức chế PCR. Nếu xảy ra, cần thiết phải pha loãng mẫu DNA.
3. Đối với mẫu vi khuẩn, dùng một que tăm và cạo răng, hoặc lau ngực, tai hoặc giữa các ngón chân cái. Hoà vật liệu vào 500 µl nước.

Đóng lạnh và giải đóng mẫu ba lần có lắc mạnh hoặc lắc rung giữa những lần làm lại để phá thành tế bào vi khuẩn. Tuy không phải tất cả DNA được giải phóng ra khỏi tế bào, nhưng cũng đủ lượng DNA cho PCR.

4. Ủ mẫu 5 phút ở 95 – 100°C để bắt hoạt DNase. Nếu không, DNase có thể cắt DNA mong muốn thành các phân đoạn không thích hợp cho PCR. Nếu mẫu DNA loãng, có thể cò bằng tủa ethanol. Mẫu này đã sẵn sàng cho PCR.

Dù tách chiết bằng cách nào, mẫu DNA dùng cho PCR đều được kiểm tra về hàm lượng và độ sạch tương đối. Ngay sau khi pha hỗn hợp chung, bổ sung một lượng nhỏ DNA vào ống đối chứng để phát hiện các chất ức chế PCR trong quá trình chuẩn bị mẫu DNA.

3.2.2. Mồi

Mồi là một phân đoạn nucleotide ngắn bổ trợ cho vùng DNA cần được khuếch đại trong PCR. Mồi bắt cặp với sợi khuôn DNA biến tính tạo ra vùng khởi động quá trình tổng hợp nên phân tử DNA mới. Mồi có thể đặc hiệu hoặc chung (universal) với trình tự nucleotide DNA đặc trưng. Mồi chung bổ trợ cho nhiều trình tự nucleotide là một bộ phân tử DNA đặc trưng nào đó. Do đó, chúng có thể liên kết rộng rãi với nhiều loại khuôn DNA.

Mồi là một trong các yếu tố quan trọng nhất quyết định sự thành công của PCR. Việc thiết kế mồi đặc hiệu và sử dụng nồng độ thích hợp sẽ dẫn đến PCR có sản phẩm hay không. Mục sau sẽ đi sâu và chi tiết hơn về thiết kế mồi tối ưu.

Kiểm tra trình tự mồi để tránh hình thành nhị phân mồi. Bổ sung một kẹp GC ở đầu 5' nếu vị trí hạn chế được đưa vào đây. Một hoặc hai nucleotide G hay C ở đầu 3' là tốt nhưng tránh nhiều quá (nó có thể cho kết quả các sản phẩm PCR không đặc hiệu). Tính bổ trợ hoàn chỉnh của 18 nt hoặc hơn là lý tưởng. Các chỉ dẫn chọn mồi :

1. Mồi PCR thường dài 15 – 30. Mồi dài hơn cho tính đặc hiệu cao hơn.
2. Hàm lượng GC phải là 40 – 60%. Nucleotide G và C phải được phân bố đều khắp trình tự mồi. Nhiều hơn ba nucleotide G hoặc C ở đầu 3' của mồi phải được tránh, do mồi không đặc hiệu có thể xảy ra.

3. Mỗi không được tự bổ trợ hoặc bổ trợ với mỗi khác trong hỗn hợp phản ứng, để tránh nhị phân mỗi và hình thành kẹp tóc.
4. Nhiệt độ nóng chảy của các mỗi biến không được khác nhau quá 5°C, như vậy hàm lượng và độ dài GC phải được chọn tương ứng.
5. Tất cả các vị trí bổ trợ giữa mỗi và khuôn DNA phải được xác nhận.
6. Mỗi suy biến phải có ít nhất 3 nucleotide bảo thủ nằm ở đầu 3'.
7. Xác định nhiệt độ nóng chảy và bắt cặp của mỗi : nếu mỗi ngắn hơn 25 nt, nhiệt độ nóng chảy T_m gần đúng được tính theo công thức : $T_m = [4(G + C) + 2(A + T)]$; G, C, A, T – số nt tương ứng trong mỗi. Nhiệt độ bắt cặp phải thấp hơn nhiệt độ nóng chảy khoảng 5°C. Nếu mỗi dài hơn 25 nt, nhiệt độ nóng chảy phải được tính toán sử dụng các chương trình máy tính chuyên dụng, ở đó các mỗi tương quan giữa các nt liền kề, ảnh hưởng của nồng độ muối được đánh giá.

Nồng độ mỗi có thể được sử dụng tối 3 μM , nhưng tỷ lệ giữa mỗi và khuôn cao có thể cho kết quả khuếch đại không đặc hiệu và tạo ra cấu trúc nhị phân mỗi. Nồng độ mỗi cần phải dưới 1 μM trừ khi có độ suy biến cao ; thường 0,2 μM là đủ cho các mỗi tương tự. Nên bảo quản mỗi ở lượng nhỏ.

Các gene ribosomal DNA vi khuẩn chứa các trình tự nt phổ biến đối với tất cả các vi khuẩn. Do đó, các mỗi chung vi khuẩn có thể được tổng hợp bằng bổ trợ cho các trình tự này. Ví dụ các trình tự mỗi chung vi khuẩn là mỗi xuôi 5'-GAT CCT GGC TCA GGA TGA AC-3' (20-mer) và mỗi ngược 5'-GGA CTA CCA GGG TAT CTA ATC-3' (21-mer).

Các dòng tế bào động vật chứa một trình tự đặc biệt được biết đến là gene *alu*. Có khoảng 900.000 bản sao gene *alu* được phân bố trên toàn bộ hệ genome người, cũng như các da bản sao được phân bố trên genome của các tế bào động vật khác. Do đó, gene *alu* có trình tự mỗi chung cho mọi dòng tế bào động vật. Mỗi *alu* có lợi đặc biệt khi nó liên kết theo cả hai chiều xuôi và ngược. Trình tự mỗi chung *alu* như sau : 5'-GTG GAT CAC CTG AGG TCA GGA GTT TC-3' (26-mer). Khi sử dụng các mỗi chung, nhiệt độ bắt cặp trên máy chu trình nhiệt hạ xuống thấp ở 40 – 55°C.

Thông thường, đơn vị molarity được tính theo độ hấp phụ quang học A_{260} ở bước sóng 260 nm. Nếu cần thiết, có thể chuyển đổi mật độ quang thành phân tử lượng (μM) sử dụng phương trình : $[A_{260} / (10.000 \times N/\text{m} \times \text{cm})] \times 10^6 = \text{nồng độ } (\mu\text{M})$; N là số base của molarity; khối lượng phân tử = $\sim 330 \times N$. Ví dụ molarity dài 20 base, $N = 20$; và $A_{260} = 10$; khối lượng phân tử = $\sim 330 \times 20 = 6.600$; nồng độ molarity là : $[10 A_{260} / (10.000 \times 20/\text{m} \times \text{cm})] \times 10^6 = 50 \mu\text{M}$.

3.2.3. dNTP

Nồng độ molarity loại dNTP (dATP, dCTP, dGTP, dTTP) trong hỗn hợp phản ứng thường là 200 μM . Một điều quan trọng là nồng độ molarity loại dNTP phải như nhau. Chỉ cần một dNTP nào đó chênh lệch về nồng độ có thể gia tăng mức độ gắn nhầm nucleotide vào sản phẩm PCR, dẫn đến kết quả trình tự có lỗi và sản phẩm biểu hiện sẽ có thể là một thảm họa.

Nồng độ nucleotide cần không quá 50 μM molarity, tuy nhiên sản phẩm dài có thể cần nhiều hơn. Đánh dấu sản phẩm PCR với digoxigenin-11-dUTP (dig) cần được thực hiện với nồng độ 50 μM molarity dNTP, với dTTP được thay thế bởi 35% với dig-11-dUTP. Sản phẩm sẽ có khối lượng phân tử cao hơn sản phẩm nguyên thủy.

Khi độ tin cậy tối đa của quy trình PCR là tối hạn, nồng độ dNTP cuối cùng phải là 10 – 50 μM , do độ tin cậy tổng hợp DNA là tối đa trong dãy nồng độ này. Ngoài ra, nồng độ MgCl_2 phải được chọn nghiêm ngặt, bắt đầu từ 1 mM và tăng lên 0,1 mM mỗi bước, cho tới khi thu được hiệu suất có đủ sản phẩm PCR. Trong các kit thương mại, người ta thường pha chế hai loại nồng độ 2 mM và 10 mM để sử dụng cho PCR.

Nồng độ dNTP có thể được sử dụng tới 1,5 mM. Do dNTP bắt kẹp các ion Mg^{2+} , nên nồng độ $[\text{Mg}^{2+}]$ sử dụng có thể cần được thay đổi. Tuy nhiên, thừa dNTP có thể tăng tỷ lệ lỗi và có khả năng ức chế *Taq* pol. Bởi vậy, giảm dNTP (10 – 50 μM) có thể cũng giảm tỷ lệ lỗi. Phân đoạn PCR kích thước lớn hơn cần nhiều dNTP hơn.

3.2.4. Polymerase

Taq DNA polymerase có một tỷ lệ lỗi cao hơn *Tli* hoặc *Pfu* do không có hoạt tính đọc kiểm tra 3' → 5' exonuclease. Sử dụng *Tli*, *Pfu* hoặc các

polymerase khác với khả năng đọc kiểm tra tốt nếu cần độ tin cậy cao. Tuy nhiên, *Taq* pol ít phức tạp hơn các polymerase khác và ít khi thất bại. Người ta có thể sử dụng *Taq* pol cùng với các enzyme khác để tăng độ tin cậy của phản ứng (bảng 1.1 và 1.2).

Taq pol cũng có khuynh hướng bổ sung thêm adenine vào đầu 3'. Đây là điểm rất thuận lợi để nhận dòng TA nhưng lại cản trở nhận dòng đầu bằng, cần loại bỏ A trước khi thực hiện phản ứng gắn dính đầu bằng. Nhiều enzyme cũng có thể được bổ sung vào PCR để tăng hiệu suất tổng hợp, do *Taq* pol có thể bị biến tính trong chu trình lặp lại, nhưng có thể tăng các sản phẩm PCR không đặc hiệu. Vent polymerase có thể phân huỷ mồi và bởi vậy không lý tưởng lắm cho việc gây đột biến bằng PCR.

Thường 1 – 1,5 unit *Taq* DNA polymerase được sử dụng trong một hỗn hợp phản ứng 50 µl. Hàm lượng *Taq* pol cao hơn có thể tổng hợp nhiều sản phẩm PCR không đặc hiệu. Tuy nhiên, nếu các chất ức chế có mặt trong hỗn hợp phản ứng (ví dụ khuôn DNA sử dụng không được tinh sạch cao), hàm lượng *Taq* pol (2-3 unit) cao hơn có thể là cần thiết để thu hiệu suất sản phẩm khuếch đại cao hơn.

3.2.5. Nồng độ Mg²⁺

Nồng độ [Mg²⁺] ảnh hưởng tới bắt cặp mồi ; T_m của sợi khuôn, sản phẩm và liên kết mồi-khuôn ; tính đặc hiệu sản phẩm ; hoạt tính enzyme và tính chính xác. [Mg²⁺] là một biến số chính, thay đổi hàm lượng bổ sung vào phản ứng nếu kết quả PCR xấu. Mg²⁺ ảnh hưởng tới sự bắt cặp oligo với khuôn DNA bằng cách bền hóa mối tương tác oligo-khuôn, cũng như phức hợp sao chép của polymerase với khuôn-mồi. Vì vậy nó cũng tăng sự bắt cặp không đặc hiệu và tổng hợp các sản phẩm PCR không mong muốn (cho nhiều băng trên gel). EDTA kẹp càng cua Mg²⁺ có thể thay đổi nồng độ [Mg²⁺].

Do ion Mg²⁺ tạo ra các phức hợp với dNTP, mồi và khuôn DNA, nồng độ tối thích của MgCl₂ được lựa chọn cho từng thí nghiệm. Quá ít ion Mg²⁺ cho hiệu suất thu nhận sản phẩm PCR thấp và quá nhiều ion tăng hiệu suất thu các sản phẩm không đặc hiệu và tăng gắn nhầm. Nồng độ [Mg²⁺] thấp hơn được mong muốn khi độ tin cậy tổng hợp DNA là tối hạn.

Taq pol cần các ion Mg^{2+} tự do. Do các thành phần như dNTP, mồi và khuôn, tất cả đều tạo càng cua và thu giữ cation ; trong đó dNTP có nồng độ cao nhất, do đó $[Mg^{2+}]$ ít nhất phải là $0,5 - 2,5$ mM cao hơn [dNTP]. Cần phải thực hiện chuẩn độ $[Mg^{2+}]$ khác nhau với các tổ hợp khuôn-mồi mới, do chúng có thể đòi hỏi rất khác nhau, ngay cả khi dưới cùng các điều kiện và thời gian/nhiệt độ chu trình.

Dãy nồng độ $[MgCl_2]$ được đề xuất là $1 - 4$ mM, dưới các điều kiện phản ứng chuẩn đặc biệt. Trong các thí nghiệm, nồng độ [dNTP] cuối cùng là $0,2$ mM, nồng độ $[MgCl_2]$ trong khoảng $1,5 \pm 0,25$ mM (trong đệm *Taq* với KCl) và $2,0 \pm 0,5$ mM [trong đệm *Taq* với $(NH_4)_2SO_4$] là thích hợp trong hầu hết các trường hợp. Nồng độ $[MgCl_2]$ trong hỗn hợp phản ứng phải được tăng lên theo tỷ lệ thuận. Nhiều loại kit PCR thương mại đã đơn giản hóa sự tối ưu hoá, chứa tất cả các hoá chất cần thiết và đưa ra hiệu suất PCR tối ưu.

3.2.6. Phụ gia

Phụ gia là những hóa chất dùng để bổ sung vào PCR nhằm thu các sản phẩm PCR mong muốn nhiều hơn. Phụ gia thường dùng cho các phân đoạn PCR rất dài, giàu GC hoặc để hình thành cấu trúc bậc hai. Các phụ gia glycerol, formamide, NMP có tác dụng hạ thấp nhiệt độ bắt cặp vài độ và giảm biến tính, trong khi DMSO giảm hình thành cấu trúc bậc hai.

Glycerol (5 – 10%), formamide (1 – 5%) hoặc DMSO (2 – 10%) có thể được bổ sung vào PCR đối với khuôn DNA có hàm lượng GC cao, chúng sẽ thay đổi nhiệt độ nóng chảy T_m của phản ứng lai mồi-khuôn và tính bền nhiệt của polymerase. Glycerol có thể bảo vệ *Taq* pol trước sự phá huỷ nhiệt, trong khi formamide có thể làm giảm khả năng chống chịu của enzyme.

Betaine với nồng độ $0,5 - 2$ M cũng thích hợp cho PCR có hàm lượng GC cao và các dài DNA dài (long PCR, LA PCR). Chuẩn độ để xác định nồng độ tối ưu, thường là $1,3$ M. Giảm nhiệt độ nóng chảy xuống $92 - 93^\circ C$ và nhiệt độ bắt cặp thấp hơn $1 - 2^\circ C$ khi sử dụng betaine. Có thể sử dụng betaine cùng với các phụ gia khác như 5% DMSO.

Nhiều kit thương mại thường giữ bí mật khi bổ sung betaine hay các phụ gia khác tuy nó không đắt.

Các loại phụ gia khác như TMAC (tetramethyl ammonium chloride), TEAC (tetraethyl ammonium chloride) và TMANO (trimethylamine N-oxide) cũng có thể được sử dụng với nồng độ cao hơn 50 mM. Albumin huyết thanh bò BSA với nồng độ tới 0,8 µg/µl cũng có thể cải thiện hiệu suất PCR.

3.2.7. Đệm PCR

Thường các kit thương mại đều cung cấp đệm đã pha sẵn. Nhưng người ta có thể sử dụng nồng độ đệm PCR tự pha chế cao hơn để nâng cao hiệu suất. Đệm tự pha chế có thể hoạt động tốt hơn đệm mua. Đệm bao gồm các thành phần : 16,6 mM ammonium sulfate ; 67,7 mM Tris-HCl, pH 8,9 ; 10 mM β-mercaptoethanol ; 170 µg/ml BSA ; 1,5 – 3 mM MgCl₂.

Đệm PCR trong kit thường chứa các thành phần sau : 10 – 50 mM Tris-HCl, pH 8,3 ; tới 50 mM KCl ; 1,5 mM MgCl₂ hoặc cao hơn ; tới 100 µg/ml gelatin hoặc BSA ; và/hoặc [0,05 – 0,10% (v/v)] chất tẩy rửa không ion như Tween-20 hoặc Nonidet P-40 hoặc Triton X-100.

Nồng độ muối KCl hoặc NaCl cao hơn 50 mM sẽ ức chế *Taq* pol, nhưng lại cần thiết để tạo điều kiện dễ dàng bắt cặp mồi.

Một số enzyme không cần bổ sung protein, một số khác lại phụ thuộc vào nó. Một số enzyme hoạt động tốt hơn với sự có mặt của chất tẩy rửa, có lẽ do nó ngăn cản xu hướng tự nhiên của enzyme liên kết với nhau.

PCR thường cũng có thể hoạt động tốt trong đệm PCR phiên mã ngược (reverse transcriptase PCR) và ngược lại, điều này có nghĩa là có thể thực hiện được một hỗn hợp PCR và PCR phiên mã ngược trong cùng một ống nghiệm (với tổng hợp cDNA và PCR tiếp theo).

3.2.8. Lớp chống bốc hơi

Các loại máy chu trình nhiệt thế hệ cũ đều chưa thiết kế nắp nhiệt (105°C) để chống các ống PCR bốc hơi. Với các loại máy thế hệ cũ này, trước khi PCR xảy ra ở các bước biến tính nhiệt độ cao thường 94°C, hồn

hợp phản ứng bắt đầu bốc hơi. Nếu không có lớp chống bốc hơi thì sau một số chu kỳ hồn hợp phản ứng sẽ bị bốc hơi hết, PCR sẽ không tiếp tục được nữa và hiệu suất phản ứng sẽ rất kém hoặc không có sản phẩm PCR.

Khi cần thiết đổi với các loại máy cũ này, hồn hợp phản ứng có thể được tráng với một lớp dầu khoáng, paraffin (nhiệt độ nóng chảy 50-60°C), sáp hoặc vaseline có độ tinh khiết đặc biệt cho PCR. Thể tích lớp chống bốc hơi thường vào khoảng một nửa thể tích phản ứng là đủ.

Cần bảo đảm dầu khoáng phủ hoàn toàn tiết diện hồn hợp phản ứng để tránh bốc hơi. Nếu máy chu trình nhiệt có nắp đốt nóng để tránh bốc hơi, thì dầu khoáng không cần thiết, nhưng thời gian chạy lâu hơn thông thường.

3.2.9. Chuẩn bị hồn hợp PCR

3.2.9.1 Chuẩn bị khuôn

PCR cho phép tổng hợp hơn 10 triệu bản sao của trình tự DNA đích chỉ từ một số ít phân tử ban đầu. Độ nhạy của kỹ thuật có nghĩa là mẫu không được lân tạp với các DNA khác hoặc các sản phẩm khuếch đại trước đó (thể khuếch đại) mà chúng có thể tồn tại trong không khí phòng thí nghiệm. Do đó, người ta phải tách chiết và tinh sạch DNA đến độ tinh khiết không lân tạp DNA khác và thể khuếch đại trước đó.

1. Nên thực hiện các bước chuẩn bị mẫu DNA, pha chế hồn hợp phản ứng, thực hiện qui trình PCR, tiếp theo phân tích sản phẩm phản ứng ở nơi dành riêng cho PCR.
2. Dùng tủ vô trùng laminar có trang bị đèn UV để pha chế hồn hợp phản ứng.
3. Mang găng tay mới để khuếch đại DNA và mỗi phản ứng khởi động.
4. Nên sử dụng chai lọ, pipet đầy dầu tip và tip lọc khí để chuẩn bị cả mẫu DNA lân hồn hợp phản ứng.
5. Các hóa chất cho PCR phải được pha chế riêng và chỉ sử dụng cho mục đích này. Nên khử trùng tất cả các dung dịch, trừ dNTP, mồi và *Taq* DNA polymerase. Các dung dịch được chia thành phần nhỏ và bảo quản ở nơi dành riêng cho PCR và tách khỏi các mẫu DNA.

- Một phản ứng đối chứng âm không khuôn DNA luôn luôn được thực hiện, để khẳng định không có mặt của tạp chất.
- Đây chỉ là các chỉ dẫn sơ đẳng. Các chỉ dẫn chi tiết về khởi động phòng thí nghiệm PCR và giữ gìn có thể được tìm thấy trong các tài liệu chuyên ngành.

3.2.9.2. Chuẩn bị hỗn hợp phản ứng

Để thực hiện nhiều phản ứng song song, pha chế một hỗn hợp chung gồm nước, đệm, dNTP, mồi và *Taq* pol trong ống to, sau đó chia nhỏ vào từng ống PCR. Tiếp theo, dung dịch mồi và khuôn DNA được bổ sung vào. Cách pha chế này hạn chế khả năng hút pipet nhầm, giảm số lần thao tác hoá chất và do đó tiết kiệm thời gian.

Tùy theo mục đích của các PCR, mà người ta pha chế hỗn hợp chung. Nếu muốn thu một số lượng lớn của một phân đoạn DNA nào đó, người ta có thể pha một hỗn hợp chung chứa tất cả các thành phần cần thiết (bảng 3.5) để tổng hợp sợi DNA mới trong quá trình PCR, sau đó chia nhỏ sang các ống PCR. Nếu muốn tối ưu hoá một thành phần PCR, người ta pha một hỗn hợp chung chứa tất cả các thành phần cần thiết trừ thành phần cần tối ưu ra. Sau khi chia nhỏ vào các ống PCR, thành phần cần tối ưu được bổ sung vào với nồng độ khác nhau. Khi thực hiện đồng thời nhiều PCR với các sản phẩm PCR và mồi khác nhau, hỗn hợp chung không chứa khuôn và mồi (bảng 3.5).

Bảng 3.5. Nồng độ và thể tích các thành phần cho 50 µl PCR.

TT	Thành phần	Nồng độ cuối cùng	Mục đích	Thể tích
1	Nước cất khử ion	–	Bù cho thể tích nhất định	Thay đổi
2	Đệm 10x PCR (± MgCl ₂)	1x	Giữ pH ổn định để PCR xảy ra	5 µl
3	25 mM MgCl ₂	1 – 4 mM		Thay đổi
4	Deoxy nucleotide 2 mM	200 µM	Cung cấp năng lượng và nucleoside để tổng hợp DNA. Nồng độ mỗi nucleotide (dATP, dTTP, dCTP, dGTP) bổ sung vào hỗn hợp chung đều nhau để ngăn sai base.	5 µl

5	Polymerase	1,25 unit/ 50 µl	Enzyme bén nhiệt gắn dNTP vào khuôn DNA	Thay đổi
6	Mỗi xuôi và ngược	0,1 – 1 µM	Đoạn DNA ngắn (20-30 base) liên kết với khuôn DNA cho phép <i>Taq</i> pol khởi động phản ứng gắn dNTP. Cả mỗi đặc hiệu lần mồi chung có thể được sử dụng	Thay đổi
7	Sợi khuôn DNA	10 pg – 1 µg	Làm khuôn để khuếch đại phân đoạn DNA mong muốn	Thay đổi

Sau khi giải đông, lắc rung nhẹ và ly tâm ngắn tất cả các dung dịch. Bổ sung vào ống PCR thành mỏng trên đá lạnh những hoá chất cần thiết của hỗn hợp chung tùy theo mục đích sử dụng. Trộn đều, ly tâm thu dịch và chia đều sang các ống PCR. Sau đó bổ sung các thành phần còn lại. Lắc rung nhẹ mẫu và ly tâm ngắn để thu hơi nước từ thành ống. Tra lên mẫu một lớp dầu khoáng hoặc bổ sung một lượng sáp thích hợp. Đối với loại máy chu trình nhiệt có trang bị nắp đốt nóng, không cần thiết phải tra dầu khoáng. Sau đó, đặt mẫu vào máy chu trình nhiệt và khởi động PCR.

Nồng độ thành phần PCR càng ít càng tốt, như thế sản phẩm PCR sẽ đặc hiệu hơn. Nồng độ cuối cùng của một số thành phần thường như sau : 0,1 – 1 µM mỗi ; 1 – 4 mM MgCl₂ (phụ thuộc vào *Taq* pol sử dụng) ; 0,2 mM mỗi loại dNTP (phụ thuộc vào *Taq* pol sử dụng). Các thành phần PCR thường nhạy với đông lạnh và giải đông cho nên người ta thường chia nhỏ để bảo quản và sử dụng. Lắc rung hoặc búng tay các hoá chất thật kỹ trước khi sử dụng.

Vật liệu và hoá chất

1. Khuôn DNA (ví dụ khuôn GAPDH) : nồng độ 0,5 – 2 ng/10 µl.
2. Mỗi xuôi và ngược (ví dụ cho GAPDH) : 50 µM mỗi loại.
3. Đệm 10x PCR. *Taq* pol : 5 unit/µl *Taq* polymerase.
4. Hỗn hợp dNTP : 10 mM mỗi loại dNTP.
5. Nước cất sạch nuclease.

Tiến hành

1. Khởi động hai phản ứng sau đây với tổng thể tích 50 µl mỗi ống.
Ống mẫu : 5 µl đệm 10x PCR ; 1 µl hỗn hợp dNTP ; 1,25 unit *Taq* pol ;

40 μ l nước cất ; búng nhẹ ống trộn đều và ly tâm 2 giây. Bổ sung 1 μ l mỗi xuôi ; 1 μ l mỗi ngược ; 2 μ l khuôn. Trong ống đối chung âm, khuôn được thay bằng nước cất. Tra 1-2 giọt (20 – 40 μ l) dầu khoáng khi cần thiết.

2. Chạy PCR : 30x (94°C/45'', 60°C/45'', 72°C/2').
3. Lấy ra 5 μ l từ mỗi ống phản ứng sau 15, 20, 25 chu kỳ và bảo quản ở 4°C. Bảo quản phần còn lại sau 30 chu kỳ ở 4°C.
4. Chạy điện di gel agarose sử dụng các hỗn hợp phản ứng bảo quản. Sau điện di gel agarose, nhìn thấy một băng có kích thước mong đợi (452 bp) trong ống mẫu, nhưng không thấy trong ống đối chung âm.

Những điều cần lưu ý

1. Đệm PCR thường được pha chế dưới dạng dung dịch gốc 10x (100 mM Tris-HCl, pH 8,3 ; 500 mM KCl ; 1,5 mM MgCl₂) và pha loãng 1x khi dùng. Cả đệm PCR và nước tinh khiết có thể được bảo quản ở nhiệt độ phòng. Bảo quản dNTP, mỗi và *Taq* pol ở -20°C.
2. Thường người ta dùng 100 μ l hỗn hợp PCR cho mỗi phản ứng, nhưng có thể rút xuống 25 hoặc 50 μ l để tiết kiệm hóa chất. Không phụ thuộc vào thể tích, nồng độ các thành phần không thay đổi.
3. Thành phần hỗn hợp chung có thể được mua từ các công ty khác nhau. Nồng độ các hóa chất ban đầu thường khác nhau tùy theo hãng sản xuất. Để dàng tính được lượng dung dịch gốc cần dùng theo công thức đơn giản sau : $C_i \times V_f = C_f \times V_i$, trong đó C_i và V_i là nồng độ và thể tích dung dịch gốc, C_f và V_f là nồng độ và thể tích dung dịch pha chế.
4. Ví dụ đệm 10x PCR ; 10 mM mỗi loại nucleotide ; 0,5 mM mỗi và *Taq* pol với 5 unit/ μ l.
5. **Đệm 10x PCR :** $(10x) \times (5 \mu\text{l}) = (1x) \times (50 \mu\text{l})$.
6. **Dung dịch nucleotide :** $(10.000 \mu\text{M}) \times (1 \mu\text{l}) = (200 \mu\text{M}) \times (50 \mu\text{l})$; $10 \text{ mM} = 10.000 \mu\text{M}$.

- 7. Dung dịch môi :** $(500 \mu\text{M}) \times (0,1 \mu\text{l}) = (1,0 \mu\text{M}) \times (50 \mu\text{l})$. Do không thể hút pipet chính xác 0,1 μl nên trước hết phải pha loãng. Bổ sung 10 μl dung dịch môi gốc vào 990 μl nước để nhận môi nồng độ 5 μM . Dung dịch môi mới này có thể được bảo quản ở 4°C. Tính dung dịch gốc 5 μM : $(5 \mu\text{M}) \times (10 \mu\text{l}) = (1,0 \mu\text{M}) \times (50 \mu\text{l})$.
- 8. Taq pol :** $(5 \text{ unit}/\mu\text{l}) \times (0,25 \mu\text{l}) = (0,025 \text{ unit}/\mu\text{l}) \times (50 \mu\text{l})$ hoặc 2,5 unit cho phản ứng 100 μl . Do không thể hút pipet chính xác 0,25 μl nên trước hết pha loãng 1,25 unit/ μl . Bổ sung 1,25 μl dung dịch gốc vào 3,75 μl nước để có nồng độ 1,25 unit/ μl . Tính dung dịch gốc 1,25 unit/ μl : $(1,25 \text{ unit}/\mu\text{l}) \times (1 \mu\text{l}) = (0,025 \text{ unit}/\mu\text{l}) \times (50 \mu\text{l})$. Bỏ dung dịch cũ và pha mới dung dịch môi khi cần.
- 9. Pha hỗn hợp chung cho một phản ứng** như sau : 5 μl dệm 10x PCR ; 4 μl môi nucleotide (1 μl môi loại dATP, dCTP, dGTP, dTTP) ; 20 μl môi (10 μl môi môi) ; 1 μl Taq pol (tổng thể tích 30 μl) ; bổ sung 15 μl nước ; 5 μl sợi khuôn (tổng thể tích 50 μl). Nếu pha cho 3 phản ứng : $3 \times 50 \mu\text{l} = 150 \mu\text{l}$. Sử dụng số này trong công thức cho thể tích mẫu.
- 10.** Chạy đối chứng âm không có khuôn DNA để kiểm tra sự lẩn tặc, còn đối chứng dương của một khuôn DNA biết trước để kiểm tra thành phần PCR. Luôn luôn chạy thang chỉ thị trên gel để xem PCR có thành công hay không đồng thời kiểm tra hệ thống phát hiện DNA.

3.2.10. Chu trình nhiệt

Các thông số khuếch đại phân lớn phụ thuộc vào khuôn, môi và thiết bị khuếch đại được sử dụng. Thời gian biến tính có thể tăng lên nếu hàm lượng GC khuôn cao. Mỗi dài với hàm lượng GC cao cần nhiệt độ bắt cặp T_m cao hơn. Để xác định nhiệt độ bắt cặp, cách tốt nhất là sử dụng thang nhiệt độ trên máy PCR gradient.

Sản phẩm PCR dài cần thời gian tổng hợp lâu hơn ; nhưng giảm xuống để hạn chế sự biến tính enzyme. Thời gian tổng hợp DNA cũng phụ thuộc vào loại enzyme sử dụng, ví dụ Taq pol tổng hợp DNA với tốc độ 1.000 nt/phút. Nên kiểm tra các chỉ dẫn của nhà cung cấp, tốc độ tổng hợp thực sự cao hơn nhiều.

Số chu kỳ có thể được tăng lên nếu số khuôn DNA rất thấp và giảm nếu hàm lượng khuôn DNA cao được sử dụng. Số khuôn DNA cao hơn sẽ hạ thấp tỷ lệ lỗi trong PCR và tối ưu cho nhân dòng PCR.

3.2.10.1. Bước biến tính khởi động

Liên kết bổ trợ đặc hiệu bằng cầu nối hydro giữa các nucleic acid sợi đơn được hiểu là **bắt cặp** hay annealing : hai trình tự bổ trợ sẽ tạo ra các liên kết hydro giữa các base bổ trợ (G với C và A với T hay U) và tạo ra một phân tử lai (hybrid) song song đối chiều, sợi kép bền.

Người ta tạo ra sợi đơn nucleic acid (NA) nhằm mục đích tái bắt cặp nếu chúng chưa ở dạng sợi đơn như hầu hết các RNA virus. Khi đun nóng lên một điểm trên nhiệt độ nóng chảy của các sợi kép hoàn toàn hoặc sợi kép từng phần và sau đó làm nguội thật nhanh thì chúng sẽ bị biến tính hoàn toàn, có nghĩa là các sợi tách nhau ra và không tái bắt cặp. Hơn nữa, nếu NA được đun nóng trong đậm với cường độ ion thấp hơn 150 mM NaCl, nhiệt độ nóng chảy thường thấp hơn 100°C, điều đó lý giải tại sao PCR hoạt động với nhiệt độ biến tính 91 – 97°C.

Biến tính khuôn DNA hoàn toàn ở bước khởi động PCR là chìa khoá quyết định khuếch đại thành công. Biến tính DNA không hoàn toàn cho kết quả sử dụng khuôn không hiệu quả trong chu kỳ khuếch đại đầu tiên và cho hiệu suất thu sản phẩm PCR thấp. Biến tính đầu tiên phải được thực hiện hơn 1 – 3 phút ở 95°C nếu hàm lượng GC là 50% hoặc ít hơn. Thời gian biến tính có khi lên tới 10 phút đối với khuôn giàu GC.

Nếu biến tính đầu tiên không dài hơn 3 phút ở 95°C, *Taq* pol có thể được bổ sung vào hỗn hợp phản ứng ban đầu. Nếu biến tính ban đầu lâu hơn hoặc nhiệt độ cao hơn là cần thiết, *Taq* pol phải được bổ sung sau khi biến tính ban đầu, do độ bền của enzyme giảm mạnh ở nhiệt độ trên 95°C.

3.2.10.2. Bước biến tính

Thường biến tính 0,5 – 2 phút ở 94 – 95°C là đủ, do sản phẩm PCR được tổng hợp ở chu kỳ khuếch đại thứ nhất ngắn hơn đáng kể so với khuôn DNA ban đầu và được biến tính hoàn toàn dưới các điều kiện này. Nếu DNA khuếch đại có hàm lượng GC rất cao, thời gian biến tính có thể được tăng lên 3 – 4 phút.

Để biến tính DNA dễ dàng, người ta cũng có thể sử dụng các phụ gia như glycerol ($\leq 10 - 15\%$), DMSO ($\leq 10\%$) hoặc formamide ($\leq 5\%$). Khi có mặt của phụ gia như vậy, nhiệt độ nóng chảy của sợi kép mồi-khuôn DNA giảm đáng kể, do đó nhiệt độ bắt cặp phải được chỉnh theo thí nghiệm. DMSO và formamide ở nồng độ sử dụng ức chế *Taq* pol khoảng 50%, do đó hàm lượng enzyme trong hỗn hợp phản ứng phải được tăng lên.

Cách thứ ba là thay dGTP bằng 7-deaza-dGTP trong hỗn hợp phản ứng sẽ giảm nhiệt độ nóng chảy của sản phẩm PCR.

Taq pol có thời gian bán hoạt động là 30 phút ở 95°C , đây là một phản lý giải tại sao người ta chỉ thực hiện PCR không quá 30 chu kỳ. Tuy nhiên, có thể giảm nhiệt độ biến tính sau khoảng 10 chu kỳ khuếch đại, do độ dài trung bình của DNA đích bị giảm xuống. Đối với khuôn 300 bp hoặc ít hơn, nhiệt độ biến tính có thể được giảm xuống thấp như 88°C đối với khuôn 50% GC. Điều này có nghĩa người ta có thể tổng hợp 40 chu kỳ mà không giảm nhiều hoạt tính enzyme.

"Đúng nhiệt độ" là nguyên nhân chính gây biến tính hoặc mất hoạt tính *Taq* pol, do vậy, nếu người ta hạn chế được điều này, thì có thể tăng số chu kỳ lên, dù nhiệt độ có được giảm hay không. Thường thời gian biến tính là 1 phút ở 94°C , đối với các trình tự ngắn có thể giảm xuống 30 giây hoặc ít hơn. Đôi khi người ta lại tăng nhiệt độ biến tính lên 96°C và giảm thời gian xuống còn 15 giây.

3.2.10.3. Bước gắn mồi

Thường nhiệt độ bắt cặp tối thích thấp hơn 5°C so với nhiệt độ nóng chảy của sợi kép mồi-khuôn DNA. Thời gian gắn mồi kéo dài 0,5 – 2 phút là đủ. Tuy nhiên, nếu sản phẩm PCR không đặc hiệu thu được ngoài sản phẩm mong đợi, nhiệt độ bắt cặp phải được tối ưu bằng cách tăng từng bước $1 - 2^{\circ}\text{C}$.

Độ dài và trình tự mồi là những yếu tố quan trọng làm thông số thiết kế để khuếch đại thành công. Nhiệt độ nóng chảy T_m của sợi kép NA tăng theo độ dài mồi và hàm lượng GC cao. Một công thức đơn giản để tính nhiệt độ nóng chảy là $T_m = [4(G + C) + 2(A + T)]^{\circ}\text{C}$.

Do vậy, nhiệt độ bắt cặp được chọn cho PCR phụ thuộc trực tiếp vào độ dài và thành phần mồi. Người ta nên sử dụng nhiệt độ bắt cặp T_a (annealing temperature) thấp hơn 5°C so với nhiệt độ nóng chảy T_m của cặp mồi được sử dụng. Có thể xử lý nhiệt độ T_a chính xác hơn, nếu T_a được tăng 1°C với mỗi chu kỳ, thì tính đặc hiệu của khuếch đại và hiệu suất thu sản phẩm có độ dài nhỏ hơn 1 kb được tăng lên.

Hệ quả của việc giảm nhiệt độ bắt cặp T_a quá thấp là một hay cả hai mồi sẽ bắt cặp với những trình tự khác đích, do đó những base đơn bắt cặp nhầm có thể xảy ra. Điều này rất tốt nếu người ta muốn khuếch đại các trình tự đích tương tự hay họ hàng. Tuy nhiên, nó có thể dẫn tới khuếch đại không đặc hiệu và giảm hiệu suất sản phẩm mong muốn, nếu base đầu 3' bắt cặp với trình tự đích.

Hệ quả của nhiệt độ bắt cặp T_a quá cao là rất ít sản phẩm được tổng hợp ra, do sự bắt cặp mồi bị giảm. Một quan sát quan trọng khác là cặp mồi với nhiệt độ bắt cặp rất khác nhau có thể không cho hiệu suất thích hợp của một sản phẩm duy nhất và có thể còn cho kết quả khuếch đại bắt đối xứng không mong muốn hoặc sợi đơn của sợi sản phẩm được mồi hiệu quả nhất.

Thời gian bắt cặp không lâu : hầu hết mồi sẽ bắt cặp hiệu quả trong 30 giây hoặc ít hơn, trừ khi T_a quá gần với T_m , hoặc chúng quá dài.

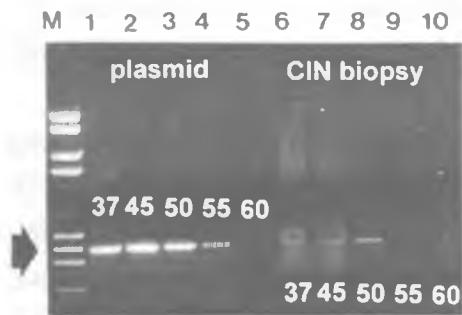
3.2.10.4. Bước tổng hợp

Thường bước tổng hợp (kéo dài mồi) được thực hiện ở 70-75°C. Tốc độ tổng hợp DNA bằng *Taq* DNA polymerase đạt cao nhất ở nhiệt độ này. Thời gian tổng hợp phân đoạn PCR dài 2 kb được đề xuất là 1 phút. Khi khuếch đại các phân đoạn PCR dài hơn, thời gian tổng hợp được tăng lên 1 phút cho mỗi kb.

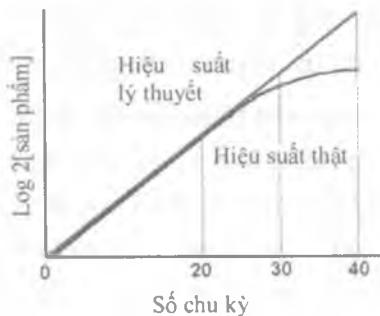
Nhiệt độ và thời gian tổng hợp chuẩn là 70 – 72°C và 0,5 – 3 phút. *Taq* pol thực sự có hoạt tính đặc hiệu ở 37°C gần với nhiệt độ của phân đoạn Klenow từ *E. coli* DNA polymerase I. Điều này giải thích một nghịch lý rõ ràng là mồi bắt cặp ở nhiệt độ tối ưu, nhưng sau đó có thể được kéo dài ở một nhiệt độ cao hơn đáng kể. Câu trả lời là tổng hợp xảy ra ngay từ lúc mới bắt cặp, ngay cả ở thời gian ngắn ngủi, kết quả là độ

bên lớn hơn đáng kể. Ở khoảng 70°C hoạt tính tối thích và kéo dài mồi xảy ra với tốc độ 100 nt/giây. Khoảng thời gian 1 phút là đủ để khuếch đại trình tự dài 2 kb. Sản phẩm dài hơn cần thời gian tổng hợp dài hơn. Ba phút là đủ cho sản phẩm 3 kb hoặc dài hơn.

Thời gian dài hơn có thể cũng có tác dụng trong những chu kỳ cuối cùng khi sản phẩm vượt qua ngưỡng nồng độ enzyme (>1 nM), khi dNTP và mồi trở nên cạn kiệt.



Hình 3.8. Tác dụng của nhiệt độ bắt cặp đến tính đặc hiệu và hiệu suất khuếch đại (theo Coyne et al. 2001).



Hình 3.9. Hiệu ứng bão hoà trong phản ứng khuếch đại PCR.

Hình 3.8 minh họa tác dụng của nhiệt độ bắt cặp đến tính đặc hiệu và hiệu suất khuếch đại papillomavirus type 16 ở người (HPV-16). Khuôn plasmid và DNA mẫu sinh thiết được khuếch đại ở những nhiệt độ bắt cặp khác nhau 37, 45, 50, 55 và 60°C. Trong khi plasmid được khuếch đại từ 37 đến 55°C, thì mẫu sinh thiết chứa HPV DNA chỉ được khuếch đại đặc hiệu ở 50°C.

3.2.10.5. Số chu kỳ

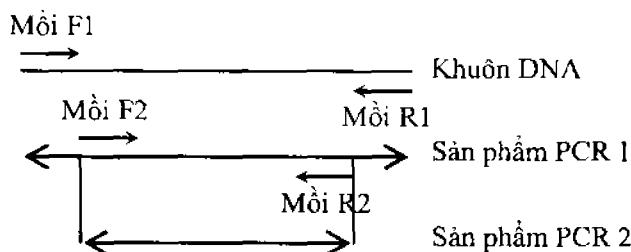
Số chu kỳ PCR phụ thuộc vào số lượng khuôn DNA trong hỗn hợp phản ứng và hiệu suất mong đợi sản phẩm PCR. Đối với số khuôn DNA ít hơn 10 bản sao, cần phải thực hiện 40 chu kỳ. Nếu số khuôn DNA ban đầu cao hơn, 25 – 35 chu kỳ thường là đủ.

Số chu kỳ khuếch đại cần thiết để tạo ra một băng nhìn thấy trên gel phụ thuộc vào nồng độ ban đầu của DNA đích. Nên chạy 40 – 45 chu kỳ

để khuếch đại 50 phân tử đích và 25 – 30 để khuếch đại 3×10^5 phân tử ở cùng nồng độ. Tính không tỷ lệ này là do một hiệu ứng được gọi là hiệu ứng bão hòa, là sự giảm tốc độ lũy thừa của tích tụ sản phẩm ở giai đoạn cuối của PCR, khi sản phẩm đạt 0,3 – 1,0 nM.

Nguyên nhân có thể là do sự phân huỷ các thành phần phản ứng (dNTP, enzyme) ; sự cạn kiệt các thành phần phản ứng (mồi, dNTP – trước tiên là vấn đề với các sản phẩm ngắn, sau đó với các sản phẩm dài) ; ức chế sản phẩm cuối (sự tạo thành pyrophosphate) ; sự cạnh tranh các chất phản ứng bởi các sản phẩm không đặc hiệu ; sự cạnh tranh liên kết mồi bởi tái bắt cặp của sản phẩm đậm đặc 10 nM.

Nếu sản phẩm mong muốn không được tạo ra sau 30 chu kỳ, lấy một lượng mẫu nhỏ ($1 \mu\text{l}$) của hỗn hợp khuếch đại và tái khuếch đại 20-30 chu kỳ trong một hỗn hợp mới sẽ tốt hơn là tiếp tục chạy phản ứng nhiều chu kỳ hơn. Trong một số trường hợp nồng độ khuôn có giới hạn, thì tái khuếch đại có thể cho sản phẩm tốt, trong khi đó sự kéo dài tới 40 chu kỳ hoặc hơn sẽ không cho sản phẩm PCR.

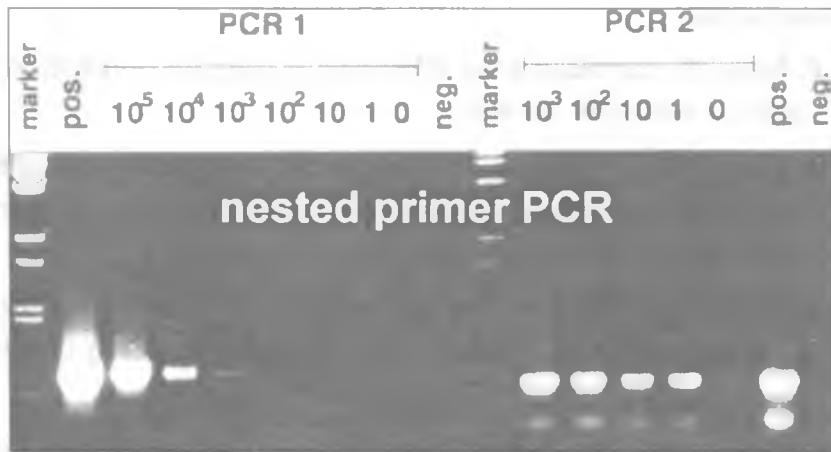


Hình 3.10. Phản ứng khuếch đại PCR lồng.

Một cải tiến của phương pháp tái khuếch đại là PCR lồng (hình 3.10). Nguyên lý là thực hiện khuếch đại PCR với một bộ mồi ngoài, sau đó dùng một số sản phẩm đã loại bỏ hoặc không loại bỏ các hoá chất để tái khuếch đại với một bộ mồi lồng nằm trong.

Quy trình này tăng độ nhạy và tính đặc hiệu thêm một mức nữa, nghĩa là tất cả các sản phẩm được khuếch đại không đặc hiệu ở vòng thứ nhất sẽ không được khuếch đại ở vòng thứ hai. Với quy trình PCR lồng, ở PCR thứ hai, người ta có thể phát hiện ra sản phẩm với số phân tử khuôn ban đầu chỉ bằng một phần nghìn nếu chỉ sử dụng PCR thứ nhất.

Hình 3.11 cho thấy tác dụng của khuếch đại PCR lồng trong việc phát hiện DNA virus gây thiếu máu gà (chicken anaemia virus : CAV) trong dãy pha loãng sau : PCR1 phát hiện 1.000 phân tử khuôn ; PCR2 khuếch đại 1 phân tử khuôn.



Hình 3.11. Điện di đồ sản phẩm phản ứng khuếch đại PCR lần đầu tiên và lần thứ hai (theo Coyne et al. 2001).

3.2.10.6. Bước tổng hợp cuối cùng

Sau chu kỳ cuối cùng, mẫu thường được ủ 5 – 15 phút ở 72°C để lấp đầy các đầu nhô của sản phẩm PCR đã được tổng hợp. Ở bước này, *Taq* pol có hoạt tính transferase đầu cuối bổ sung một nucleotide A vào đầu 3' của các sản phẩm PCR. Nếu các phân đoạn PCR được nhân dòng vào T/A vector, bước này được kéo dài tới 30 phút.

3.2.11. Phân tích sản phẩm PCR

Sản phẩm PCR phải là một hoặc nhiều phân đoạn DNA có độ dài xác định. Cách tốt nhất để kiểm tra sự có mặt của những phân đoạn này là nạp mẫu từ sản phẩm phản ứng, cùng với chỉ thị khối lượng phân tử thích hợp, lên gel agarose chứa 0,8 – 4,0% ethidium bromide. Các băng DNA trên gel có thể được nhìn thấy dưới tia cực tím. So sánh các băng sản phẩm với các băng có khối lượng phân tử đã biết, có thể xác định được phân đoạn sản phẩm có kích thước thích hợp.

Những điều cần lưu ý

Khi phân tích sản phẩm PCR, nếu không có sản phẩm đặc hiệu, sản phẩm nhoè, nhiều sản phẩm phụ, có sản phẩm phụ mà không có sản phẩm đặc hiệu thì phải tìm ra các yếu tố quyết định đến sự thành công của PCR. Các yếu tố quyết định này đã được phân tích kỹ ở trên. Có thể tóm tắt như sau :

1. Nồng độ các thành phần PCR càng ít càng tốt, tối thiểu để tổng hợp các sản phẩm đặc hiệu.
2. Thành phần PCR được chia lượng nhỏ, bảo quản ở -20°C. Không nên đông lạnh và giải đông nhiều lần, trước khi dùng phải trộn kỹ và ly tâm thu hơi nước để bảo đảm đúng nồng độ.
3. Thiết kế mồi là rất quan trọng, nhiệt độ bắt cặp càng cao càng tốt (ít nhất là 50°C) và cặp mồi phải có nhiệt độ nóng chảy chênh nhau không quá 5°C.
4. Nhiệt độ bắt cặp và thời gian các bước biến tính, bắt cặp và tổng hợp rất quan trọng.
5. Cần phải chuẩn độ các yếu tố, ít nhất là nhiệt độ và MgCl₂. Có thể sử dụng máy chu trình thang nhiệt độ để chuẩn độ.
6. Bước biến tính khởi động rất quan trọng, cải thiện hiệu suất phản ứng và ít tạo ra nhị phân mồi hơn. Có thể bổ sung *Taq* pol vào ống trong máy chu trình nhiệt ở 94°C (hoặc MgCl₂ hoặc dNTP).
7. Luôn luôn kiểm tra các chương trình trên máy chu trình nhiệt.
8. Luôn luôn chạy PCR âm không khuôn và dương tính có khuôn DNA đã biết để kiểm tra sự lắn tạp cũng như thành phần PCR. Trên gel agarose luôn luôn có chỉ thị khối lượng phân tử DNA để ước lượng kích thước sản phẩm PCR và kiểm tra hệ thống soi gel.
9. Nếu khuếch đại yếu hoặc không có sản phẩm PCR thì có thể :
a) tăng số chu kỳ lên ; b) giảm thời gian ở các bước ; c) tăng thời gian tổng hợp cuối cùng lên 10 phút thậm chí 45 phút ở 60°C ; d) tăng số lượng khuôn ; e) tăng nồng độ *Taq* pol lên hai lần ; f) thử *Taq* pol khác (ví dụ Faststart) ; g) tăng MgCl₂, mỗi, dNTP ; h) tái khuếch đại sản phẩm PCR từ phản ứng trước.

10. Nếu sản phẩm PCR không đặc hiệu thì có thể : a) giảm nồng độ *Taq* pol đối với các băng khối lượng phân tử cao ; b) giảm nồng độ MgCl₂ đối với các băng khối lượng phân tử thấp ; c) giảm số chu kỳ ; d) tăng nhiệt độ bắt cặp ; e) giảm nồng độ mồi ; f) giảm dNTP.
11. Nếu PCR bị ức chế có thể : a) pha loãng khuôn 1 : 5 với 10 mM Tris-HCl (pH 8,0) hoặc nước cất ; b) bổ sung BSA – nồng độ cuối cùng 1x.
12. Nếu có vệt nhòe trên băng thì giảm số chu kỳ. Nếu có thảm thì giảm nồng độ các hoá chất.

3.3. XỬ LÝ SAU PCR

Sản phẩm PCR có thể được tinh sạch sử dụng nhiều phương pháp khác nhau như các loại kit thương mại, tinh sạch gel nếu cần phải loại bỏ khuôn. Sản phẩm PCR tinh sạch có thể được đọc trình tự.

Quy trình này được sử dụng để đọc trình tự trực tiếp sản phẩm PCR đặc hiệu sau khi khuếch đại số lượng lớn. Phản ứng được thực hiện theo cách thông thường, ngoại trừ sau khi tối ưu nhiệt độ bắt cặp và các điều kiện khác trong một thể tích 5 µl. Một thể tích phản ứng lớn 50 µl (10x) được sử dụng để tổng hợp đủ sản phẩm PCR cho đọc trình tự. Các phản ứng thể tích lớn được thực hiện chính xác như mô tả chi tiết trong phương pháp phản ứng chuỗi trùng hợp, ngoại trừ phản ứng được tăng lên 10 lần. Một mồi PCR hay mồi mới nội phân tử có thể làm mồi đọc trình tự. Tuy nhiên, cần thiết phải tinh sạch sản phẩm PCR ra khỏi mồi để không phản ứng nào còn lại trong PCR trước khi đọc trình tự.

Tinh sạch sản phẩm PCR được thực hiện bằng cách nạp sản phẩm PCR lên cột tách Sepharose. Nếu sản phẩm PCR cần đọc trình tự rất đặc hiệu (chỉ có một băng), có thể nạp sản phẩm trực tiếp lên cột để tinh sạch. Tuy nhiên, nếu PCR tạo ra một số băng không đặc hiệu, băng quan tâm cần phải được cắt ra khỏi gel và thôi điện (electroelute) hoặc cách khác là thu hồi qua cột.

Để đọc trình tự tối đa, người ta phải tách băng DNA ra khỏi mỗi dạng nhị phân. Sau khi tinh sạch sản phẩm PCR, nên dùng que nhúng để định lượng, không nên đo nồng độ bằng quang phổ. Không cần thiết phải biết

chính xác nồng độ sản phẩm PCR để đọc trình tự. Xác định nồng độ sản phẩm bằng quang phổ cũng chấp nhận được nhưng que nhúng đơn giản hơn và lưu giữ số liệu nồng độ lâu hơn. Khuôn có nồng độ vào khoảng 3 ng/ μ l/100bp. Ví dụ, sản phẩm 300 bp, cần 5 μ l dung dịch 9 ng/ μ l. Khuôn và mồi đều ở trong nước và có nồng độ tương đương.

Khi đọc trình tự một sản phẩm PCR trên máy đọc trình tự tự động, nên tinh sạch theo quy trình sau.

Tùa ethanol sản phẩm PCR trong một PCR 50 hoặc 100 μ l.

Hoà tüa vào 10 μ l đệm nạp. Nạp lên gel 2-3% agarose Nusieve tinh khiết. Không nên sử dụng agarose còn nhiều tạp chất vì nó sẽ ảnh hưởng tới phản ứng đọc trình tự. Chạy điện di và chụp ảnh.

Cắt băng thích hợp ra. Chiết DNA sử dụng cột.

Có thể tinh sạch đồng thời 96 sản phẩm PCR sử dụng đĩa 96 giếng theo một số quy trình tinh sạch khác nhau bao gồm : tinh sạch bằng Sepahcrysyl S-300 hoặc S-500 trong đĩa lọc với kích thước lỗ 0,45 hoặc 1,2 μ m ; tinh sạch bằng guanidine-HCl trong đệm MES, trong đĩa lọc thuỷ tinh 96 giếng ; tinh sạch bằng guanidine-HCl trong đệm KOAc, cũng trong đĩa lọc thuỷ tinh 96 giếng như trên.

Tất cả các quy trình này đều có thể cho độ tinh sạch sản phẩm tương tự nhau và chấp nhận được. Hiệu suất thu hồi các sản phẩm nhỏ hơn 700 bp có thể cao hơn một ít ở cột S-300 so với cột S-500. Giá thiết là tất cả các phương pháp được mô tả dưới đây đều cho các kết quả tương tự nhau, thì giá cột tinh sạch là một yếu tố. Giá cột hiện nay ước tính là 16 dollar đối với đĩa lọc 96 giếng thuỷ tinh và 26 dollar cho đĩa lọc 96 giếng Sephadex, bao gồm tất cả các hoá chất, đĩa lọc và đĩa thu sản phẩm khi ly tâm.

4

LẬP BẢN ĐỒ GEN

4.1. GIẢI TRÌNH TỰ NUCLEOTIDE

4.1.1. Các phương pháp xác định trình tự nucleotide

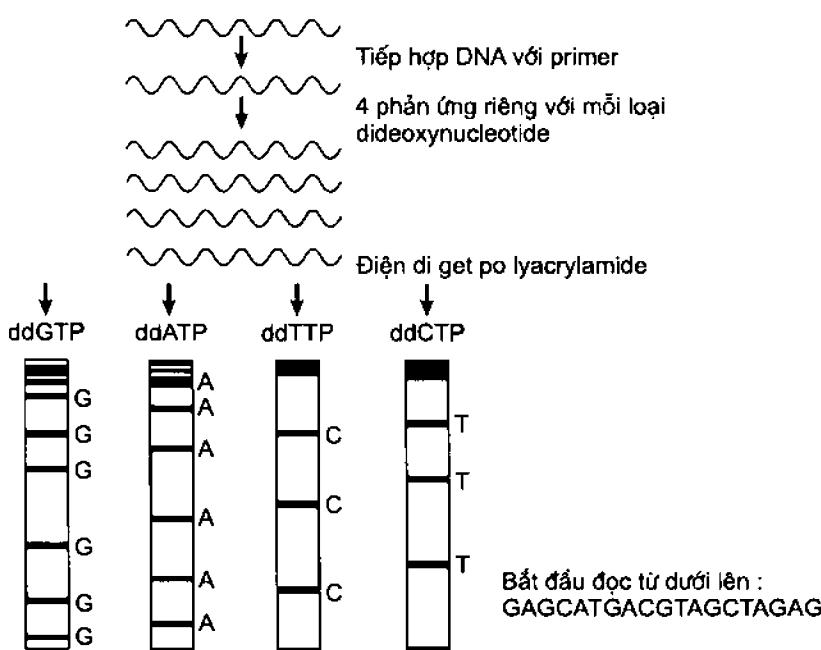
Có nhiều kỹ thuật xác định chính xác trình tự các cặp base trong một chuỗi DNA.

1. Kỹ thuật điện di phân đoạn đánh dấu đầu cuối cho phép xác định trình tự chuỗi DNA sợi đơn.
2. Xác định trình tự nucleotide thông qua trình tự amino acid. Có một vấn đề nảy sinh là bộ ba mã hoá của 20 amino acid thường có một số nucleotide trùng lặp, cho nên trình tự DNA thu được chỉ đúng tương đối.
3. Xác định trực tiếp thông qua phản ứng kết thúc đầu cuối.

Nguyên tắc chính của mọi quy trình xác định trực tiếp là tạo ra các phân tử DNA sợi đơn, mà phân tử sau có cùng trình tự và chỉ dài hơn phân tử trước một nucleotide ở một đầu cuối. Băng điện di polyacrylamide có thể tách các phân tử thành một thang dài tới 300 cặp nucleotide. Để làm được điều đó, kỹ thuật hoá học (Maxam-Gilbert) dùng phương pháp cắt từng base một, phương pháp enzyme học thì tổng hợp *in vitro*. Thông thường, phải tiến hành 4 phản ứng riêng rẽ cho 4 loại base và khi điện di cũng chạy trên 4 giếng riêng rẽ. Khi đọc, cứ so sánh 4 hàng cột đọc để xác định ra phân tử tiếp theo là phân tử có băng điện di gần nhất so với băng vừa đọc xong. Để có được các băng hiển thị, phương pháp hoá học đánh dấu các phân tử bằng phóng xạ ^{32}P .

Phương pháp enzyme học (phương pháp Sanger) tiến hành phản ứng tổng hợp các phân tử DNA với một hàm lượng nhỏ dẫn xuất dideoxy của các nucleotide (dideoxynucleotide triphosphate, ddNTP). Các nucleotide

dẫn xuất này bị khử mất thêm một nhóm OH nên phản ứng polymerase không tiếp tục mà kết thúc tại nucleotide đó. Như vậy, tất cả các phân tử được tạo ra trong phản ứng đó đều mang đuôi là A, nếu ddNTP trộn vào là ddATP, tương tự như vậy là T, C hoặc G khi trộn ddTTP, ddCTP hoặc ddGTP. Hình 4.1 mô tả điện di đồ trên gel polyacrylamide và cách đọc trình tự. Phân tử ngắn nhất kết thúc bằng G (cuối cột dọc ngoài cùng bên trái), tiếp theo là A (cột dọc thứ 2 từ trái sang phải), sau đó là G, theo đường zigzag. Trình tự đầu 5' (vì polynuclease nối các nucleotide dài ra ở đầu 5' của sợi DNA) của chuỗi DNA đọc được là 5'-GAGCATGACGTAGCTAGAG-3'.



Hình 4.1. Xác định trình tự nucleotide trong phân tử DNA bằng phản ứng kết thúc với ddNTP và điện di trên gel polyacrylamide.

Trong thực tế, những phương pháp xác định trình tự nucleotide thủ công nhờ trợ giúp của máy điện di, đánh dấu phóng xạ và phóng xạ tự chụp đã không còn được sử dụng vì cần phải bảo đảm mức độ an toàn phóng xạ, rất tốn thời gian và hiệu quả không cao.

Hiện nay, thiết bị đọc trình tự nucleotide hoạt động phần lớn tự động dựa trên nguyên lý sắc ký cột mao dẫn (capillary), đánh dấu nucleotide bằng huỳnh quang và đầu dò laser. Kết quả đọc được trực tiếp xử lý trên thiết bị đọc nhờ liên kết với máy tính có khả năng xử lý nhanh (hình 4.2).



Hình 4.2. Thiết bị đọc trình tự nucleotide ABI PRISM 3100 Genetic Analyzer thuộc thế hệ 2003 có thể đọc 16 phản ứng.

4.1.2. Khung đọc của gene

So sánh trình tự nucleotide trong DNA và trình tự nucleotide trong mRNA tương ứng hoặc trình tự amino acid protein do gene thuộc tế bào nhân sơ cho thấy hai trình tự này hoàn toàn tương đồng và liên tục, không có sự gián đoạn nào cả. Cũng so sánh như vậy đối với gene của tế bào nhân chuẩn, có khi phát hiện thấy giữa trình tự nucleotide của mRNA và trình tự nucleotide của gene không có sự trùng khớp về tính liên tục, ví dụ sau một trình tự mã hoá vài chục amino acid có đoạn polynucleotide không mã hoá gì cả và các amino acid tiếp theo được mã hoá ở một quãng sau đó. Hiện tượng gene bị gián đoạn này được phát hiện ra năm 1977. Đoạn nucleotide mã hoá protein được gọi là exon và khi phiên mã được phiên thành mRNA, còn đoạn không mã hoá đó được gọi là intron

sẽ không được phiên mã thành mRNA. Độ dài và số lượng intron nằm trên một gene rất khác nhau. Có những gene không có intron nhưng có những gene có tới vài trăm intron lớn nhỏ khác nhau. Đoạn intron sẽ không được phiên mã thành mRNA mà bị bỏ qua. Hiện tượng phiên mã bỏ qua các intron được gọi là bỏ vòng khuyên (splicing).

4.2. ENZYME HẠN CHẾ VÀ ÚNG DỤNG TRONG LẬP BẢN ĐỒ GENE

Enzyme hạn chế (restriction enzyme, RE) là các nuclease do vi khuẩn sản sinh ra, có khả năng nhận biết những đoạn DNA có trình tự nucleotide xác định, bám vào đoạn DNA đó và cắt cả hai sợi của phân tử DNA vào giữa 2 nucleotide đối xứng nhau tại điểm cắt.

4.2.1. Sự phát hiện enzyme hạn chế

Sự phát hiện ra RE là một phát minh quan trọng, tạo nên một loại công cụ chính của kỹ thuật DNA tái tổ hợp. Đó là những chiếc kéo phân tử cho phép các nhà nghiên cứu cắt xén phân tử DNA và thiết kế lại theo ý muốn trong thao tác gene. Nhờ có những phát minh này mà Albert, Smith và Nathans được nhận giải thưởng Nobel năm 1978.

Quan sát khoa học đầu tiên dẫn đến phát hiện ra RE được bắt đầu từ những năm đầu của thập kỷ 50. Khi đó người ta phát hiện thấy một số chủng vi khuẩn có khả năng miễn dịch với thực khuẩn thể. Cơ chế của tính kháng thực khuẩn thể là do tế bào vi khuẩn tạo ra một loại enzyme cắt phân tử DNA genome của thực khuẩn thể thành nhiều đoạn trước khi phân tử DNA kịp nhân bản và tạo ra thể thực khuẩn thể mới. Hiện tượng cắt DNA như thế được gọi chung là cắt hạn chế (hạn chế thực khuẩn thể sinh sản) và enzyme thực hiện việc cắt đó được gọi enzyme hạn chế. RE chỉ cắt các phân tử DNA của thực khuẩn thể mới xâm nhập vào tế bào, còn DNA của chính tế bào lại được bảo vệ bằng cơ chế methyl hoá gốc Adenin ở vùng nhận biết làm cho enzyme không gắn được vào phân tử DNA ở vị trí cắt.

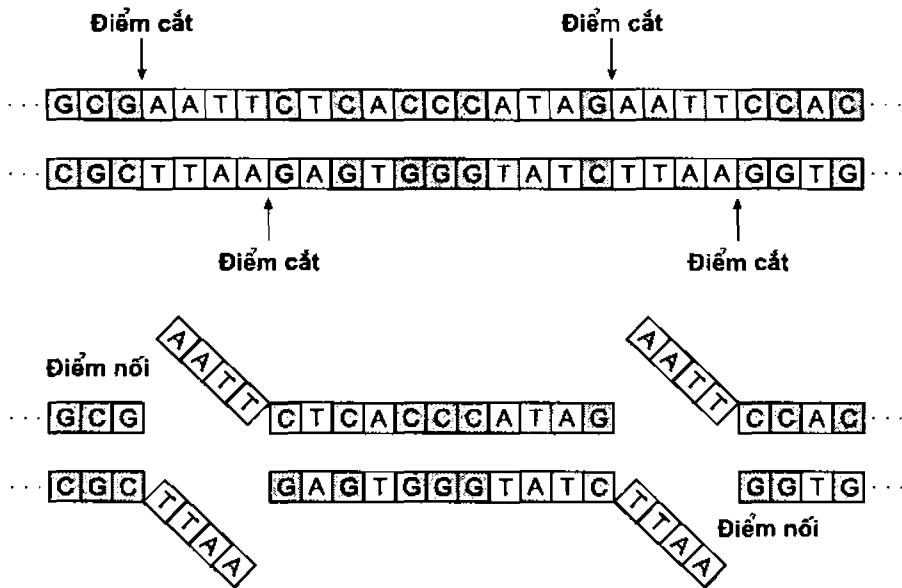
Bảng 4.1. Sự phân loại các RE theo 3 nhóm khác nhau.

Đặc điểm	Loại I	Loại II	Loại III
Điểm cắt	Cách xa điểm nhận biết trên 1000 bp	Nằm trong điểm nhận biết	Nằm không xa ngoài điểm nhận biết
Khả năng methyl hoá gốc adenyl	Có	Không	Có
Điều kiện để cắt	ATP, Mg ²⁺ , S-AdoMet	Mg ²⁺ hoặc Mn ²⁺	Mg ²⁺ , S-AdoMet
Cấu trúc của enzyme : số polypeptide	3 chuỗi khác nhau	2 chuỗi giống nhau	2 chuỗi khác nhau

Đến nay, có trên 1200 loài vi khuẩn được phát hiện là có khả năng tạo ra một hay nhiều loại RE khác nhau. RE được đặt tên bằng cách lấy chữ cái viết in hoa đầu tiên của tên chi, ghép với hai chữ cái đầu của tên loài và một chữ cái của tên dòng. Trong trường hợp một dòng có nhiều enzyme thì thêm vào các chữ số I, II, III, IV, V. Ví dụ EcoRV là enzyme thứ V do vi khuẩn *Escherichia coli* dòng R tạo ra ; hoặc HindIII là một trong 3 enzyme do vi khuẩn *Haemophilus influenzae* dòng d sinh ra. Bảng 4.1 trình bày nguyên tắc phân loại enzyme hạn chế. Enzyme hạn chế được chia thành 3 nhóm theo đặc điểm của sự nhận biết và cách thức cắt phân tử DNA.

Trong ba nhóm thì chỉ có nhóm I mỗi phân tử enzyme gồm ba chuỗi polypeptide khác nhau, trong đó có một chuỗi thực hiện chức năng nhận biết, một chuỗi cắt và một chuỗi methyl hoá. Điểm cắt của các enzyme nhóm này không đặc trưng nên không được sử dụng nhiều. Nhóm III gồm các enzyme có hai chuỗi polypeptide trong phân tử, một chuỗi làm chức năng cắt, còn một chuỗi đảm nhận cả chức năng nhận biết và methyl hoá. Cũng như nhóm I, nhóm III có điểm cắt không đặc trưng nên rất khó sử dụng trong kỹ thuật DNA tái tổ hợp.

So sánh giá trị thực tiễn của ba nhóm thì nhóm II có giá trị lớn nhất vì (1) mỗi enzyme có điểm nhận biết và điểm cắt đặc trưng ; (2) không cần năng lượng (ATP) trong quá trình cắt ; (3) Không có hoạt tính methyl hoá. Các enzyme nhóm này có đoạn nhận biết gồm những nucleotide đối xứng nhau.



Hình 4.3. Sơ đồ cắt của một số enzyme giới hạn.

Đường cắt của các enzyme hạn chế cũng khác nhau. Có những enzyme cắt đứt hai sợi của phân tử DNA xoắn kép tại một điểm giữa 2 cặp nucleotide và tạo ra hai đầu bằng (blunt end). Phân lớn các enzyme hạn chế cắt hai sợi polynucleotide tại hai điểm đối xứng nhau trên vùng nhận biết và tạo ra hai đầu phân tử DNA không bằng, mà có một dải đơn ngắn từ 2 đến 4 nucleotide (hình 4.3). Khi cắt lệch như vậy hai đầu cắt được gọi là đầu dính (sticky end), chúng có thể "khớp mộng" với nhau để nối hai đoạn DNA với nhau theo nguyên tắc cặp base bổ sung thông thường.

Đáng chú ý là mặc dù có vùng nhận biết khác nhau nhưng có nhiều enzyme lại tạo ra các đầu cắt có đầu dính giống nhau. Ví dụ *Bam*HI có vùng nhận biết là GGATCC và *Bgl*III có vùng nhận biết là AGATCT tạo ra đầu dính có chuỗi GATC giống nhau. Sau 3A chỉ nhận biết GATC cũng tạo ra đầu dính GATC như hai enzyme trên. Các đầu dính được tạo ra khi dùng ba enzyme trên để cắt có thể nối lại được với nhau, tuy nhiên, khi nối lại như vậy, điểm cắt bị thay đổi, không còn nguyên như ban đầu.

4.3. LẬP BẢN ĐỒ CẮT HẠN CHẾ

4.3.1. Sự cần thiết của việc lập bản đồ hạn chế

Nhiễm sắc thể chứa một phân tử DNA dài, liên tục bao gồm nhiều gene. Đột biến trên phân tử DNA làm cho trình tự nucleotide thay đổi. Khi lập bản đồ về mức độ liên kết các gene sẽ phát hiện được tần số tái tổ hợp khác đối với những gene có khoảng cách vật lý khác nhau. Một nhiễm sắc thể tế bào nhân chuẩn mang hàng ngàn gene, vì thế đơn vị bản đồ xác định được thường rất lớn.

Chẳng hạn, khi gấp nhóm gene cách nhau 40 đơn vị bản đồ, tức là cứ khoảng 40 đơn vị bản đồ mới có một chỉ thị phân tử được xác định, để tìm một gene trong genome của người cần ít nhất 100 chỉ thị phân tử. Trong thực tế, để xác định đúng vị trí của một gene cần phải có mật độ chỉ thị phân tử dày đặc hơn nhiều.

Việc lập bản đồ gene không những góp phần chính xác hoá phân tích di truyền chủ yếu là nhằm vào mục đích phân lập gene để xác định trình tự và lại dùng chúng như chỉ thị phân tử cho việc lập bản đồ.

So sánh chiều dài của các nhóm gene liên kết trong genome với chiều dài tổng thể của DNA và lấy chỉ số tái tổ hợp là 50% cho nhóm gene có khoảng cách trung bình thì sẽ thu được số liệu về lập bản đồ gene như trình bày ở bảng 4.2.

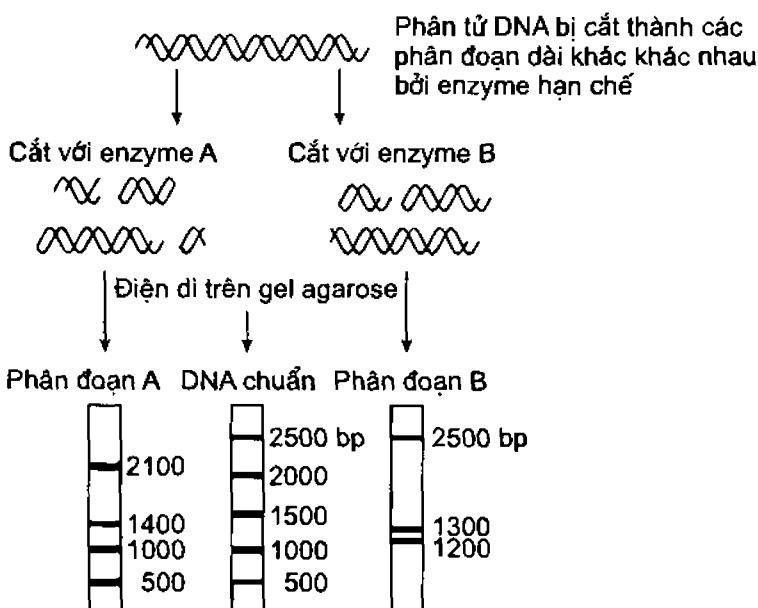
Bảng 4.2. Tần số tái tổ hợp thông tin di truyền xảy ra ở sinh vật bậc cao với tế bào nhân chuẩn thấp hơn nhiều so với sinh vật bậc thấp với tế bào nhân sơ.

Loài	Độ lớn của genome đơn bội (bp)	Đơn vị trên bản đồ gene	Kích thước đơn vị bản đồ gene (bp)	Khoảng cách trung bình giữa các trao đổi chéo (bp)
Thực khuẩn T4	$1,6 \times 10^5$	800	200	$1,0 \times 10^4$
<i>E. coli</i>	$1,6 \times 10^5$	800	200	$1,0 \times 10^4$
Nấm men	$1,0 \times 10^5$	4200	200	$1,0 \times 10^4$
Nấm thương đắng	$2,7 \times 10^7$	1000	200	$1,0 \times 10^4$
Tuyên trùng	$2,7 \times 10^7$	320	250000	$1,0 \times 10^4$
Ruồi giấm	$1,6 \times 10^5$	320	250000	$1,0 \times 10^4$
<i>E. coli</i>	$1,6 \times 10^5$	320	1800000	$9,0 \times 10^7$
Người (Nam giới)	$1,6 \times 10^5$	2809	1800000	$6,0 \times 10^7$
Người (Nữ giới)	$1,6 \times 10^5$	4782	250000	$1,0 \times 10^4$

4.3.2. Lập bản đồ hạn chế một loại phân tử DNA

Sử dụng enzyme hạn chế để phân chia phân tử DNA tại các điểm cắt. Do điểm nhận biết mang tính đặc trưng của enzyme và cự ly giữa các điểm nhận biết phụ thuộc vào trình tự nucleotide trong phân tử DNA, vì thế enzyme hạn chế sẽ chia phân tử DNA thành những phân đoạn có độ dài đặc trưng. Tất cả những vị trí trên phân tử DNA có đúng trình tự đó đều được enzyme cắt. Các enzyme hạn chế khác nhau thì có đoạn nhận biết với trình tự base đặc trưng khác nhau. Sơ đồ mô tả các vị trí cắt khác nhau của một phân tử DNA được gọi là *bản đồ hạn chế*. Cự ly giữa hai điểm cắt gần được tính bằng số cặp base (bp), xa hơn được tính bằng kilobase ($1\ kb = 1000\ bp$). Ở mức nhiễm sắc thể, cự ly được tính bằng megabase ($1\ Mb = 106\ bp$).

Khi phân tử DNA được cắt bằng một hay nhiều enzyme nhất định thành những phân đoạn dài ngắn khác nhau và phân loại chúng bằng điện di trên gel agarose hay gel polyacrylamide thì sẽ thu được một loạt các băng. Phân đoạn càng nhỏ thì chuyển động càng nhanh và càng xa điểm xuất phát (hình 4.4).

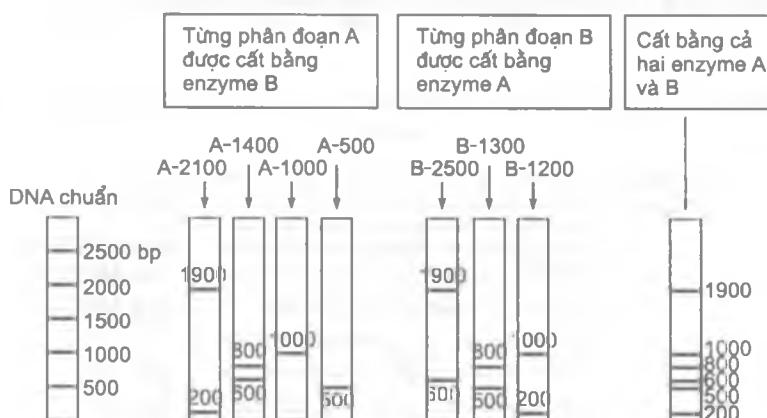


Hình 4.4. Cắt DNA bằng enzyme hạn chế và tách các phân đoạn của chúng trên điện di gel agarose.

Hình 4.4 nêu một ví dụ cắt một phân tử DNA dài 5000 bp với 2 loại enzyme hạn chế khác nhau là A và B và chạy điện di. Enzyme A cắt phân tử DNA thành 4 phân đoạn dài 2100, 1400, 1000 và 500 bp (cột dọc bên trái), trong khi đó enzyme B cắt phân tử DNA đó thành 3 phân đoạn 2500, 1300 và 1200 bp (cột dọc bên phải). Từ kết quả này có thể xây dựng được bản đồ hạn chế của phân tử DNA đó không?

Sau khi điện di các phân đoạn cắt đơn bằng enzyme A và B cần tiến hành thu nhận các phân đoạn riêng rẽ gồm hai nhóm: Nhóm A thu được nhờ enzyme A bao gồm A-2100 ; A-1400 ; A-1000 và A-500. Nhóm B thu được nhờ enzyme B gồm có B-2500 ; B-1300 và B-1200. Tiến hành cắt các phân đoạn của nhóm A với enzyme B và các phân đoạn của nhóm B với enzyme A. Hình 4.5 trình bày bước phân tích tiếp theo bằng cách dùng enzyme thứ 2 cắt kép để có kết quả tổng thể.

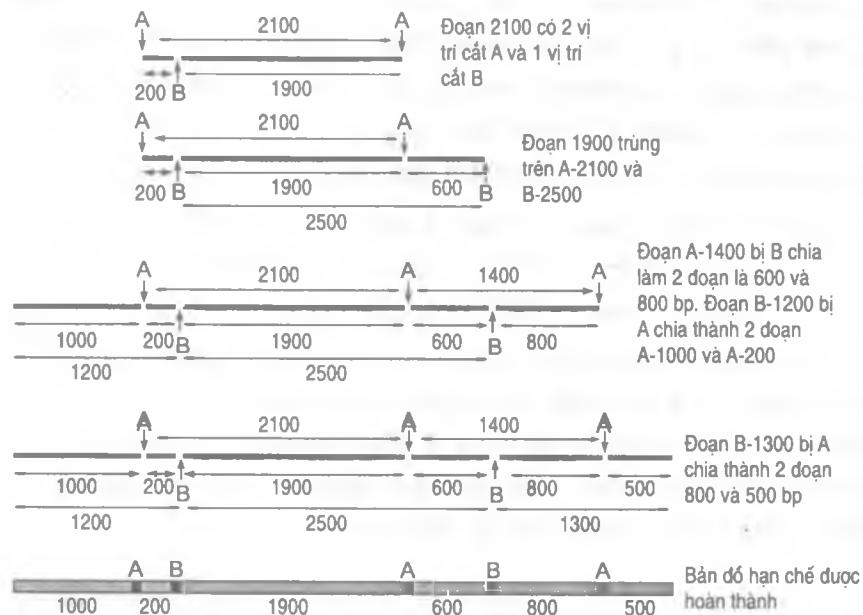
Trên hình 4.5 ta thấy ở nhóm A có đoạn A-2100 bị enzyme B cắt thành 2 đoạn là 1900 bp và 200 bp ; đoạn A-1400 bị cắt thành 2 đoạn là 800 bp và 600 bp ; đoạn A-1000 không bị cắt ; đoạn A-500 cũng không bị cắt. Ở nhóm B có đoạn B-2500 bị enzyme A cắt thành 2 đoạn là 1900 bp và 600 bp ; đoạn B-1300 thành 800 bp và 500 bp ; đoạn B-1200 thành 1000 bp và 200 bp. Bức tranh điện di tổng thể nếu cùng một lúc cắt kép với enzyme A và B, ta thấy phân tử DNA ban đầu bị tách cắt thành 6 băng là 1900 ; 1000 ; 800 ; 600 ; 500 và 200 bp.



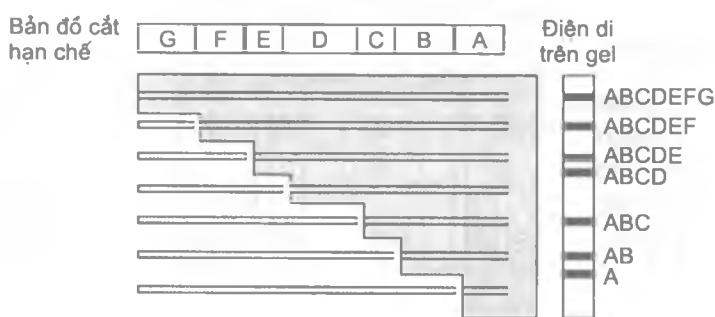
Hình 4.5. Cắt kép bằng hai enzyme hạn chế để xác định vị trí cắt của từng loại enzyme.

Kỹ thuật cắt không hoàn toàn (dở dang) nhiều khi cung cấp được những thông tin chính xác về trình tự sắp xếp của các phân đoạn. Ví dụ, enzyme A cắt dở dang sẽ cho ra phân đoạn 3500 bp khi nhát cắt giữa phân đoạn 2100 bp và 1400 bp chưa hoàn thành.

Từ kết quả hình 4.3 có thể tiến hành xây dựng bản đồ hạn chế theo từng bước như trình bày ở hình 4.6. Với các loại phân đoạn có chiều dài khác nhau thu được thông qua xử lý khác nhau dùng phương pháp xếp chồng để tìm ra trình tự hợp lý cho bản đồ hạn chế như trình bày trên hình 4.6.



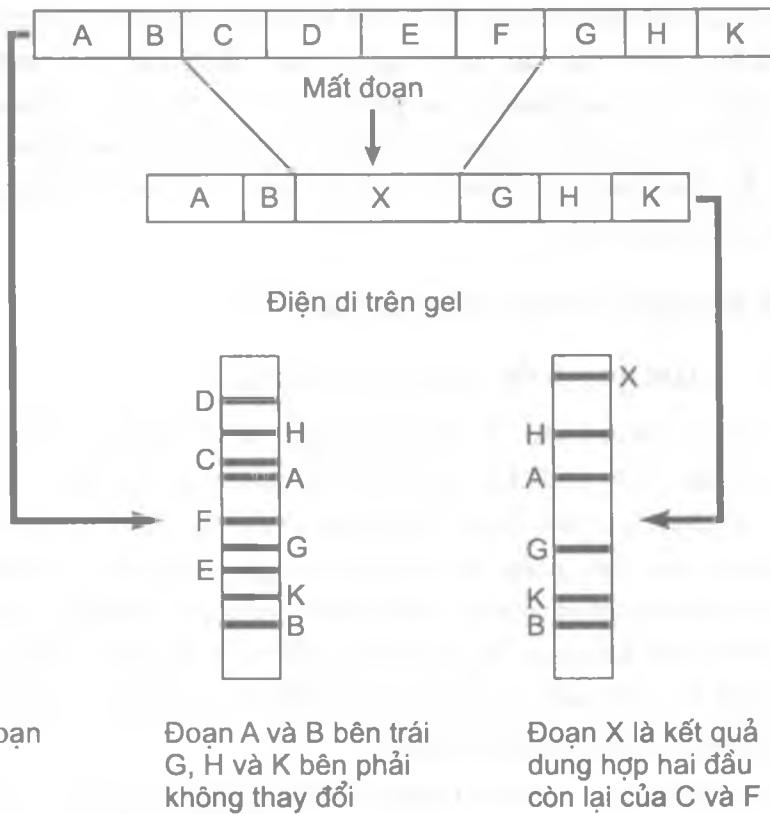
Hình 4.6. Lập bản đồ hạn chế một phân tử DNA trên cơ sở các dữ liệu cắt đơn và cắt kép.



Hình 4.7. Đánh dấu đầu cuối giúp xác định nhanh vị trí cắt của các enzyme hạn chế.

Để thuận tiện hơn trong lúc sắp xếp trình tự các phân đoạn, người ta dùng kỹ thuật đánh dấu đầu cuối. Kỹ thuật này cho phép xác định khá chính xác từng vị trí cắt (hình 4.7).

4.3.3. Sử dụng điểm cắt hạn chế làm chỉ thị di truyền



Hình 4.8. Thay đổi trong phân tử DNA làm thay đổi bản đồ hạn chế

Bản đồ di truyền và bản đồ hạn chế phản ánh các khía cạnh khác nhau của vật liệu di truyền. Về nguyên tắc có thể xây dựng bản đồ hạn chế cho bất kỳ chuỗi DNA nào. Khi chuỗi DNA có những thay đổi do đột biến hay do quá trình tái tổ hợp xảy ra thì những vị trí cắt có thể bị chuyển dịch hoặc thay đổi, lúc đó bản đồ hạn chế cũng bị thay đổi (hình 4.8). Sự sai khác trên bản đồ cắt hạn chế của hai cá thể được gọi là tính đa dạng về chiều dài phân đoạn hạn chế (restriction fragment length polymorphism = RFLP), sẽ được trình bày kỹ ở phần sau.

4.4. LAI PHÂN TỬ TRONG PHÂN TÍCH SOUTHERN, NORTHERN VÀ WESTERN

Khi chưa thể phân tích trình tự nucleotide của phân tử DNA và RNA hoặc trình tự amino acid của phân tử protein, người ta có thể so sánh các đại phân tử như DNA, RNA và protein có nguồn gốc khác nhau của cùng nhóm bằng cách tiến hành lai phân tử. Kỹ thuật lai phân tử nucleic acid dựa vào sự kết hợp tương đồng của các base bổ sung trên các chuỗi nucleotide của phân tử DNA và RNA. Trong khi đó, muốn đánh giá phân tử protein, người ta phải dựa vào phản ứng miễn dịch đặc hiệu giữa hai phân tử protein, trong đó một là kháng nguyên và một là kháng thể.

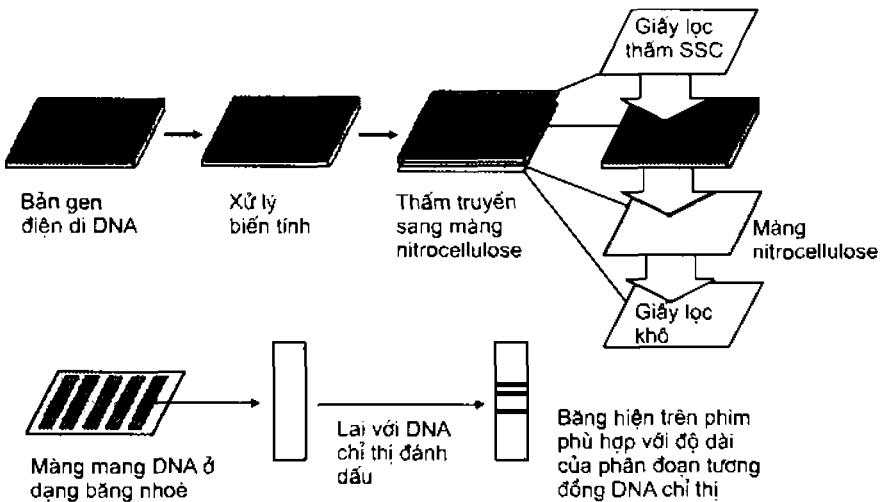
4.4.1. Nguyên lý chung của lai phân tử

4.4.1.1. Thẩm truyền đại phân tử lên màng lọc

Phương pháp phân tích thẩm truyền là kỹ thuật có hiệu quả cao trong nghiên cứu phát hiện các loại đại phân tử sinh học có cấu trúc phân tử phức tạp. Kỹ thuật có thể áp dụng cho mọi loại đại phân tử sinh học có khả năng kết gắn vào giá thể lưu giữ như màng nitrocellulose, màng nilon, màng giấy cellulose nhưng vẫn duy trì được khả năng kết gắn với các nhóm chất hiển thị khác. Thông thường, các phân tử sinh học được điện di để phân đoạn theo kích thước lớn hoặc nhỏ rồi mới chuyển lên màng.

Về nguyên tắc, kỹ thuật thẩm truyền không thay đổi nhiều từ khi được phát minh. Các đại phân tử sau khi được tách thành từng loại thông qua điện di trên bản gel được truyền sang màng nhờ lực mao dẫn (capillary) của dung dịch đậm (hình 4.9). Càng ngày, chủng loại các màng và kỹ thuật phát hiện các loại đại phân tử thông qua lai đánh dấu hay miễn dịch đánh dấu kèm các bộ kit luôn được cải tiến cho tiện lợi và an toàn. Màng nitrocellulose được dùng rất nhiều nhưng gần đây màng nylon đang thay thế bởi nó có sức mang gắn và độ dai lớn hơn, ví dụ như màng Hybond (+) (Hàng Amesham). Trong hai loại màng nylon trung tính và màng tích điện thì loại màng tích điện do tích điện nên có sức gắn DNA cao hơn. đương nhiên, trong kỹ thuật Western, chất liều

nitrocellulose vẫn đang được sử dụng rộng rãi, ngoài ra còn có màng polyvinylidene diflouride (PVDF) và màng nylon.



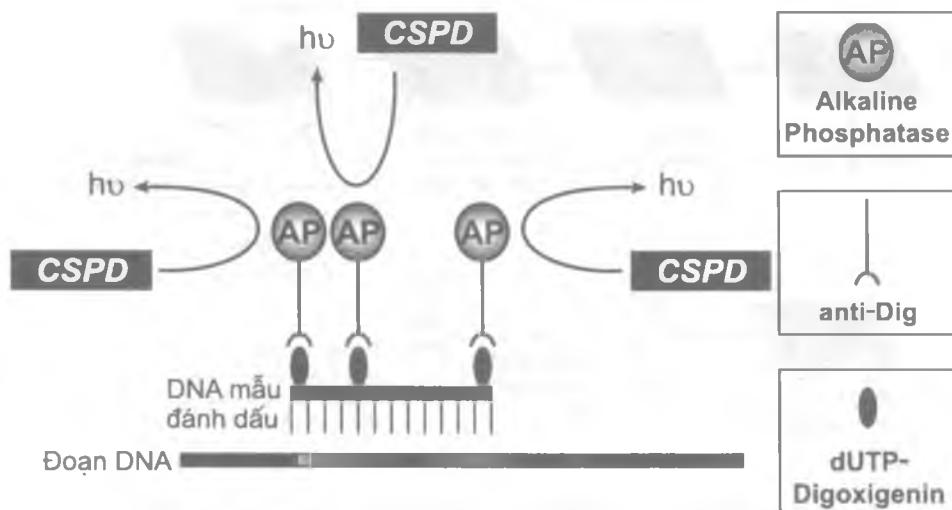
Hình 4.9. Minh họa kỹ thuật thấm truyền DNA lên màng nitrocellulose.

4.4.1.2. Đánh dấu và phát hiện phân tử

Việc phát hiện các phân tử trên màng được tiến hành thông qua đánh dấu : Qua nhiều năm, kỹ thuật đánh dấu bằng đồng vị phóng xạ đã được ứng dụng một cách rộng rãi, phổ biến nhất là các nucleotide gắn đồng vị phóng xạ ^{32}P dạng ^{32}P -dNTP. Ưu điểm lớn của phương pháp phóng xạ là có độ nhạy rất cao, có thể phát hiện được vết DNA ở mức nanogram và thấp hơn. Tuy nhiên, phương pháp phóng xạ đòi hỏi phòng thí nghiệm có những thiết bị kiểm tra và bảo đảm an toàn phóng xạ đồng bộ. Những năm gần đây, kỹ thuật phi phóng xạ (non-radioactive) hay kỹ thuật đánh dấu lạnh đang từng bước thay thế kỹ thuật phóng xạ. Trong đó, nổi bật là kỹ thuật dạ quang hóa học kích hoạt (enhancing chemical luminescence, ECL) và kỹ thuật sử dụng đánh dấu với dioxygenein (DIG) (hình 4.10).

Hình 4.10 minh họa nguyên tắc đánh dấu phân tử không dùng đồng vị phóng xạ. Sơ đồ trình bày kỹ thuật đánh dấu bằng dioxygenein : dUTP có gắn dioxygenein được dùng để tổng hợp đoạn DNA mẫu. Sau khi lai mẫu được dò bằng phản ứng miễn dịch với anti-Dig gắn sẵn với

phosphatase kiềm, cơ chất CSPD (Lumi-Phos) giải phóng ra năng lượng ánh sáng ($h\nu$) dưới tác dụng của phosphatase làm hiện hình trên phim X-quang.

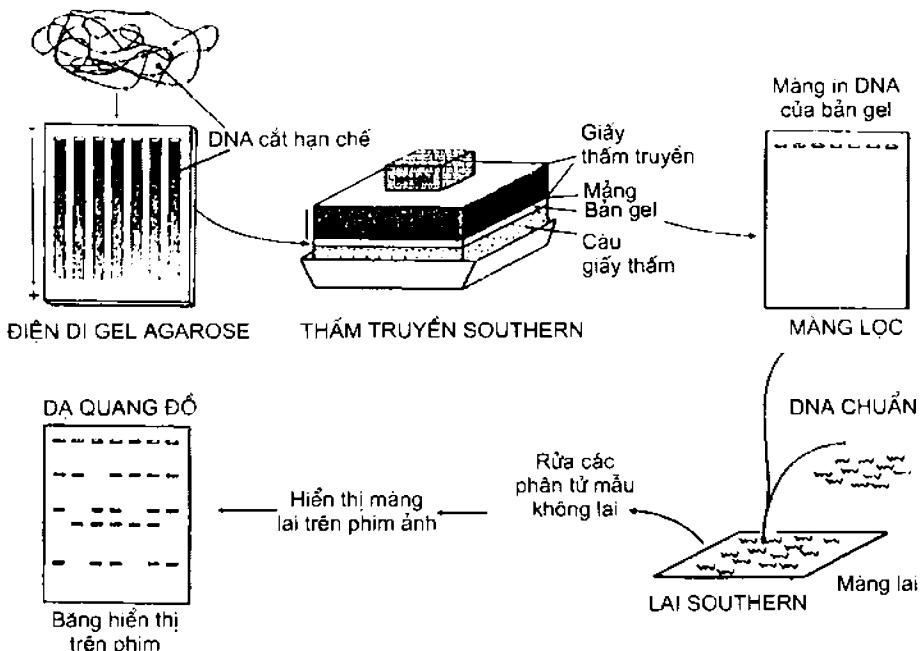


Hình 4.10. Nguyên tắc đánh dấu phân tử bằng dioxygenin.

4.4.2. Lai Southern

Đối với DNA, phương pháp được Edwin M. Southern (1975) xây dựng và thử nghiệm thành công. Lai Southern dựa trên nguyên lý các cặp base bổ sung của hai phân tử DNA bắt cặp và tạo ra sợi DNA kép. Trước hết DNA được cắt bằng enzyme hạn chế, sau đó chạy điện di trên gel agarose, rồi thấm truyền lên màng nitrocellulose, cuối cùng thì lai với mẫu DNA tương đồng và hiển thị trên phim thành từng băng dạng vạch ngang. Kỹ thuật này được gọi là *kỹ thuật phân tích màng thấm truyền Southern* (Southern blot analysis), gọi ngắn là Southern (hình 4.11).

Lai Southern được sử dụng trong việc phát hiện sự có mặt của gene cần tìm trong nguyên liệu nghiên cứu, xác định số lượng copy của một gene nhất định trong genome của đối tượng nghiên cứu, đặc biệt là trong sinh vật chuyển gene.



Hình 4.11. Sơ đồ tiến hành kỹ thuật Southern bao gồm các bước xử lý bản gel, thám truyền DNA lên màng, lai phân tử và phát hiện các vệt lai.

4.4.3. Lai Northern

Khi ứng dụng kỹ thuật Southern đối với RNA, kỹ thuật này được đặt tên là *kỹ thuật Northern*, bước phát hiện các băng RNA được tiến hành thông qua lai với mẫu DNA hoặc RNA đánh dấu. Nguyên lý của lai Northern cũng dựa trên sự bắt cặp của các đôi base tương đồng, tạo nên phân tử lai RNA/DNA hay RNA/RNA. Ứng dụng phép lai Northern là nhằm phát hiện ra các bản sao mRNA của gene trong quá trình phiên mã, nói cách khác là chứng minh gene đang nghiên cứu được hoạt hoá và phiên mã thành mRNA.

4.4.4. Lai Western

Khi tiến hành kỹ thuật kết tủa miến dịch giữa với protein đóng vai trò kháng nguyên với kháng thể đặc trưng của nó thì phép phân tích được gọi là *kỹ thuật Western*. Nguyên tắc của phép lai Western là phản ứng kết tủa miến dịch của kháng thể và kháng nguyên, mà một trong hai loại protein sẽ được đánh dấu. Tới nay đã có nhiều cải tiến phương pháp, hình thành

nhiều kỹ thuật phối hợp như *Southwestern*, phát hiện bằng protein bằng DNA đánh dấu hoặc *Western xa*, không dùng mẫu protein không phải kháng thể (non-antibody).

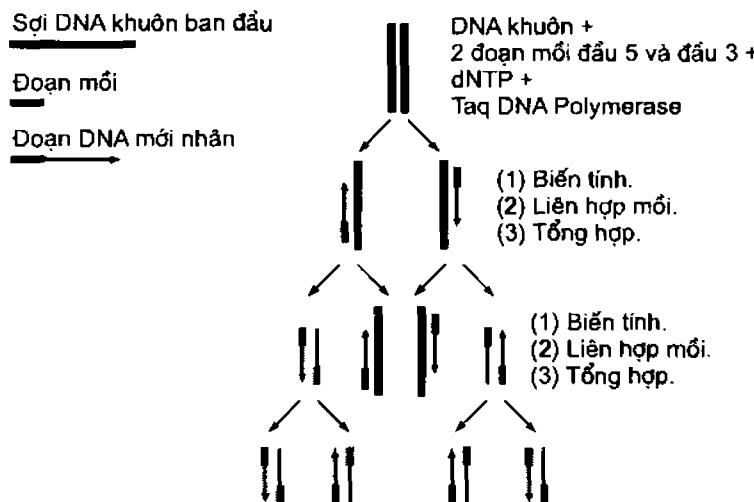
Trong nhiều trường hợp không cần chạy điện di trước mà đưa mẫu phân tích trực tiếp lên màng dưới dạng các vết hình tròn (dot blot) hoặc vết hình vạch ngang (slot blot). Những thiết bị dạng khuôn hỗ trợ để đảm bảo sự đồng đều của các vết và khoảng cách giữa chúng khi ta truyền mẫu lên màng thường gắn với các bơm hút chân không.

4.4.5. Kỹ thuật PCR

PCR là tên gọi tắt của phản ứng chuỗi polymerase (polymerase chain reaction = PCR) để nhân bản các đoạn DNA trong ống nghiệm.

Phản ứng chuỗi polymerase (PCR) được Mullis và cộng sự mô tả đầu tiên và đã góp phần đem lại cho Tiến sĩ Mullis giải thưởng Nobel Y học và Sinh lý năm 1993. PCR là quá trình tương đối đơn giản cho phép nhân nhanh một số lượng không hạn chế nguyên bản một đoạn DNA nhất định trong một khoảng thời gian ngắn. Quá trình nhân bản bằng enzyme này bao gồm 3 bước được lặp lại nhiều lần : (1) Biến tính DNA từ dạng sợi kép thành dạng sợi đơn bằng cách nâng cao nhiệt độ lên 94 – 95°C trong khoảng thời gian ngắn ; (2) Tiếp hợp đoạn mồi (primer) thông qua hạ nhiệt độ xuống 40 – 60°C. Hai đoạn mồi là các oligonucleotide dài khoảng 18 – 24 base sẽ tiếp hợp theo nguyên tắc bổ sung với đoạn DNA tương đồng ở hai đầu của đoạn DNA cần nhân. (3) Phản ứng polymerase tổng hợp phân tử DNA kéo dài từ đoạn mồi. Hai phân tử DNA mới được tổng hợp theo nguyên tắc bổ sung từ hai đầu phân tử khuôn ban đầu (hình 4.12). Để tránh hiện tượng tiếp hợp không đặc thù, phản ứng tổng hợp được thực hiện ở 72°C. Loại polymerase dùng trong PCR là loại chịu nhiệt (*Taq* DNA polymerase) được tách chiết từ vi khuẩn *Thermus aquaticus* hoặc chủ yếu là enzyme tái tổ hợp.

Như vậy, mỗi vòng phản ứng sẽ có 2 lần số lượng DNA của đoạn DNA cần nhân được tổng hợp, hệ số nhân bản tính theo số chu kỳ n là 2^n . Ví dụ, sau 20 chu kỳ có tới trên 30 triệu bản của một đoạn DNA được nhân lên.



Hình 4.12. Phản ứng chuỗi polymerase để nhân bản một đoạn DNA nhất định.

4.5. LẬP BẢN ĐỒ RFLP

RFLP là tên gọi tắt của kỹ thuật đánh giá tính đa dạng về chiều dài của các phân đoạn DNA được cắt bằng enzyme giới hạn (restriction fragment length polymorphism = RFLP).

4.5.1. Giới thiệu chung về RFLP

Trong những năm gần đây, kỹ thuật phân lập gene đã tạo nên những tiến bộ cơ bản trên nhiều lĩnh vực của sinh học hiện đại. Một trong những tiến bộ đáng kể nhất là việc sử dụng các đoạn DNA nhiễm sắc thể làm phân tử đánh dấu di truyền. Kỹ thuật này được gọi là kỹ thuật lập bản đồ về tính đa dạng độ dài các phân đoạn hạn chế (restriction fragment length polymorphism = RFLP). Về nguyên lý, kỹ thuật này dựa trên sự khác biệt tự nhiên của các chuỗi nucleotide trong phân tử DNA và khi sợi DNA bị cắt thành nhiều phân đoạn nhỏ bởi các enzyme hạn chế thì những phân đoạn này có thể khác nhau về kích thước hay chiều dài. Sự khác nhau đó được khai thác để theo dõi quá trình phân ly của các đoạn nhiễm sắc khi lai tạo.

DNA là một loại đai phân tử sinh học. Hàm lượng của nó trong mỗi tế bào được tính theo giá trị tuyệt đối bằng đơn vị picogram (pg) hoặc số cặp base (bp, bội số 1000 là kb và 1000000 là mb) (bảng 4.3). Có thể

dùng công thức sau để chuyển đổi khối lượng và độ dài của phân tử DNA : 1 picogram DNA = $0,965 \times 10^9$ bp = $6,1 \times 10^{11}$ dalton = 29 cm.

Bảng 4.3. Hàm lượng DNA có trong một số cơ quan tử và cơ thể điển hình

Cơ quan tử/ Cơ thể	Hàm lượng DNA (một genome)	
	picogram	kb
<i>E. coli</i>	0,0047	$4,2 \times 10^3$
Lục lạp (Ngô, <i>Zea mays</i>)	0,0017	$4,2 \times 10^3$
Ty thể (Ngô, <i>Zea mays</i>)	0,0017	$4,2 \times 10^3$
<i>Arabidopsis thaliana</i>	0,0047	$4,2 \times 10^3$
Lúa (<i>Oryzae sativa</i>)	0,0047	$4,2 \times 10^3$
Cà chua (<i>Lycopersicon esculentum</i>)	0,07	0,07
Ngô (<i>Zea mays</i>)	0,0047	$4,2 \times 10^3$
Người (<i>Homo sapiens</i>)	<i>E. coli</i>	$4,2 \times 10^3$

Thông tin di truyền làm nên các gene của thực vật bậc cao được chứa đựng trong các chuỗi DNA của nhiễm sắc thể nhân và của các cơ quan tử. Tế bào thực vật có khả năng nhân bản toàn bộ DNA của chúng với tốc độ và tính chính xác cao. Tuy nhiên, trong quá trình nhân bản DNA có khá nhiều khâu hoạt động sai lệch gây nên những thay đổi trong phân tử DNA. Có thể có thay đổi ở một cặp base, cũng có thể có thay đổi lớn hơn do đảo ngược, hoán vị hay mất đoạn. Kết quả là khó có thể có hai cá thể có các chuỗi DNA tuyệt đối giống nhau trong trình tự các base. Vì vậy, có thể nói DNA của thế giới thực vật có mức độ biến dị rất lớn, nhưng không có một phương pháp trực tiếp để khai thác cho nghiên cứu di truyền thực vật.

Có một số cách để xác định những biến dị của chuỗi DNA. Người ta có thể xác định trình tự của chuỗi DNA và so sánh tỉ mỉ từng đoạn. Phương pháp này thường quá tốn kém và đòi hỏi nhiều thời gian. Cũng có thể sử dụng các *enzyme hạn chế*. Đó là những endonuclease do vi khuẩn sinh ra và có khả năng nhận biết những vị trí *hạn chế* để cắt chuỗi DNA dài thành những phân đoạn có chiều dài khác nhau. Số lượng phân đoạn

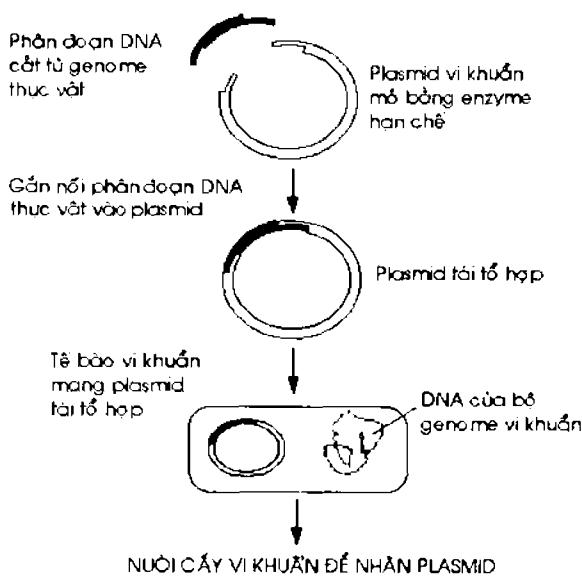
và chiều dài của các phân đoạn phản ánh số lượng các vị trí hạn chế có trên chuỗi DNA mà ta quan tâm. Từ đó, nếu chọn được tổ hợp DNA/loại enzyme hạn chế, ta có thể thiết lập được *ván tay* đặc trưng cho từng loại phân tử DNA (cũng đặc trưng cho cá thể chứa loại phân tử DNA đó).

Các phân tử DNA tương đối nhỏ của lục lạp (chẳng hạn thường bị cắt thành 40 đoạn khác nhau) khi xử lý với enzyme hạn chế như *EcoRI*, thì các phân đoạn này có thể được tách riêng theo kích thước lớn nhỏ bằng kỹ thuật điện di trên gel agarose. Khi nhuộm bản gel với ethidium bromide, cách thức phân bố của các đoạn DNA trên bản gel hiện rõ dưới ánh sáng cực tím, có thể quan sát hoặc chụp ảnh trực tiếp. Các phân tử DNA của lục lạp có những khác biệt trong trình tự base hay đã được xấp xếp lại thông qua các biến đổi dạng *thêm vào, mất đi* hoặc *xâm nhập* sẽ tạo ra những phân đoạn cắt hạn chế có độ dài khác nhau. Những khác nhau khi so sánh mẫu cắt hạn chế DNA của nhân, cơ quan tử hoặc DNA toàn phần được gọi là *sự đa dạng về chiều dài phân đoạn cắt hạn chế (RFLP)* và được sử dụng để đánh giá tính đa dạng di truyền. Đã có nhiều công trình sử dụng RFLP của lục lạp để nghiên cứu quá trình tiến hoá và phân loại các loài của nhiều nhóm thực vật. Các công trình này chứng tỏ rằng phân tích RFLP của DNA lục lạp rất hữu hiệu trong việc làm sáng tỏ quan hệ phân loại của nhiều nhóm thực vật. Rất tiếc là khả năng ứng dụng DNA lục lạp trong công tác tạo giống rất hạn chế. Hầu hết các gene có ý nghĩa kinh tế quan trọng đều nằm trong nhân ; chỉ có một số ít gene được tìm thấy trong DNA lục lạp.

Phân tích RFLP hoàn toàn có thể ứng dụng đối với DNA nhân, nhưng phức tạp hơn vì lượng DNA của nhân lớn hơn hàng trăm ngàn lần so với lục lạp. Khi xử lý DNA nhân với enzyme hạn chế sẽ có hàng triệu phân đoạn được tạo thành với mọi kích thước biến thiên liên tục từ nhỏ đến lớn. Chạy điện di trên gel agarose và nhuộm bằng ethium bromid sẽ không nhận ra các băng riêng biệt mà là một dải sáng nhòe liên tục. Phải cần có kỹ thuật tinh vi hơn mới xác định được sự khác nhau trong tính đa dạng về chiều dài các phân đoạn cắt hạn chế của DNA nhân khi muốn so sánh nhiều loại tế bào của nhiều cá thể khác nhau. Kỹ thuật đó bao gồm việc sử dụng các phân tử đánh dấu và kỹ thuật lai DNA.

4.5.2. Thư viện chỉ thị phân tử

Do RFLP của DNA nhân không thể nhận biết bằng mắt thường được nên người ta phải dùng những đoạn nhiễm sắc thể làm phân tử đánh dấu để nhận biết các phân đoạn cắt hạn chế. Khi sử dụng phương pháp lai DNA - DNA có tính đặc hiệu cao, người ta có thể xác định được những phân đoạn cắt hạn chế riêng biệt từ một hỗn hợp các phân đoạn của DNA nhân trong chế phẩm xử lý cắt hạn chế. Để thực hiện tốt kỹ thuật này cần phải có một loạt các đoạn DNA mẫu. Tập hợp các mẫu đó được gọi là thư viện. Thông thường, DNA của một loài nhất định được phân lập và cắt thành những phân đoạn dài khoảng 2 đến 5 kb được dùng làm mẫu đánh dấu để lai phân tử. Nói chung các phân đoạn DNA đều có thể dùng làm mẫu nhưng cần phải đạt trạng thái tinh khiết. Sẽ rất khó khăn nếu như phải tinh chế các phân đoạn riêng rẽ từ hỗn hợp cắt hạn chế, nhưng may mắn là không cần phải làm việc đó vì nhờ có kỹ thuật nhân gene (gene cloning). Có thể tóm tắt kỹ thuật nhân gene như sau : Đoạn DNA cần nhân được gắn vào plasmid rồi biến nạp vào tế bào vi khuẩn (như *E. coli*). Tế bào vi khuẩn sẽ nhân và tạo nhiều phiên bản của đoạn DNA đó cùng với quá trình sinh trưởng và phân bào. Thông qua nuôi cấy và phân lập plasmid từ vi khuẩn có thể thu được khối lượng đáng kể các đoạn DNA thực vật dùng làm mẫu cho lai DNA (hình 4.13). Chúng vi khuẩn mang đoạn DNA mẫu đó có thể được bảo quản dài hạn và chuyển giao một cách thuận tiện cho nhiều nhà nghiên cứu có nhu cầu sử dụng.



Hình 4.13. Tạo plasmid tái tổ hợp mang đoạn DNA của genome thực vật nhằm xây dựng thư viện các đoạn DNA chỉ thị dùng trong RFLP.

Với những mẫu đánh dấu quy trình xác định RFLP của DNA nhân được tiến hành như sau : phân lập DNA nhân của thực vật cần xác định ; xử lý DNA với enzyme hạn chế ; chạy điện di DNA đã cắt trên gel agarose. Kết quả là hàng triệu các đoạn DNA sẽ được phân đoạn trong bản gel theo khối lượng phân tử. Để có thể tiến hành lai phân tử DNA đánh dấu và phân tử DNA trên gel phải ở trạng thái sợi đơn, đó là trạng thái biến tính. Điều này xảy ra khi ngâm bản gel trong dung dịch NaOH để biến tính DNA và thấm chuyển chúng lên màng lọc nitrocellulose. Quá trình này được gọi là Southern. Vì mảnh màng được cắt cùng kích cỡ và đặt trực tiếp lên bản gel, nên khi thấm chuyển từ gel lên màng, toàn bộ hình thái sắp xếp của các phân đoạn cắt hạn chế vẫn được giữ nguyên như một bản in. Do tính chất của loại màng nitrocellulose là có khả năng giữ rất chặt các phân tử DNA trên đó cho nên các mảnh màng lọc này có thể được sử dụng nhiều lần cho nhiều phép lai với nhiều loại mẫu chuẩn khác nhau.

Sau khi chuẩn bị xong màng lọc mang DNA cắt hạn chế và phân đoạn có thể tiến hành lai với mẫu chuẩn. Nguyên lý chung bao gồm việc biến tính mẫu chuẩn và tạo điều kiện thích hợp về nhiệt độ và nồng độ muối sao cho phân tử mẫu chuẩn lai với phân tử tương đồng có trên màng theo nguyên tắc kết hợp của các cặp base bổ sung. Để có thể đánh giá được kết quả lai cần phải đánh dấu các phân tử mẫu. Thông thường, có thể dùng đồng vị phóng xạ như ^{32}P (kỹ thuật đánh dấu nóng). Hiện nay, có thể không cần dùng chất phóng xạ mà vẫn đánh dấu được. Đó là kỹ thuật *kích hoạt phát quang hóa học* (Enhanced Chemical Luminescence = ECL = kỹ thuật đánh dấu lạnh). Sau khi rửa bỏ những phân tử mẫu đánh dấu còn dư thừa, màng lọc được đưa lên phim âm bản để hiện hình. Đối với kỹ thuật sử dụng phóng xạ để đánh dấu thì công đoạn này được gọi là *phóng xạ tự chụp* (autoradiography), cần từ một vài phút đến hàng chục ngày ; còn trường hợp dùng kỹ thuật ECL thì phim bị hoạt hoá trong thời gian từ một vài phút đến một vài giờ. Bằng cách thức như vừa trình bày, mẫu RFLP của DNA nhân thực vật hoàn toàn có thể được xác định và so sánh.

4.5.3. Xây dựng bản đồ di truyền RFLP

Để thuận lợi cho việc giải thích khả năng ứng dụng của phân tử đánh dấu RFLP đối với lập bản đồ di truyền, chúng ta thử thảo luận một ví dụ

sau : Lai hữu tính hai bố mẹ nhị bội để thu được thế hệ F_1 và F_1 tự thụ phấn sẽ cho ra đời F_2 . Giả sử cây bố mẹ đều đồng hợp tử tới mọi loci do tự thụ phấn nhiều thế hệ. Rõ ràng là con lai sẽ có một bộ nhiễm sắc thể của bố và một bộ nhiễm sắc thể của mẹ. Như vậy, hai nhiễm sắc thể của một đôi nhiễm sắc thể tương đồng sẽ khác nhau ở mức độ mà bố mẹ chúng khác nhau trong trình tự base của các chuỗi DNA.

Khi trong các cây F_1 xảy ra phân bào giảm nhiễm để hình thành các giao tử thì nhiễm sắc thể của chúng sẽ tái tổ hợp thông qua trao đổi chéo. Quá trình tái tổ hợp sẽ tạo cho nhiễm sắc thể ở giao tử là một thể khám của các phân đoạn DNA từ bố và từ mẹ. Điều đáng lưu tâm là không có hai nhiễm sắc thể được tổ hợp từ những phân đoạn giống hệt nhau. Ở thế hệ F_2 được tạo ra do tự thụ phấn các cây F_1 , cặp nhiễm sắc thể tương đồng vẫn giữ được trạng thái khám bằng các đoạn nhiễm sắc thể của bố mẹ.

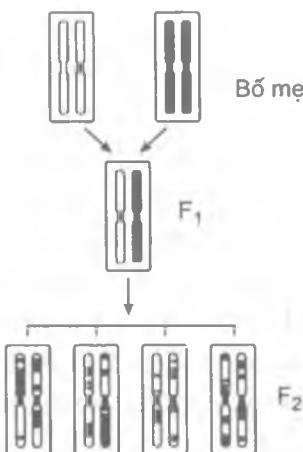
Quá trình tái tổ hợp đó chính là cơ sở cho việc thiết lập bản đồ di truyền theo phương pháp truyền thống. Việc làm này phụ thuộc vào hai quan sát :

1. Các đoạn nhiễm sắc thể không nằm trên cùng một nhiễm sắc thể sẽ tham gia tái tổ hợp ngẫu nhiên.
2. Các đoạn của cùng một nhiễm sắc thể sẽ tham gia tái tổ hợp phù hợp khoảng cách lý học của chúng. Đoạn nhiễm sắc liên kết chặt chẽ (gần nhau) thường khó tham gia tái tổ hợp hơn những đoạn xa nhau. Vì thế, khoảng cách di truyền hoặc khoảng cách trên bản đồ gene là một hàm số của quá trình tái tổ hợp xảy ra trong quá trình hình thành giao tử.

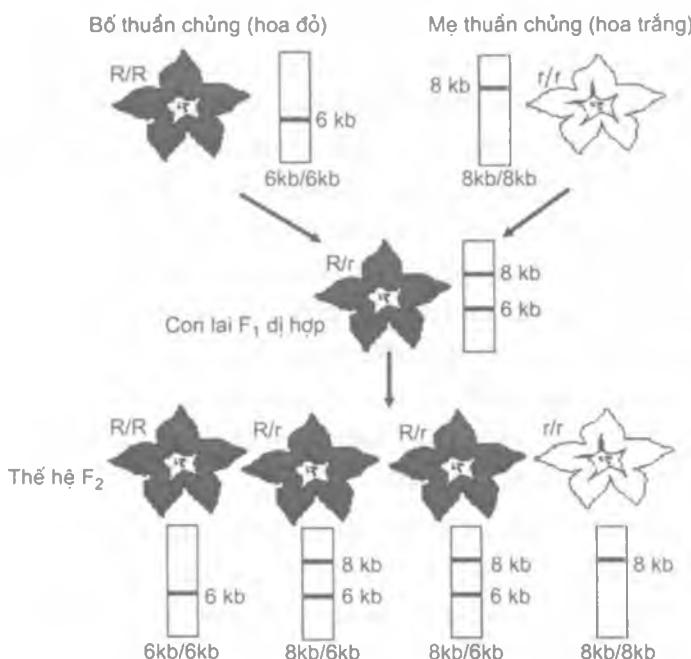
Trước đây, người ta chỉ có mỗi một cách theo dõi quá trình tái tổ hợp các đoạn nhiễm sắc thể trong quá trình hình thành giao tử là quan sát kiểu hình quyết định bởi các gene nằm trên các đoạn nhiễm sắc thể tham gia tái tổ hợp. Thông qua quan sát kiểu hình như màu hoa, chiều cao cây, tính kháng côn trùng hoặc tính trạng của nội nhũ, người ta có thể liên hệ tới biểu hiện của các đoạn nhiễm sắc thể liên đới trong quá trình phân bào giảm nhiễm và tái tổ hợp. Theo dõi các đoạn nhiễm sắc thể theo phương thức gián tiếp như vậy đã được ứng dụng để thiết lập nên bản đồ di truyền khá chi tiết của nhiều loài thực vật. Tuy nhiên, phương pháp này khá tốn kém công sức và thời gian.

Rất may, kỹ thuật RFLP mở ra cách thức mới để theo dõi sự phân ly và tái tổ hợp của các đoạn nhiễm sắc thể và đơn giản hóa khá nhiều việc lập bản đồ di truyền. Thay cho việc quan sát các kiểu hình, người ta quan sát trực tiếp các phân tử chỉ thị trên đoạn nhiễm sắc thể. Bằng cách này, người ta quan sát trực tiếp kiểu gene chứ không phải quan sát gián tiếp kiểu hình như trước đây (hình 4.14). Phương pháp này có khá nhiều ưu điểm sẽ được thảo luận dưới đây.

Nếu quan sát cách thức di truyền các chỉ thị RFLP sẽ nhận thấy chúng phân ly hoàn toàn chính xác như các chỉ thị gene và tuân thủ các quy luật Mendel một cách nghiêm ngặt. Vì thế, bản đồ của RFLP được thiết lập theo nguyên tắc như đối với các chỉ thị truyền thống (so sánh hình 4.15).



Hình 4.14. Di truyền các đoạn nhiễm sắc thể.



Hình 4.15. Di truyền một chỉ thị phân tử liên quan đến màu sắc hoa.

4.5.4. So sánh chỉ thị di truyền truyền thống và RFLP

Khi so sánh chỉ thị di truyền truyền thống với chỉ thị RFLP, ta sẽ thấy các chỉ thị RFLP có nhiều ưu điểm đối với việc thiết lập bản đồ di truyền. Những ưu điểm đó là :

Sự đa dạng tự nhiên : Để thiết lập bản đồ di truyền thông qua theo dõi phân ly các đoạn nhiễm sắc thể, người ta phải phân biệt được các đoạn của hai nhiễm sắc thể bố mẹ. Nếu muốn theo dõi các đoạn nhiễm sắc thông qua các gene truyền thống thì mỗi bố mẹ phải có các allele khác nhau (thông thường là thể bình thường và thể đột biến). Nhiều khi không thể trông chờ vào sự đa dạng tự nhiên mà phải tạo ra sự sai khác thông qua đột biến thực nghiệm. Kỹ thuật RFLP lợi dụng tính biến động tự nhiên rất cao trong phân tử DNA mà không cần phải gây đột biến. Số lượng các chỉ thị RFLP đối với một loài thực vật có thể rất lớn, tối mức di truyền học thông thường không cần sử dụng hết.

Đặc trưng giai đoạn phát triển và đặc trưng cơ quan : Để dự đoán quá trình phân ly của các đoạn nhiễm sắc thể bằng phương pháp theo dõi kiểu hình truyền thống, người ta hoàn toàn phụ thuộc vào kết quả của quá trình biểu hiện gene. Điều này đôi lúc không dễ dàng, chẳng hạn có nhiều gene chỉ biểu hiện trên những loại mô nhất định (ví dụ như màu hoa chỉ có thể đánh giá khi cây ra hoa thực sự) ; có những gene chỉ biểu hiện ở một giai đoạn phát triển nhất định hoặc có những gene khác lại chỉ thể hiện dưới điều kiện ngoại cảnh đặc trưng (gene chống chịu điều kiện ngoại cảnh bất lợi như phèn, mặn, nóng, lạnh). Vì vậy, việc đánh giá một gene theo phương pháp truyền thống đôi lúc rất tốn kém tiền của và thời gian. Ví dụ, muốn đánh giá gene kháng côn trùng thường phải bố trí trồng cây trong khu vực có lưới che chắn và phải thả côn trùng vào đúng chủng loài, đúng kiểu sinh học, đúng giai đoạn phát triển và đúng mật độ. Trong khi đó, kỹ thuật RFLP chỉ cần một lượng nhỏ của bất kỳ loại mô nào, bất kỳ trạng thái phát triển nào miễn là từ đó có thể tách được một lượng nhỏ DNA đủ cho phân tích.

Hiệu quả kiểu hình : Chỉ thị RFLP phản ánh mức độ biến động tự nhiên tồn tại trong chuỗi DNA của thực vật. Phần lớn những sai khác này không biểu hiện ra kiểu hình. Như vậy, trong thực tế, số lượng các chỉ thị RFLP là rất lớn. Ngược lại, chỉ có một số lượng rất hạn chế các chỉ thị kiểu hình có thể đồng thời tồn tại ở một loài thực vật bởi vì đột biến

thường gây hại hoặc đột biến xảy ra không tập trung vào một cá thể mà ở nhiều cá thể khác nhau và khó có thể dùng phương pháp lai tạo để tích luỹ nhiều đột biến vào một cá thể được. Thực tế này có một hệ quả quan trọng đối với việc thiết lập bản đồ gene. Vì mỗi một cá thể thực vật đều mang trong nó một lượng đầy đủ thông tin về độ biến động DNA cho nên kỹ thuật RFLP chỉ cần một nhóm nghiên cứu vừa phải tập trung phân tích một cặp lai với một quần thể lặp thế hệ tiếp theo là có thể lập được bản đồ di truyền, trong khi đó kỹ thuật truyền thống cần tập trung nhiều nhóm nghiên cứu làm việc với hàng trăm cặp lai trong nhiều năm.

Hiệu quả môi trường : Quá trình thể hiện gene bắt đầu bằng chuỗi DNA có chức năng gene được phiên mã thành RNA thông tin (mRNA), từ đó dịch mã thành protein. Phân tử protein phải biến đổi để có cấu trúc không gian thích hợp và được vận chuyển đến vùng hoạt động thể hiện chức năng hoặc dưới dạng enzyme xúc tác hay tạo ra cấu trúc tế bào. Tất cả các bước trong chuỗi phản ứng thể hiện hoạt tính gene đều phụ thuộc vào điều kiện môi trường. Kỹ thuật RFLP không phụ thuộc vào sự biểu hiện gene nên không phụ thuộc vào yếu tố môi trường, nó chỉ chứng minh sự có mặt hay vắng mặt của một chuỗi base đặc trưng nào đó.

Mối tương tác gene : Hoạt tính gene thường phụ thuộc sự tương tác giữa các gene. Ví dụ đơn giản nhất là biểu hiện tính trội ; hoặc nhiều đột biến lại thể hiện tính lặn. Những tính trạng lặn như vậy chỉ được nhận biết trong cơ thể đồng hợp tử đối với allele lặn. Còn cơ thể dị hợp không thể phân biệt được với thể gốc bình thường. Vì chỉ thị RFLP phản ánh trực tiếp kiểu gene cho nên có thể phân biệt rất rõ ràng cả ba khả năng tổ hợp của một allele trội với một allele lặn đột biến.

4.5.5. Lập bản đồ di truyền bằng chỉ thị RFLP

Lược trình các bước cần thực hiện khi xây dựng bản đồ di truyền thông qua kỹ thuật RFLP như sau :

4.5.5.1. Chọn cây bố mẹ

Trong cặp lai để phân tích RFLP cần phải có cây bố và cây mẹ xa nhau về mặt di truyền đủ để thu được sự đa dạng về độ dài phân đoạn cắt hạn chế, nhưng cũng không quá xa để đến mức gây ra hiện tượng bất dục của con lai. Thuận lợi nhất là chọn một cặp lai có vài đặc điểm nông học

đáng giá mà chúng sẽ phân ly sau này. Trong hầu hết các cây đã được sử dụng để lập bản đồ di truyền, người ta tìm thấy mức độ biến động đủ lớn giữa các giống của cùng một loài. Ví dụ trong đề án lập bản đồ RFLP của cây lúa, người ta đã sử dụng các cặp lai giữa giống indica và giống javanica. Ở cà chua, cơ sở di truyền của các giống trồng trọt lại quá gần cho nên cần phải lai chúng với loài hoang dã.

Vì người ta không thể biết trước mức độ biến động RFLP có thể có trong một loài cây trồng, cho nên cần phải thăm dò các mẫu chỉ thị thông qua chọn lọc ngẫu nhiên. Việc thăm dò như vậy có thể tiến hành với các giống cây trồng có nguồn gốc xa nhau về phân loại. Những loài hoang dại có thể lai với giống trồng trọt. Nên sử dụng mọi nguồn thông tin có thể có được vào việc đánh giá, thăm dò, trong đó nên lưu ý tới nguồn gốc tiến hoá hoặc phân loại và các phân tích về isozyme.

Sau khi chọn được một loạt cá thể để thăm dò thì tiến hành tách chiết DNA của từng cá thể, xử lý DNA với enzyme hạn chế và sàng lọc tính đa dạng bằng phương pháp thông dụng. Mỗi một lô thí nghiệm được đánh giá theo allele lựa chọn. Cặp cá thể nào có mức độ đa dạng thích hợp thì được chọn làm cặp lai.

4.5.5.2. Tạo quần thể lập bản đồ

Lai hai cá thể bố mẹ để thu được một hay nhiều cây F₁. Trong trường hợp cây tự thụ phấn như lúa tất cả các cây F₁ đều giống nhau và là dạng đị hợp tử đối với tất cả các chỉ thị RFLP của bố mẹ. Một quần thể cho việc lập bản đồ di truyền có thể được tạo ra bằng cách tự phôi các cá thể F₁ để thu được thế hệ F₂ có phân ly hoặc phải lai ngược với một trong hai bố mẹ để thu được phân ly trong thế hệ lai ngược đầu tiên. Nên sử dụng thế hệ F₂, vì như thế sẽ thu được nhiều thông tin hơn so với tiến hành lai ngược. Một quần thể F₂ hoặc lai ngược có 50 cá thể là đủ để xây dựng được bản đồ di truyền chi tiết. Đối với những loài cây được nghiên cứu tạo giống nhiều, không cần phải tiến hành lai mà có thể sử dụng ngay hạt của cặp lai trước đây để rút ngắn thời gian chuẩn bị.

4.5.5.3. Đánh giá RFLP trong quần thể lập bản đồ

Khi có được quần thể lập bản đồ, tiến hành tách chiết DNA ở từng cá thể. Nên lưu ý rằng DNA của từng cá thể đều có một bộ riêng tổ hợp các

đoạn nhiễm sắc thể của bố mẹ. Cần phải xác định rõ mỗi cá thể trong quần thể mang những phân đoạn nhiễm sắc thể nào của bố mẹ. Nếu để các cá thể F_2 tự phối thì quá trình tái tổ hợp các đoạn DNA tiếp tục xảy ra, lúc đó bức tranh về RFLP sẽ thay đổi. Vì vậy, phải tìm cách duy trì các cá thể trong quần thể lập bản đồ ở trạng thái sống càng lâu càng tốt để có thể thu được thật nhiều vật liệu sống cho việc tách chiết đủ lượng DNA cần thiết cho phân tích. Trong điều kiện cho phép có thể bảo quản nguyên liệu như lá cây ở nhiệt độ -80°C trong một thời gian dài. Quá trình lập bản đồ được tiến hành theo quy trình sau :

a) *Sàng lọc RFLP* : Mẫu được thử lần lượt đối với cây bố mẹ nhằm phát hiện ra loại enzyme hạn chế thích hợp tạo ra được sự đa dạng RFLP rõ rệt nhất giữa hai cá thể bố và mẹ. Số lượng enzyme cần thiết phụ thuộc vào mức độ biến động tồn tại trong cặp bố mẹ. Những giống cây có tính đa dạng cao như ngô dòng thuần thì chỉ cần 1 đến 2 enzyme là đủ nhưng ở lúa cần tới 11 enzyme mới phát hiện được tính đa dạng ở 75% mẫu đánh dấu.

b) *Chọn cặp chỉ thị/enzyme* : Trước hết cần chọn cặp chỉ thị/enzyme thích hợp, đó là loại enzyme để cắt DNA và loại chỉ thị cho ra kết quả khác nhau giữa cây bố và cây mẹ. Có như vậy mới lập được bản đồ trên quần thể con. Công việc cần tiến hành như sau :

Tách chiết DNA của cây bố, cây mẹ và từng cây con trong quần thể con lập bản đồ. Cắt bằng từng loại enzyme nhất định. Chạy điện di và thẩm truyền lên màng lai. Mỗi một enzyme dùng một bộ màng lai riêng rẽ (ví dụ bộ thứ nhất dùng *EcoRI*, bộ thứ hai dùng *PstI*), sau đó tiến hành lai với từng loại phân tử chỉ thị. Kết quả thu được ví dụ chỉ thị số 36 trên bộ màng *PstI* cho sự sai khác rõ ràng thì cặp #36/*PstI* được chọn cho việc lập bản đồ tiếp theo. Những chỉ thị khác có thể không phù hợp với *PstI* nhưng có thể có một chỉ thị nào đó lại phù hợp với bộ màng *EcoRI*. Cứ như thế, ta có các cặp chỉ thị/enzyme thể hiện được sự đa dạng về kiểu gene sẽ được dùng để lập bản đồ.

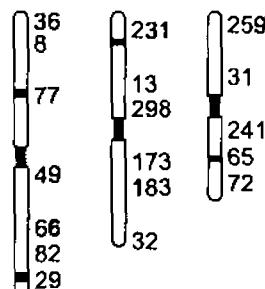
c) *Đánh giá mức độ liên kết* : Khi so sánh hai loại chỉ thị ta thu được hai kết quả : Một là chúng có xu hướng cùng phân ly, ví dụ F_2 dị hợp với chỉ thị số một và cũng dị hợp với chỉ thị số hai, trong khi đó bố mẹ đều đồng hợp với cả hai chỉ thị. Như vậy, hai chỉ thị này có mối liên kết. Hai là sự phân bố tính đồng hợp và dị hợp của hai chỉ thị này không có quy luật, mà hoàn toàn ngẫu nhiên. Như vậy, chúng không liên kết. Khi

mở rộng số lượng chỉ thị, ta sẽ thu được mức độ liên kết khác nhau của từng chỉ thị hay nhóm chỉ thị. Dùng chương trình máy vi tính để tính toán và biểu diễn mức độ liên kết dưới dạng biểu đồ.

4.5.5.4. Ứng dụng RFLP trong chọn, tạo giống cây trồng

Chọn lọc gián tiếp một gene thông qua sử dụng bản đồ RFLP sẽ làm giảm giá thành, đơn giản hoá và tiết kiệm được thời gian. Tuy nhiên, khi sử dụng bản đồ RFLP, người ta không thể chọn lọc trực tiếp gene quan tâm mà chỉ chọn thông qua những chỉ thị liên kết chặt chẽ với gene đó. Hình 4.16 trình bày một ví dụ phân tích tính trạng di truyền số lượng: Giả thuyết cây có 3 nhiễm sắc thể, các số ghi bên phải từng nhiễm sắc thể chỉ vị trí của các chỉ thị, vùng bôi đen liên quan đến tính trạng cần quan tâm. Có thể tiến hành lai và phân tích RFLP về sự di truyền và phân ly tính trạng đó một cách gián tiếp qua các chỉ thị. Trong ví dụ này, các chỉ thị số 77, 29, 231 và 65 liên quan chặt chẽ với tính trạng quan tâm.

Tóm lại, chỉ thị RFLP được dùng trực tiếp như mẫu chuẩn để xác định sự có mặt hay thiếu vắng các đoạn nhiễm sắc thể nhất định. Khả năng theo dõi trực tiếp quá trình di truyền các đoạn nhiễm sắc thể cho phép ứng dụng kỹ thuật RFLP trong tạo giống cây trồng và sinh học phân tử. Sắp tới, RFLP trở thành một công cụ thông dụng cho các nhà tạo giống trong việc tạo ra các giống cây trồng siêu trội mới.



Hình 4.16. Liên hệ giữa chỉ thị RFLP và tính trạng số lượng.

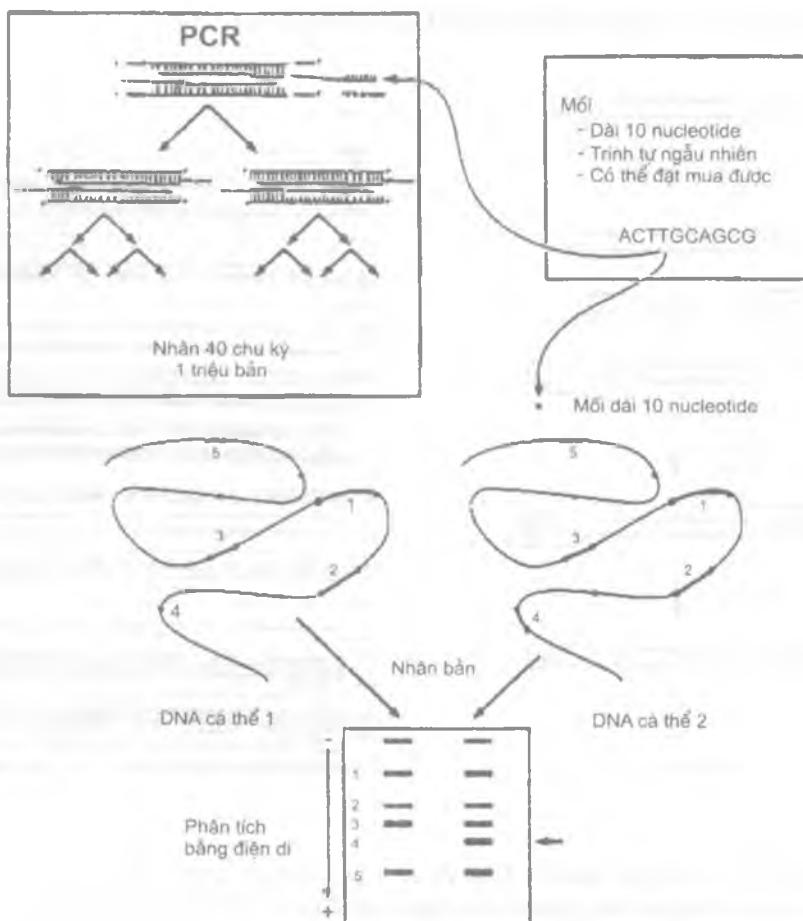
4.6. ĐÁNH GIÁ TÍNH ĐA DẠNG DI TRUYỀN BẰNG KỸ THUẬT RAPD

RAPD là tên gọi tắt của kỹ thuật phân tích DNA đa hình được nhân bản ngẫu nhiên (Random Amplified Polymorphic DNA = RAPD).

Trong khi tiến hành phân tích RFLP phải dùng đến kỹ thuật đánh dấu đối với mẫu DNA chuẩn. Tuy nhiên, hiện nay đã có nhiều kỹ thuật phi phóng xạ bên cạnh kỹ thuật phóng xạ, nhưng vẫn có những tìm tòi làm giản tiện hơn nữa kỹ thuật phân tích đánh giá genome thực vật. Một trong những kỹ thuật đó là kỹ thuật phát hiện tính đa dạng của DNA nhân bản

ngẫu nhiên dựa trên nguyên lý của kỹ thuật PCR. Khác với PCR thông thường phải cần biết thông tin về trình tự DNA của đoạn gene cần nhân bản để thiết kế đoạn mồi. So với RFLP thì kỹ thuật RAPD chỉ cần một lượng nhỏ DNA có nguồn gốc genome là đủ. Các đoạn mồi được tổng hợp dưới dạng decamer, tức là các oligonucleotide chỉ dài 10 nucleotide theo những trình tự nhất định. Đoạn mồi được tiếp hợp với DNA khuôn ở nhiệt độ không cao (36°C) cho nên sẽ có nhiều đoạn DNA khác nhau được nhân bản. Sau PCR, sản phẩm được tách bằng điện di trên gel agarose và hiển thị trực tiếp bằng tia tử ngoại và ethidium bromide.

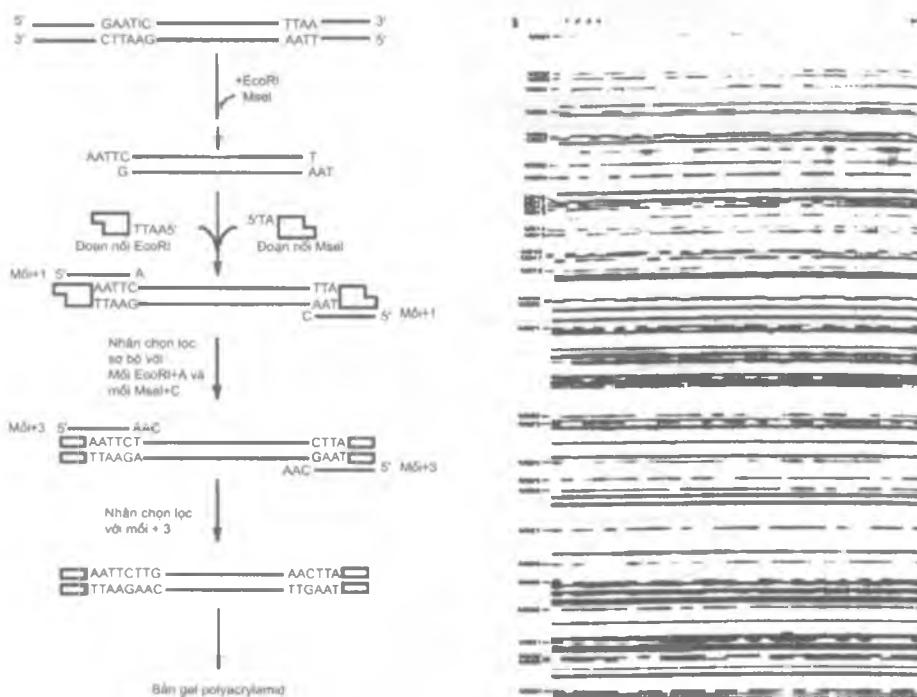
Thông thường, đối với một đối tượng thực vật phải tiến hành sàng lọc để chọn được một số loại mồi thích hợp.



Hình 4.17. Sơ đồ kỹ thuật RAPD để so sánh các tính đa hình của các đoạn DNA được nhân bản ngẫu nhiên.

4.7. LẤY DẤU DNA BẰNG KỸ THUẬT AFLP

AFLP là tên gọi tắt của kỹ thuật lấy dấu DNA bằng cách so sánh tính đa dạng chiều dài các phân đoạn gene được nhân bản (Amplified Fragment Length Polymorphism = AFLP). Nguyên tắc cơ bản của phương pháp AFLP là cắt nhỏ DNA tổng số thành những đoạn bằng hai enzyme hạn chế, sau đó gắn các đoạn DNA này với các đầu nối. Mỗi dùng trong nhân bản là các trình tự thích ứng với đầu một, cộng thêm 3 nucleotide ngẫu nhiên cho phép giảm số đoạn cần nhân xuống 64 lần, để chỉ còn khoảng vài chục đến hàng trăm đoạn DNA được nhân bản, đủ điều kiện để nhận biết chúng trên bản gel poly acrylamide dài khoảng 50 cm. AFLP là kỹ thuật lấy dấu vân DNA hiệu quả nhất bởi vì bằng cách này, khó có 2 cá thể có cùng kiểu hình AFLP như nhau.



Hình 4.18. Sơ đồ tiến hành kỹ thuật AFLP trên cơ sở sử dụng EcoR1 và Mse1 (hình bên trái) và bản điện di sản phẩm AFLP của các dòng cây cà chua cho thấy có nhiều băng vạch đa hình, nghĩa là chúng có xuất hiện ở cá thể này nhưng lại không có ở cá thể kia (hình bên phải).

5

KỸ THUẬT DNA TÁI TỔ HỢP

5.1. CÁC KHÁI NIỆM CƠ BẢN

- Kỹ thuật DNA tái tổ hợp được tạo ra để cắt dán một phân đoạn DNA lạ vào một phân tử nhỏ có khả năng sao chép (như plasmid vi khuẩn), sau đó nó sẽ khuếch đại phân đoạn này cùng với nó và cho kết quả một dòng phân tử phân đoạn chèn DNA.
- Các enzyme hạn chế cắt DNA ở những vị trí đích đặc trưng, cho kết quả các phân đoạn với các đầu dính nhất định thích hợp để chèn vào một vector mà cũng được cắt mở vòng bằng chính enzyme đó.
- Một tập hợp các dòng DNA bao gồm toàn bộ genome của một sinh vật được gọi là một thư viện genome.
- Một dòng DNA riêng biệt có thể được chọn từ một thư viện bằng cách sử dụng một mẫu dò đặc hiệu cho DNA hoặc sản phẩm protein của nó hoặc bằng khả năng biến đổi thể đột biến của nó.
- Các phân đoạn DNA kích thước khác nhau được sản sinh ra bằng cách cắt enzyme hạn chế có thể được phân tách vì chúng di chuyển tới các vị trí khác nhau trên một gel điện di.
- Các phân tử RNA hoặc DNA cắt enzyme hạn chế được phân tách theo kích thước trên một gel điện di, có thể được dò tìm để phát hiện phân tử đặc trưng.
- Các vị trí đích enzyme hạn chế có thể được lập bản đồ, cung cấp các chỉ thị tiện lợi để thao tác DNA.
- Một gene có thể được tìm thấy bằng việc thử các dòng gối nhau phân tán từ một chỉ thị liên kết.
- Sau khi một gene được nhận dòng, trình tự nucleotide của nó được xác định, và có thể được sử dụng để nghiên cứu chức năng và tiến hoá gene.

- Có thể sử dụng một cặp mồi sao chép trải dài qua một trình tự DNA để khuếch đại và phân lập trình tự đó.

Mục đích của di truyền học là nghiên cứu cấu trúc và chức năng của gene và genome. Kể từ thời Mendel, nhiều gene đã được nhận dạng qua quan sát tỷ lệ kiểu hình chuẩn trong các dòng lai kiểm tra. Manh mối về chức năng gene xuất hiện trước tiên từ các đột biến đặc biệt liên quan tới sự thiếu hụt enzyme và protein khác. Mỗi quan hệ của các vùng đột biến trong một gene với sự thay thế amino acid trong protein tương ứng dẫn đến sự hiểu biết hơn về cấu trúc và chức năng gene. Thêm vào những ý tưởng này là những phát hiện về bản chất DNA và mã di truyền, dẫn tới sự hiểu biết toàn diện hơn về bản chất cơ bản của gene. Tuy nhiên, tất cả đều là những điều suy ra gián tiếp từ gene ; không có gene nào được phân lập và trình tự DNA của nó được kiểm tra trực tiếp. Thực tế, điều này không thể làm được để phân lập một gene cụ thể nào đó từ genome.

Mặc dù phân lập DNA từ mô sống tương đối dễ dàng, trong một ống nghiệm, DNA giống như một cục nhầy. Nhưng làm thế nào có thể phân lập một gene đơn lẻ từ một mớ DNA lộn xộn đó ? Kỹ thuật DNA tái tổ hợp cung cấp các kỹ thuật để thực hiện được điều đó, ngày nay, các gene đơn lẻ và các phần của genome được phân lập hàng ngày.

Tại sao phân lập gene lại quan trọng đến thế ? Trước tiên, phân lập gene tạo ra khả năng xác định trình tự nucleotide. Từ thông tin này, những mốc giới bên trong gene có thể được xác định, ví dụ số intron và vị trí. So sánh các trình tự DNA giữa các gene, cũng có thể dẫn tới hiểu biết sâu sắc trong tiến hoá gene. Chuyển đổi trình tự DNA của một gene thành trình tự amino acid bằng cách sử dụng mã di truyền dẫn tới xác định trực tiếp cấu trúc ở sản phẩm protein của gene, và từ hiểu biết này, chức năng của gene có thể được suy đoán ra. Chức năng cũng có thể được nghiên cứu bằng cách cài biến trực tiếp một phân trình tự DNA của gene, tiếp theo là đưa gene vào genome. Hơn nữa, một gene có thể được lấy ra từ một sinh vật này và đưa vào một sinh vật khác. Một sinh vật chứa một gene lạ được gọi là *sinh vật chuyển gene*. Các sinh vật chuyển gene có thể được sử dụng cho nghiên cứu cơ bản hoặc cho các ứng dụng thương mại chuyên dụng. Một ứng dụng đã tạo ra các sản phẩm gene người có giá trị như insulin trong vi khuẩn chuyển gene mang gene người thích hợp.

Từ tổng quan ngắn này, chúng ta xem việc phân lập gene đã trở thành một công cụ không thể thiếu được trong công nghệ sinh học hiện đại.

Một số ví dụ thú vị về các gene được phân lập là gì ? Câu trả lời phụ thuộc vào rất nhiều quá trình sinh học được nghiên cứu. Hãy xem một số trường hợp. Một nhà di truyền về nấm nghiên cứu con đường tế bào tổng hợp tryptophan có thể quan tâm tới các gene mà khi bị đột biến, tạo ra nhu cầu dinh dưỡng thụ động đối với tryptophan, do mỗi gene có thể đại diện một bước trên con đường tổng hợp. Những gene này có thể được nhận dạng qua phân tích đột biến, phân ly và lập bản đồ. Chúng có thể được được đặt tên *trp1*, *trp2*, *trp3*.... Nhà di truyền này rất quan tâm tới việc phân lập và xác định tính chất một hoặc nhiều gene như thế.

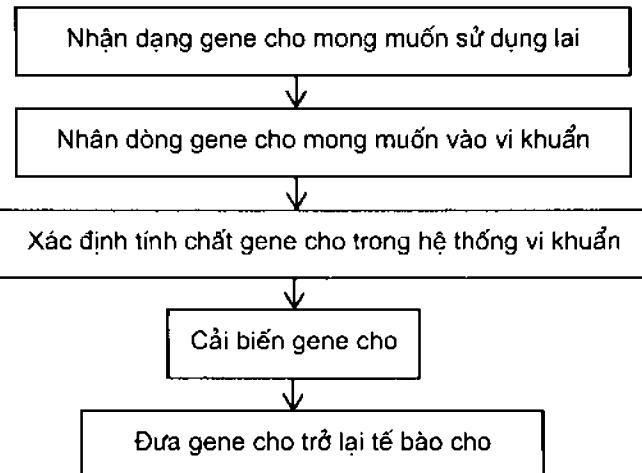
Cũng như vậy, các gene người có các allele đột biến tạo kiểu hình rối loạn chức năng rất bổ ích cho xác định các nguyên nhân sinh học và y học. Chúng ta đã thấy những gene này được nhận dạng bằng phân tích phả hệ. Chúng bạch tạng và chứng nước tiểu chứa alcapton (alkaptonuria) là hai ví dụ về nhiễm sắc thể diễn hình lặn. Trong những trường hợp như thế này, bản chất chung của sự thiếu hụt đã được hiểu là thiếu hụt enzyme nhưng nó sẽ rất ích lợi để phân lập chính các gene đó. Các gene khác ở người được biết từ phân tích phả hệ, nhưng không có chức năng hoá sinh nào của chúng được biết đến. Phân lập các gene như thế sẽ là bổ ích đặc biệt bởi vì việc đánh giá cấu trúc gene có thể dẫn tới xác định chức năng gene và bản chất bệnh tật. Một ví dụ là bệnh xơ nang, một bệnh được biết từ phân tích phả hệ gây ra bởi một allele lặn trên nhiễm sắc thể diễn hình của một gene mà không biết chức năng cho tới khi gene được phân lập và đọc trình tự. Các trường hợp như thế này có thể được đưa ra trong tất cả các sinh vật được sử dụng trong nghiên cứu di truyền.

Trong việc xem xét tách chiết gene, trước hết chúng ta cần phải kiểm tra bản chất của DNA tái tổ hợp và nguyên lý mà kỹ thuật DNA tái tổ hợp có thể được sử dụng để tách chiết gene. Tiếp theo, chúng ta cần kiểm tra các phương pháp phân lập gene đặc trưng như những gene đã thảo luận.

5.2. TẠO DNA TÁI TỔ HỢP

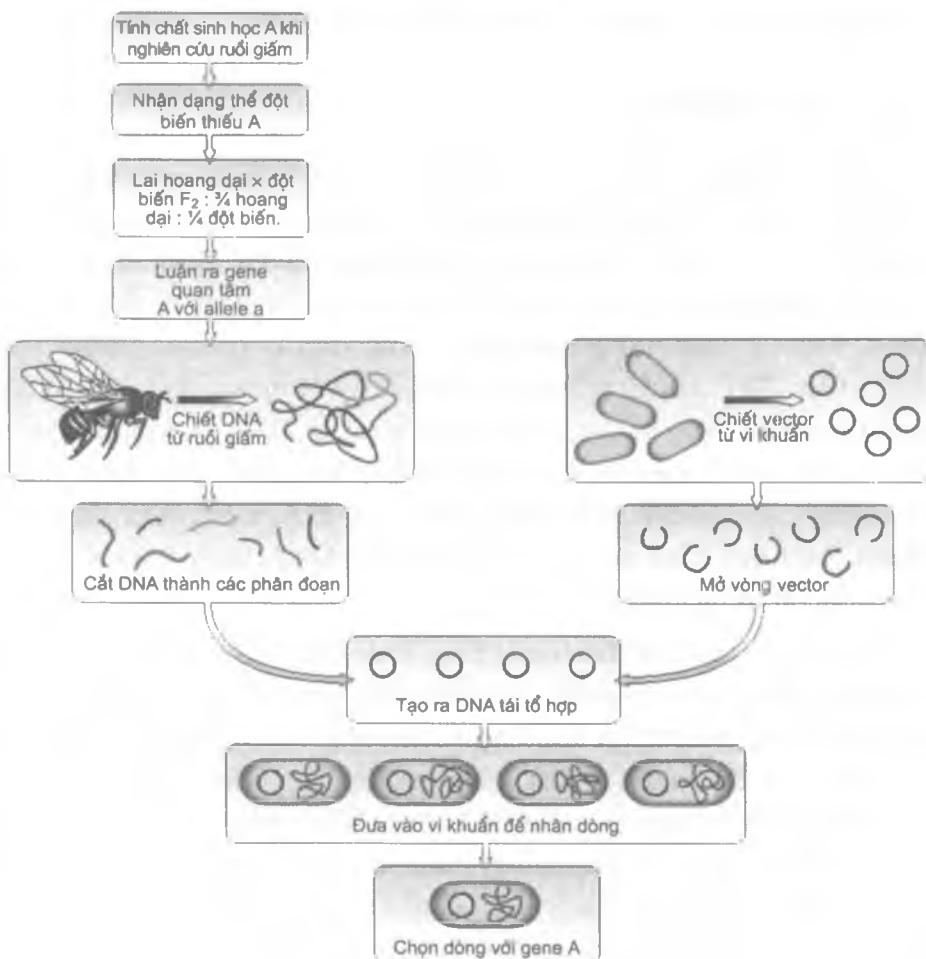
Kỹ thuật DNA tái tổ hợp xảy ra như thế nào ? Sinh vật được sử dụng để cho DNA phân tích, được gọi là *sinh vật cho (donor organism)*. Quy trình cơ bản bao gồm tách chiết và cắt dán DNA từ một genome cho

thành các phân đoạn chứa một hoặc vài gene và cho phép các phân đoạn này được chèn vào các phân tử DNA nhỏ mờ vòng có khả năng sao chép tự động như các plasmid vi khuẩn. Các phân tử mạch vòng nhỏ này có nhiệm vụ như thể dẫn truyền hay *vector*, cho những phân đoạn DNA. Các phân tử vector với các phân đoạn chèn của chúng được gọi là DNA tái tổ hợp do chúng hợp thành từ các tổ hợp mới của DNA từ genome cho (có thể từ mọi sinh vật) với vector DNA từ một nguồn hoàn toàn khác (thường là một plasmid vi khuẩn hoặc một virus). Hỗn hợp DNA tái tổ hợp sau đó được sử dụng để biến nạp vào tế bào vi khuẩn, và nói chung là để các phân tử vector tái tổ hợp đơn lẻ tìm đường chui vào những tế bào vi khuẩn đơn lẻ. Các tế bào vi khuẩn được trang lên đĩa thạch và cho phép phát triển thành các khuẩn lạc. Một tế bào biến nạp đơn với một vector tái tổ hợp đơn sẽ phân chia thành một khuẩn lạc với hàng triệu tế bào, tất cả đều mang cùng một vector tái tổ hợp. Bởi vậy, một khuẩn lạc đơn chứa một quần thể lớn những phân đoạn chèn DNA giống nhau và quần thể này được gọi là một dòng DNA (DNA clone). Một công việc lớn phân tích phân đoạn DNA nhân dòng có thể được thực hiện ở bước khi nó ở trong tế bào vi khuẩn chủ. Tuy nhiên, sau đó người ta thường mong muốn đưa DNA nhân dòng trở lại tế bào của sinh vật cho ban đầu để mang những cải biến đặc biệt của cấu trúc và chức năng genome. Bởi vậy, quy trình thường như sau :



Thông điệp : Nhân dòng cho phép khuếch đại và thu hồi một phân đoạn DNA đặc trưng từ một mẫu DNA lớn và phức tạp như một genome.

Vì DNA cho được cắt thành nhiều phân đoạn khác nhau, hầu hết các khuẩn lạc sẽ mang một DNA tái tổ hợp khác nhau (nghĩa là, một phân đoạn chèn nhân dòng khác nhau). Bởi vậy, bước tiếp theo là tìm cách sàng lọc dòng với phân đoạn chèn chứa gene đặc trưng mà chúng ta quan tâm. Khi nhận được dòng này, DNA được tinh sạch với lượng lớn và gene nhân dòng quan tâm có thể được sử dụng cho nhiều phân tích, chúng ta sẽ xem xét ở những nội dung sau của chương này. Phương pháp nhân dòng hoạt động được do các phân tử DNA tái tổ hợp đơn lẻ đi vào các tế bào chủ vi khuẩn đơn lẻ và sau đó, những tế bào này thực hiện công việc khuếch đại các phân tử đơn lẻ thành một lượng phân tử lớn mà có thể được xử lý như hóa chất. Hình 5.1 mô tả sơ đồ chung của quy trình này.



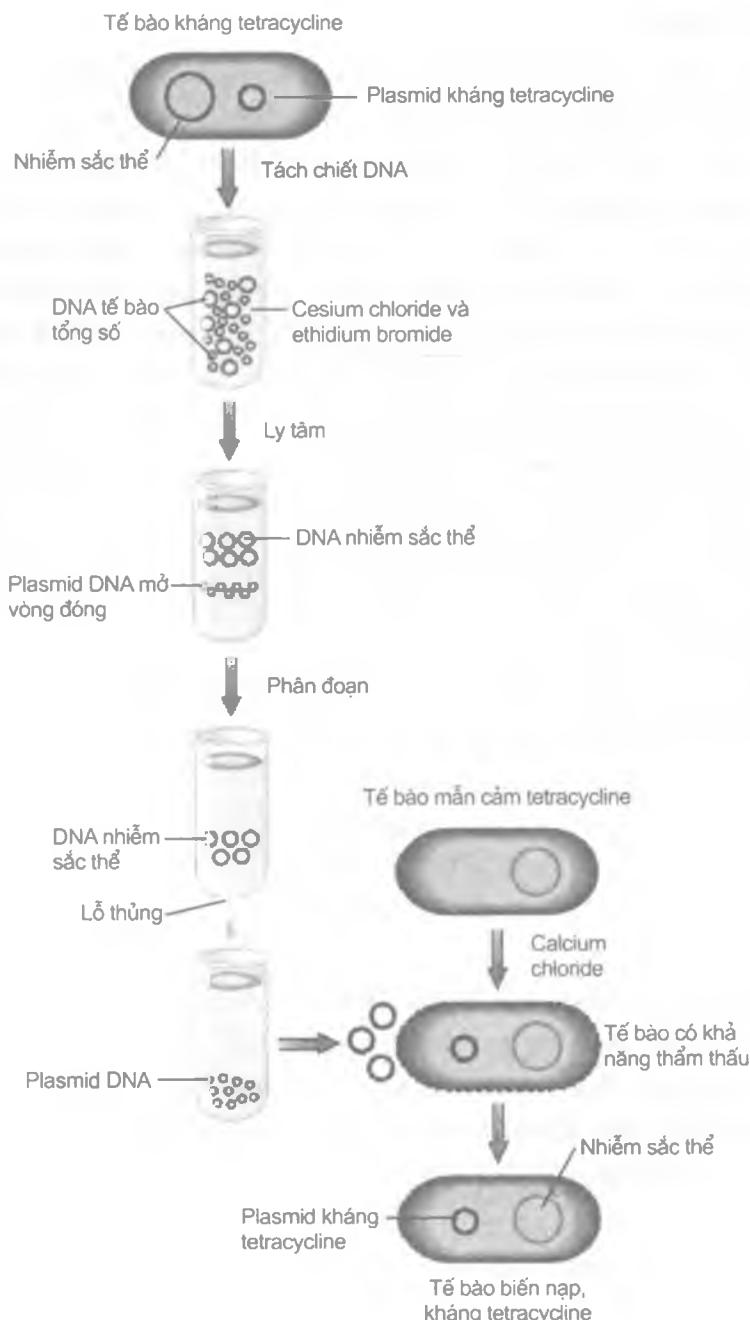
Hình 5.1. Kỹ thuật DNA tái tổ hợp tạo khả năng cho các phân đoạn DNA đơn lẻ từ mọi genome có thể được chèn vào các phân tử DNA vector như plasmid và khuếch đại trong vi khuẩn. Mỗi phân đoạn khuếch đại được gọi là một dòng (DNA clone).

Khái niệm DNA tái tổ hợp cần được phân biệt với tái tổ hợp DNA tự nhiên là kết quả từ lai giống giữa các nhiễm sắc thể tương đồng ở cả eukaryote và prokaryote. DNA tái tổ hợp theo nghĩa được sử dụng trong chương này là một tổ hợp không tự nhiên của các DNA từ các nguồn không tương đồng, thường từ các sinh vật khác nhau. Một số nhà di truyền thích sử dụng khái niệm khác là DNA quái dị (chimeric DNA), theo tên quý thần thoại Hy Lạp Chimera. Qua nhiều thế kỷ, Chimera như là biểu tượng của một thể hợp nhất sinh học không thể có, một tổ hợp của nhiều bộ phận từ các động vật khác nhau. Cũng như vậy, DNA tái tổ hợp là một DNA chimera và sẽ không thể có nếu không có thao tác thực nghiệm mà chúng ta gọi là kỹ thuật DNA tái tổ hợp.

5.2.1. Tách chiết DNA

Bước đầu tiên trong việc tạo ra DNA tái tổ hợp là tách chiết DNA cho và DNA vector. Các quy trình thông dụng để tách chiết DNA có từ mấy thập kỷ nay, trước khi kỹ thuật DNA tái tổ hợp ra đời. Bằng cách sử dụng những phương pháp như thế, mớ DNA được tách chiết từ thể cho sê là DNA nhân tế bào trong eukaryote hoặc DNA genome chính trong prokaryote. Nói chung, những loại DNA này là những loại cần để phân tích. Quy trình được sử dụng để thu nhận vector DNA phụ thuộc vào bản chất của vector. Plasmid ở vi khuẩn là những vector sử dụng thông thường và những plasmid này phải được tinh sạch khỏi DNA nhân tế bào vi khuẩn. Một quy trình để tách chiết plasmid DNA bằng ly tâm siêu tốc được tóm tắt trong hình 5.2.

Plasmid DNA hình thành một băng khác sau khi ly tâm siêu tốc trong thang mật độ cesium chloride chứa ethidium bromide. Băng plasmid được thu nhận bằng cách châm kim ống ly tâm nhựa. Một quy trình khác dựa trên quan sát thấy rằng ở một pH kiểm đặc biệt, DNA nhân tế bào vi khuẩn biến tính nhưng plasmid thì không. Bước trung hòa tiếp theo gây kết tủa DNA nhân tế bào, nhưng plasmid vẫn ở trong dung dịch. Các phage như λ cũng có thể được sử dụng làm vector để nhân dòng DNA vào các hệ thống vi khuẩn. Phage DNA được tinh sạch từ dịch huyền phù sạch của phage thu nhận từ dịch ly giải phage.

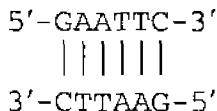


Hình 5.2. Plasmid mang gene kháng sinh tetracycline (trên cùng bên phải) có thể được tách từ DNA nhiễm sắc thể vi khuẩn. Do liên kết khác nhau với ethidium bromide bởi hai loại DNA tạo cho plasmid DNA vòng đặc hơn DNA nhiễm sắc thể, các plasmid tạo ra một băng khác khi ly tâm trong thang mật độ cesium chloride và có thể được tách ra (đáy ống bên trái). Sau đó, chúng có thể được đưa vào tế bào vi khuẩn bằng biến nạp (bên phải).

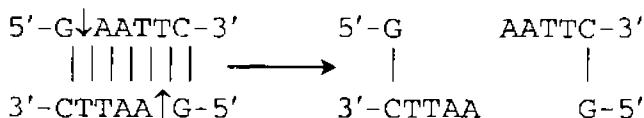
5.2.2. Cắt DNA

Điểm chọc thủng phòng tuyến tạo ra khả năng cho kỹ thuật DNA tái tổ hợp là sự phát hiện và xác định tính chất của các *enzyme hạn chế*. Các enzyme hạn chế được sinh tổng hợp bởi vi khuẩn là một cơ chế bảo vệ chống lại phage. Các enzyme hoạt động như những cái kéo, cắt đứt DNA của phage và bằng cách ấy bắt hoạt nó. Quan trọng là các enzyme hạn chế không cắt ngẫu nhiên ; chính xác hơn, chúng cắt ở những trình tự đích DNA đặc trưng, đây là một trong những tính chất then chốt tạo cho chúng thích hợp cho thao tác DNA. Mọi phân tử DNA, từ virus cho tới con người, đều chứa các vị trí đích enzyme hạn chế hoàn toàn ngẫu nhiên và bởi vậy có thể được cắt thành các phân đoạn có kích thước nhất định thích hợp để nhân dòng. Các vị trí hạn chế không liên quan tới chức năng của sinh vật và chúng cũng không bị cắt nội trong cơ thể (*in vivo*), bởi vì hầu hết các sinh vật không có các enzyme hạn chế.

Hãy xem xét một ví dụ về enzyme hạn chế *EcoRI* (từ *E.coli*) nhận biết trình tự sáu cặp nucleotide sau trong DNA của mọi sinh vật :

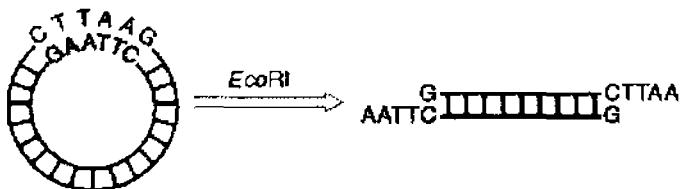


Kiểu đoạn này được gọi là một DNA *palindrome*, có nghĩa là cả hai sợi có cùng một trình tự nucleotide nhưng theo chiều đối song. Nhiều enzyme hạn chế khác nhau nhưng cùng nhận biết và cùng cắt một palindrome đặc trưng. Enzyme *EcoRI* cắt nội trong trình tự này nhưng cắt một cặp so le giữa nucleotide G và A :



Cắt so le như vậy đã để lại một cặp "đầu dính" giống nhau. Các đầu này được gọi là dính do chúng có thể liên kết hydrogen (dính) với một

trình tự bổ trợ. Hình 5.3 cho thấy *Eco*RI tạo ra nhát cắt đơn trong một phân tử DNA vòng như plasmid : nhát cắt mở vòng và phân tử mạch thẳng được tạo ra có hai đầu đính. Sinh ra hai đầu đính này là một tính chất khác của enzyme hạn chế mà tạo cho chúng thích hợp với kỹ thuật DNA tái tổ hợp. Nguyên lý đơn giản là nếu hai phân tử DNA khác nhau được cắt với cùng một enzyme hạn chế, cả hai sẽ tạo ra các phân đoạn với cùng đầu đính bổ trợ, tạo cho chúng khả năng hình thành nên DNA chimera. Bởi vậy, nếu cả hai DNA vector và DNA cho được cắt cùng với *Eco*RI, các đầu đính của vector có thể liên kết với các đầu đính của phân đoạn cho khi cả hai được trộn với nhau.



Hình 5.3. Enzyme hạn chế EcoRI cắt một phân tử DNA vòng mang một trình tự đích, cho kết quả một phân tử mạch thẳng với các đầu đính sợi đơn.

Enzyme hạn chế có hai tính chất tiện ích trong kỹ thuật DNA tái tổ hợp. Thứ nhất, chúng cắt DNA thành các phân đoạn có kích thước thích hợp để nhân dòng. Thứ hai, nhiều enzyme hạn chế tạo ra phân đoạn cắt so le sinh ra các đầu đính sợi đơn dẫn đến sự hình thành DNA tái tổ hợp.

Hàng trăm enzyme hạn chế với tính đặc hiệu trình tự khác nhau đã được nhận dạng, một số được trình bày trong bảng 5.1. Cần nhớ là tất cả các trình tự đích đều phải là palindrome, nhưng giống như EcoRI, một số enzyme tạo ra nhát cắt so le, trong khi đó những enzyme khác tạo ra nhát cắt bằng. Ngay cả các nhát cắt bằng, thiếu đầu so le, có thể được sử dụng tạo DNA tái tổ hợp.

DNA cũng có thể được cắt bằng lực cơ học. Ví dụ, khuấy DNA trong một máy pha trộn sẽ bẻ gãy các phân tử kích thước nhiễm sắc thể thành những phân đoạn đầu bằng có thể nhân dòng.

Bảng 5.1. Vị trí hạn chế, cắt và cải biến của các enzyme hạn chế khác nhau.

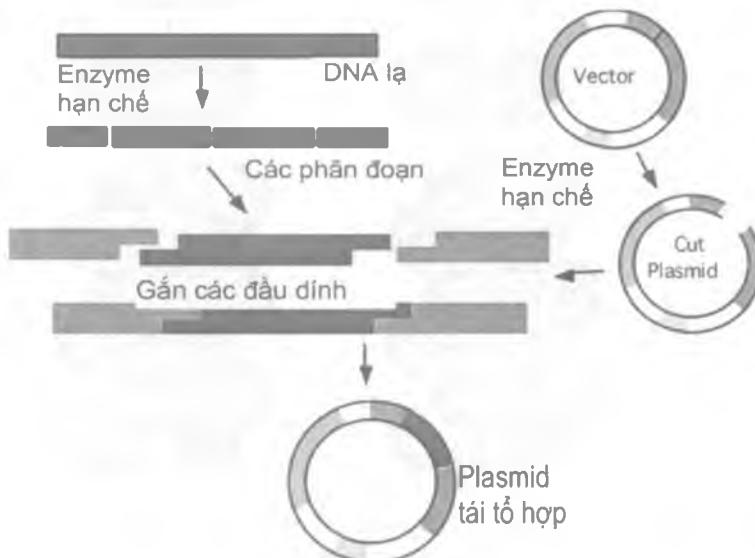
Enzyme	Nguồn	Vùng hạn chế và cắt trên sợi kép DNA
EcoRI	<i>Escherichia coli</i>	↓m 5'-G-A-A-T-T-C-3' 3'-C-T-T-A-A-G-5' m↑
EcoRI	<i>Escherichia coli</i>	↓m 5'-G-C-C-T-G-G-C-3' 3'-C-G-G-A-C-C-G-5' m↑
EcoRI	<i>Escherichia coli</i>	↓m 5'-G-A-A-T-T-C-3' 3'-C-T-T-A-A-G-5' m↑
EcoRI	<i>Escherichia coli</i>	m↓ 5'-A-A-G-C-T-T-3' 3'-T-T-C-G-A-A-5' m↑
EcoRI	<i>H. egyptius</i>	↓ 5'-G-G-C-C-3' 3'-C-C-G-5' ↑
HpaII	<i>H. parainfluenzae</i>	↓ 5'-C-C-G-G-3' 3'-G-G-C-C-5' ↑
EcoRI	<i>Escherichia coli</i>	↓ 5'-C-T-G-C-A-G-3' 3'-G-A-C-G-T-C-5' ↑
EcoRI	<i>Escherichia coli</i>	↓ 5'-C-C-G-G-3' 3'-G-G-C-C-5' ↑
EcoRI	<i>Escherichia coli</i>	↓ 5'-C-T-G-C-A-G-3' 3'-G-A-C-G-T-C-5' ↑
EcoRI	<i>Escherichia coli</i>	↓m 5'-G-A-A-T-T-C-3' 3'-C-T-T-A-A-G-5' m↑

m : có nghĩa là base được methyl hoá. Thông thường, người ta ký hiệu là * để chỉ sự methyl hoá nhưng ở đây lại ký hiệu là m để tránh nhầm sang đánh dấu phóng xạ.

5.2.3. Gắn dính DNA

Thông thường nhất, cả hai DNA cho và DNA vector được cắt với cùng một enzyme hạn chế, tạo ra các đầu dính và sau đó trộn vào một ống nghiệm để các đầu dính của DNA cho và DNA vector liên kết với nhau, hình thành các phân tử tái tổ hợp. Hình 5.4 biểu diễn một vector plasmid mang một vị trí hạn chế *Eco*RI đơn, như thế việc cắt với enzyme hạn chế *Eco*RI chuyển đổi DNA mạch vòng thành phân tử mạch thẳng với các đầu dính. DNA cho từ mọi nguồn (ví dụ *Drosophila*) cũng được xử lý với enzyme *Eco*RI để tạo ra một tập hợp các phân đoạn mang cùng các đầu dính như nhau. Khi hai tập hợp được trộn với nhau, các phân đoạn DNA từ hai nguồn có thể hợp nhất lại, do các cấu trúc xoắn helix đôi hình thành giữa các đầu dính của chúng.

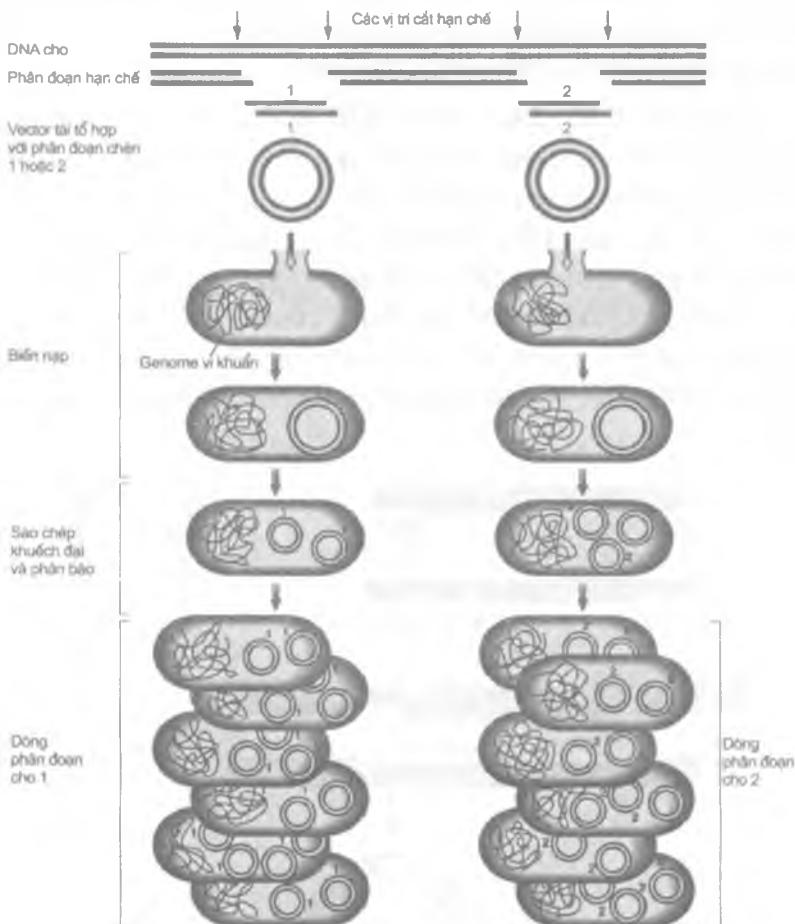
Có nhiều phân tử vector mở vòng trong dung dịch và nhiều phân đoạn *Eco*RI khác nhau của DNA cho. Bởi vậy, một tập hợp khác nhau của các vector mang những phân đoạn chèn cho khác nhau sẽ được tạo ra. Ở bước này, mặc dù các đầu dính đã hợp nhất với nhau tạo ra một tập hợp các phân tử chimera, khung đường-phosphate vẫn chưa được hoàn thiện ở hai vị trí mỗi đầu nối. Tuy nhiên, các khung có thể được gắn lại bằng cách bổ sung enzyme *DNA ligase*, nó tạo ra liên kết phosphodiester ở điểm nối. Các ligase nhất định có khả năng gắn dính các phân đoạn DNA với đầu cắt bằng.



Hình 5.4. Phương pháp tạo ra plasmid DNA quái chứa các gene dẫn xuất từ DNA lỏng.

5.2.4. Khuếch đại DNA tái tổ hợp

DNA tái tổ hợp đã gắn dính chui vào một tế bào vi khuẩn bằng biến nạp. Sau khi nó đã ở trong tế bào chủ, vector plasmid có thể sao chép do các plasmid thường có vùng khởi động sao chép. Tuy nhiên, bây giờ phân đoạn chèn DNA cho là một phần của nó, DNA cho sẽ được sao chép tự động cùng với vector. Mỗi plasmid tái tổ hợp đi vào một tế bào sẽ tạo ra đa phiên bản của chính nó trong tế bào này. Tiếp theo, nhiều chu kỳ phân chia tế bào sẽ xảy ra và các vector tái tổ hợp sẽ theo nhiều vòng sao chép. Khuẩn lạc vi khuẩn hình thành sẽ chứa hàng tỷ bản sao phân đoạn chèn DNA cho đơn lẻ. Bộ bản sao khuếch đại của phân đoạn DNA đơn lẻ là dòng DNA (hình 5.5).



Hình 5.5. Khuếch đại in vivo hoạt động như thế nào ? Xử lý enzyme hạn chế DNA cho và vector nhận cho phép chèn các phân đoạn đơn vào vector. Một vector đơn sẽ chui vào một tế bào vi khuẩn, ở đó quá trình sao chép và phân bào cho kết quả một số lượng bản sao lớn phân đoạn cho.

5.3. NHÂN DÒNG MỘT GENE ĐẶC HIỆU

Những mô tả trên đây là các phương pháp tiếp cận chung để tạo ra DNA tái tổ hợp. Tuy nhiên, một nhà di truyền học quan tâm tới phân lập và đánh giá một gene mong muốn đặc trưng nào đó, như vậy các quy trình phải được đáp ứng để phân lập một dòng DNA tái tổ hợp chứa gene đặc trưng. Những chi tiết của quy trình khác nhau từ sinh vật này tới sinh vật khác và từ gene này tới gene khác. Một yếu tố quan trọng đầu tiên là sự lựa chọn một vector thích hợp cho công việc này.

5.3.1. Lựa chọn vector nhân dòng

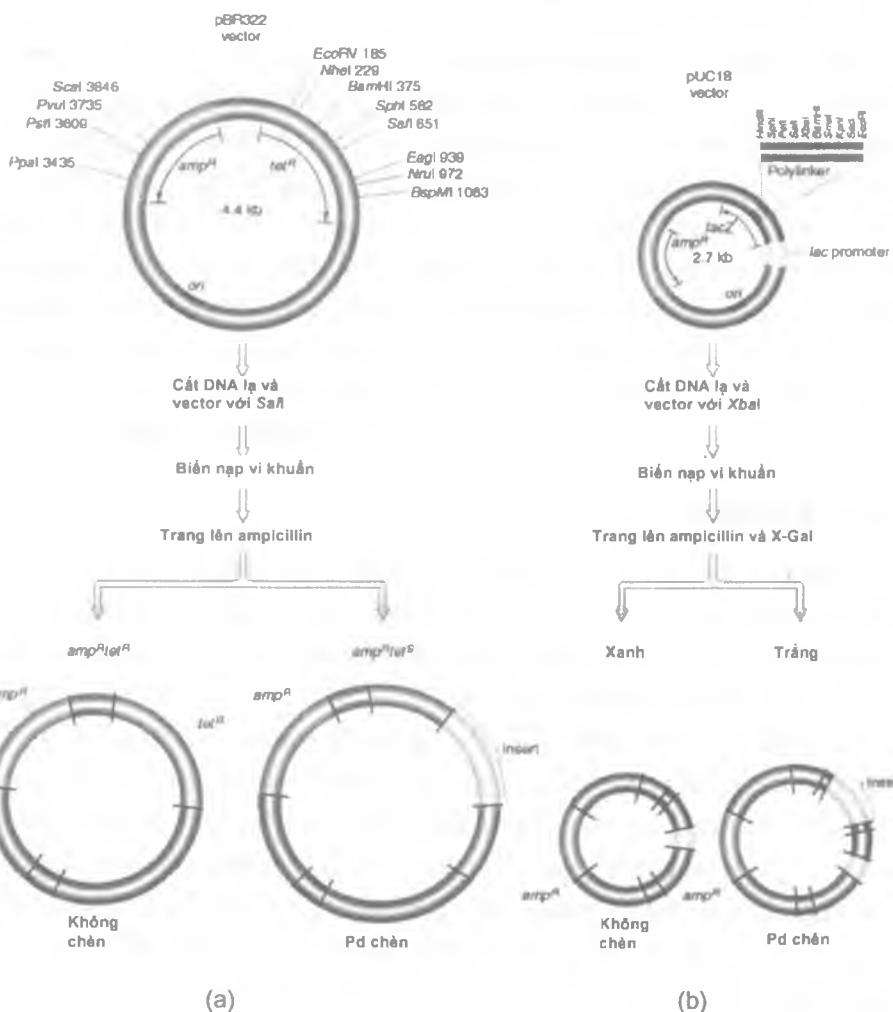
Vector phải là những phân tử tương đối nhỏ để dễ dàng thao tác. Chúng phải có khả năng sao chép nhiều trong một tế bào sống, bằng cách đó tạo ra khả năng khuếch đại phân đoạn cho đã chèn vào. Một yêu cầu quan trọng khác là các vị trí hạn chế thông dụng có thể được sử dụng để chèn DNA cần nhân dòng. Các vị trí duy nhất là tốt nhất bởi vì sau đó, phân đoạn chèn được đưa vào chỉ một vị trí trên vector. Điều cũng quan trọng là có cơ chế nhận dạng dễ dàng và thu nhận phân tử tái tổ hợp. Có một số vector nhân dòng đang sử dụng hiện nay và sự lựa chọn giữa chúng thường phụ thuộc vào kích thước phân đoạn DNA cần được nhân dòng. Chúng ta sẽ xem xét một vài dạng sử dụng thông thường.

5.3.1.1. Plasmid

Như mô tả ở trên, các plasmid vi khuẩn là những phân tử DNA nhỏ mảnh vòng khác với nhiễm sắc thể chính của vi khuẩn. Chúng sao chép DNA của chúng không phụ thuộc vào nhiễm sắc thể vi khuẩn. Nhiều dạng plasmid khác nhau được tìm thấy ở vi khuẩn. Sự phân bố của plasmid bất kỳ trong một loài nói chung là rải rác ; một số tế bào có plasmid, trong khi đó các tế bào khác không có. Chúng ta đã gặp F plasmid tạo ra các dạng nhất định của quan hệ tiếp hợp cho tế bào *E. coli*. F plasmid có thể được sử dụng làm một vector để mang các phân đoạn chèn DNA cho số lượng lớn. Tuy nhiên, các plasmid được sử dụng hàng ngày làm vector là những plasmid mang các gen kháng thuốc. Các gen kháng thuốc rất tiện ích do kiểu hình kháng thuốc có thể được sử dụng để chọn lọc không chỉ tế bào được biến nạp bởi plasmid, mà còn cho các vector chứa DNA tái tổ hợp. Plasmid cũng là một phương tiện

hiệu quả để khuếch đại DNA nhân dòng bởi vì có nhiều bản sao trong mỗi tế bào (có thể là vài trăm cho một số plasmid).

Hai plasmid được thiết kế làm các vector nhân dòng DNA, cho thấy cấu trúc chung và các vị trí hạn chế. Việc chèn vào pBR322 được phát hiện bởi sự bất hoạt của một gene kháng thuốc (*tetR*), được chỉ ra bởi kiểu hình *tetS* (mẩn cảm). Chèn vào pUC18 được phát hiện bởi sự bất hoạt của chức năng β -galactosidase của *Z'*, cho kết quả mất khả năng chuyển hóa cơ chất nhân tạo X-Gal thành màu xanh.



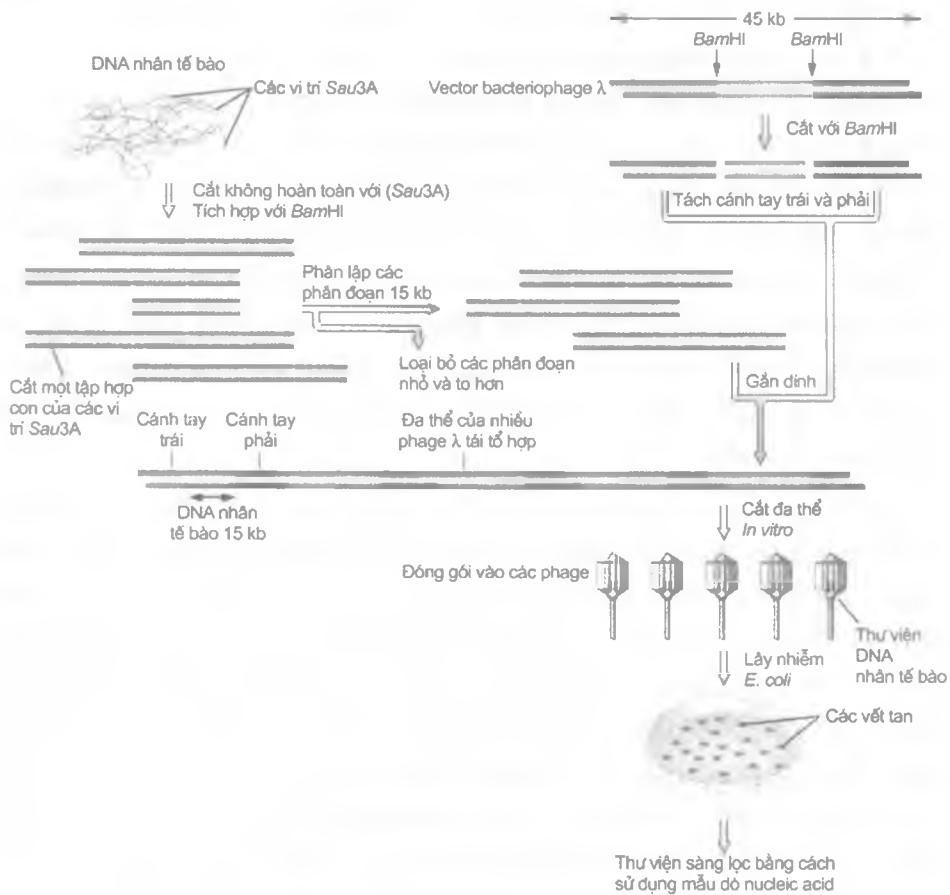
Hình 5.6. Cấu trúc hai vector nhân dòng pBR322(a) và pUC18(b).

Hai vector plasmid sử dụng nhiều nhất trong di truyền được biểu thị trên hình 5.6. Các vector này có nguồn gốc từ các plasmid tự nhiên nhưng cả hai đã được cải biến di truyền để dễ dàng sử dụng làm các vector DNA tái tổ hợp. Plasmid pBR322 có cấu trúc đơn giản, nó có hai gene kháng thuốc *ter^R* và *amp^R*. Cả hai gene đều chứa các vị trí đích hạn chế duy nhất mà tiện lợi cho nhân dòng. Ví dụ, DNA cho có thể được chèn vào gene *ter^R*. Một sự chèn thành công sẽ phân chia và bất hoạt gene *ter^R*, sau đó nó sẽ không còn khả năng kháng tetracycline nữa và tế bào sẽ mẫn cảm với kháng sinh. Bởi vậy, quy trình nhân dòng sẽ trộn các mẫu của plasmid và DNA cho đã cắt, biến nạp vi khuẩn và trước tiên sàng lọc các khuẩn lạc kháng ampicillin, chúng phải được biến nạp thành công bởi phân tử plasmid. Trong số các khuẩn lạc Amp^R, chỉ những khuẩn lạc chứng tỏ mẫn cảm tetracycline mới mang phân đoạn chèn, nói cách khác, các khuẩn lạc Amp^R Tet^S là những khuẩn lạc chứa DNA tái tổ hợp. Các thí nghiệm tiếp theo là cần thiết để tìm ra các khuẩn lạc với phân đoạn chèn đặc trưng cần đến.

pUC plasmid là một vector nhiều ưu thế, cấu trúc của nó cho phép chọn lọc quan sát các khuẩn lạc chứa vector với phân đoạn chèn DNA cho. Yếu tố then chốt là một phần nhỏ của gene β-galactosidase chủng *E. coli*. Vùng này được chèn một đoạn DNA được gọi là polylinker, vùng nhân dòng đa vị, chứa các vị trí đích hạn chế duy nhất tiện ích cho việc chèn các phân đoạn cho. Polylinker nằm trong khung dịch mã với phân đoạn β-galactosidase và không ảnh hưởng tới dịch mã. Quy trình biến nạp sử dụng các tế bào nhận chứa gene β-galactosidase thiếu mất phân đoạn có mặt trên plasmid. Một dạng bổ trợ thông thường xảy ra, ở đó, các tiểu phân protein được mã hoá bởi hai phân đoạn, hợp nhất với nhau tạo ra β-galactosidase chức năng. Một cơ chất không màu cho β-galactosidase được gọi là X-Gal được bổ sung vào môi trường và enzyme chức năng chuyển hoá cơ chất này thành thuốc màu xanh, nhuộm xanh khuẩn lạc. Nếu DNA cho được chèn vào polylinker, phân đoạn enzyme được mang trên vector bị gián đoạn, không có protein hoàn chỉnh β-galactosidase nào được hình thành và khuẩn lạc trở nên trắng. Do đó, chọn lọc các khuẩn lạc Amp^R trắng sẽ chọn trực tiếp vector mang phân đoạn chèn, và các khuẩn lạc như thế được phân lập để nghiên cứu tiếp theo.

Các plasmid chứa các phân đoạn chèn lớn của DNA lạ có xu hướng mất đi phân đoạn chèn tự phát ; bởi vậy, các plasmid không tiện dụng để nhân dòng các phân đoạn DNA lớn hơn 20 kb.

5.3.1.2. Phage lambda



Hình 5.7. Nhân dòng trong phage λ.

Phage λ là một vector nhân dòng thông dụng do nhiều nghiên cứu. Trước hết, đầu phage λ sẽ đóng gói chọn lọc một nhiễm sắc thể dài khoảng 50 kb và tính chất này có thể được áp dụng để chọn lọc các phân tử λ với phân đoạn chèn DNA cho. Phần trung tâm của genome phage không cần cho sao chép hoặc đóng gói các phân tử λ DNA ở *E. coli*.

Như vậy, phân trung tâm có thể được cắt ra (bằng cách sử dụng các enzyme hạn chế) và bỏ đi. Hai "cánh tay" được gắn đính với DNA cho đã được cắt enzyme hạn chế. Phân tử quái chimera có thể được đưa vào *E.coli* trực tiếp bằng biến nạp hoặc đóng gói *in vitro* vào đầu phage. Trong các hệ thống *in vitro*, DNA và các thành phần đầu phage được trộn lại với nhau và các phage λ lây nhiễm hình thành tự phát. Trong phương pháp khác, các phân tử tái tổ hợp với các phân đoạn chèn 10 – 15 kb là những phân tử sẽ được đóng gói có hiệu quả vào các đầu phage, bởi vì kích thước phân đoạn chèn thay thế cho phân trung tâm đã bị cắt khỏi genome phage và mang lại một kích thước phân tử tổng số tới 50 kb. Bởi vậy, sự có mặt của vết tan phage trên thảm vi khuẩn báo hiệu tự động sự có mặt của phage tái tổ hợp mang một phân đoạn chèn (hình 5.7). Một tính chất tiện lợi thứ hai của một vector phage là các phân tử tái tổ hợp được đóng gói tự động vào các hạt phage lây nhiễm, chúng có thể được bảo quản thông thường và với thao tác thí nghiệm.

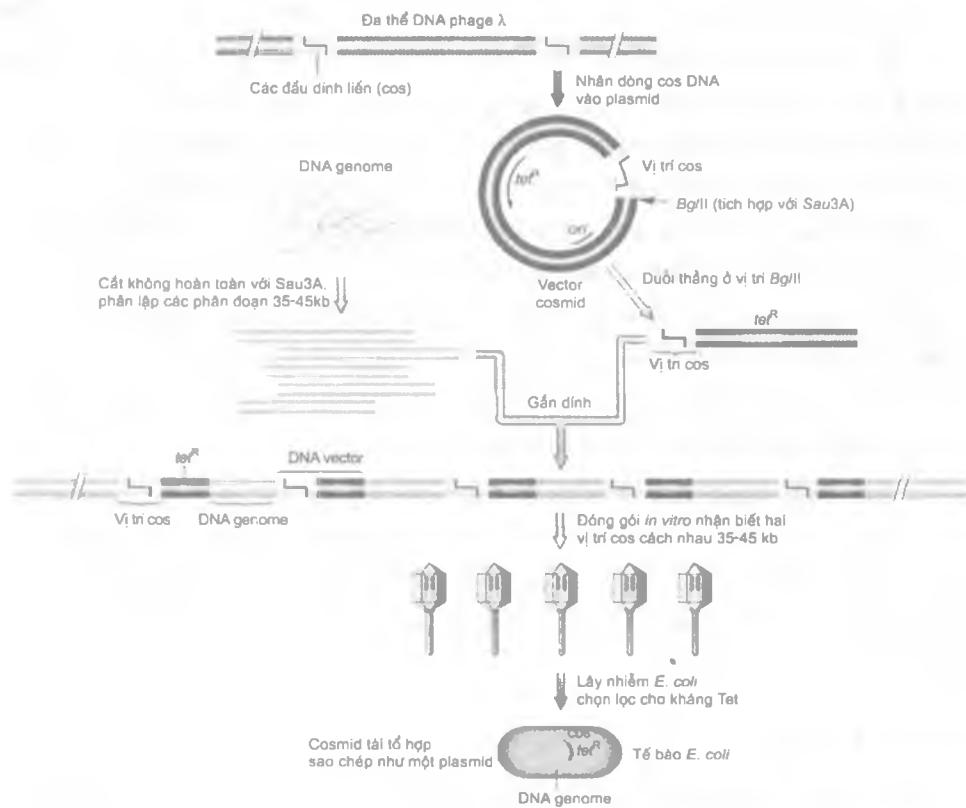
Một vùng trung tâm không cần thiết của nhiễm sắc thể phage được bỏ đi và các đầu được gắn đính với các phân đoạn 15 kb ngẫu nhiên của DNA cho. Một da thể mạch thẳng hình thành, sau đó nó được nhồi vào đầu phage mỗi lần một đơn phân bằng cách sử dụng một hệ thống đóng gói *in vitro*.

5.3.1.3. Cosmid

Cosmid là những vector, phân tử lai hybrid giữa phage λ và plasmid, mà DNA của nó có thể sao chép như trong tế bào của plasmid hoặc được đóng gói như của phage. Tuy nhiên, các cosmid có thể mang phân đoạn DNA lớn gấp khoảng ba lần những phân đoạn DNA do chính λ mang (lớn vào khoảng 45 kb). Chìa khóa là hầu hết cấu trúc phage λ bị cắt đi, nhưng các trình tự tín hiệu xúc tiến việc nhồi đầu phage (các vùng *cos*) vẫn giữ lại. Cấu trúc cải biến này tạo ra khả năng cho các đầu phage được nhồi với hầu hết tất cả DNA cho. Cosmid DNA có thể được đóng gói vào các hạt phage bằng cách sử dụng hệ thống *in vitro*. Nhân dòng bằng cosmid được trình bày trong hình 5.8.

Cosmid được cắt ở vị trí *Bgl*II gần điểm *cos*. DNA nhân tế bào cho được cắt bằng *Sau*3A, tạo ra các đầu đính tương thích với *Bgl*II. Một dãy

liên tiếp của DNA vector và DNA cho được tạo ra từ hỗn hợp. Phage được đóng gói *in vitro* bằng cách cắt tại vùng *cos*. Cosmid với phân đoạn chèn quay vòng sau khi nó đã ở tế bào vi khuẩn.



Hình 5.8. Nhân dòng bằng cosmid.

5.3.1.4. Các phage sợi đơn

Một số phage chỉ chứa các phân tử DNA sợi đơn. Khi xâm nhiễm vào vi khuẩn, sợi xâm nhiễm đơn được chuyển thành dạng sao chép sợi kép, mà có thể được phân lập và sử dụng cho nhân dòng. Ưu điểm của những phage này như các vector nhân dòng là DNA sợi đơn là cơ chất cần thiết cho kỹ thuật đọc trình tự DNA theo phương pháp Sanger đang được sử dụng rộng rãi hiện nay. Phage M13 là một trong những phage được sử dụng rộng rãi nhất cho mục đích này.

5.3.1.5. Các vector biểu hiện

Một cách để phát hiện gene nhân dòng đặc hiệu là phát hiện sản phẩm protein của nó được biểu hiện trong tế bào vi khuẩn. Bởi vậy, trong những trường hợp này, điều cần thiết là tạo ra khả năng biểu hiện gene trong vi khuẩn ; có nghĩa là phiên mã và dịch mã mRNA thành protein. Hầu hết các vector nhân dòng không cho phép biểu hiện các gene nhân dòng, nhưng có thể biểu hiện như vậy nếu sử dụng các vector đặc biệt. Tuy nhiên, do vi khuẩn không tạo ra các intron, các trình tự nhân dòng phải được tháo bỏ các intron. Gene nhân dòng được chèn ngay vào các tín hiệu khởi động phiên mã và dịch mã. Một số vector biểu hiện được thiết kế với các vị trí hạn chế nằm ngay vùng điều khiển *lac*. Các vị trí hạn chế này cho phép DNA lạ được cắt dán vào vector để biểu hiện dưới sự kiểm soát của hệ thống điều khiển *lac*.

5.3.2. Tạo thư viện DNA

Chúng ta đã biết mục đích quan trọng nhất của kỹ thuật DNA tái tổ hợp là nhân dòng một gene đặc biệt hoặc một phân đoạn genome quan tâm đối với nhà nghiên cứu. Cách tiếp cận được sử dụng để nhân dòng một gene đặc biệt phụ thuộc vào độ lớn của gene đang được nói đến và vào thông tin về nó. Nói chung, các quy trình khởi động với một mẫu DNA như DNA nhân tế bào eukaryote. Bước tiếp theo là nhận được một tập hợp lớn các khuẩn lạc được tạo ra từ mẫu DNA gốc. Tập hợp các khuẩn lạc được gọi là một *thư viện DNA*. Bước này đôi khi được gọi là nhân dòng "bắn súng" bởi vì người làm thí nghiệm nhân dòng một mẫu lớn các phân đoạn và hy vọng rằng một trong số các dòng sẽ chứa "đúng" gene mong muốn. Sau đó, nhiệm vụ là tìm ra dòng đặc biệt đó.

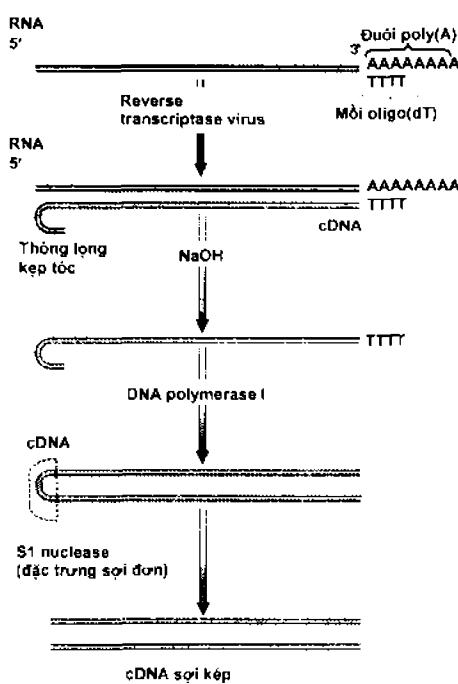
Có nhiều dạng thư viện khác nhau, được phân loại trước tiên là theo vector nào được sử dụng và thứ hai là theo nguồn DNA. Các vector nhân dòng khác nhau mang các hàm lượng DNA khác nhau, như vậy sự lựa chọn vector để xây dựng thư viện phụ thuộc vào kích thước genome (hoặc mẫu DNA khác) được làm thư viện. Vector plasmid và phage mang hàm lượng DNA nhỏ, như vậy những vector này là thích hợp cho các gene dịch từ các sinh vật với genome nhỏ. Các cosmid mang hàm lượng DNA lớn và những vector khác như YAC và BAC mang các hàm lượng lớn nhất trong số đó. Sự thuận tiện trong thao tác là một yếu tố quan trọng

khoa học trong việc chọn vector. Một thư viện phage là một dịch phage. Một thư viện plasmid hoặc cosmid là một dịch vi khuẩn hoặc một bộ dịch nuôi cấy vi khuẩn nhất định được bảo quản trong ống nghiệm nuôi cấy hoặc đĩa microtiter.

Quyết định quan trọng thứ hai là liệu có tạo ra một thư viện genome hay một thư viện *cDNA* không. *cDNA* hay DNA bổ trợ (viết tắt từ complementary DNA) là DNA tổng hợp được tạo ra từ mRNA có sử dụng một enzyme đặc biệt được gọi là *reverse transcriptase* phân lập gốc từ các retrovirus. Sử dụng mRNA làm khuôn, *reverse transcriptase* tổng hợp một phân tử DNA sợi đơn, sau đó được sử dụng làm khuôn để tổng hợp DNA sợi kép (hình 5.9). Được tạo ra từ mRNA, *cDNA* không có các trình tự điều hoà trên, dưới và các intron. Điều này có nghĩa là *cDNA* từ eukaryote có thể được dịch mã thành protein chức năng ở vi khuẩn, một tính chất quan trọng khi biểu hiện các gene eukaryote trong các tế bào chủ vi khuẩn.

Một chuỗi oligo(dT) ngắn được lai với đuôi poly(A) và một sợi mRNA. Phân đoạn oligo(dT) đóng vai trò làm môi cho *reverse transcriptase* hoạt động, sử dụng mRNA làm khuôn để tổng hợp sợi DNA bổ trợ. *cDNA* thu được kết thúc ở thông lọng kép tóc. Khi sợi mRNA được phân huỷ bằng xử lý với NaOH, thông lọng kép tóc trở thành môi cho DNA polymerase I, nó hoàn chỉnh sợi DNA cặp đôi. Thông lọng sau đó được S1 nuclease cắt di (nó hoạt động chỉ trên mộc sợi đơn) để tạo ra phân tử *cDNA* sợi kép.

Sự lựa chọn giữa DNA nhân tế bào và *cDNA* phụ thuộc vào hoàn cảnh. Nếu một gene đặc trưng hoạt động ở mô đặc trưng trong thực vật



Hình 5.9. Tổng hợp *cDNA* sợi kép từ mRNA.

hoặc động vật đang được tìm kiếm, thì có nghĩa là sử dụng mô đó vào việc tách chiết mRNA để chuyển thành cDNA và sau đó tạo ra một thư viện cDNA. Thư viện này cần phải chứa nhiều gene đang nói tới. Thư viện cDNA dựa trên các vùng của genome được phiên mã, không thể tránh khỏi, nhỏ hơn thư viện DNA đầy đủ, chứa toàn bộ genome. Tuy nhiên, thư viện genome lớn hơn, có lợi thế chứa các gene ở dạng nguyên vẹn, gồm cả các trình tự intron và điều hoà. Nếu mục đích xây dựng thư viện là hàng đầu để nhân dòng toàn bộ một genome, thì thư viện genome là cần thiết ở một phạm vi nào đó.

Trong một số trường hợp, có thể sử dụng một phần nhỏ genome để xây dựng thư viện và như vậy dễ dàng phát hiện gene mong muốn hơn. Có thể tiếp cận cách này nếu người làm thí nghiệm biết trước nhiễm sắc thể nào chứa gene. Một kỹ thuật được sử dụng trong di truyền phân tử động vật là phân loại nhiễm sắc thể bằng *máy đếm tế bào chạy qua* (flow cytometer). Một dịch nhiễm sắc thể đi qua một thiết bị, phân loại các nhiễm sắc thể theo kích thước. Phân nhiễm sắc thể thích hợp sau đó được sử dụng để tạo ra thư viện.

Một kỹ thuật khác có thể áp dụng cho các sinh vật có nhiễm sắc thể nhỏ là phân chia toàn bộ nhiễm sắc thể bằng cách sử dụng *diện di gel xung trường* (pulsed field gel electrophoresis, PFGE). Diện di là một kỹ thuật chung để phân loại các nucleic acid hoặc protein theo kích thước trên các gel dưới sự ảnh hưởng của một điện trường mạnh. Phương pháp này phân tách các phân đoạn DNA nhỏ hơn. PFGE là một dạng điện di chuyên biệt có lợi thế cho các phân tử DNA rất dài. Nó sử dụng một vài điện trường dao động định hướng theo một số hướng khác nhau. Điều này tạo khả năng cho các phân tử DNA lớn như toàn bộ nhiễm sắc thể để chui qua gel tới những vị trí khác nhau theo kích thước của chúng. Nhiễm sắc thể thích hợp có thể được nhận dạng trên gel bằng cách dò tìm với mẫu dò đặc hiệu nhiễm sắc thể. Sau đó, nhiễm sắc thể mong muốn có thể được cắt ra, thôi gel và sử dụng để tạo ra một thư viện chuyên biệt nhiễm sắc thể.

Người làm thí nghiệm có thể xác định xem có một thư viện nào đủ lớn, chứa mọi trình tự đơn quan tâm duy nhất với một mức độ chắc chắn hợp lý không? Có những công thức để tính số tối thiểu các dòng cần thiết

nhưng có một cách tính thô như sau : độ lớn của thư viện có thể được thu nhận đơn giản bằng cách lấy kích thước toàn bộ genome chia cho kích thước trung bình của các phân đoạn chèn vào vector sử dụng. Nói chung, số này ít nhất sẽ được nhân đôi nhưng nó cung cấp một sự ước lượng thô về độ lớn của việc xây dựng thư viện.

Nhiệm vụ phân lập một dòng gene đặc trưng bắt đầu với việc xây dựng một thư viện DNA genome hoặc cDNA nếu có thể, được làm giàu cho các trình tự chứa gene đang quan tâm.

5.3.3. Tìm các dòng đặc trưng sử dụng mẫu đò

Thư viện, có thể chứa hàng trăm đến hàng nghìn phân đoạn nhân dòng nên phải được sàng lọc để tìm ra phân tử DNA tái tổ hợp chứa gene quan tâm. Sự sàng lọc như thế được thực hiện bằng cách sử dụng một mẫu đò đặc biệt để dò tìm và đánh dấu dòng giúp cho nhà nghiên cứu nhận dạng. Nói rộng ra, có hai dạng mẫu đò : những mẫu nhận ra DNA và những mẫu nhận ra protein.

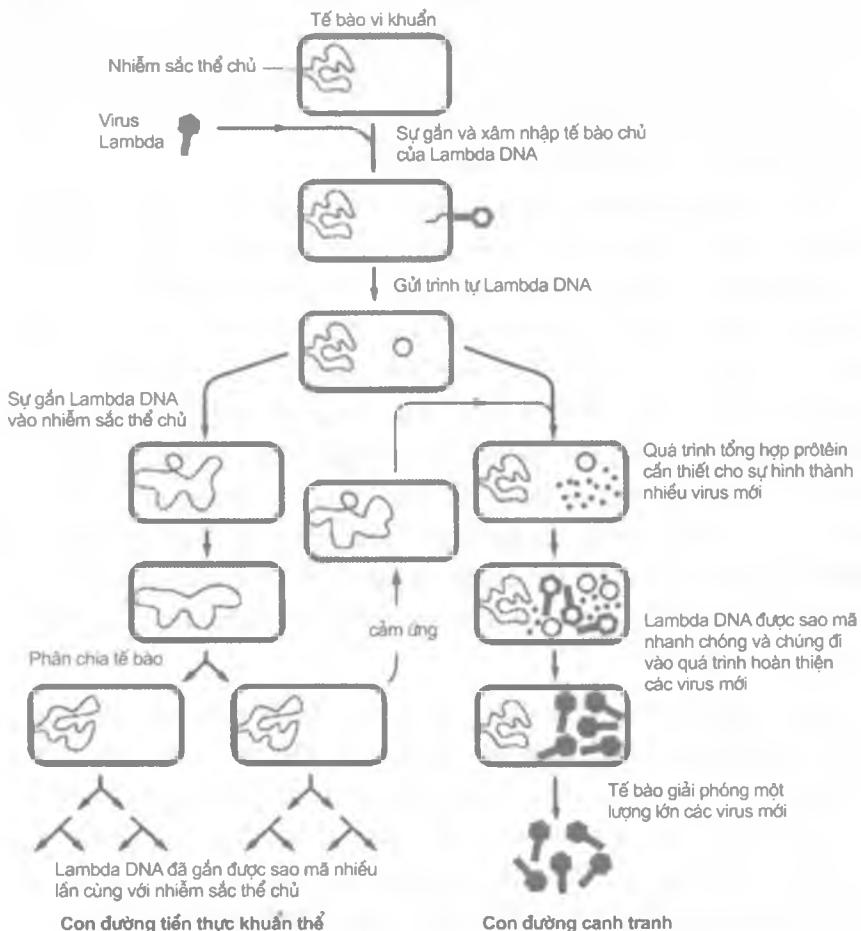
5.3.3.1. Mẫu đò để tìm DNA

Những mẫu đò này phụ thuộc vào xu hướng tự nhiên của một sợi đơn của nucleic acid để tìm và lai với một sợi đơn khác với trình tự base bổ trợ. Một mẫu đò mà chính là DNA, khi bị biến tính (tạo ra sợi đơn bằng cách tháo xoắn hai nửa của helix kép), bởi vậy sẽ tìm ra và liên kết với các DNA bị biến tính tương tự khác trong thư viện.

Mẫu đò	3'-AAGCCTATTATGGGCAAT-5'
Dòng	5'...AGTAGGGATCTCGGATAAAACCCGTTACGTACTTGGAAG...3'

Sự nhận dạng của một dòng đặc biệt trong một thư viện là một quy trình gồm hai bước. Trước tiên, các khuôn lắc hoặc vết tan của thư viện trên đĩa thạch được chuyển sang một màng hấp thụ (thường là nitrocellulose) đơn giản bằng cách đặt màng lên bề mặt môi trường.

Màng được tách ra và các khuẩn lạc hoặc các vết tan bám vào bề mặt được ly giải tại chỗ (*in situ*) và DNA bị biến tính. Bước tiếp theo là nhúng màng vào một dung dịch của mẫu dò đặc hiệu để tìm DNA. Mẫu dò được đánh dấu phóng xạ hay huỳnh quang. Nói chung, mẫu dò chính là một đoạn nhân dòng DNA có một trình tự tương đồng với gene mong muốn. Mẫu dò DNA được biến tính ; sau đó chỉ liên kết với DNA của dòng đang cần tìm. Vị trí của dòng dương tính trở nên trong từ vị trí của đánh dấu đậm, thường là một chấm trên phóng xạ tự ký.



Hình 5.10. Tạo thư viện genome

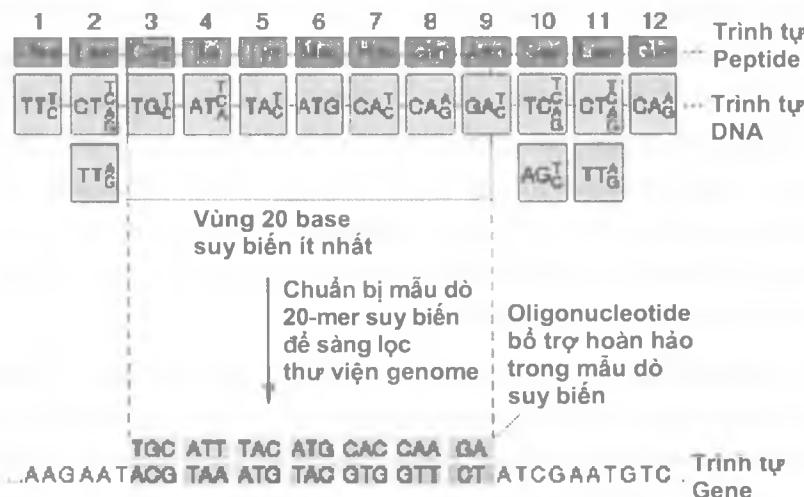
Hình 5.10 mô tả một thư viện genome có thể được tạo ra bằng nhân dòng gene trong bacteriophage λ . Khi một lớp mỏng vi khuẩn trên đĩa

thạch bị xâm nhiễm bởi một số lượng lớn các phage lai khác nhau, mỗi vết tan trên lớp mỏng vi khuẩn do một dòng phage đơn chiếm cứ có nguồn gốc từ phage xâm nhiễm ban đầu. Mỗi dòng mang một phân đoạn khác nhau của DNA tế bào. Nhận dạng dòng mang gene đặc biệt quan tâm bằng cách dò tìm các dòng với DNA hoặc RNA đã biết có họ hàng với gene mong muốn. Kiểu vết tan được chuyển sang giấy lọc nitrocellulose và protein phage được hòa tan, để lại DNA tái tổ hợp, sau đó nó bị biến tính sao cho nó sẽ gắn dính vào giấy lọc. Giấy lọc được ủ với một mẫu dò đánh dấu phóng xạ ; có nghĩa là, một bản sao DNA của RNA thông tin điện di cho gene mong muốn. Mẫu dò lai với mọi DNA tái tổ hợp gắn một trình tự DNA phù hợp và vị trí của dòng có DNA được bày ra bằng ký phóng xạ. Bây giờ, dòng mong muốn có thể được chọn lọc từ môi trường nuôi cấy và chuyển sang một tế bào chủ vi khuẩn mới, sao cho gene tinh sạch có thể được thao tác.

DNA để tạo ra mẫu dò có từ đâu ? DNA có thể từ một số nguồn. Một nguồn là cDNA từ mô biểu hiện gene quan tâm. Do mRNA của một gene có dư nên nhiều cDNA được tạo ra từ mô này và chèn từng cái vào vector sẽ có rất nhiều khả năng cho gene mong muốn. Ví dụ, trong buồng cầu lưỡi ở động vật có vú, 90% mRNA đã biết được phiên mã từ gene β -globin, như vậy các buồng cầu lưỡi sẽ là một nguồn mRNA tốt tạo ra mẫu dò cDNA để dò tìm gene globin ở genome. Trong trường hợp này, một thư viện genome sẽ được dò tìm. Nhu cầu phân tích kiểu này phụ thuộc vào vấn đề đặt ra về gene này như thế nào. Nếu chỉ quan tâm tới trình tự phiên mã thôi, thì chính dòng cDNA có thể cung cấp đầy đủ thông tin. Tuy nhiên, nếu cần đến các intron và vùng điều hoà thì phải thu nhận dòng genome.

Một nguồn DNA khác cho mẫu dò có thể là gene tương đồng từ một sinh vật họ hàng gần. Ví dụ, nếu một gene nào đó được nhân dòng trong nấm nang *Neurospora*, sau đó, có thể sử dụng gene này làm mẫu dò để tìm gene tương đồng trong nấm họ hàng *Podospora*. Phương pháp này phụ thuộc vào độ bảo thủ tiến hoá của các trình tự DNA qua thời gian. Ngay cả DNA mẫu dò và DNA của dòng mong muốn không thể giống hệt nhau nhưng chúng cũng đủ tương đồng để tiến hành lai. Phương pháp này được gọi là nhân dòng biết sẵn bởi vì, nếu đã biết một dòng gene quan tâm từ một sinh vật họ hàng, thì công việc nhân dòng sẽ trở nên tương đối dễ dàng.

Mẫu dò DNA có thể được tổng hợp nếu sản phẩm protein của gene quan tâm được biết và một trình tự amino acid được thu nhận. Các mẫu dò DNA tổng hợp được thiết kế dựa trên cơ sở biết mã di truyền, như vậy một trình tự amino acid chỉ đơn thuần là phải được dịch mã ngược lại để thu nhận trình tự DNA mã hoá nó. Tuy nhiên, do có nhiều bộ code (nói cách khác, sự thực là hầu hết các amino acid được mã hoá bởi nhiều hơn một codon) nên một số trình tự DNA khả năng có thể mã hoá cho một protein đang nói tới. Để tránh vấn đề này, một chuỗi amino acid ngắn với độ dư tối thiểu được chọn lọc. Trình tự nucleotide được tính toán bằng cách sử dụng từ điển codon. Phản ứng tổng hợp DNA hoá học là một quá trình từng bước một, như vậy ở bất kỳ đâu trên trình tự đều có các nucleotide lựa chọn. Một hỗn hợp của các nucleotide lựa chọn đó được cho vào phản ứng và tất cả các sợi DNA có thể có được tổng hợp.



Hình 5.11. Một trình tự ngắn của protein được sử dụng để thiết kế một bộ oligonucleotide có dư để sử dụng làm mẫu dò phát hiện gene mã hoá protein. Một trong số bộ mẫu dò sẽ bắt cặp hoàn chỉnh với gene.

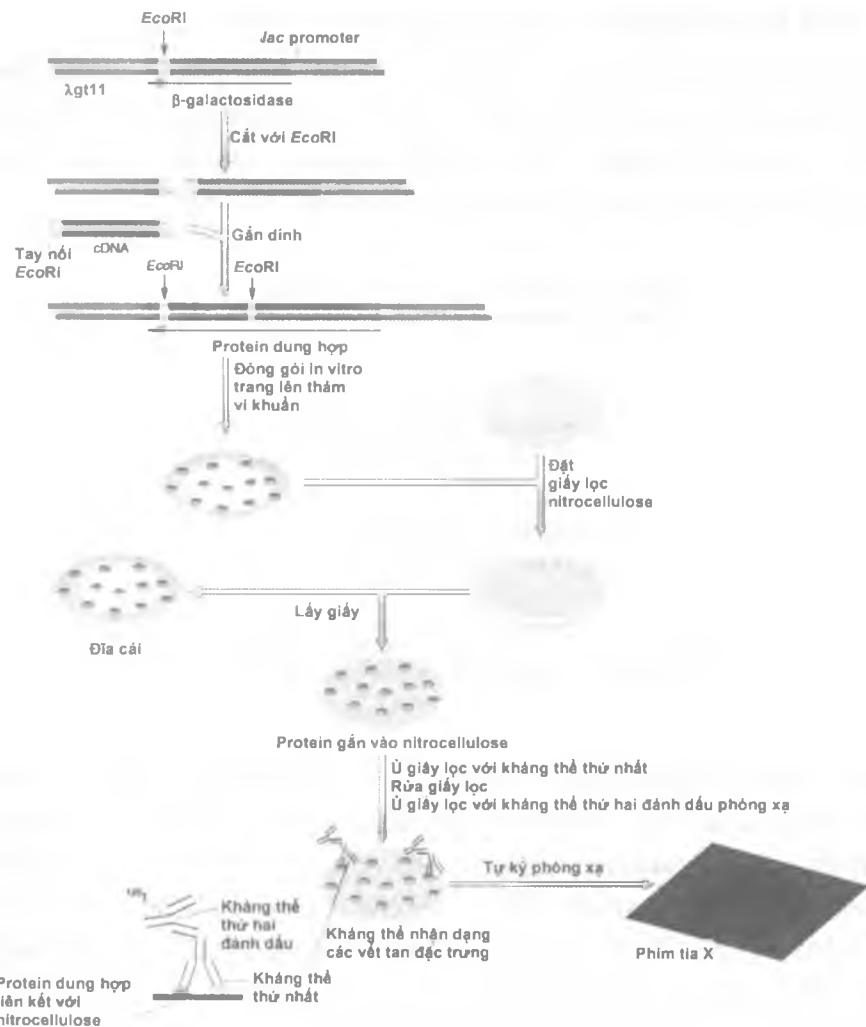
Hình 5.11 biểu diễn một ví dụ, trong đó có 5 vị trí cho 2, 3, 2, 2, 2 sự lựa chọn tương ứng. Phản ứng có thể tạo ra $2 \times 3 \times 2 \times 2 \times 2 = 48$ sợi oligonucleotide cùng một lúc. Món "cocktail" của oligonucleotide này có thể được sử dụng làm một mẫu dò. Sợi đúng trong cocktail này sẽ tìm ra gene quan tâm. Hai mươi nucleotide là hiện thân tính đặc hiệu đủ để tìm một trình tự DNA duy nhất trong thư viện.

Hơn nữa, RNA tự do có thể được đánh dấu phóng xạ và sử dụng làm mẫu dò. Điều này chỉ có thể khi một quản thể tương đối thuần khiết của các phân tử nhận dạng của RNA được tách chiết, như rRNA hoặc tRNA phân loại.

5.3.3.2. Mẫu dò tìm các protein

Nếu sản phẩm protein của một gene đã biết và tinh sạch ở dạng đồng nhất, thì protein này có thể được sử dụng để phát hiện trực tiếp dòng gene tương ứng trong thư viện. Quá trình được mô tả trong hình 5.12. Một kháng thể kháng protein được chuẩn bị và kháng thể này được sử dụng để sàng lọc thư viện biểu hiện. Các thư viện này được tạo ra bằng cách sử dụng các vector biểu hiện được thiết kế để biểu hiện mức độ cao của một protein vi khuẩn đặc biệt. Để tạo ra thư viện, cDNA được chèn vào vector trong khung với protein vi khuẩn và các tế bào sẽ tạo ra một protein dung hợp. Một màng được đặt lên bề mặt môi trường và lấy ra với một dấu in của các khuẩn lạc. Nó được làm khô và nhúng vào một dung dịch kháng thể. Các khuẩn lạc dương tính được hiện lên bằng cách tạo ra một kháng thể kháng lại protein đầu tiên ; kháng thể thứ hai được đánh dấu bằng isotope phóng xạ hoặc hoá chất phát quang hoặc trở thành một thuốc hiện màu. Bằng cách phát hiện protein đúng, kháng thể nhận dạng có hiệu quả ra dòng chứa gene tổng hợp protein đó.

Ở đầu chương này, chúng ta đã đặt ra câu hỏi làm thế nào để tìm ra gene người bạch tạng. Thực tế nó được nhân dòng bằng cách sử dụng một kháng thể kháng enzyme mà được biết có thể được phát hiện trong điều kiện này – enzyme tyrosinase. Enzyme này cũng như mọi protein, có thể được tinh sạch bằng các quy trình hoá sinh chuẩn, và tiếp theo một kháng thể kháng enzyme được sản xuất ở thỏ. Từ các tế bào sinh tyrosinase, mRNA được tinh sạch và sử dụng để tạo ra cDNA. cDNA này được sử dụng để tạo ra một thư viện vector biểu hiện. Thư viện được dò tìm với kháng thể kháng tyrosinase và một vài dòng dương tính được phát hiện. cDNA trong các dòng dương tính được đọc trình tự và chứa một gene có các exons với tổng số 1590 cặp nucleotide. cDNA được sử dụng để dò tìm một thư viện DNA genome người, và trong quá trình này, gene tyrosinase nguyên vẹn được tìm thấy. Nó xác nhận có 5 exon và 4 intron.

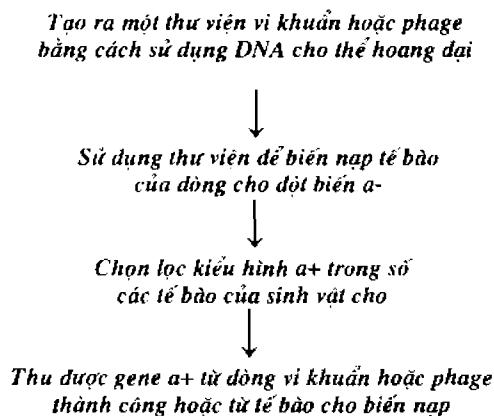


Hình 5.12. Tím dòng quan tâm bằng cách sử dụng kháng thể. Một thư viện biểu hiện được tạo ra với dẫn xuất phage λgt11 được sàng lọc với một kháng thể đặc hiệu-protein. Sau khi các kháng thể không liên kết được rửa trôi khỏi giá lọc, các kháng thể liên kết được nhìn thấy qua việc liên kết với một kháng thể thứ hai phóng xạ.

Một gene nhân dòng có thể được chọn lọc từ một thư viện bằng cách sử dụng các mẫu dò cho trình tự DNA của gene hoặc cho sản phẩm protein của gene.

5.3.4. Tìm các dòng đặc trưng bằng bổ trợ chức năng

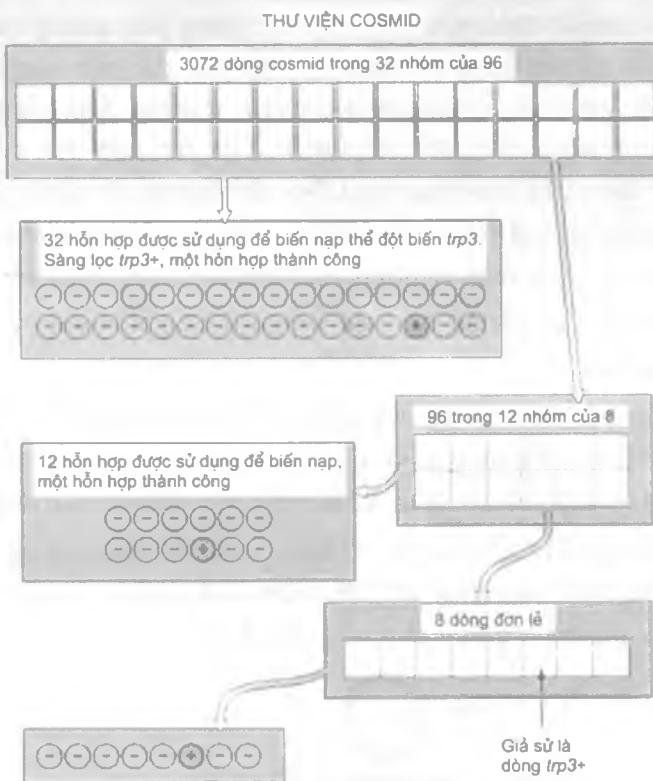
Các dòng đặc biệt trong thư viện vi khuẩn hoặc phage có thể được phát hiện qua khả năng chúng tạo ra một chức năng không có mặt ở dòng đột biến của sinh vật cho, sẽ làm nhiệm vụ như thể nhận biến nạp. Quy trình này được gọi là bổ trợ chức năng. Quy trình như sau :



Quy trình này phụ thuộc vào khả năng để biến nạp sinh vật cho, thường một eukaryote. Chúng ta đã từng xem xét biến nạp trong prokaryote, nhưng eukaryote cũng có thể được biến nạp. Quy trình này khác nhau giữa các eukaryote, nhưng nói chung cần xử lý đặc biệt đối với tế bào nhận. Ví dụ, để biến nạp nấm nói chung, thành tế bào phải được loại bỏ bằng enzyme. Giả sử chúng ta đã phân lập một thể đột biến có liên quan tới quá trình sinh học nào đó mà chúng ta quan tâm. Đối với mục đích hiện đang xem xét, chúng ta giả sử rằng nó là một đột biến dinh dưỡng thụ động ở nấm. Chúng ta sẽ phải sử dụng DNA từ thư viện để biến nạp chủng đột biến dinh dưỡng thụ động và sau đó trang lên đĩa thạch những tế bào nấm nhận này ở môi trường tối thiểu. Các tế bào nấm chứa allele hoang dại (từ dịch nuôi cấy hoang dại được sử dụng để tạo ra thư viện) sẽ biến đổi dinh dưỡng thụ động thành dinh dưỡng chủ động và cho phép phát triển trong môi trường tối thiểu.

Lý do mà phương pháp biến đổi này hoạt động được là phân đoạn biến nạp bổ trợ chức năng cho sự thiếu hụt do allele đột biến trong thể nhận gây ra. Đây không phải là kiểu hình hoang dại được tạo ra từ hai bộ

genome. Tuy nhiên, vector biến nạp mang một số gene mà genome thể nhận thiếu (allele kiểu hoang dại cần tìm) và genome thể nhận mang một số gene mà vector thiếu (tất cả bộ phận còn lại của genome). Như vậy, đây là một dạng bổ trợ chức năng cho nhau.

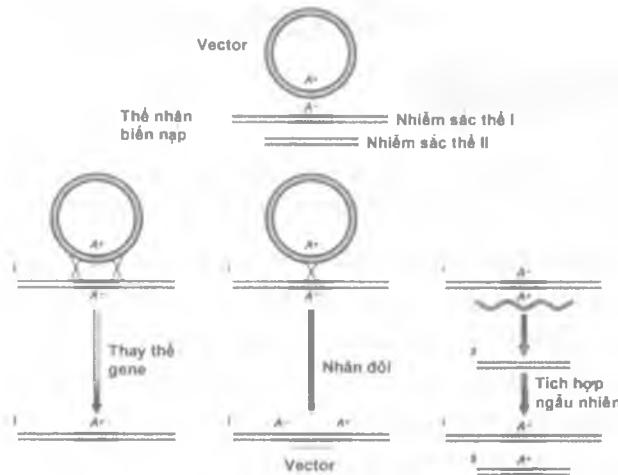


Hình 5.13. Tìm kiếm một gene nhân dòng sử dụng các mẫu DNA nhỏ dần hợp lại trong biến nạp. Trong ví dụ này là *trp3* gene của *Neurospora*.

Nếu thể nhận biến nạp là một sinh vật mà ở đó, các vector plasmid sao chép tự động (chủ yếu là vi khuẩn và nấm men) thì sau đó phân đoạn chèn biến nạp có thể được thu lại đơn giản bằng cách tách chiết plasmid. Tuy nhiên, như chúng ta sẽ thấy trong hầu hết các sinh vật eukaryote, vector vi khuẩn hoặc phage không thể sao chép và phải chèn vào genome để tạo ra biến nạp bền vững. Trong những trường hợp như thế, phân đoạn biến nạp tương đối không tiếp cận được và phải được lấy lại từ dòng thành công trong thư viện. Phương pháp này sử dụng một thư viện ở đó các dòng được trải ra như một tập hợp các dịch nuôi cấy vi khuẩn đánh số trong ống nghiệm hoặc trên đĩa microtiter. DNA được phân lập từ mỗi hòn

hợp tất cả các chủng trong các tập hợp con đặc trưng của thư viện và biến nạp được kiểm tra. Bằng một quy trình giảm các tập hợp con thư viện, biến nạp thành công, dòng với allele hoang dại có thể được nhận dạng. Quy trình được biểu thi trong hình 5.13, sử dụng một ví dụ gene *Neurospora trp3*, được thảo luận ở đầu chương. Trong trường hợp này, một thư viện cosmid được sử dụng. Cosmid cũng phải mang một gene chỉ thị mà có thể được sử dụng để chọn lọc cho các thể biến nạp thành công của nấm. Một gene cho kháng hygromycin thường được sử dụng trong nấm, mà chúng nhạy cảm đối với thuốc. Các tập hợp con của thư viện cosmid được tạo ra từ *Neurospora* DNA thể hoang dại được sử dụng để biến thái các tế bào đột biến *trp3*, và các dòng *trp3⁺* được chọn lọc bằng cách trang các tế bào biến nạp lên môi trường chứa hygromycin nhưng thiếu tryptophan. Các khuẩn lạc mọc lên có thể chứa allele *trp3⁺* và được phân lập từ đĩa thạch.

Trong hầu hết các trường hợp, các thể biến nạp được tìm thấy chứa vector mang allele kiểu hoang dại được chèn vào một trong những nhiễm sắc thể của sinh vật nhận ở vị trí khác với locus đột biến trong thể nhận. Điều này được gọi là chèn *lạc vị trí* (hình 5.14). Ít thông dụng hơn, allele kiểu hoang dại biến nạp, thay thế đột biến dinh dưỡng thụ động ngay tại chỗ bằng một quá trình giống như lai giống kép.



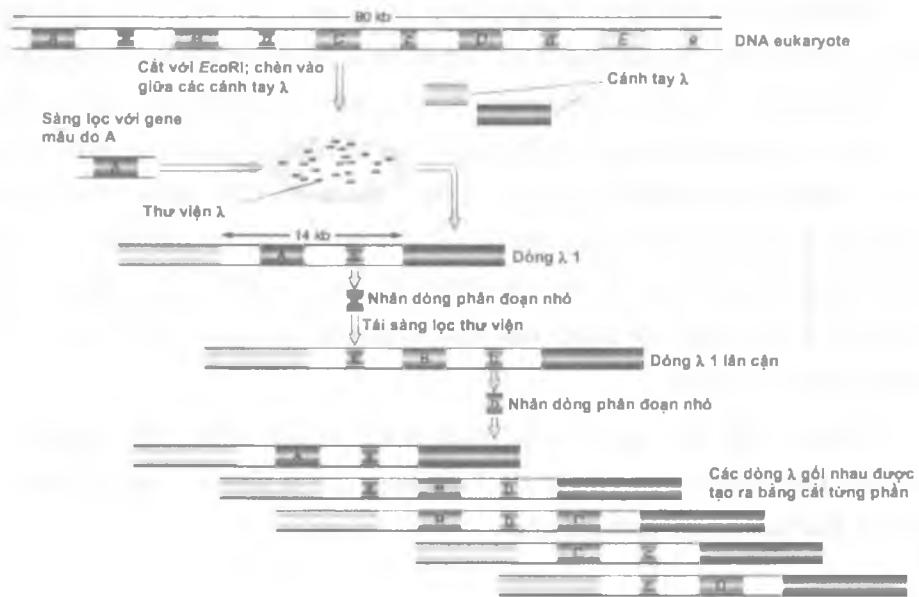
Hình 5.14. Các con đường có thể biến nạp DNA. Một allele kiểu hoang dại A⁺ (được nhân dòng trong một vector vi khuẩn) biến nạp một thể nhận A⁻ bằng một trong ba kiểu chèn khác nhau. Ghi nhớ : Thể nhận thường là cùng loài như DNA cho, hoặc prokaryote hoặc eukaryote. Hai nhiễm sắc thể nhận được biểu thị.

Nếu một gene eukaryote được nhân dòng trên một vector prokaryote nhưng một trình tự eukaryote đặc trưng được biết mà có thể hoạt động như một điểm sao chép, trình tự này có thể được bổ sung vào vector. Sau đó, vector sẽ có khả năng sao chép ở cả hai tế bào vi khuẩn và eukaryote và việc chèn vào nhiễm sắc thể là không cần thiết. Các dạng vector này được gọi là vector con thoi (*shuttle vector*) do chúng có thể được lấy từ vật chủ này chuyển sang vật chủ khác. Không có điểm khởi động sao chép, DNA cho phải hội nhập vào nhiễm sắc thể eukaryote để thực hiện sự biến nạp bền vững.

Các gene cho đặc trưng được nhân dòng có thể được chọn lọc bằng cách sử dụng DNA của chúng để biến nạp và bù vào các allele không trong tế bào nhận của sinh vật cho.

5.3.5. Nhân dòng vị trí

Thông tin về một vị trí của gene trong genome có thể được sử dụng để tháo gỡ khó khăn trong việc phân tích một thư viện hoàn chỉnh để tìm ra dòng quan tâm. Nhân dòng vị trí là một khái niệm có thể được áp dụng vào mọi phương pháp có sử dụng thông tin neutrôn. Thông thường, việc dò tìm và hỗ trợ là một phần của nhân dòng vị trí. Một điểm khởi động thông thường là khả năng có được một gene nhân dòng khác hoặc chỉ thị khác được biết liên kết gần với gene đang cần tìm. Chỉ thị liên kết hoạt động như điểm xuất phát trong một quá trình, được gọi là cuộn chiếu nhiễm sắc thể (chromosome walking), mà nó sẽ kết thúc ở gene đích. Hình 5.15 tóm tắt quy trình cuộn chiếu nhiễm sắc thể. Các phân đoạn cuối của một dòng của chỉ thị liên kết được sử dụng làm mẫu dò để chọn lọc các dòng khác từ thư viện. Các mẫu dò này sẽ phát hiện các dòng của các vùng DNA mà gối với dòng ban đầu. Các bản đồ han chế được tạo ra của DNA ở tập hợp dòng thứ hai này và một lần nữa, các phân đoạn hướng ngoại được sử dụng cho vòng mới của chọn lọc các dòng gối nhau từ thư viện. Do đó, quá trình cuộn chiếu chuyển động hướng ngoại theo hai hướng từ điểm xuất phát. Mỗi dòng có thể được đọc trình tự hoặc được thử theo cách khác, phụ thuộc vào mục đích khai thác.



Hình 5.15. Cuốn chiếu nhiễm sắc thể. Một phage tái tổ hợp nhận được từ thư viện phage bằng các cắt EcoRI không hoàn toàn của genome tế bào nhân chuẩn có thể được sử dụng để phân lập một phage tái tổ hợp khác chứa phân đoạn bên cạnh của DNA tế bào nhân chuẩn.

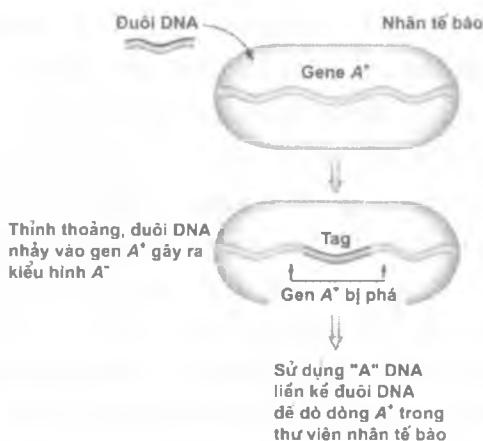
Thỉnh thoảng, một phân đoạn chèn lớn được biết chứa chỉ thị liên kết cũng sẽ may mắn chứa gene cần tìm và việc nhân dòng phụ và biến nạp sẽ làm hẹp vùng thích hợp của cosmid. Khả năng có được số lớn của các chỉ thị DNA trung tính (đa hình độ dài phân đoạn hạn chế – restriction fragment length polymorphism) rải rác trên hầu hết genome đã cung cấp nhiều điểm xuất phát tiện dụng. Nhân dòng vị trí đặc biệt tiện lợi cho nhân dòng các gene người, nhiều gene chưa có chức năng hoá sinh và kháng thể chọn lọc dễ dàng bằng bổ trợ chức năng. Gene người cho xơ nang, đã kể ở phần đầu của chương, được nhân dòng bằng cuốn chiếu nhiễm sắc thể. Đối với mọi trường hợp cuốn chiếu nhiễm sắc thể, cần phải có một số tiêu chí để đánh giá mỗi bước tiến theo gene quan tâm và các tiêu chí này phụ thuộc vào gene đơn lẻ được quan tâm.

5.3.6. Nhân dòng gene bằng gắn đuôi

Gắn đuôi là một phương pháp nhân dòng xuất phát tại điểm số không trên gene mong muốn trực tiếp bằng cách chèn một đoạn DNA đặc trưng làm thẻ đột biến chèn. Sau đó, trình tự đặc trưng được sử dụng làm đuôi

để tìm lại gene. Cách tiếp cận này được tóm tắt trong hình 5.16. Một dạng gắn đuôi biến nạp DNA. Khi DNA ngoại lai được bổ sung vào bằng biến nạp hoặc bằng các phương pháp khác như tiêm, nó có thể tích hợp vào genome và trở thành một phần của nhiễm sắc thể. Hội nhập lèch lạc là ngẫu nhiên trên toàn genome và rõ ràng không một vùng DNA nhiễm sắc thể nào có khả năng kháng lại sự tích hợp. Khi tích hợp xảy ra trong hoặc gần một gene, phân đoạn tích hợp hoạt động như là một thẻ gây đột biến, phá hỏng chức năng gene. Tính chất này có thể được sử dụng cho ưu thế tốt. Giả sử rằng chúng ta sử dụng một gene nhân dòng đặc biệt x^+ và biến nạp tế bào x^- của sinh vật cho thành x^+ . Nhiều thẻ biến nạp x^+ sẽ bị đột biến cho các gene trong đó DNA biến nạp chèn vào ngẫu nhiên. Một tập hợp con như các tế bào x^+ sẽ đột biến cho gene đích A^+ , gene quan tâm, và sẽ là kiểu hình A^- . Do đó, trong số các thẻ biến nạp x^+ , các kiểu hình A^- được nhận dạng. Bước tiếp theo là lai các thẻ biến nạp để xác định liệu kiểu hình A^- có phân ly với x^+ không. Nếu có, đột biến có thể là đã gây ra tích hợp của phân đoạn chứa x^+ . DNA của dòng đột biến được sử dụng để xây dựng thư viện, và gene x^+ có thể được sử dụng làm mẫu dò để phát hiện dòng của gene bị phá hỏng A . Để phát hiện ra A gene thể hoang dại nguyên vẹn, một phân đoạn của trình tự gene A bị phá hỏng được sử dụng trong một vòng dò tìm khác, lần này với thư viện kiểu hoang dại.

Một cách tiếp cận tương tự sử dụng các thẻ nhảy transposon làm đuôi. Transposon là các phân đoạn DNA có khả năng chuyển động tự nhiên được tìm thấy ở rất nhiều sinh vật. Khi chúng chuyển động, chúng có thể chèn vào bất kỳ đâu trên genome. Nếu chúng chèn vào hoặc gần một gene, chúng tạo



Hình 5.16. Sử dụng phân đoạn chèn DNA làm đuôi để đánh dấu và thu hồi một gene từ genome. Đuôi DNA có thể biến hình DNA hoặc một transposon nội sinh (yếu tố nhảy).

ra một dột biến không. Trong một dòng chứa một transposon hoạt động, các dột biến cho gene mong muốn được chọn lọc. Nhiều thể dột biến được tạo ra do transposon chèn vào. Dòng dột biến này được sử dụng để tạo ra một thư viện. Sau đó, một phần nhân dòng của DNA transposon có thể được sử dụng làm đuôi gắn để phát hiện gene, theo cách tương tự như đã chỉ trong Hình 5-16.

Gây dột biến gene bằng cách chèn DNA biến nạp hoặc một transposon cho phép gene được gắn đuôi như một việc mở đầu để phân lập nó.

5.4. SỬ DỤNG DNA NHÂN DÒNG

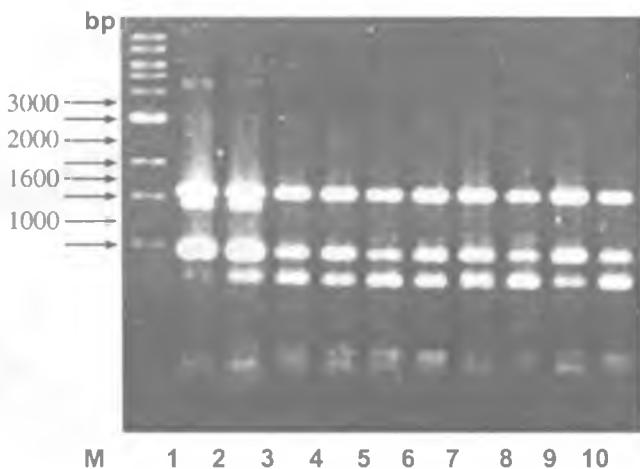
DNA nhân dòng có thể được sử dụng theo nhiều nhu cầu thí nghiệm. Trong mục này, chúng ta xem xét một số ứng dụng cơ bản mà có thể áp dụng trong một phạm vi rộng của các hoàn cảnh thí nghiệm.

5.4.1. Sử dụng DNA nhân dòng làm mẫu dò

Chúng ta đã xem xét một số ví dụ của dạng ứng dụng này, như sử dụng một gene nhân dòng từ một sinh vật để chọn lọc dòng thích hợp từ sinh vật khác.

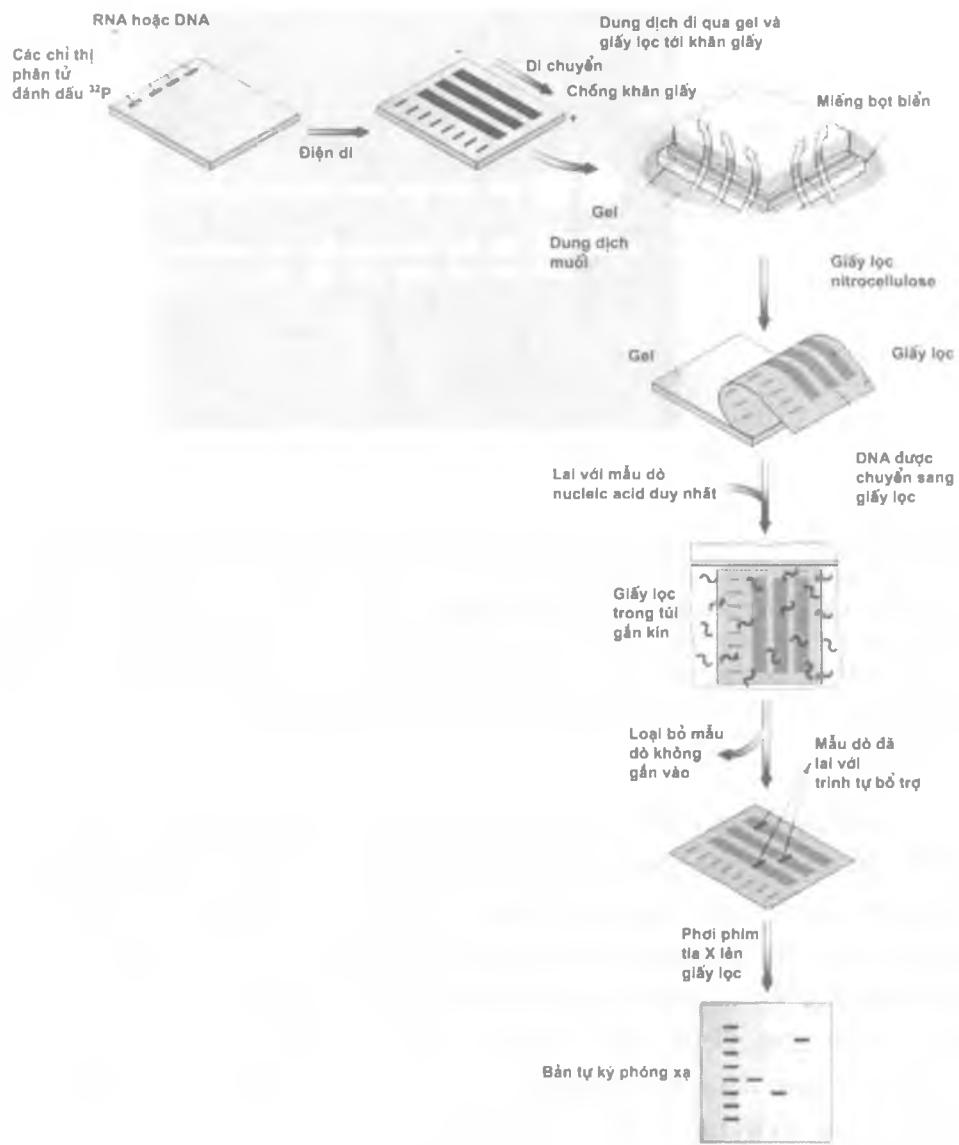
5.4.2. Dò tìm nucleic acid đặc trưng trong hỗn hợp

Trong khi thao tác gene và genome, thường cần phải phát hiện và phân lập các phân tử DNA từ hỗn hợp. Ví dụ, trong nhân dòng sử dụng phage λ , điều cần thiết là tách hai cánh nhiễm sắc thể từ vùng trung tâm không mong muốn. Có vài cách để phân tách DNA, nhưng phương pháp được sử dụng nhiều nhất là điện di. Nếu hỗn hợp các phân tử DNA mạch thẳng được tra vào giếng gel agarose được đặt gần cathode của một điện trường, các phân tử sẽ di chuyển qua gel tới anode, với tốc độ phụ thuộc vào kích thước của chúng (hình 5.17). Do đó, nếu có các nhóm kích thước khác nhau trong hỗn hợp, chúng sẽ hình thành các băng trên gel. Các băng có thể được nhìn thấy khi nhuộm DNA với ethidium bromide, làm cho DNA phát quang dưới tia UV. Các băng phân tách tốt có thể được cắt ra từ gel và mẫu DNA tinh sạch có thể được lấy ra từ bản gel. Bởi vậy có thể sử dụng điện di DNA để chẩn đoán phân đoạn DNA nào có mặt hoặc nghiên cứu phân lập các phân đoạn DNA đặc trưng.



Hình 5.17. Hỗn hợp của các phân đoạn DNA kích thước khác nhau được phân tách điện di trên gel agarose. Trong trường hợp này, các mẫu là mười sản phẩm RAPD-PCR. Hỗn hợp được đưa vào giếng trên đầu gel, các phân đoạn chuyển động dưới tác dụng của một điện trường tới các vị trí khác nhau phụ thuộc vào kích thước (và do đó là số điện tích). Các băng DNA được nhìn thấy bằng nhuộm với ethidium bromide và chụp ảnh dưới ánh sáng UV. (M điện di cho các kênh chứa các phân đoạn chuẩn hoạt động như là chỉ thị cho ước lượng độ dài DNA).

Cắt enzyme hạn chế DNA nhân tế bào cho kết quả nhiều phân đoạn DNA tạo ra một vệt nhoè trên gel điện di. Một mẫu dò có thể được nhận dạng một phân đoạn trong hỗn hợp, với sự sử dụng kỹ thuật được phát triển bởi E. M. Southern được gọi là lai *Southern blotting*. Sau đó, các phân đoạn DNA được phân chia trên gel, một màng hấp phụ được đặt lên gel và các băng DNA được chuyển ("thấm" "blotted") sang màng bằng nguyên lý hoạt động mao dẫn. Khi đã chuyển sang màng, các băng DNA lưu lại ở các vị trí tương đối như cũ ở trên gel. Màng được nhúng vào mẫu dò đánh dấu. Bán tự ký phóng xạ cho thấy sự có mặt của mọi băng trên gel mà tương đồng với mẫu dò. Nếu tái tổ hợp, những băng đó có thể được cắt ra khỏi gel và tiếp tục xử lý. Kích thước phân đoạn DNA trên gel có thể được xác định bằng cách chạy một thang chuẩn những phân đoạn đã biết kích thước trên cùng gel. Do đó, các kích thước của mọi phân đoạn quan tâm trong mẫu thí nghiệm có thể được suy ra. Hình 5-18 trình bày quy trình thấm Southern chung.



Hình 5.18. Quy trình lai Southern.

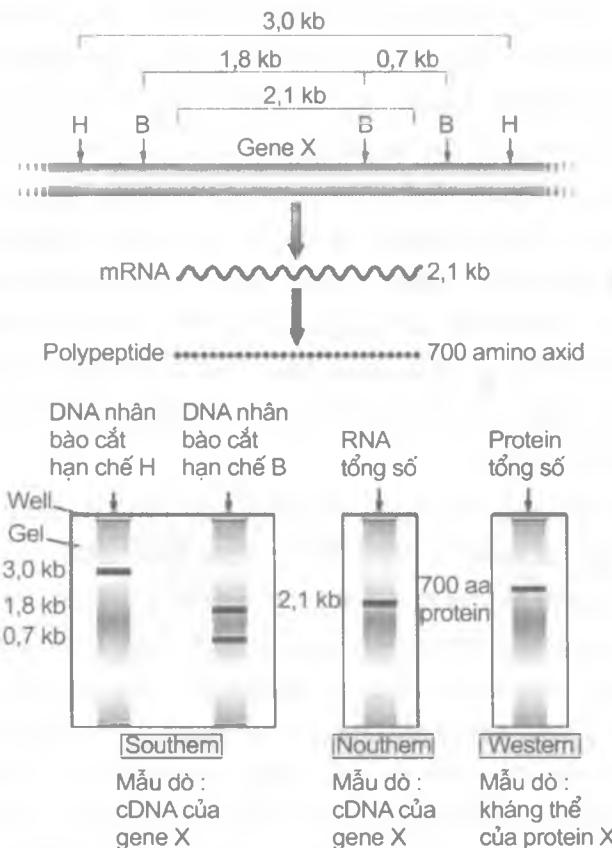
Điện di gel và thẩm trong việc nhận dạng các gene nhân dòng đặc trưng. Các phân đoạn hạn chế DNA và RNA được tra vào gel agarose và chạy điện di. Các phân đoạn khác nhau di chuyển với các tốc độ khác nhau tới các kích thước tương ứng của chúng. Gel được đặt vào đệm và dập bằng giấy lọc nitrocellulose và một chồng khăn giấy. Các phân đoạn được biến tính thành các sợi đơn, như vậy chúng có thể dính vào giấy lọc.

Chúng được chuyển sang giấy lọc nhờ đệm, được hút lên bởi khăn giấy. Giấy lọc sau đó được lấy ra và ủ với mẫu dò sợi đơn đánh dấu bổ trợ cho trình tự đích. Mẫu dò không liên kết được rửa trôi và phim tia X được phơi lên giấy lọc. Do mẫu dò phóng xạ đã lai chỉ với các phản ứng hạn chế bổ trợ của nó, phim sẽ được phơi chỉ ở các băng tương ứng với những phân đoạn đó. So sánh những băng này với các chỉ thị đánh dấu bày ra số và kích thước của những phân đoạn ở đó các trình tự đích được tìm thấy. Thông tin này có thể được sử dụng để định vị các trình tự đọc theo bản đồ hạn chế. Quy trình này được gọi là thẩm Southern (Southern blotting) khi chuyển DNA sang nitrocellulose và thẩm Northern (Northern blotting) khi chuyển RNA sang.

Kỹ thuật thẩm Southern có thể được mở rộng để phát hiện phân tử RNA đặc trưng từ một hỗn hợp RNA phân tách trên gel. Kỹ thuật này được gọi là thẩm Northern (*Northern blotting*) ngược lại với kỹ thuật Southern để phân tích DNA. RNA phân tách được thẩm lên màng và dò tìm theo cùng một cách như lai Southern. Một ứng dụng phân tích Northern trong việc xác định liệu có một gene đặc trưng được phiên mã trong mô nào đó dưới các điều kiện môi trường nào đó hay không. RNA được chiết từ mẫu tế bào thích hợp, sau đó điện di, thẩm và dò tìm với gene nhân dòng đang nói tới. Một tín hiệu dương tính cho thấy sự có mặt của thể phiên mã.

Do đó, chúng ta xem DNA nhân dòng tìm ra ứng dụng rộng rãi như một mẫu dò, để phân đoạn hoặc dòng cụ thể hoặc phân đoạn DNA cụ thể hoặc RNA cụ thể. Trong tất cả các trường hợp, khả năng của các nucleic acid với các trình tự nucleotide bổ trợ để tìm và liên kết với nhau trong dung dịch được khai thác.

Một kỹ thuật song hành được gọi là thẩm Western (*Western blotting*) được phát triển để chuyển protein phân tách điện di từ gel và sau đó quan sát các protein đặc trưng với kháng thể. Tuy nhiên, nên nhớ là mẫu dò ở đây không phải là DNA nhưng là kháng thể đánh dấu. Các kỹ thuật này là những công cụ thí nghiệm hữu hiệu trong việc phát hiện và xác định kích thước các đại phân tử đặc biệt trong di truyền phân tử. Chúng được so sánh trong hình 5.19.



Hình 5.19. So sánh phân tích Southern, Northern và Western của gene X. B và H đại diện các vị trí đích cho các enzyme hạn chế B và H. Các vị trí của các băng trên thấm Southern không thể được so sánh với các vị trí băng trên gel Northern và Western.

DNA hoặc RNA được phân tách theo kích thước trên một gel điện di có thể được thấm vào một màng hấp phụ và được dò để phát hiện ra vị trí của các phân đoạn đặc hiệu.

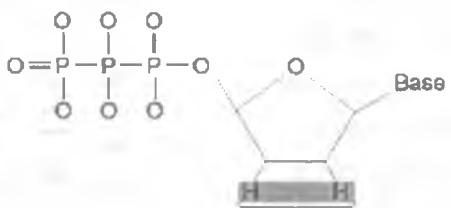
5.5. XÁC ĐỊNH TRÌNH TỰ DNA

Thành công trong nhân dòng gene mong muốn chỉ là sự khởi đầu của một vòng phân tích thứ hai, ở đó mục đích chính là xác định cấu trúc và chức năng của gene đó. Trình tự nucleotide của gene là sự xác định tính chất sơ đẳng của một cấu trúc di truyền. Một trong những yêu cầu để đọc trình tự DNA là khả năng thu nhận các phân đoạn DNA xác định. Bởi vậy, có sự phụ thuộc lẫn nhau giữa kỹ thuật nhân dòng DNA và kỹ thuật đọc

trình tự DNA, bởi vì nhân dòng DNA cung cấp các mẫu khuếch đại của các phân đoạn DNA xác định.

Mọi phương pháp đọc trình tự DNA đều bắt đầu với một số phân đoạn DNA xác định có một dấu đánh dấu. Từ số phân đoạn này, một loạt các phân tử được tạo có kích thước khác nhau chỉ một base ở đầu kia. Sau đó, các phân tử này được phân tách. Sự phân tách các phân tử DNA sơ đơn nhờ điện di gel acrylamide hoặc gel agarose. Tính di chuyển của một sợi tỷ lệ nghịch với logarithm của độ dài. Kỹ thuật này nhạy đến mức các phân đoạn dài ngắn khác nhau chỉ một nucleotide đơn có thể được tách nhau ra. Base ở đầu bị cắt của các phân tử phân loại được xác định và do vậy trình tự nucleotide được thiết lập.

Phương pháp đọc trình tự được sử dụng rộng rãi nhất hiện nay do Fred Sanger phát triển. Phương pháp Sanger dựa trên sự tổng hợp DNA có mặt của dideoxy nucleotide, chúng khác với deoxynucleotide thông thường ở chỗ chúng thiếu nhóm 3' hydroxyl (hình 5.20).

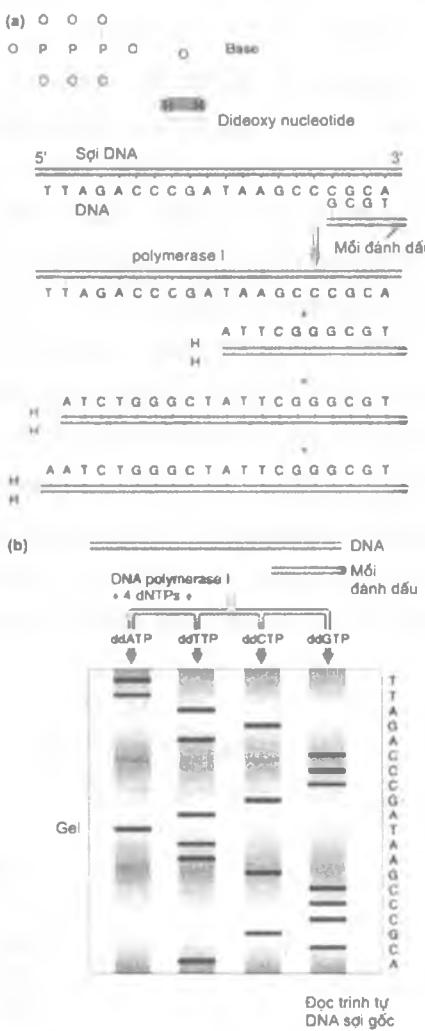


Không thể hình thành
một liên kết
phosphodiester với
dNTP tiếp theo

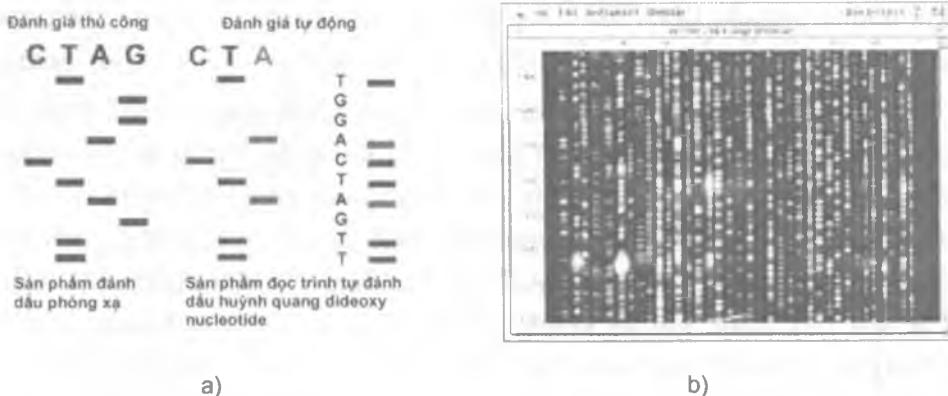
Hình 5.20. Cấu trúc 2',3'-dideoxy nucleotide, chúng được dùng trong phương pháp đọc trình tự DNA Sanger.

Các dideoxy nucleotide triphosphate (ddNTP) tương ứng có thể được gắn vào chuỗi đang phát triển, nhưng khi đã gắn vào, chúng lại kết thúc sự tổng hợp do chúng thiếu nhóm 3' hydroxyl cần thiết để liên kết với nucleotide triphosphate tiếp theo. Mỗi một phản ứng trong 4 ống được phân tách với một khuôn DNA sơ đơn cho trình tự quan tâm, cộng với DNA polymerase và một phân môi đánh dấu. Một ống nhận được một hàm lượng nhỏ của một ddNTP khác nhau (ddATP, ddTTP, ddCTP

hoặc ddGTP), cùng với 4 deoxynucleotide triphosphate bình thường (dNTP). Một dideoxy nucleotide sẽ được gắn ngẫu nhiên vào các vị trí khác nhau trong các phân ứng tổng hợp khác nhau trong ống phản ứng. Bởi vậy, trong từng ống, độ dài chuỗi bị ngắt khác nhau sẽ được tạo ra, tương ứng với điểm mà tại đó ddNTP đại diện cho ống được gắn vào và kết thúc sự phát triển của chuỗi. Những độ dài này, lần lượt là một chỉ thị rõ ràng của vị trí ở đó các base bổ trợ cho ddNTP trên sợi khuôn. Bởi vì sự gắn vào là ngẫu nhiên, nên tất cả các phân đoạn bị ngắt có thể có sẽ được tạo ra, tương ứng với tất cả các vị trí khác nhau của base đặc biệt đó. Các phân đoạn có thể được nhìn thấy qua điện di 4 mẫu trong 4 kênh trên một gel acrylamide, ở đó các phân đoạn hình thành các băng. Trình tự base có thể được xác định bằng soi quét gel, vây quanh tất cả bốn kênh và ghi nhận bất kỳ base nào chiếm giữ đầu cuối trong băng tiếp theo. Quy trình được minh họa trong hình 5.21 và 5.22.

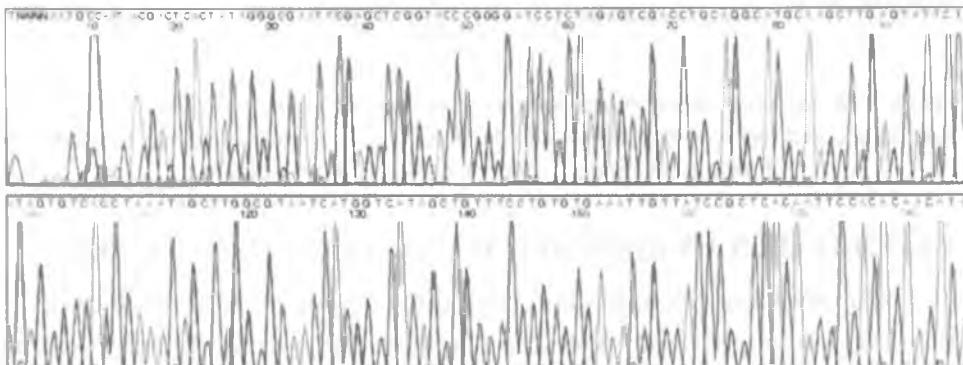


Hình 5.21. Phương pháp đọc trình tự dideoxy. (a) Mồi đánh dấu (được thiết kế từ trình tự biến vector) được sử dụng để khởi động tổng hợp DNA. Thêm vào 4 dideoxy nucleotide khác nhau (ở đây là ddATP) ngừng ngẫu nhiên phản ứng tổng hợp. (b) Các phân đoạn thu nhận được phân tách điện di và được đưa ra tự ký phỏng xá.



Hình 5.22. Gel đọc trình tự Sanger. Trình tự suy ra được trình bày ở bên phải. a) Đánh dấu phóng xa ; b) Đánh dấu huỳnh quang dideoxy nucleotide ; Đánh dấu huỳnh quang mồi.

Thay thế đánh dấu phóng xạ, các đuôi phát huỳnh quang được gắn vào mỗi oligonucleotide. Theo phương pháp trước đây, mỗi một phản ứng sử dụng một chất phát huỳnh quang khác nhau và bốn hỗn hợp phản ứng được điện di cùng nhau. Theo phương pháp hiện nay, chỉ cần một phản ứng đọc trình tự nhưng chứa cả 4 loại chất phát huỳnh quang. Sự phát hiện huỳnh quang được sử dụng trong máy đọc trình tự DNA tự động, nó có thể đọc được hàng nghìn base trong một lần phân tách. Hình 5.23 minh họa một hình ảnh đọc trình tự tự động.

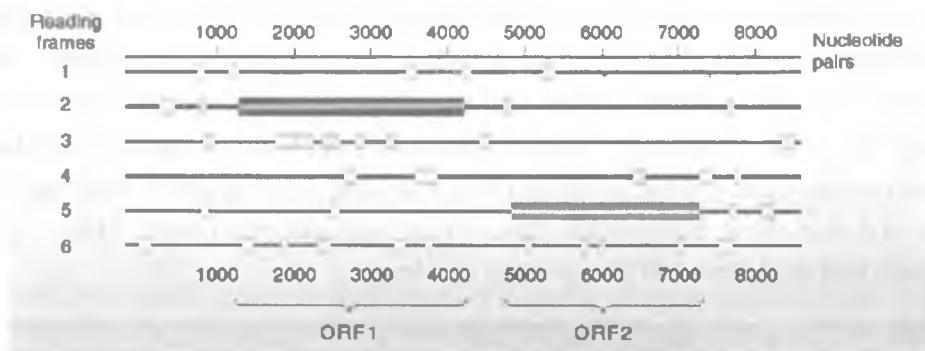


Hình 5.23. Ảnh đọc trình tự từ máy đọc trình tự tự động sử dụng chất phát huỳnh quang. N là base không ấn định được.

Một phân đoạn DNA nhân dòng có thể được đọc trình tự bằng cách tạo ra một tập hợp DNA sợi đơn khác nhau về độ dài bởi một nucleotide ;

khi những phân tử DNA này được điện di, trình tự nucleotide có thể được đọc trực tiếp từ tự ký phóng xạ trên gel, hoặc được suy ra từ huỳnh quang.

Trình tự nucleotide của phân đoạn DNA nhân dòng có thể được sử dụng để tìm ra (các) gene nhân dòng chứa trong đó. Trình tự nucleotide được đưa vào một máy tính, sau đó nó soi quét tất cả 6 khung đọc (ba khung mỗi chiều) để tìm kiếm một dải DNA có độ lớn của một gene, bắt đầu với một codon khởi động ATG và kết thúc với một codon kết thúc. Các dải này được gọi là khung đọc mở (*open reading frame, ORF*). Chúng là các trình tự gene ứng cử. Hình 5.24 cho thấy một phân tích như thế ở đó hai gene ứng cử được nhận dạng là ORF. Điều thú vị đáng ghi nhớ là việc phát hiện ORF và phân tích tỷ lệ phân ly theo Mendel đều nhằm nhận dạng gene, mặc dù hai cách tiếp cận khác nhau.



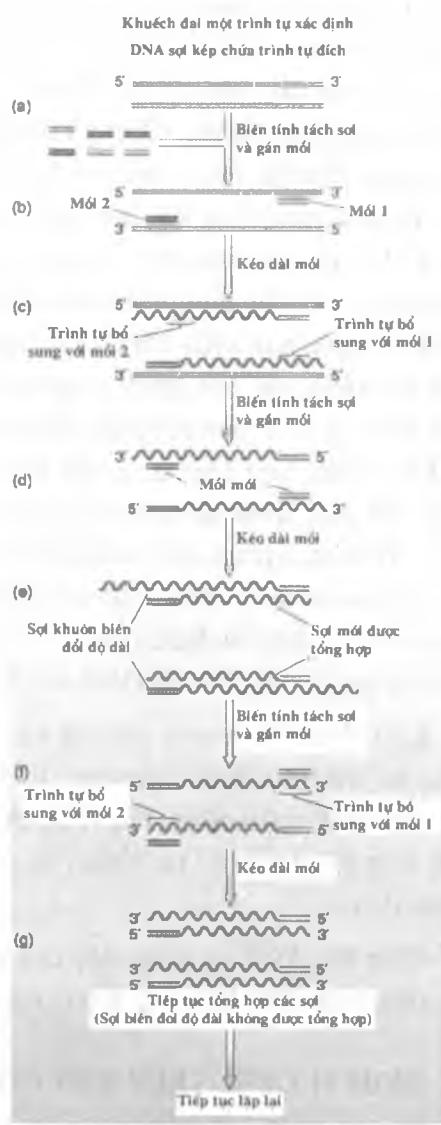
Hình 5.24. Mọi phân đoạn DNA có thể có 6 khung đọc, 3 khung mỗi chiều. Ở đây, máy tính đã soi quét một trình tự plasmid 9 kb để tìm ORF (các gene tiềm năng). Hai ORF lớn 1 và 2 là những ứng cử có thể là những gene tiềm năng.

5.6. PHÁT HIỆN VÀ KHUẾCH ĐẠI CÁC TRÌNH TỰ BẰNG PCR

Nếu một vùng DNA đã được nhân dòng và đọc trình tự, trình tự đó có thể được sử dụng để khôi phục các phần của vùng tương đương từ các cá thể đặc trưng hoặc từ các loài khác, tất cả đều không cần nhân dòng. Kỹ thuật sử dụng một quy trình được gọi là phản ứng chuỗi trùng hợp polymerase (*polymerase chain reaction, PCR*). Hình 5.25 minh họa nguyên lý của kỹ thuật. Một DNA polymerase bền nhiệt, *Taq* polymerase, từ vi khuẩn *Thermus aquaticus* được sử dụng để xúc tác sự phát triển của mỗi DNA. Các cặp mồi trên các sợi đối nhau được kéo dài về hai hướng

khác nhau, như trình bày trong Hình 5.25. Sau khi hoàn thành sao chép phân đoạn giữa hai mồi (không phải mạch vòng), hai sợi kép mới được biến tính nhiệt để tạo ra các khuôn sợi đơn, và chu kỳ sao chép thứ hai được thực hiện bằng cách hạ nhiệt độ xuống với sự có mặt của tất cả các thành phần cần thiết cho phản ứng trùng hợp. Các chu kỳ tổng hợp và biến tính lặp lại cho kết quả tăng luỹ thừa số phân đoạn cần sao chép. Sự nhân bội gấp hàng triệu lần có thể được tạo ra. Sử dụng PCR có thể khuếch đại một gene đơn bản từ một mẫu genome, nếu tổng hợp được các mồi tương ứng với các trình tự đã biết của gene. Do khuếch đại luỹ thừa, PCR rất nhạy và có thể phát hiện các trình tự đích chỉ với số bản sao cực thấp trong mẫu. Ví dụ, các phân đoạn DNA người có thể được khuếch đại bằng cách sử dụng chỉ một vài tế bào nang bao quanh một sợi tóc được nhổ ra. Nói chung, PCR rất hữu hiệu trong chẩn đoán DNA, nói cách khác, trong việc kiểm tra sự có mặt của gene hoặc trạng thái đột biến của một gene đặc trưng hoặc trong khuếch đại một vùng nhất định nào đó.

(a) DNA sợi kép chứa trình tự đích. (b) Hai mồi chọn hoặc tạo ra có các trình tự bổ sung các vị trí liên kết ở đầu 3' của gene đích trên hai sợi. Các sợi được phân tách bằng hâm nóng, cho phép hai mồi gắn vào



Hình 5.25. Phản ứng PCR (polymerase chain reaction).

các vị trí liên kết mồi. Cùng với, các mồi do đó tại biên trình tự đích. (c) *Taq* polymerase sau đó tổng hợp bộ thứ nhất của các sợi bổ trợ trong phản ứng. Hai sợi đầu tiên này có diện di biến thiên, do chúng không có một tín hiệu kết thúc chung. Chúng kéo dài ở xa phía các đầu của trình tự đích như phác họa bởi các vị trí liên kết mồi (d). Hai sợi kép được hâm nóng lại, lộ ra 4 vùng liên kết (Để đơn giản hoá, chỉ hai sợi mới được minh họa). Hai mồi lại liên kết với các sợi tương ứng của chúng ở đầu 3' của vùng đích. (e) *Taq* polymerase lại tổng hợp hai sợi bổ trợ. Mặc dù các sợi khuôn ở bước này biến đổi độ dài, hai sợi vừa được tổng hợp từ chúng có diện di chính xác của trình tự đích mong muốn. Đây là do mỗi sợi mới khởi động ở vị trí liên kết mồi, ở mỗi đầu của trình tự đích và tiếp tục cho tới khi không còn khuôn, ở đâu khác của trình tự. (f) Mỗi sợi mới khởi động với một trình tự mồi và các đầu với trình tự liên kết mồi cho mồi khác. Tiếp theo phân tách sợi, các mồi lại gắn và các sợi được kéo dài tới diện di của trình tự đích. Các sợi dài biến thiên từ phần (c) cũng sản sinh ra các sợi có diện di đích. (g) Quy trình có thể được nhắc lại vô hạn định, mồi lần tạo ra hai phân tử DNA sợi kép giống như trình tự đích.

Kỹ thuật PCR không cần cắt hạn chế cơ chất DNA, bởi vì các mồi sẽ đóng tại trình tự thích hợp của DNA nguyên bản. Hơn nữa cũng không cần quy trình nhân dòng, bởi vì DNA đủ được khuếch đại mà một băng rõ ràng trên gel được tạo ra. Ngoài ra, chỉ một hàm lượng nhỏ cơ chất DNA là cần đến.

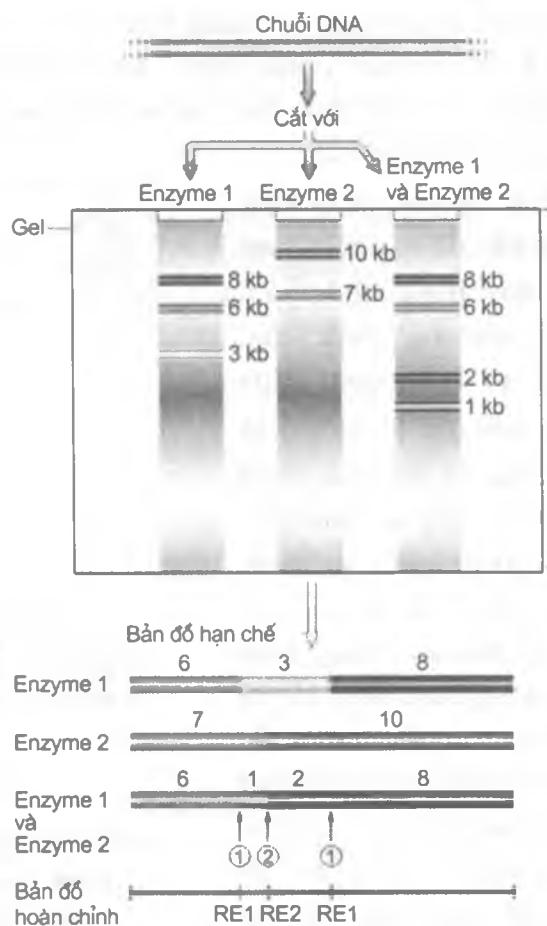
Phản ứng PCR sử dụng các mồi được thiết kế đặc biệt để khuếch đại các vùng ngắn đặc hiệu của DNA mà không cần nhân dòng.

5.7. ĐỊNH VỊ GENE TRÊN BẢN ĐỒ HẠN CHẾ

Chúng ta đã xem xét tầm quan trọng của các enzyme hạn chế trong việc cắt DNA để sử dụng nhân dòng. Một tính chất ích lợi khác của các enzyme hạn chế là các vị trí của các vùng đích của chúng đọc theo phân tử DNA (có nghĩa là đọc theo nhiễm sắc thể) có thể được sử dụng làm các chỉ thị DNA. Mặc dù các vùng tồn tại ngẫu nhiên, nói chung chúng ở các vị trí nhu nhau trên các nhiễm sắc thể tương đồng. Bởi vậy, các vị trí của các vùng đích có thể được sử dụng giống như các cột mốc đọc theo quốc lộ, mặc dù chúng không quan trọng nhưng chúng có thể được sử dụng làm các điểm đối chiếu để định vị các tính chất quan trọng hơn nữa đọc

theo con đường. Do đó, một bản đồ của các vị trí của các vùng hạn chế là một công cụ có giá trị trong phân tích genome nói chung. Thông thường, một bản đồ hạn chế được tạo ra cho một vùng nhiễm sắc thể định vị mà được quan tâm đặc biệt hoặc cho một nhiễm sắc thể tương đối nhỏ như plasmid hoặc của cơ quan tử.

Một cách vẽ bản đồ hạn chế là so sánh các điểm cắt enzyme hạn chế đơn với các điểm cắt đôi. Hai enzyme hạn chế được sử dụng trong các phản ứng cắt tách nhau và sau đó, hai enzyme được sử dụng cùng nhau. Sau khi cắt, các kích thước phân đoạn được xây dựng bằng điện di. Cắt hai lần cho thấy liệu một phân đoạn được tạo ra bởi một enzyme có chứa các vị trí của enzyme khác hay không? nếu như vậy, phân đoạn biến mất và được thay thế bằng hai hoặc nhiều hơn phân đoạn nhỏ hơn. So sánh kích thước của các phân



Hình 5.26. Vẽ bản đồ hạn chế.

đoạn được tạo ra trong các phản ứng cắt khác nhau cho phép định vị các vị trí đích hạn chế. Quy trình được minh họa trong hình 5.26.

Vẽ bản đồ bằng cách so sánh sự phân tách điện di của các dịch cắt đơn và đa enzyme. Trong ví dụ đơn giản này, cắt với enzyme 1 cho thấy

là có hai vị trí hạn chế đối với enzyme này nhưng không lộ ra là liệu phân đoạn 3 kb nằm ở giữa hay đầu của trình tự cắt dài 17 kb. Cắt tổ hợp với hai enzyme 1 và enzyme 2 để lại các phân đoạn 6kb và 8 kb nguyên vẹn nhưng cắt phân đoạn 3 kb, cho thấy rằng enzyme 2 cắt ở một vị trí nằm trong phân đoạn của enzyme 1. Nếu phân 3 kb nằm ngoài trình tự đang nghiên cứu, chỉ cắt với enzyme 2 mang lại một phân đoạn 1kb hoặc 2 kb. Bởi vì đây không phải là trường hợp này, từ ba phân đoạn hạn chế được sinh ra bởi enzyme 1, phân đoạn 3 kb phải nằm ở trong. Vị trí RE2 nằm gần phân 6 kb hơn phân 8 kb có thể được suy ra từ độ dài 7kb và 10 kb của dịch cắt enzyme 2.

Một phương pháp khác để vẽ bản đồ hạn chế được trình bày trong hình 5.27. Trong phương pháp này, một đoạn DNA được đánh dấu ở mỗi đầu 5' với ^{32}P và cắt để tạo ra các phân đoạn với đánh dấu ở một đầu. Phân đoạn dài nhất được phân lập và cắt với enzyme hạn chế thứ hai nhưng phản ứng không được phép xảy ra hoàn toàn ; bởi vậy một tập hợp các phân đoạn đánh dấu kích thước khác nhau được tạo ra, mỗi cái với một vị trí hạn chế khác nhau ở đầu không đánh dấu. Các phân đoạn này được phân tách điện di và khoảng cách của mỗi vị trí hạn chế từ đầu đánh dấu có thể được suy ra.

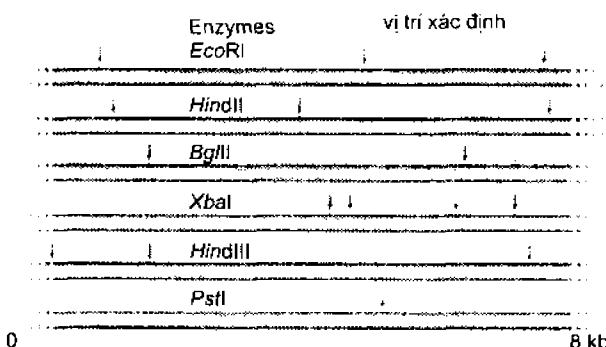
Trong ví dụ này, vòng DNA với vị trí EcoRI đơn được mở thành một phân tử mạch thẳng bởi nhát cắt với EcoRI. Phân tử được đánh dấu ở đầu



Hình 5.27. Nhận dạng trình tự của các vị trí enzyme hạn chế trong phân tử DNA mạch vòng.

5' của nó với ^{32}P và sau đó được cắt không hoàn toàn với *Alu* (enzyme hạn chế từ *Arthrobacter lutens*, có thể dễ dàng như *HaeIII*, enzyme hạn chế từ *Haemophilus aegyptius*) để sản sinh ra các phân đoạn có độ dài biến thiên. Mỗi phân đoạn bây giờ mang dấu phóng xạ ở một đầu. Phân đoạn dài nhất được chọn lọc (sau khi thử chạy trên gel để so sánh độ dài) và được chia thành hai mẫu. Một mẫu được cắt không hoàn toàn với *Alu* và mẫu khác được cắt không hoàn toàn với *HaeIII*. Tập hợp mới của các phân đoạn được phân tách trên gel. Sự định vị của phân đoạn đánh dấu có thể được xác định bằng tự ký phóng xạ trên gel. Trình tự các vị trí hạn chế có thể bây giờ được đọc theo thứ tự như hai gel được so sánh.

Một ví dụ cho bản đồ hạn chế được chỉ ra trong hình 5.28. Bản đồ này là DNA của plasmid mạch thẳng từ một chủng nấm.



Hình 5.28. Bản đồ hạn chế của một plasmid chủng nấm mạch thẳng, cho thấy các vị trí của sáu enzyme khác nhau.

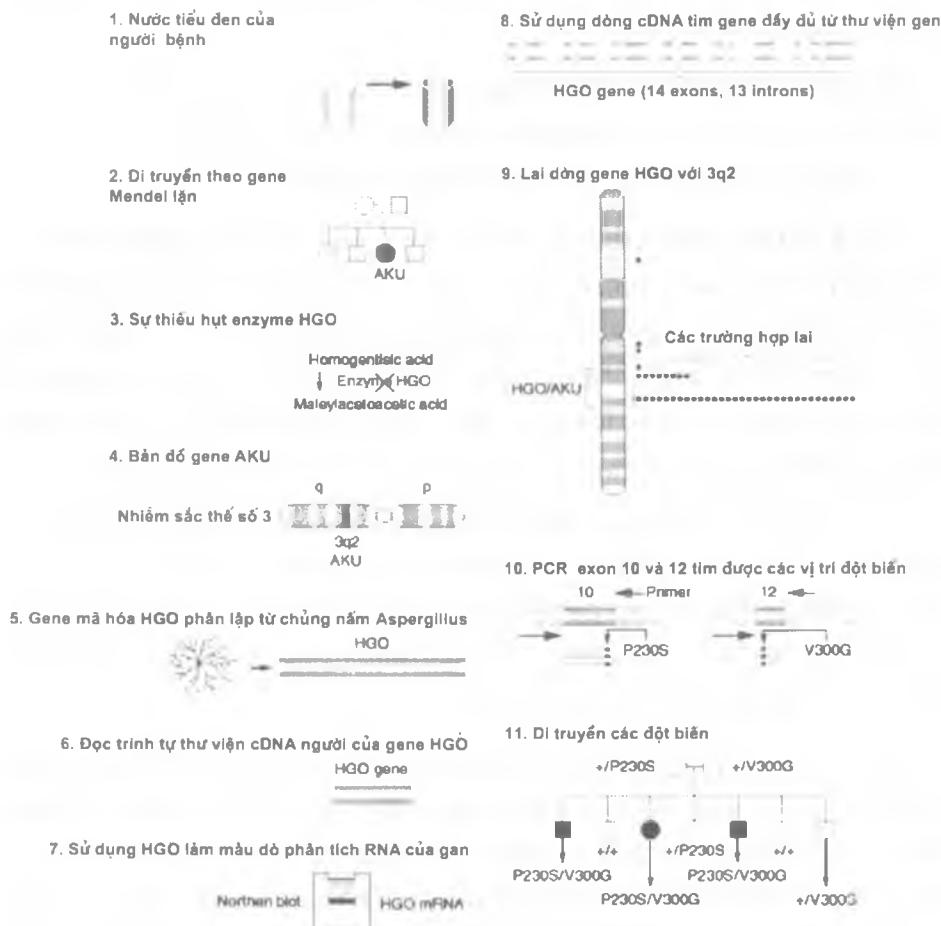
Sau khi một bản đồ hạn chế được tạo ra, các gene có thể được đặt vào bản đồ sử dụng phân tích Southern. Ví dụ, nếu một gene nhân dòng lai với phân đoạn A được tạo ra từ enzyme hạn chế 1 và với phân đoạn B được tạo ra từ enzyme hạn chế 2, sau đó gene phải được định vị tại vùng mà ở đó các phân đoạn A và B gối nhau.

Trong chương này giới thiệu các kỹ thuật nền tảng đã làm cuộc cách mạng di truyền học. Di truyền học mô tả phân tích di truyền của bệnh alkaptonuria, nó gắn và tích hợp nhiều kỹ thuật được đưa vào ở chương này.

Gene không còn là những thực thể giả thuyết nữa. Nhờ có kỹ thuật DNA tái tổ hợp và các kỹ thuật liên quan, các gene có thể được phân lập trong ống nghiệm và xác định tính chất như là các trình tự nucleotide cụ thể. Nhưng điều này vẫn chưa phải là chương cuối của cuốn tiểu thuyết. Ở chương tiếp theo cho thấy kiến thức về trình tự thường là sự khởi động của một chu kỳ mới của thao tác di truyền của trình tự đó, nguồn gốc sinh ra các cách tiếp cận mới thông minh để thay đổi một kiểu hình của sinh vật.

5.8. MỘT THẾ KỶ TRUY TÌM DI TRUYỀN ALKAPTONURIA

Alkaptonuria là một bệnh của người với một số triệu chứng, trong đó triệu chứng rõ ràng nhất là nước tiểu trở thành đen khi tiếp xúc với không khí. Vào năm 1898, một bác sĩ người Anh tên là Archibald Garrod đã chỉ ra rằng cơ chất gây ra hiện tượng này là homogentisic acid, nó được bài tiết với hàm lượng lớn bất thường vào nước tiểu của người gây bệnh alkaptonuria. Vào năm 1902, ngay sau thời kỳ Mendel, Garrod đã đề xuất rằng, dựa theo phả hệ cha mẹ, alkaptonuria được di truyền như một gene Mendel lặn. Ngay sau đó vào năm 1908, ông ta đã đề nghị rằng sự rối loạn bị gây ra do thiếu một enzyme mà bình thường phân cắt vòng thơm của homogentisic acid chuyển thành maleylacetoacetic acid. Do sự thiếu hụt enzyme này, ông ta suy ra rằng homogentisic acid tích tụ. Do đó alkaptonuria là một trong những trường hợp đề xuất sớm nhất của "lỗi bẩm sinh về trao đổi chất", một sự thiếu hụt enzyme gây ra do gene hỏng hóc. Sau đó 50 năm, một số người khác cho rằng trong gan của bệnh nhân alkaptonuria, hoạt tính enzyme mà bình thường phân cắt homogentisic acid (enzyme được gọi là homogentisate 1,2-dioxygenase (HGO)) thực sự hoàn toàn vắng mặt. Điều đó dường như enzyme HGO bình thường được mã hoá bởi alkaptonuria gene. Vào năm 1992, gene alkaptonuria được vẽ bản đồ di truyền ở băng số 2 trên cánh tay dài của nhiễm sắc thể số 3 (band 3q2).



Hình 5.29. Tìm tác nhân gây bệnh akaptunoria bằng kỹ thuật di truyền.

Vào năm 1995, Jose Fernandez-Canon và các cộng sự đã đánh giá một gene mã hoá hoạt tính HGO trong nấm *Aspergillus nidulans* và vào năm 1996, họ đã sử dụng trình tự amino acid suy diễn của gene này để nhờ máy tính tìm một số lượng lớn các phân đoạn đọc trình tự của một thư viện cDNA người. Họ nhận dạng một dòng dương tính mà chứa một gene người có 445 amino acid, nó cho thấy 52% tương đồng với gene *Aspergillus*. Khi gene người được biểu hiện trong một vector biểu hiện *E. coli*, nó cho hoạt tính HGO. Hơn nữa, khi gene được sử dụng làm mẫu

dò cho phân tích Northern của RNA gan, một RNA duy nhất của kích thước mong đợi được lai.

Khi gene nhân dòng được sử dụng làm mẫu dò trên các nhiễm sắc thể, ở đó DNA bị biến tính từng phần (lai *in situ*) mẫu dò liên kết với băng 3q2, cho thấy rằng đó thực sự là gene cho alkaptonuria.

Dòng cDNA được sử dụng để thu lấy gene đầy đủ từ một thư viện nhân bào. Gene được tìm ra có 14 exon và trải dài tổng cộng 60 kb. Một gia đình có 7 người con trong đó 3 người con bị chứng bệnh alkaptonuria được kiểm tra đột biến. Phân tích PCR được sử dụng để khuếch đại tất cả exon riêng lẻ. Các sản phẩm khuếch đại được đọc trình tự. Một người bố hoặc mẹ là dị hợp tử cho thay thế proline → serine ở vị trí 230 trong exon 10 (đột biến P230S). Người còn lại là dị hợp tử cho thay thế valine → glycine ở vị trí 300 trong exon 12 (đột biến V300G). Tất cả các người con bị bệnh alkaptonuria đều có cấu trúc P230S/V300G, như mong đợi nếu những đột biến này là những vùng đột biến bất hoạt HGO enzyme.

Di truyền học tập trung vào bản chất của gene và mục đích chính là đánh giá cấu trúc và chức năng của chúng. Kỹ thuật DNA tái tổ hợp đã cho phép các gene đơn lẻ được phân lập trong ống nghiệm và sau đó đánh giá tính chất ở mức độ phân tử. Công nghệ này dựa trên các enzyme hạn chế, mà chúng cắt DNA thành các phân đoạn nhất định. Các vùng đích hạn chế có thể được lập bản đồ và hoạt động như là các cột mốc DNA. Các phân đoạn hạn chế thường có các đầu dính, tạo cho chúng khả năng được chèn vào một vector có khả năng sao chép trong tế bào vi khuẩn. Các thể lai phân tử như vậy được biết là DNA tái tổ hợp. Vì khuẩn kiai một phân tử DNA tái tổ hợp đơn để tạo ra một dòng DNA. Các vector thông thường là các plasmid, phage và cosmid. Một bộ genome có thể được nhân dòng thành một tập hợp dòng được gọi là thư viện. Một dòng đặc hiệu có thể được tìm thấy ở trong một thư viện bằng cách sử dụng một mẫu dò mà liên kết đặc hiệu với DNA hoặc protein của dòng mong muốn. Các dòng đặc hiệu cũng có thể được phân lập bởi khả năng của chúng để biến đổi các thể đột biến

không. Gắn đuôi cũng rất tiện lợi cho nhân dòng gene : biến nạp DNA hoặc transposon được sử dụng để gây đột biến bằng cách chèn và DNA gắn đuôi được phân lập. Tiếp theo, nhiễm sắc thể tạo ra một cách phân lập gene bằng cách phân lập tiếp theo của các dòng gối nhau, khởi động từ một chỉ thị liên kết với gene mong muốn. DNA nhân dòng có thể được đọc trình tự bởi một số phương pháp, bao gồm việc ngừng phát triển chuỗi DNA bằng các dideoxy nucleotide. Phản ứng trùng hợp polymerase (polymerase chain reaction) sử dụng các mồi để khuếch đại trình tự DNA. Đây là một phương pháp phân lập DNA nhanh mà cấu trúc của nó vừa được đọc trình tự một phần và phát hiện hàm lượng nhỏ của một kiểu đặc hiệu DNA. Diện di phân tách các phân tử DNA hoặc RNA kích thước khác nhau từ một hỗn hợp, các mẫu dò có thể phát hiện các phân tử DNA hoặc RNA đặc hiệu trên gel, trong các quy trình được biết như phân tích Southern và Northern tương ứng.

6

ỨNG DỤNG KỸ THUẬT DI TRUYỀN

6.1. SẢN XUẤT PROTEIN TÁI TỔ HỢP

Phần này chủ yếu giới thiệu về các kỹ thuật trong việc sản xuất và tinh chế protein tái tổ hợp có các mục đích sử dụng khác nhau.

Việc sử dụng protein tái tổ hợp gia tăng nhanh trong những năm qua, chính yêu cầu này khiến cho các kỹ thuật sản xuất và tinh chế protein tái tổ hợp cũng phát triển rất nhanh, trong đó phải kể đến tính ưu việt được công nhận rộng rãi của kỹ thuật *protein dung hợp* (fusion protein) trong khâu tinh chế và phát hiện protein tái tổ hợp. Đã có nhiều công trình giới thiệu những hướng dẫn chung về việc thu hoạch, chiết xuất, xử lý các protein dạng hạt, loại bỏ các thành phần kèm, các muối và phân tử nhỏ không mong muốn.

Có thể khẳng định càng có nhiều hiểu biết về đặc tính của một protein thì càng dễ dàng phân lập và tinh chế nó. Vậy theo nguyên tắc này thì các protein dung hợp thuộc loại những protein đơn giản và thuận tiện trong thao tác. Trong nhiều trường hợp, chỉ cần có một bước tinh chế trên cột sắc ký ái lực thông thường là đủ để có protein tinh sạch. Bên cạnh kỹ thuật thao tác với protein dung hợp được trình bày sau đây gồm GST và $(\text{His})_6$ là hai loại protein dung hợp phổ biến nhất, phần này cũng đề cập thêm những loại protein dung hợp khác.

Phần cuối chương sẽ trình bày về kỹ thuật tinh chế nhiều bước dành cho một vài trường hợp, khi một bước tinh chế là chưa đủ hoặc yêu cầu độ tinh khiết cao hơn.

6.1.1. Lựa chọn tế bào chủ

Hiện nay có những hệ thống tế bào chủ dùng cho sản xuất protein tái tổ hợp như thực bào, vi khuẩn, nấm men, nấm sợi, thực vật, tế bào côn trùng, tế bào động vật và động vật chuyển gene. Kết quả chọn lựa hệ

thống sản xuất phụ thuộc vào những yêu cầu đặc thù của mục đích sử dụng loại protein tái tổ hợp được sản xuất. Bảng 6.1 trình bày những hệ thống tế bào chủ thông dụng với ưu và nhược điểm của chúng.

Bảng 6.1. Các hệ thống tế bào chủ dùng trong sản xuất protein tái tổ hợp.

Hệ thống tế bào chủ	Ưu điểm	Nhược điểm
Ví khuẩn <i>Escherichia coli</i>	<ul style="list-style-type: none"> Rất nhiều công bố và kinh nghiệm rất phong phú Nhiều loại vector để chọn lựa Dễ điều khiển quá trình biểu hiện gene Dễ nuôi cấy, năng suất cao (đạt đến 50% protein tế bào) Sản phẩm có thể thiết kế dạng tiết ra môi trường nuôi 	<ul style="list-style-type: none"> Không biến đổi sau dịch mã Hoạt tính sinh học và miễn dịch có thể khác với protein tự nhiên Hàm lượng endotoxin cao đối với vi khuẩn Gram âm
Ví khuẩn <i>Staphylococcus aureus</i>	<ul style="list-style-type: none"> Tiết protein dung hợp ra môi trường 	<ul style="list-style-type: none"> Không biểu hiện mạnh như ở <i>E. coli</i> Là vi khuẩn gây bệnh
Tế bào động vật có vú	<ul style="list-style-type: none"> Có cùng hoạt tính như protein tự nhiên Có các loại vector biểu hiện gene Có thể nuôi cấy quy mô lớn 	<ul style="list-style-type: none"> Nuôi cấy tế bào khó và đắt Tế bào mọc chậm Tế bào chuyển gene không ổn định Năng suất thấp hơn so với vi khuẩn
Tế bào nấm men	<ul style="list-style-type: none"> Không có endotoxin Được công nhận là an toàn (GRAS) Lên men không đắt Có thể glycosyl hóa và tạo liên kết disulfite Chỉ có 0,5% protein tự nhiên tiết ra môi trường nên dễ tinh chế Hệ thống sản xuất quy mô lớn và chế biến sau đó đã được thiết lập 	

Tế bào côn trùng nuôi cấy Baculovirus vector	<ul style="list-style-type: none"> Nhiều quy trình chế biến giống như hệ thống tế bào nhân chuẩn An toàn vì nhiều loài tiết túc cũng là vật chủ của Baculovirus Baculovirus vector được phép thử lâm sàng Virus ức chế sinh tổng hợp protein tế bào chủ Nồng độ sản phẩm protein tái tổ hợp đạt mức cao 	<ul style="list-style-type: none"> Thiếu thông tin về khả năng glycosyl hoá Sản phẩm thường không có hoạt tính đầy đủ Có những điểm khác về chức năng và tính antigen giữa sản phẩm và protein tự nhiên
Nấm Aspergillus sp	<ul style="list-style-type: none"> Có hệ thống lên men nấm sợi hoàn thiện Nuôi cấy không đắt Niger được xếp vào nhóm an toàn GRAS Có khả năng tiết lượng lớn sản phẩm ra môi trường nuôi cấy, thường để sản xuất enzyme công nghiệp 	<ul style="list-style-type: none"> Hiệu quả biểu hiện chưa thật cao Đoạn dẫn thường gây ảnh hưởng cấu trúc 3 chiều, xếp gấp sản phẩm và hoạt tính sinh học Cắt bỏ đoạn dẫn (tag) không thường xuyên đạt 100%
Nấm Aspergillus sp	<ul style="list-style-type: none"> Nuôi trồng không bị đắt tiền Quy mô sản xuất mở rộng tùy theo yêu cầu, không bị hạn chế Bảo quản sản phẩm chưa chế biến dạng hạt hay củ rất thuận tiện Hiệu quả kinh tế cao khi sử dụng sản phẩm dạng không tinh chế 	<ul style="list-style-type: none"> Hiệu quả chuyển gene chưa cao Thời gian dài mới tạo được cài chuyển gene

6.1.2. Lựa chọn vector

Để nhân dòng một gene quan tâm thì tất cả các vector phải có một vị trí cắt độc nhất tại vùng sau promoter. Việc lựa chọn vector hoàn toàn do hệ thống tế bào chủ quyết định. Sau khi chọn được hệ thống tế bào chủ thì sẽ xem xét lựa chọn đến loại vector thích hợp. Những loại vector để chọn lựa có thể là loại chỉ để biểu hiện gene đơn giản đến loại có khả năng tiết sản phẩm ra môi trường.

Tuy nhiên, điều quan trọng nhất trong quá trình chọn lựa hệ thống và vector biểu hiện là phải chú ý đến các yêu cầu đặc biệt do việc sử dụng sản phẩm protein tái tổ hợp đặt ra. Một yếu tố quyết định khiến cho việc sử dụng hệ thống protein dung hợp ngày càng trở nên phổ biến là quá

trình sản xuất một protein tái tổ hợp có mang một đoạn dẫn sẽ làm đơn giản hoá được các khâu tiếp theo như tách chiết, tinh chế và xác định sản phẩm cần thiết. Tiếp đến là chú ý đến năng suất protein. Bảng 6.2 giới thiệu một vài đặc tính của protein dung hợp giúp cho việc chọn lựa vector được đúng hơn. Việc duy trì và nhân bản các vector (bảng 6.3) phải tuân thủ những quy trình riêng rẽ đặc thù cho từng loại.

Bảng 6.2. So sánh ưu điểm và nhược điểm của một số protein dung hợp và không dung hợp.

Loại protein	Ưu điểm	Nhược điểm
Protein dung hợp	<ul style="list-style-type: none"> Có thể chọn khoang tế bào cho việc biểu hiện Có được chỉ thị trong khi biểu hiện Dễ tinh chế bằng sắc ký ái lực Dễ định tính Có thể tái xếp gấp trên cột sắc ký Lý tưởng cho protein tiết ra môi trường, dễ tách chiết 	<ul style="list-style-type: none"> Đoạn dẫn có thể ảnh hưởng đến cấu trúc và hoạt tính sinh học của protein đích Khi cắt bỏ đoạn dẫn có thể không đạt 100%
Protein không dung hợp	<ul style="list-style-type: none"> Không cần cắt bỏ đoạn dẫn 	<ul style="list-style-type: none"> Tinh chế và định tính không đơn giản Khó khăn trong việc hòa tan protein sản phẩm, làm giảm năng suất biểu hiện

Bảng 6.3. Các loại vector dùng trong sản xuất protein dung hợp và không dung hợp.

Vector để sản xuất protein không dung hợp	Đặc tính
pTrc 99 A	<ul style="list-style-type: none"> Vector cho tế bào nhân sơ biểu hiện các protein không có mã khởi đầu Cảm ứng bằng IPTG
pKK223-3	<ul style="list-style-type: none"> Để biểu hiện dư thừa protein nhờ sự kiểm soát của tac promoter mạnh trong tế bào nhân sơ
pSVK3	<ul style="list-style-type: none"> Để biểu hiện <i>in vivo</i> bằng các dòng tế bào động vật có vú
pKK223-3	<ul style="list-style-type: none"> Để biểu hiện tạm thời mức độ cao trong tế bào nhân chuẩn
pKK223-3	<ul style="list-style-type: none"> Để biểu hiện có khả năng cảm ứng trong tế bào động vật có vú

Vector để sản xuất protein dung hợp	Tag	Tinh chế sản phẩm
pGEX	Glutathione-S-transferase	<ul style="list-style-type: none"> GST MicroSpin TM module tinh chế GSTrapTM Glutathione SepharoseTM Fast Flow
pQE	6x Histidine	<ul style="list-style-type: none"> His MicroSpin Purification Module His TrapTM HiTrapTM Chelating Chelating Sepharose Fast Flow
pET	6x Histidine	<ul style="list-style-type: none"> IgG Sepharose 6 Fast Flow
pEZ 18 (biểu hiện không cảm ứng được)	Vùng liên kết IgG của protein A	<ul style="list-style-type: none"> IgG Sepharose 6 Fast Flow
pRIT2 (cảm ứng được bằng thay đổi nhiệt độ)	Vùng liên kết IgG của protein A	

6.1.3. Lựa chọn tag-dung hợp

Có hai loại tag thường hay được dùng là glutathione S-transferase (GST tag) và 6x histidine (His_6 -tag). Việc lựa chọn loại tag nào cho protein cần sản xuất phụ thuộc vào mục đích sử dụng protein tái tổ hợp. Bảng 6.4 trình bày những đặc điểm chính của hai loại tag này.

Bảng 6.4. Các đặc điểm chính của hai loại tag phổ biến dùng cho protein dung hợp.

Tag GST	Tag (His_6)
<ul style="list-style-type: none"> Có thể sử dụng cho mọi hệ thống biểu hiện 	<ul style="list-style-type: none"> Có thể sử dụng cho mọi hệ thống biểu hiện
<ul style="list-style-type: none"> Quy trình tinh chế cho năng suất cao 	<ul style="list-style-type: none"> Quy trình tinh chế cho năng suất cao
<ul style="list-style-type: none"> Có thể lựa chọn sản phẩm tinh chế ở mọi quy mô 	<ul style="list-style-type: none"> Có thể lựa chọn sản phẩm tinh chế ở mọi quy mô
<ul style="list-style-type: none"> Promoter của pGEX6P PreScissionTM protease cho phép cắt bỏ GST và tinh chế trong một bước duy nhất 	<ul style="list-style-type: none"> Tag nhỏ có thể không cần loại bỏ vì đặc tính gây miễn dịch của nó yếu tới mức có thể dùng cả protein dung hợp làm kháng nguyên để tạo kháng thể
<ul style="list-style-type: none"> Việc cắt bỏ tag có thể thực hiện dễ dàng nhờ có vị trí cắt đặc hiệu 	<ul style="list-style-type: none"> Việc cắt bỏ tag có thể thực hiện dễ dàng nhờ có vị trí cắt đặc hiệu. Điểm cắt enterokinase cho phép loại bỏ tag không để lại một gốc amino acid nào

<ul style="list-style-type: none"> Dễ dàng định tính được tag GST thông qua phép thử enzyme hoặc miễn dịch 	<ul style="list-style-type: none"> Dễ dàng định tính tag $(\text{His})_6$ bằng phản ứng miễn dịch
<ul style="list-style-type: none"> Tinh chế đơn giản, điều kiện dung môi dùng khi thỏi cột sắc ký khá yếu làm giảm thiểu nguy cơ bị hư hại đến hoạt tính sinh học và hiệu lực gây miễn dịch của protein đích 	<ul style="list-style-type: none"> Tinh chế đơn giản, nhưng điều kiện dung môi dùng khi thỏi cột sắc ký không yếu như khi dùng tag GST. Nếu cần, tinh chế có thể được tiến hành dưới điều kiện biến tính. Tuy pH trung tính nhưng imidazole có thể gây nên kết tủa. Cần thiết phải khử muối để loại imidazole
<ul style="list-style-type: none"> Tag GST có thể làm ổn định cấu trúc xếp gấp của protein tái tổ hợp 	<ul style="list-style-type: none"> Tag $(\text{His})_6$-dihydrofolate reductase làm bền vững các peptide nhỏ khi biểu hiện
<ul style="list-style-type: none"> Protein dung hợp tạo thể phức dimer 	<ul style="list-style-type: none"> Tag nhỏ thường có ít tác động lên cấu trúc và chức năng của protein đối tác Xác định khối lượng bằng khối phổ đối với vài loại protein dung hợp với tag $(\text{His})_6$ không đảm bảo

6.1.4. Lập kế hoạch sản xuất protein tái tổ hợp

Trước khi tiến hành sản xuất quy mô lớn cần bố trí một mẻ sản xuất quy mô nhỏ để thử về đặc điểm của hệ thống tế bào chủ, theo dõi mức độ biểu hiện gene trong suốt quá trình sinh trưởng và cảm ứng biểu hiện gene, thử nghiệm mức độ phù hợp của phương pháp tinh chế protein. Từ số liệu thí nghiệm quy mô nhỏ này có thể tối ưu hóa các điều kiện để mở rộng quy mô sản xuất.

Năng suất protein tái tổ hợp phụ thuộc những yếu tố sau : (1) Đặc tính của protein dung hợp ; (2) Tế bào chủ ; (3) Điều kiện nuôi cấy. Có thể dùng số liệu trình bày ở bảng 6.5 làm căn cứ tính toán cho quy mô sản xuất.

Bảng 6.5. Mối tương quan đại lượng giữa năng suất protein, thể tích dịch nuôi tế bào và thể tích phá tế bào.

Protein (μg)	12,5	50	1000	10000	50000
Thể tích dịch nuôi tế bào (ml)	5	400	400	4000	20000
Thể tích dịch phá tế bào bằng siêu âm (ml)	200	200	200	200	1000

Khi tiến hành sản xuất protein thường có thể áp dụng một số kinh nghiệm sau :

1. Bổ sung glucose (2%) vào môi trường không ảnh hưởng đến hiệu ứng cảm ứng của IPTG, mà làm giảm hiện tượng biểu hiện nền, tức là ức chế quá trình sinh tổng hợp những protein cơ bản của tế bào, chứ không phải protein tái tổ hợp mà ta quan tâm.
2. Biểu hiện nền, đó là quá trình biểu hiện gen không có mặt chất gây cảm ứng (như IPTG) xảy ra với hầu hết các loại promoter cảm ứng, có thể gây ảnh hưởng đến kết quả thí nghiệm nhân gene đối với các gene độc tố, qua đó có thể chọn lọc loại bỏ những đoạn gene không thích hợp. Biểu hiện cơ bản có thể được giảm thiểu thông qua ức chế bằng chất trao đổi (chẳng hạn như nuôi cấy có glucose). Đoạn promoter của gene *lac* không chịu tác động của quá trình ức chế do chất trao đổi. Tuy nhiên, còn có một promoter của gene *lac* lại nằm ở đầu trên giữa đầu 3' của gene *lac* và promoter *lac*. Đoạn promoter *lac* này có thể tham gia vào quá trình biểu hiện cơ bản của đoạn gene chèn vào vị trí ghép nối đa điểm và chịu tác động của hiện tượng ức chế do chất trao đổi gây ra.
3. Khi không phát hiện được protein tái tổ hợp trong dịch phá tế bào thì phải :
 - Kiểm tra lại trình tự DNA xem đoạn gene đã được chèn vào đúng khung đọc hay chưa.
 - Tối ưu hoá điều kiện nuôi cấy nhằm cải thiện năng suất. Nghiên cứu hiệu quả của chủng giống, thành phần môi trường, nhiệt độ nuôi cấy và điều kiện cảm ứng. Các điều kiện nuôi cấy chuẩn xác sẽ làm thay đổi đáng kể năng suất sản sinh protein dung hợp.
 - Phân tích một phần nhỏ nuôi cấy qua đêm bằng SDS-PAGE. Nhìn chung, một protein được biểu hiện mạnh sẽ cho băng nhìn thấy với cách nhuộm Coomassie bluc khi dùng khoảng 5-10 µl dịch nuôi có A₆₀₀ ≈ 1,0. Có thể dùng dịch nuôi tế bào không biến nạp hay tế bào được biến nạp với plasmid gốc để làm đối chứng. Khi thấy hoặc không thấy băng protein của dịch phá tế bào hay dịch nguyên tế bào, ta có thể kết luận về đặc tính hoà tan hay tạo thể hạt của protein dung hợp.

– Kiểm tra sản phẩm biểu hiện bằng phương pháp lai miễn dịch. Một vài protein dung hợp có thể bị che khuất bởi một protein của tế bào chủ có cùng kích thước phân tử. Trong trường hợp này thì lai miễn dịch có thể giúp ta xác định được protein dung hợp. Tiến hành điện di trên gel SDS-polyacrylamide dịch chiết tế bào được cắm ứng và truyền protein sang màng nitroso-cellulose hay màng PVDF (ví dụ màng HybondTM hoặc Hubond-P). Phát hiện protein bằng phương pháp lai màng với kháng thể anti-GST hoặc anti-His.

4. Hầu hết các protein dung hợp còn lại trong cặn sau khi phá tế bào

- Kiểm tra lại quá trình nghiên phá tế bào. Hiện tượng tế bào bị phá huỷ có thể nhìn thấy dịch tế bào trở nên trong hơn hoặc soi kính hiển vi. Bổ sung lysozyme (0,1 thể tích dung dịch 10 mg/ml lysozyme pha trong 25 mM Tris-HCl, pH 8,0) trước khi siêu âm phá tế bào có thể cải thiện được kết quả. Tránh hiện tượng tạo bọt vì nó gây biến tính protein dung hợp.
- Giảm cường độ siêu âm vì nếu siêu âm quá sẽ làm cho protein dung hợp cùng kết tủa với protein của tế bào chủ.
- Protein dung hợp có thể được tạo ra ở dạng hạt không tan khi biểu hiện quá mạnh, lúc đó có thể thay đổi điều kiện nuôi cấy để giảm tốc độ biểu hiện. Ví dụ như : (1) Giảm nhiệt độ nuôi cấy trong phạm vi từ 20°C đến 30°C để cải thiện tính hòa tan của protein dung hợp. (2) Giảm nồng độ IPTG xuống thấp hơn 0,1 mM. (3) Thay đổi thời gian cắm ứng. (4) Cắm ứng với thời gian ngắn hơn. (5) Cắm ứng ở nồng độ tế bào cao hơn, thời gian ngắn hơn. (6) Tăng cường sục khí. Sự vận chuyển nhiều oxygen vào tế bào cũng ngăn cản quá trình tạo thể hạt của protein dung hợp.
- Có thể cần phải phối hợp các phương cách khác nhau vừa nêu trên. Cách thức tốt nhất là phải thông qua kinh nghiệm thực nghiệm để tìm ra điều kiện tối ưu cho từng loại protein dung hợp.
- Thay đổi điều kiện chiết xuất protein cũng cải thiện được tính tan của protein dạng hạt.

5. Định lượng protein dung hợp

- Protein dung hợp cần được tinh chế và định lượng bằng phương pháp định lượng chuẩn.

- Có thể tính năng suất tương đối bằng cách đo độ hấp phụ ánh sáng ở 280 nm, đều dùng được cho cả hai loại protein dung hợp là GST và (His)₆.
- Năng suất protein có thể xác định bằng phương pháp quang phổ chuẩn như phương pháp Lowry, BCA và Bradford.

Phương pháp miến dịch cũng có thể dùng trong định lượng nếu xây dựng được đồ thị chuẩn. Phương pháp này rất thích hợp cho việc đánh giá sàng lọc nhiều mẫu, chẳng hạn như đánh giá sự có/không của protein dung hợp trong các phân đoạn sắc ký.

6.1.5. Protein dung hợp GST

6.1.5.1. Nhân gene và sản xuất protein

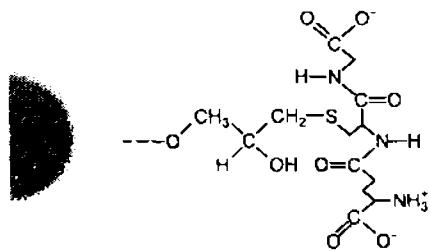
Hệ thống dung hợp gene glutathione S-transferase (GST) là một bộ sản phẩm tích hợp dùng trong sản xuất, tinh chế và định lượng các protein dung hợp GST trong tế bào *E. coli*. Những đặc tính của hệ thống GST được trình bày trong bảng 6.6.

Bảng 6.6. Các thông số kỹ thuật mô tả đặc tính của hệ thống GST.

Glutathione-S-Transferase	Protein tồn tại tự nhiên, MW 26000. Có thể biểu hiện ở <i>E. coli</i> có hoạt tính enzyme
Tính chất được xác định trong pGEX-1N	
Khối lượng phân tử dimer	MW 58000
KM (glutathione)	$0,3 \pm 0,07$ mM
KM (CDNB)	$0,3 \pm 0,07$ mM
pI (điểm đẳng điện)	MW 58000
Họ GST	Tinh chất lai của nhóm protein alpha và Mu

Hệ thống gene dung hợp GST được kiểm soát bởi promoter lac chịu tác động cảm ứng của đồng chất lactose là isopropyl-β-D-thiogalactoside (IPTG). Tế bào bị cảm ứng cho phép biểu hiện protein dung hợp GST kéo dài vài giờ trước khi thu hoạch.

Hình 6.1. Gắn glutathione vào hạt sepharose thông qua gốc oxirane. Cấu trúc của glutathione tương đồng với điểm gắn của glutathione-S-transferase.



6.1.5.2. Tế bào chủ

Dòng tế bào *E. coli* BL21 được lựa chọn làm tế bào chủ để sản xuất protein dung hợp GST vì dòng tế bào này bị thiếu protease (bảng 6.7).

Bảng 6.7. Một số đặc điểm của dòng tế bào *E. coli* BL21.

Kiểu gene	<i>F</i> -, <i>ampT</i> , <i>hsdS</i> (<i>r_B</i> ⁻ , <i>m_B</i> ⁻), <i>gal</i> (52,53)
Điều kiện nuôi	Huyền phù hoá nuôi cấy đông khô trong 1 ml môi trường L. Nuôi qua đêm trước khi trải lên đĩa thạch chứa môi trường L
Bảo quản lâu dài	Trộn glycerol với một thể tích tương đương huyền phù tế bào đã đạt pha ổn định nuôi bằng môi trường L. Bảo quản ở -70°C. Khi dùng lại hoạt hoá tế bào bằng nuôi vạch trên đĩa chứa môi trường L

Khi nhân dòng và bảo quản vector thì dùng dòng tế bào chủ khác, ví dụ như JM105.

Nếu dùng dòng tế bào không thiếu protease thì sản phẩm protein dung hợp có thể bị phân giải, tạo ra nhiều vạch trên bản điện di SDS-PAGE hoặc màng Western blot.

6.1.5.3. Vector

Vector pGEX bao gồm các phiên bản với cả 3 khung đọc kèm theo hàng loạt vị trí nhận biết cho các enzyme cắt (bảng 6.8). Vị trí ghép nối tương thích của các vector cho phép chuyển các đoạn chèn từ vector này sang vector khác dễ dàng. Các vector đều mang gene *lacZ^q* cho phép biểu hiện ở bất kỳ chủng tế bào chủ nào.

Bảng 6.8. Thông số kỹ thuật của các vector dùng trong sản xuất protein dung hợp GST.

pGEX-6P-1, pGEX-6P-2, pGEX-6P-3	PreScission protease
pGEX-4T-1, pGEX-4T-2, pGEX-4T-3	Thrombin
pGEX-5X-1, pGEX-5X-2, pGEX-5X-3	Factor Xa
pGEX-2TK cho phép xác định protein biểu hiện bằng đánh dấu <i>in vitro</i> trực tiếp	Thrombin, protein kinase phụ thuộc c-AMP

Hệ vector pGEX6P preScission protease là hệ vector cho phép cắt gọt và tinh chế protein dung hợp GST một cách hiệu quả nhất hiện nay. Có thể gắn cố định protease vào cột sắc ký để cắt gọt protein dung hợp đồng thời với quá trình tinh chế. Loại protease này có hoạt tính cao ở nhiệt độ thấp, vì thế có thể tiến hành tất cả quá trình tinh chiết cắt gọt trong phòng lạnh. Thông qua một bước duy nhất, enzyme cắt gọt và đoạn protein dẫn đều được loại bỏ.

6.1.5.4. Tinh chế

Bảng 6.9. Thông số kỹ thuật của các chế phẩm dùng trong tinh chế protein dung hợp GST.

Cột và chất nhồi	Hiệu suất thu protein	Khuyên cáo
GST MicroSpin Purification Module	Đến 400 µg	<ul style="list-style-type: none"> Dùng được liền, cột, đệm và hoà chất gói sẵn. Tốc độ nhanh nếu sử dụng cùng. MicroPlexTM 24 vacuum (tối 48 mẫu cùng lúc)
GSTrap 1 ml	10-12 mg	Cột gói sẵn dùng được ngay
GSTrap 5 ml	50-60 mg	Cột gói sẵn dùng được ngay
Glutathione Sepharose 4B	8 mg/ml	Để nhồi cột nhỏ
Glutathione Sepharose 4 Fast Flow	10-12 mg/ml	Để nhồi cột chảy nhanh nhằm mở lớn quy mô

Để phục vụ cho quy trình tinh chế một bước và đơn giản đối với protein dung hợp GST thì nhiều loại chế phẩm được thiết kế chế tạo như trình bày ở bảng 6.9.

Những điều cần lưu ý :

- Việc tái sử dụng cột phụ thuộc vào tính chất của mẫu, chỉ nên dùng với màu hoàn hoàn giống nhau để tránh nhiễm lẩn mẫu.

- Dùng Glutathione sepharose với sản phẩm cột gói sẵn hoặc tự nhồi cột để có tốc độ chảy nhanh được coi là thuận lợi hơn cả.
- Đối với việc tinh chế lượng nhỏ thì nên dùng cột MicroSpin là thuận lợi hơn cả.
- Nhưng đối với việc tinh chế lượng lớn nên dùng cột GSTrap liên kết với xilanh bơm, máy bơm nhu động hoặc hệ sắc ký cột.

6.1.5.5. Định lượng protein

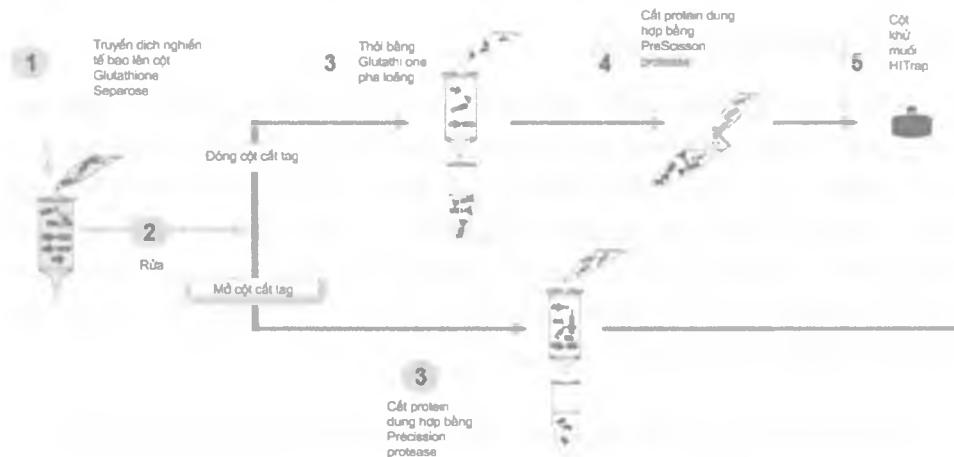
Các phương pháp định lượng protein dung hợp GST được so sánh trên bảng 6.10. Việc lựa chọn phương pháp nào là tuỳ thuộc vào yêu cầu của thí nghiệm, ví dụ phân tích SDS-PAGE thường dùng để theo dõi kết quả nhân gene và tinh chất protein dung hợp lại tỏ ra không thích hợp cho việc theo dõi thường quy đối với các thí nghiệm sàng lọc nhanh hàng loạt mẫu. Phương pháp xác định chức năng protein chỉ được xây dựng cho từng protein cụ thể đã được tinh sạch.

Bảng 6.10. So sánh các phương pháp đánh giá protein dung hợp GST.

Phương pháp	Khuyến cáo
GST ELISA module với 96 giếng	<ul style="list-style-type: none"> Lý tưởng cho việc sàng lọc hệ thống biểu hiện và các phân đoạn sắc ký Thuận tiện khi không biết lượng protein
GST module bằng phương pháp enzyme	<ul style="list-style-type: none"> Nhanh và thích hợp cho sàng lọc
Phân tích Western blot với kháng thể kháng GST và hệ ECLTM	<ul style="list-style-type: none"> Rất đặc hiệu, chỉ xác định protein dung hợp GST Hầu như không có hoặc có rất ít hiện tượng mờ nên khi dùng nồng độ tối ưu của kháng thể gắn với HRP Hệ thống phát hiện bằng ECL làm tăng độ nhạy của lai Western blot Hệ thống phát hiện bằng ECL cho độ nhạy phù hợp với mọi protein dung hợp Tăng độ nhạy thêm bằng hệ ECL Plus
Nhuộm bản SDS-PAGE bằng bạc hay Coomasie Blue	<ul style="list-style-type: none"> Cho thông tin về kích thước phân tử protein dung hợp và % tinh sạch Phát hiện được protein dung hợp và protein nhiễm tạp
Phân tích chức năng	<ul style="list-style-type: none"> Thích hợp cho việc xác định chức năng của protein dung hợp GST nhưng đối với mỗi loại cần có quá trình hoàn thiện và tối ưu

6.1.5.6. Cắt bỏ đoạn tag bằng enzyme

Trong hầu hết các trường hợp thử chức năng có thể dùng cả protein dung hợp GST còn đủ đầu tag. Nhưng trong những trường hợp phải loại bỏ tag thì nên dùng loại protein dung hợp có điểm cắt cho PreScission Protease. Đoạn tag GST có thể được loại bỏ cùng với quá trình tinh sạch protein thông qua một bước duy nhất trên cột (hình 6.2).



Hình 6.2. Sơ đồ quy trình cắt bỏ đoạn tag một bước trên cột tinh chế protein dung hợp GST.

Loại protease được sử dụng ở đây có tính chịu lạnh quý giá là có hoạt tính cực đại ở 4°C, như thế quá trình cắt bỏ đoạn tag có thể tiến hành ở nhiệt độ thấp khiến cho độ bền của protein đích được duy trì.

Vị trí nhận biết thrombin hay yếu tố Xa được cắt hoặc là khi protein dung hợp gắn kết trên cột hoặc trong dung dịch sau khi thô ra từ cột. Tuy nhiên, trong trường hợp này thì enzyme cắt phải được loại bỏ bằng sắc kí trao đổi.

Những khuyến cáo chung :

- Cắt trên cột thường được khuyến cáo sử dụng như là phương án chọn lựa vì nhiều chất pha tạp có thể được rửa qua cột và protein đích khi thô ra khỏi cột có độ tinh sạch cao.
- Lượng enzyme, nhiệt độ và thời gian ủ để thu được việc cắt bỏ toàn phần phụ thuộc vào từng loại protein dung hợp riêng, cần tiến hành thí nghiệm pilot để tối ưu hóa quá trình này.
- Nên lấy mẫu tại nhiều thời điểm khác nhau để xác định năng suất, độ sạch và thời gian xử lý cắt bỏ tag thích hợp.

6.1.6. Protein dung hợp (His)₆

6.1.6.1. Nhân gene và sản xuất protein

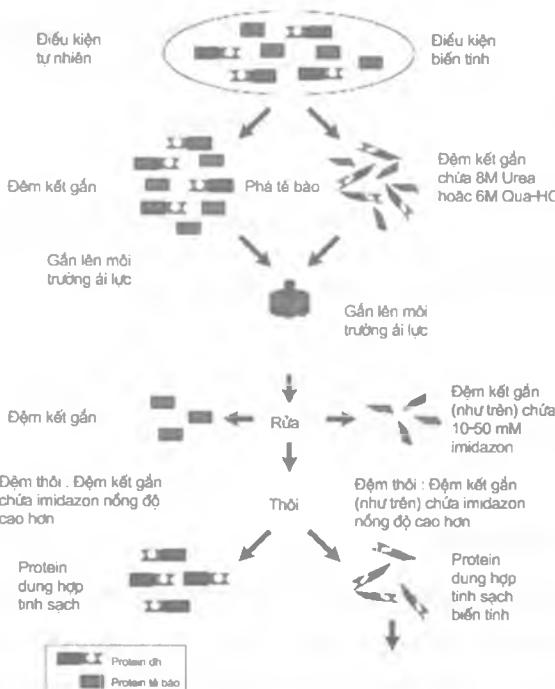
Những nguyên tắc chung trong nhân dòng gen, thiết kế vector biểu hiện đã được thảo luận ở các mục trên. Đối với sản phẩm protein dung hợp (His)₆ cần lưu ý về vector và tế bào chủ theo những hướng dẫn đã trình bày ở trên.

6.1.6.2. Tinh chế

Quy trình tinh chế điển hình của protein dung hợp (His)₆ dưới điều kiện tự nhiên và điều kiện biến tính được trình bày ở hình 6.3. Sau khi tinh sạch protein biến tính cần được tái xếp gấp lại để có cấu trúc không gian tự nhiên.

Các chế phẩm phục vụ việc tinh chế protein dung hợp (His)₆ theo quy trình một bước được trình bày ở bảng 6.11.

Sephadex chelate hóa khi tích điện với ion Ni²⁺ gắn kết một cách lựa chọn các protein có histidine ở bề mặt, sau đó protein dung hợp (His)₆ dễ dàng được thải ra bằng đệm thải chứa imidazole.



Hình 6.3. Sơ đồ quy trình tinh chế protein dung hợp (His)₆ ở điều kiện tự nhiên (trái) và điều kiện biến tính (phải).

Khuyến cáo lựa chọn phương pháp :

- Để tinh chế một lượng nhỏ hay phải sàng lọc nhanh nhiều mẫu thì nên dùng cột MicroSpind kết hợp với ly tâm hay máy hút chân không MicroPlex 24.
- Để tinh chế lượng lớn protein dung hợp (His)₆ nên lựa chọn bộ kit HisTrap hoặc cột HisTrap Chelating với ống tiêm bơm bay, bơm nhu động hay hệ thống sắc ký.
- Để mở rộng quy mô sản phẩm cần chọn nhiều cột HiTrap Chelating hoặc nhồi Chelating Sepharose Fast Flow vào cột thích hợp.

Bảng 6.11. Các chế phẩm cần dùng cho quy trình tinh chế protein dung hợp (His)₆.

Cột và chất nhồi	Hiệu suất thu protein	Khuyến cáo
His MicroSpin Purification Module	Đến 400 µg	Dùng được liền, cột bufer và hoá chất gói sẵn. Tốc độ nhanh nếu sử dụng cùng MicroPlexTM 24 vacuum (tới 48 mẫu cùng lúc)
HiTrap Chelating 1 ml	Đến 12 mg	Cột gói sẵn dùng được ngay
HisTrap Kit	Đến 12 mg	Như trên kèm đệm và ống tiêm cho 12 lượt dùng
HiTrap Chelating 5 ml	Đến 12 mg	Cột gói sẵn dùng được ngay
Chelating Sepharose Fast Flow	12 mg protein dung hợp (His) ₆	Để nhồi cột chảy nhanh nhằm mở lớn quy mô

6.1.6.3. Phát hiện protein

Bảng 6.12 trình bày tổng quát về các phương pháp hiện có dùng để phát hiện các protein dung hợp (His)₆. Việc lựa chọn phương pháp thích hợp phụ thuộc yêu cầu thí nghiệm, ví dụ phân tích SDS-PAGE thường được tiến hành nhiều lần trong quá trình nhân bản, chọn dòng và tinh chế để kiểm soát kết quả nhưng không thể dùng cho công việc thường quy khi

tiến hành sàng lọc lượng mẫu lớn. Phân tích chức năng đặc thù cho một loại protein rất có giá trị nhưng không phải lúc nào cũng đã có sẵn mà phải tự nghiên cứu hoàn thiện.

Bảng 6.12. So sánh các phương pháp đánh giá protein dung hợp (His)₆.

Phương pháp	Khuyến cáo
Phương pháp ELISA với IgG liên hợp	<ul style="list-style-type: none"> Rất đặc hiệu, chỉ xác định protein dung hợp (His)₆
Phân tích Western blot với kháng thể kháng His và hệ ECLTM	<ul style="list-style-type: none"> Rất đặc hiệu, chỉ xác định protein dung hợp (His)₆ Hầu như không có hoặc có rất ít hiện tượng mờ nền khi dùng nồng độ tối ưu của kháng thể gắn với HRP Hệ thống phát hiện bằng ECL làm tăng độ nhạy của lai Western blot Hệ thống phát hiện bằng ECL cho độ nhạy phù hợp với mọi protein dung hợp Tăng độ nhạy thêm bằng hệ ECL Plus
Nhuộm bản SDS-PAGE bằng bạc hay Coomassie Blue	<ul style="list-style-type: none"> Cho thông tin về kích thước phân tử protein dung hợp và % tinh sạch Phát hiện được protein dung hợp và protein nhiễm tạp
Phân tích chức năng	<ul style="list-style-type: none"> Thích hợp cho việc xác định chức năng của protein dung hợp (His)₆ nhưng đối với mỗi loại cần có quá trình hoàn thiện và tối ưu

6.1.7. Xử lý thể protein vùi

Sản xuất protein tái tổ hợp thường được kiểm soát thông qua quá trình tích tụ protein ở khu vực gian bào hay tiết ra vùng ngoài tế bào. Tiết protein ra ngoài tế bào có nhiều ưu điểm liên quan đến xếp gấp cấu trúc protein, tính tan và quá trình oxy hoá cysteine, trong khi đó năng suất protein tái tổ hợp được biểu hiện nội bào lại cao hơn.

Tuy nhiên, protein tái tổ hợp khi tích tụ nội bào thường có dạng hạt (vùi), không tan, tủa cùng các protein xếp gấp bị lỗi, mất hoạt tính sinh học. Vì thế với sự có mặt của protein dạng hạt, quá trình phân lập sơ bộ

có thể tiến hành rất dễ dàng. Khó khăn chủ yếu liên quan đến tái xếp gấp, sửa chữa lỗi trong cấu trúc không gian 3 chiều thông qua các cầu nối disulfide và khôi phục toàn bộ hoạt tính sinh học của protein đó. Bảng 6.13 thống kê những đặc điểm ưu việt và nhược điểm của protein tái tổ hợp được biểu hiện và tích tụ thành dạng hạt.

Bảng 6.13. Những ưu điểm và nhược điểm của protein tái tổ hợp được biểu hiện và tích tụ thành dạng hạt trong tế bào.

Ưu điểm	Nhược điểm
Mức độ biểu hiện cao góp phần làm giảm chi phí lên men	Tái xếp gấp khó khăn và chi phí tái chế sau khu hoạch cao
Dễ theo dõi bằng SDS-PAGE hay lai miễn dịch	Quá trình sản xuất không thể theo dõi bằng phương pháp phát hiện chức năng
Dễ dàng loại bỏ protein nội bào khác bằng cách rửa	Chất tạp nhiễm thứ yếu thường là các protein kị nước, màng khó tan và mảnh vỡ của thành tế bào
Tạp nhiễm chính chỉ là các dạng protein tiểu phân oligo, xếp gấp lỗi hay bị phân huỷ	Khó tách các dạng phức hợp khác nhau của một loại protein
Đoạn promoter pL có đầu T vẫn cho thu hoạch protein, trong khi đó ở hệ thống khác thì không thể	Nếu protein không xếp gấp hoàn hảo thì phải sử dụng hệ thống sản xuất khác

6.1.8. Làm tan các dạng protein hạt

Thay đổi điều kiện nuôi cấy là cách thúc đẩy tiên cần thử để tăng tính tan của protein tái tổ hợp bị đóng hạt.

Nếu thay đổi điều kiện nuôi cấy không mang lại hiệu quả thì có thể tiến hành biến tính bằng các hóa chất thông dụng như 4-6 M guanidinium hydrochloride, 4-8 M urea, chất tẩy rửa, chất kiềm pH >9, dung môi hữu cơ hoặc N-lauryl-sarcosine.

Điều kiện biến tính tùy thuộc thời gian, nhiệt độ, cường độ ion, tỷ lệ chất biến tính và sự có mặt của hóa chất phản ứng thiolt.

Protein khi hoà tan rồi có thể được tinh chế ngay bằng kỹ thuật thích hợp nhằm loại bỏ những tác nhân gây biến tính và tạo điều kiện để cho protein tái tổ hợp tái xếp gấp trở lại.

6.1.9. Tái thiết cấu trúc 3D

Sau khi hoà tan protein tái tổ hợp cần được tái thiết trạng thái cấu trúc không gian ba chiều để trở lại tình trạng có hoạt tính sinh học. Hoá chất sử dụng để biến tính protein cần được loại để phân tử protein tái tạo cấu trúc và tương tác trong phân tử. Các thông số có ảnh hưởng lên quá trình tái tạo cấu trúc gồm : pH, hoá chất thiol, tốc độ loại bỏ chất biến tính, tỷ lệ tương đối giữa protein của tế bào chủ và protein tái tổ hợp (bảng 6.14).

Bảng 6.14. So sánh các kỹ thuật tái tạo cấu trúc tự nhiên của protein tái tổ hợp biến tính hoà tan.

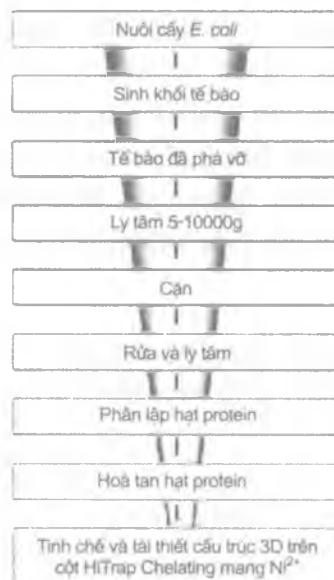
Kỹ thuật	Ưu điểm/Nhược điểm
Thẩm tích nhiều bước	<ul style="list-style-type: none">Mất vài ngàyCần nhiều dịch đậm
Pha loãng về pH trung tính	<ul style="list-style-type: none">Làm loãng nồng độ protein đích
Lọc gel	<ul style="list-style-type: none">ChậmCần có một cột thứ 2Mỗi cột chỉ xử lý được một thể tích nhỏ
Tái thiết cấu trúc trên cột	<ul style="list-style-type: none">Nhanh, đơn giảnKhông hạn chế thể tích mẫu

6.1.10. Tái tạo cấu trúc 3D trên cột

Với loại protein dung hợp hai (His_6) có thể dùng phương pháp đơn giản nhưng hiệu quả cao trong việc tinh chế, tái tạo cấu trúc xếp gấp 3D và hồi phục hoạt tính sinh học (hình 6.4).

Nồng độ cao của các chất biến tính như urea hoặc guanidinium hydrochloride mối liên kết giữa histidine và ion kim loại cố định trên cột được tăng cường. Kết quả là protein dung hợp (His_6) có thể được hoà tan trong dịch chiết biến tính và gắn vào chất mang Sepharose chelate hoá.

Loại bỏ các protein tạp chất và tái tạo cấu trúc 3D khi chuyển sang điều kiện không biến tính ở bước thô protein từ cột. Khi tái xếp gấp phân tử xong, protein dung hợp có thể được tinh sạch lần nữa nếu cần thông qua một kỹ thuật thích hợp.



Hình 6.4. Sơ đồ quy trình tái thiết cấu trúc 3D của protein tái tổ hợp dạng hạt sau khi hòa tan.

6.1.11. Thu hoạch và chiết suất protein tái tổ hợp

Phản này trình bày tổng quát các quy trình thu hoạch và chiết rút protein tái tổ hợp. Mẫu nuôi tế bào phải sạch, không vón cục mới được dùng để thu hoạch và tách chiết protein. Quy trình tách chiết có thể lựa chọn tùy theo nguồn protein như vi khuẩn, thực vật hay động vật, nội bào hay ngoại bào. Lựa chọn phương pháp căn cứ theo thiết bị sẵn có, quy mô thí nghiệm và kiểu cách của mẫu. Bảng 6.15 trình bày một số quy trình tách chiết đã được mô tả.

Cần lưu ý :

- Sử dụng quy trình càng nhẹ càng tốt vì quá trình phá huỷ tế bào, thường dẫn đến tình trạng giải phóng enzyme phân huỷ protein và tăng độ chua của dung dịch.

- Tách chiết phải tiến hành càng nhanh càng tốt ở nhiệt độ lạnh với dịch đậm thích hợp nhằm duy trì độ ổn định của pH, cường độ ion và độ bền của mẫu.
- Mẫu nhớt quá do nucleic acid được giải phóng thì nên xử lý với DNase + RNase.

Bảng 6.15. Các quy trình tách chiết protein tái tổ hợp.

Quy trình tách chiết	Điều kiện điển hình	Nguồn protein	Khuyến cáo
Nhe Lâm vỡ tế bào bằng shock thẩm thấu	2 thể tích nước cho 1 thể tích sinh khối tế bào đã rửa	<ul style="list-style-type: none"> • <i>E. coli</i> • Protein nội bào 	Năng suất thấp nhưng hạn chế được hiện tượng giải phóng protease
Phá tế bào bằng enzyme	Lysozyme 0,2 mg/ml, 37°C, 15 phút	<ul style="list-style-type: none"> • Vi khuẩn • Protein nội bào 	Quy mô phòng thí nghiệm, thường kết hợp với phá tế bào cơ học
Trung bình Nghiền với cát sắc cạnh	Theo hướng dẫn đính kèm thiết bị	<ul style="list-style-type: none"> • Vi khuẩn • Mô thực vật 	
Mạnh Siêu âm, nghiền vi sóng	Theo hướng dẫn đính kèm thiết bị	<ul style="list-style-type: none"> • Huyền phù tế bào, protein nội bào, trong vùng tế bào chất • Dạng hạt 	<ul style="list-style-type: none"> • Quy mô nhỏ • Giải phóng nucleic acid gây hiện tượng tăng độ nhớt • Thể hạt phải được hòa tan
Manton-Gaulin homogeneiser	Theo hướng dẫn đính kèm thiết bị	Huyền phù tế bào	Quy mô lớn
French press	Theo hướng dẫn đính kèm thiết bị	<ul style="list-style-type: none"> • Vi khuẩn • Tế bào thực vật 	
Túi phân đoạn		<ul style="list-style-type: none"> • Protein ngoại bào(tiết ra môi trường) • Dịch nghiên tế bào 	Chất túi phải được hòa tan

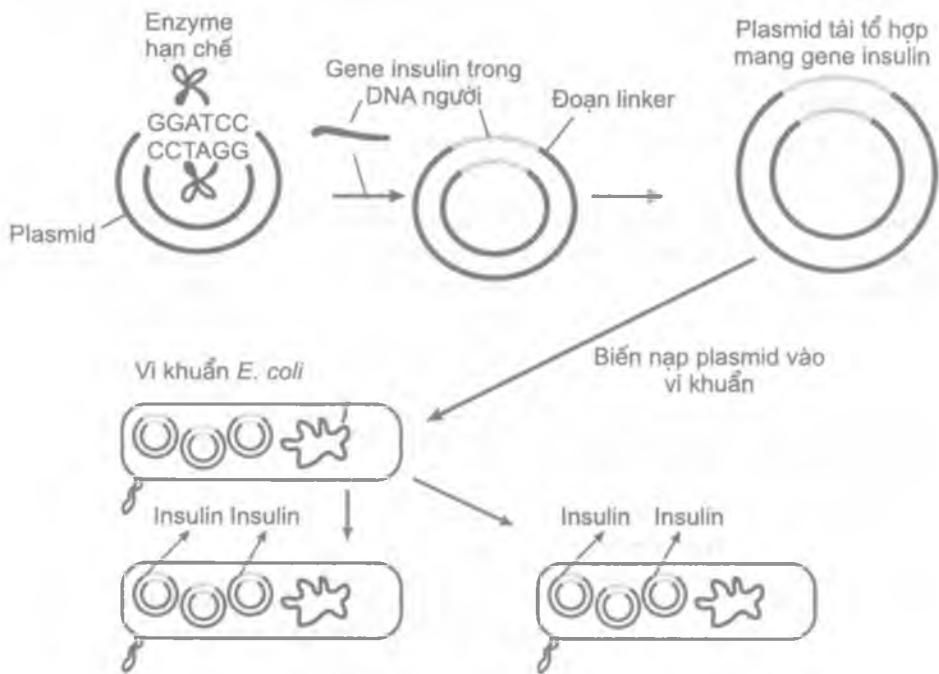
6.1.12. Sản xuất insulin tái tổ hợp

Insulin (Công thức hoá học : $C_{257}H_{383}N_{65}O_{77}S_6$; khối lượng phân tử : 5808) là một loại hormone do các tế bào đảo tụy của tuyến tụy tiết ra với

tác dụng chuyển hoá carbohydrate. Ngoài ra, insulin còn tác dụng đến chuyển hoá mô mỡ và gan thành năng lượng ATP cung cấp cho các hoạt động sống của cơ thể.

Cho đến năm 1920 không có phương pháp sản xuất insulin được biết đến vì hormone mới được phát minh. Nhưng ngay sau khi được phát minh, quá trình sản xuất insulin được giải quyết nhanh chóng vì người ta phát hiện thấy hormone này của bò, lợn kể cả của một số loài cá cũng có hoạt tính trên cơ thể con người. Phương pháp thay thế này được coi là giải pháp đầu tiên cho bệnh tiểu đường type I. Công nghệ tách chiết insulin từ tuyến tụy động vật ngày càng được cải tiến và hoàn thiện.

Tuy nhiên, việc sản xuất và cung cấp insulin vào thị trường thuốc thực sự chuyển sang một thời kỳ mới là nhờ có sự đột phá trong kỹ thuật DNA tái tổ hợp với thành công đầu tiên là insulin tổng hợp của người được sản xuất.



Hình 6.5. Sơ đồ quy trình nhân dòng và biến nạp gene mã hoá phân tử hormone insulin của người vào vi khuẩn *E. coli*.

Năm 1978, đoạn gene mã hoá cho phân tử insulin người được các nhà khoa học thuộc Công ty Genentech tách dòng và chuyển vào tế bào

vì khuẩn *E. coli*. Dòng tế bào vi khuẩn sản sinh insulin này (hình 6.5) được dùng trong nhà máy sản xuất insulin công nghiệp do Công ty Eli Lilly xây dựng đầu tiên tại Anh với mức đầu tư là 40 triệu đôla Úc. Đến năm 1982 thì sản phẩm insulin người sản xuất bằng công nghệ DNA tái tổ hợp được công nhận và thương mại hóa trên thị trường với tên thương phẩm là Humulin.

Hiện nay trên quy mô toàn cầu, Humulin được trên 4 triệu bệnh nhân bệnh tiểu đường sử dụng hằng ngày và một sản phẩm mới có tên là Humalog rất gần gũi với insulin tự nhiên, hoạt động ngay sau bữa ăn đang trở thành dược phẩm nhóm insulin bán chạy nhất.

6.2. TẠO CIỐNG CÂY TRỒNG CHUYỂN GENE

6.2.1. Lựa chọn đối tượng cây trồng

Trên thế giới, người ta tập trung chọn những đối tượng có tính khả thi cao để tiến hành tạo cây trồng chuyển gene theo những tiêu chí sau : (1) Có khả năng thích ứng và thành công cao trong nuôi cấy *in vitro* ; (2) Có gene đáp ứng nhu cầu của thực tiễn như tạo tính kháng bệnh virus, kháng chất diệt cỏ, kháng côn trùng, tăng chất lượng dinh dưỡng của hạt ; (3) Có giá trị thương mại cao, nghĩa là có thị trường tiêu thụ trong nước và xuất khẩu ra nước ngoài và (4) Có khả năng bảo vệ quyền sở hữu trí tuệ khi đưa ra thị trường (như giống ngô lai, giống bông lai).

Ở Việt Nam, nền nông nghiệp nước ta đang phụ thuộc vào thị trường nông sản quốc tế, chủ yếu là thị trường châu Âu, vì thế việc lựa chọn đối tượng triển khai công nghệ chuyển gene đòi hỏi những yêu cầu khác : (1) Đối với những cây trồng qui mô nhỏ, chưa xuất khẩu như khoai lang, đu đủ, có thể hợp tác với các tổ chức nước ngoài để tiếp nhận chuyển giao công nghệ, rút ngắn thời gian từ nghiên cứu đến triển khai ; (2) Đối với những cây trồng quan trọng có giá trị kinh tế cao thì ưu tiên nhóm cây trồng không làm thực phẩm (non food crop) như cây bông vải, cây hoa cảnh và cây lâm nghiệp ; (3) Đối với cây lương thực, thực phẩm cần có sự ưu tiên cho các gene tăng cường chất lượng, các công nghệ chuyển gene sạch, nghĩa là thận trọng trong việc sử dụng các loại gene kháng sinh.

Nội dung mục này sẽ đề cập sự chọn lựa thích hợp với hoàn cảnh kinh tế thương mại của đất nước ta trong giai đoạn hiện nay, đồng thời cũng lưu ý đến khả năng chuyển giao công nghệ từ các tổ chức khoa học công nghệ và doanh nghiệp của các nước phát triển.

6.2.1.1. Các đối tượng cây trồng chuyển gene

Các giống cây trồng thuộc các nhóm cây không làm lương thực như bông vải được dùng là các giống trong tập đoàn giống, các giống đang dùng trong sản xuất đại trà (TM1, VN36P, 254, SB1, C118, LRA, SSR, Cooker) và các dòng bông chuyển gene kháng sâu *cryIA(c)* được sử dụng để xây dựng phương pháp đánh giá cây chuyển gene do Viện Nghiên cứu và Phát triển cây Bông Nha Hố, Ninh Thuận cung cấp.

Cây trồng rừng : chỉ có cây hông được chọn vì đây là loài cây mọc nhanh, có chất lượng gỗ tốt nhưng lại dễ bị sâu ăn lá phá hoại, cụ thể là giống cây gỗ hông (*Paulownia*).

Cây cảnh : chỉ chọn đối tượng là hoa cúc được nuôi cấy trong điều kiện *in vitro*.

Cây lương thực : lấy đối tượng là các giống lúa thuộc loài phụ *indica* (IR64, KDML105, C71) và được sử dụng chuyển gene bằng máy bắn gene và thông qua *Agrobacterium* và các dòng lúa chuyển gene *OAT*, *P5CS*, *TPS*, *nhaA* giống Taipci 309 để thử tính kháng sâu đục thân.

Dù dù là các giống đang được trồng phổ biến, chủ yếu là các giống thuộc giống polo có quả to, ruột vàng.

6.2.1.2. Các loại sinh phẩm, vật tư và hóa chất

Bộ sưu tập các gene và các chủng *Agrobacterium* được sử dụng cho thí nghiệm chuyển gene vào cây trồng có nguồn gốc từ các tổ chức khoa học và công nghệ quốc tế được chuyển giao cho Viện Công nghệ sinh học thông qua *Thỏa thuận chuyển giao nguyên vật liệu* (Materia Transfer Agreement, MTA).

Chủng Bt AB51 có tính kháng sâu non bọ hà đã được Phòng Công nghệ tế bào thực vật, Viện Công nghệ sinh học sàng lọc.

Các kháng thể kháng protein *Cry*, *Vip2,3* có nguồn gốc từ ICGEB, New Dehli và Viện Công nghệ sinh học sản xuất.

Các mồi được tổng hợp dựa trên trình tự các gene nghiên cứu.

Tiến hành phân lập và thiết kế các gene có ích từ nguồn trong và ngoài nước để thực hiện việc chuyển các gene trên vào một số giống cây trồng như cây bông vải, cây hoa cúc, cây trồng rừng như cây hông và cây lương thực như cây lúa nhằm nâng cao tính chống chịu đối với sâu, bệnh (nấm, vi khuẩn và virus) hoặc ngoại cảnh bất lợi (hạn và mặn) và nâng cao chất lượng sản phẩm ; cây đủ đú kháng virus đốm vòng.

6.2.2. Sưu tập nguồn gene khác nhau

Thông qua các tổ chức phi chính phủ, các nước đang phát triển trong đó có Việt Nam, từng bước tiếp cận, ký các thoả thuận được phép sử dụng bản quyền thông qua ký Thỏa thuận chuyển giao nguyên vật liệu MTA về công nghệ và gene với CAMBIA (Australia), Mosanto (Mỹ), Syngenta và Zeneca (Anh), Đại học Ottawa (Canada), Viện Max Plank Golm (Đức). Trong khuôn khổ nghiên cứu các đề tài, Viện Công nghệ sinh học đã sưu tập được các gene có chức năng khác nhau từ nhiều nguồn khác nhau như trình bày ở bảng 6.16.

Bảng 6.16. Danh sách các nhóm gene thu thập từ các nguồn khác nhau.

Nhóm gene	Chức năng sinh học	Nguồn gốc
<i>cry1A(a), cry1A(b), cry1A(c) và cry1A(d)</i>	Protein độc tố diệt sâu	Đại học Ottawa, Canada
ACO antisense	Ức chế tạo ethylene làm chín chậm	MARDI, Malaysia
Xa21	Kháng bạc lá vi khuẩn	UC Davis, Mỹ
Chitinase gene	Kháng bệnh nấm	NFRI Tsukuba, Nhật Bản
<i>P5CS, OAT, TPS, nhaA, HAL</i>	Chống chịu mặn và hạn	FU Brussel, Bỉ ; Đại học Valencia, Tây Ban Nha
CgS, SAT	Tăng hàm lượng amino acid chứa lưu huỳnh	MPIMP Golm, Đức
Công nghệ chín chậm	Qua ACO antisense	Zeneca, Anh
Công nghệ tạo tính kháng virus bằng chuyển gene CP	Dùng cho tạo giống đủ đú kháng virus đốm vòng	Monsanto, Mỹ

6.2.3. Phân lập gene

Có nhiều phương pháp phân lập gene được trình bày trong các báo cáo khoa học cụ thể. Mọi phương pháp phân lập gene đều có mục tiêu là có được phân tử DNA mang trọn vẹn tính trạng đang quan tâm và có thể sử dụng được vào việc chuyển gene để cải tạo giống cây trồng. Trước đây, phần nhiều các phòng thí nghiệm đều bắt đầu bằng việc phân lập protein với chức năng sinh học nhất định, sau đó xác định trình tự amino acid, rồi suy ra trình tự nucleotide của gene. Trên cơ sở đó, người ta tổng hợp một đoạn DNA ngắn làm mồi và dùng kỹ thuật lai phân tử để tìm ra đoạn gene từ genome của đối tượng sinh vật quan tâm.

Trong những năm gần đây các kỹ thuật của phòng thí nghiệm về sinh học phân tử đã được phát triển một cách nhanh chóng và thuận tiện, đặc biệt là từ khi các bộ gene được giải mã hoàn chỉnh và ngành genome học ra đời thì phương pháp phân lập gene có những bước tiến mang tính đột phá với những lĩnh vực như genome học chức năng (functional genomics) ; lập phổ biểu hiện gene (expression profiling)... Dựa trên cơ sở khoa học và nguyên lý hoạt động có thể phân chia các phương pháp phân lập gene thành 3 nhóm chính.

6.2.3.1. Phân lập gene bằng PCR

Đây là phương pháp mới ra đời nhờ hai dữ kiện tiên quyết là : (1) Các Ngân hàng dữ liệu gene trên mạng được công bố mà bất kỳ người nghiên cứu nào cũng có thể truy cập tìm kiếm thông tin về gene được. (2) Phương pháp này cũng chỉ mới được thực hiện từ khi kỹ thuật PCR trở thành thông dụng cho mọi phòng thí nghiệm sinh học phân tử.

Có thể tóm tắt cách thức phân lập gene bằng PCR như sau :

1. Tìm kiếm thông tin về trình tự nucleotide của gene quan tâm trên các Ngân hàng dữ liệu gene quốc tế như : Ngân hàng EMBL : www.ebi.ac.uk/embl ; World Gene Bank : www.ncbi.nlm.nih.gov ; hay DNA Data Bank of Japan (DDJB) www.ddbj.nig.ac.jp.
2. Khi có thông tin về trình tự nucleotide của gene quan tâm thì dùng phần mềm chuyên dụng để thiết kế các cặp mồi cần thiết cho việc nhân bản gene đó. Tiến hành đặt mua hay tổng hợp mồi.
3. Tách chiết DNA tổng số từ đối tượng cần nghiên cứu, trong trường hợp khẳng định được trạng thái biểu hiện của gene quan tâm thì có thể tách mRNA rồi tổng hợp thành cDNA để làm khuôn cho phản ứng PCR.

4. Tiến hành phản ứng PCR với mồi thu được ở bước 2 và DNA khuôn ở bước 3. Khẳng định độ dài của sản phẩm PCR thông qua điện di trên gel agarose có chuẩn phân tử.
5. Dòng hoá sản phẩm PCR vào vector thích hợp như pCR2.1 TOPO hoặc bất kỳ plasmid tương thích khác.
6. Đọc trình tự đoạn gene nhân được, so sánh với số liệu của ngân hàng tìm kiếm được để khẳng định mức độ giống nhau và khác nhau của gene cần phân lập.

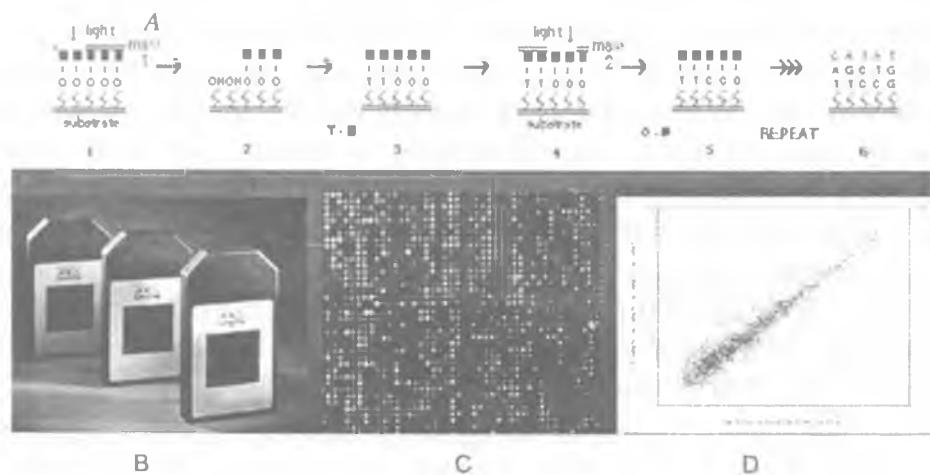
Nhìn chung, phương pháp phân lập gene bằng PCR đơn giản có thể tiến hành ở hầu hết các phòng thí nghiệm sinh học phân tử trang bị tối thiểu. Kết quả phân lập tuy chưa phải là phát minh mới nhưng là sự bổ sung rất quan trọng cho GenBank và điều thiết yếu nhất là người nghiên cứu có được bản gene cần thiết cho công việc tiếp theo.

6.2.3.2. Phân lập gene từ thư viện cDNA

Vào thời điểm chưa có kỹ thuật phân tích toàn bộ sản phẩm phiên mã gene (transcriptomics), các nhà nghiên cứu phải sử dụng kỹ thuật sàng lọc thư viện cDNA (screening cDNA library) là tiền thân của phương pháp phân tích transcriptomic. Nguyên lý của phương pháp sàng lọc thư viện phân tử cDNA là những gene chịu trách nhiệm cho một tính trạng nhất định sẽ hoạt động trong điều kiện nhất định thông qua quá trình phiên mã và dịch mã. Sản phẩm phiên mã là những phân tử mRNA, nếu như sàng lọc và định lượng được chúng thì có xác định được chức năng gene. Đây cũng là một trong hai nguyên lý cơ bản của lĩnh vực nghiên cứu *Genome học chức năng* (Functional Genomics). Dù kích thước genome của các sinh vật rất khác nhau, biến động từ 4000 gene ở *Bacillus* đến 40000 ở con người thì cũng chỉ có khoảng 200 – 6000 gene ở trạng thái hoạt động. Như vậy, thay vì sàng lọc 40000 gene, người ta chỉ cần sàng lọc tối đa từ 6000 phân tử mRNA dưới dạng cDNA sẽ phát hiện được chức năng của những gene đang hoạt động. Các bước tiến hành được trình bày tóm tắt như sau :

1. Chọn hay xử lý sinh vật thí nghiệm ở điều kiện sinh trưởng, phát triển, bệnh lý, stress đang quan tâm. Luôn có đối chứng là không xử lý hoặc không phản ứng.
2. Tách chiết mRNA tổng số.
3. Tổng hợp mRNA thành cDNA bằng bộ kit dùng reverse transcriptase (RT).

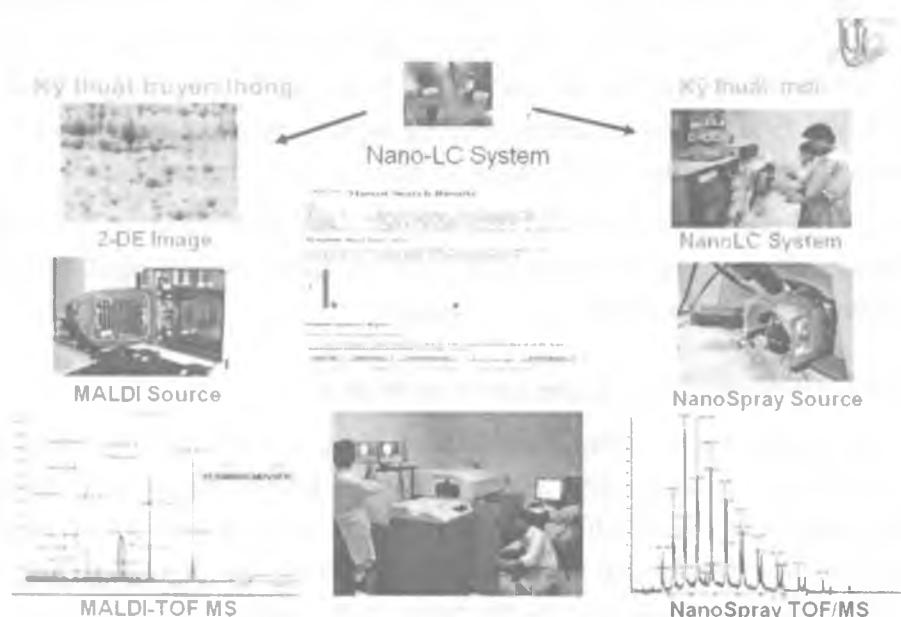
4. Đến bước này nếu theo cách sàng lọc thư viện cDNA thì thực hiện bước 5 → 7 còn nếu có điều kiện thực hiện kỹ thuật microarray thì thực hiện bước 8.
5. Dòng hoá cDNA thành thư viện cDNA bằng plasmid thích hợp.
6. Dùng phương pháp lai phân biệt với môi là toàn bộ mẫu mRNA đối chứng sẽ thu được những dòng không lai, tức là những dòng mới xuất hiện khi xử lý.
7. Đọc trình tự nucleotide các dòng này và kiểm tra chức năng để tìm ra gene.
8. Microarray do các công ty chuyên sản xuất (ví dụ như Affymetrix, Aligent Technologies, Amersham Biosciences, Axon Instruments, BioDiscovery, Clontech, Genomic Solutions, Mergene, Motorola Life Sciences, Nanogene, Partek, Perkin Elmer, Rosetta Inpharmatics, Spotfire, Virtek Vision International) với qui mô từ 1-2000 gene đến 30000 gene giúp phân biệt được những gene đang hoạt động thông qua lai toàn bộ mRNA với array. Máy đọc (Microarray Scaner) cùng phần mềm chuyên dụng sẽ xác định những gene cần tìm đang hoạt động (hình 6.6).



Hình 6.6. Minh họa hoạt động của microarray trong nghiên cứu tìm kiếm chức năng hoạt động của gene và phân lập gene. A : Các bước lai phân tử DNA phối hợp với đánh dấu huỳnh quang. B : Sản phẩm DNA Array bán trên thị trường của hãng Affymetrix ; C : Hình ảnh máy Scanner thu được sau khi lai ; D : Biểu đồ hoá kết quả lai.

6.2.3.3. Phân lập gene thông qua phân tích proteomic

Một trong những lĩnh vực của công nghệ sinh học thời kỳ sau genomic là lĩnh vực proteomic. Ngành Proteome học nghiên cứu cấu trúc và chức năng của toàn bộ hệ protein có trong tế bào, mô hay cả cơ thể sinh vật trong những trạng thái phát triển hay điều kiện môi trường nhất định. Hai tiền đề quan trọng quyết định việc phát triển proteome học là : (1) Trình tự nucleotide toàn bộ genome của nhiều nhóm đối tượng nghiên cứu (trên 100 loài) đã được hoàn tất và (2) Có tổ hợp thiết bị phân tích gồm : thiết bị sắc ký lỏng cao áp (High Pressure Liquid Chromatographer – LC) hoặc cao hơn là sắc ký lỏng nano (Nano LC) kết nối với khai phô MS (Mass Spectrometer) hoạt động theo nguyên lý TOF (Time-of-Flight Mass Analyser) đẩy và tách các phân đoạn chất phân tích qua vi cột capillary. Toàn bộ thiết bị tùy theo cấu hình mà được gọi tắt là MALDI/MS/TOF hay Nano-LC/MS/MS/TOF.



Hình 6.7. So sánh quy trình phân tích proteomic cũ sử dụng điện di gel polyacrylamide hai chiều (2D-PAGE) và mới dùng sắc ký lỏng nano (Nano-LC).

Hình 6.7 so sánh nguyên lý hoạt động của phương pháp phân tích proteomic truyền thống và phương pháp mới cho thấy với kỹ thuật sắc ký

lòng nano, lượng mẫu cần thiết cho phân tích là cực nhỏ, số lượng protein phân biệt được trong một hỗn hợp protein toàn phần là rất lớn. Hiện tại có thể xác định được 1400 loại protein khác nhau có trong một mẫu huyết thanh có dung tích vài trăm microlit.

Bằng cách thức này, những protein mới, lạ là sản phẩm hoạt động của một hay nhiều gene được hoạt hoá dưới điều kiện hoặc trạng thái nhất định sẽ được phát hiện. Qua đó không những trình tự gene được tìm thấy, mà ít nhiều đã có thể khẳng định được chức năng sinh học của những gene đó.

Một số ví dụ cụ thể về phương pháp tiến hành phân lập gene *Pi-2t* và *Pi-1t* từ cây lúa kháng đao ôn hay phân lập gene vip kháng sâu từ *Bacillus thuringiensis* được trình bày trong chương 3.

6.2.4. Các nhóm gene có giá trị đang được sử dụng

Nguồn gene ngoại lai có giá trị để biến nạp vào thực vật có nguồn gốc từ vi khuẩn, siêu vi khuẩn, động vật và từ thực vật khác (Uchimiya et al., 1989). Gene tổng hợp hoá học cũng được biến nạp vào cây trồng (Perlak et al., 1990, 1991). Điều đó chứng tỏ công nghệ gene có sức mạnh vượt xa phương pháp tạo giống truyền thống về phương diện chuyển gene.

Có thể phân loại các nhóm gene chính đang được sử dụng theo chức năng của chúng trong quá trình phát triển cây trồng chuyển gene. Sau đây là những nhóm gene chính.

6.2.4.1. Gene công cụ

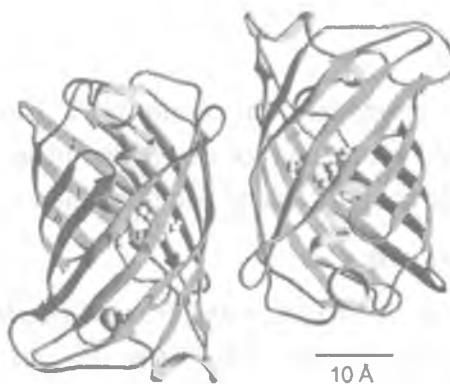
Gene công cụ là những gene được sử dụng làm phương tiện cho việc thực hiện quá trình chuyển gene và tạo cây trồng chuyển gene nhưng không tham gia vào việc hình thành tính trạng mong muốn. Trong số các gene công cụ phải kể đến các gene làm nhiệm vụ chỉ thị và các gene hỗ trợ quá trình chọn lọc tế bào và cây mang gene chuyển.

(I) Gene chỉ thị

Gene chỉ thị là gene có khả năng giúp người nghiên cứu nhận biết những tế bào nào, loại mô nào, dòng cây nào đã tiếp nhận được gene chuyển. Hiện nay có ba loại gene chỉ thị với hai phương thức chỉ thị khác nhau. Đó là :

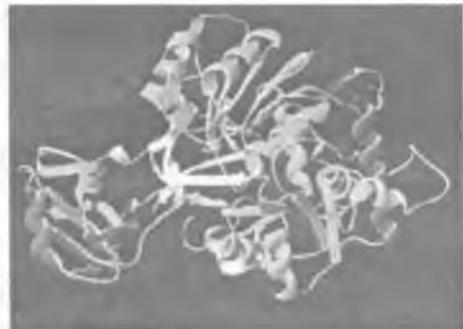
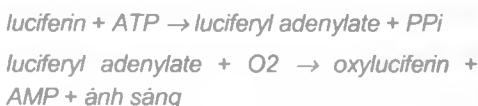
– Gene *gusA* chỉ thị nhuộm màu mô tế bào thông qua phản ứng ôxy hoá cơ chất X-Gluc không màu thành màu xanh lam dưới tác dụng xúc tác của β -glucuronidase (GUS). Hoạt tính GUS trong tế bào, dịch chiết, cặn tế bào còn có thể định lượng được thông qua phản ứng xúc tác cơ chất huỳnh quang MUG (4-methylumbelliferyl- β -D-glucuronide, Research Organics Inc., Cleveland, Ohio) (Jefferson 1987) và đo bằng RF-Mini 150 Fluorometer (Schimadzu ; Columbia, MD) với chất chuẩn là MU (4-methylumbelliferone, Sigma) ở bước sóng 360 và 460 nm.

– Gene *GFP* (Green Flourescent Protein) phát quang màu xanh lục khi soi dưới kính hiển vi huỳnh quang (hình 6.8).



Hình 6.8. Cấu trúc phân tử của GFP (Green Flourescent Protein) với 11 chuỗi tạo thành cấu trúc vỏ hình ống tròn và hai chuỗi (khác màu) tạo thành một cái lồng chứa các nhân phát huỳnh quang flourophore (theo Yang et al. 1995).

– Gene *luciferase* – tên của enzyme xúc tác quá trình phát ra ánh sáng dạ quang sinh học (bioluminescence). Tên gọi này có nghĩa là "Thân mang ánh sáng" (Lucifer) và luciferase có mã số là EC 1.13.12.7 được tách chiết từ con đom đóm (*Photinus pyralis*). Phản ứng dạ quang xảy ra như sau : Sắc tố luciferin bị ôxy hoá với sự có mặt của adenosine triphosphate (ATP) dưới vai trò xúc tác của luciferase tạo ra luciferyl adenylate. Sau đó luciferyl adenylate tác dụng với ôxy để tạo ra oxyluciferin, AMP và ánh sáng (hình 6.9).



Hình 6.9. Cấu trúc phân tử của luciferase và phản ứng ôxy hoá luciferin để tạo ra ánh sáng dạ quang.

Hiện nay, gene *luciferase* rất hay được dùng để chuyển vào động vật và thực vật, coi đó như một bằng chứng cho sự thành công của kỹ thuật di truyền.

(II) Gene chọn lọc

Trong quá trình biến nạp, gene chỉ có một vài tế bào nhận được gene ngoại lai, vì thế phải tiến hành chọn lọc các tế bào nhận gene lạ đó. Hiện nay, có 3 cách cơ bản để chọn lọc các tế bào được biến nạp.

– *Gene kháng kháng sinh :*

Gene aminoglycoside 3'-phosphotransferase II (nptII) : Diễn hình là việc sử dụng những gene mã hoá tính kháng đối với kháng sinh như kanamycin (Herrera-Estrella et al., 1983) hoặc hygromycin (van den Elzen et al., 1985). Tính kháng kanamycin được biết có nguồn gốc từ transposome của vi khuẩn do gene *nptII* quyết định. Các kháng sinh thuộc nhóm aminoglycoside như : neomycin, geneticin và kanamycin đều bị sản phẩm NPTII của gene *nptII* bắt hoạt.

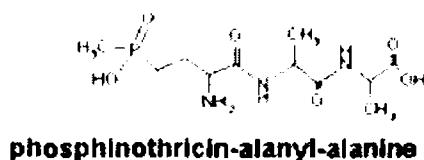
Gene hygromycin phosphotransferase (hpt) : Tính kháng hygromycin B do gene *hpt* quyết định cũng có nguồn gốc từ *Escherichia coli* được cài biến để có thể biểu hiện trong thực vật. Các gene kháng kháng sinh thường không tồn tại trong thực vật, kể cả tảo, vì thế khi bổ sung các kháng sinh này vào môi trường nuôi cấy, chỉ những tế bào đã nhận được gene kháng kháng sinh biến nạp kèm gene tính trạng mới thì chúng có thể sống sót và phát triển, những tế bào còn lại sẽ bị chết.

– Gene kháng thuốc diệt cỏ :

Gene kháng glyphosphate : Thuốc trừ cỏ nhóm glyphosate gồm Roundup, Rodeo, Accord... tác động lên quá trình trao đổi chất của thực vật thông qua ức chế hoạt tính enzyme 3-enolpyruvylshikimate-5-phosphate synthase (EPSPS). Khi hoạt tính của EPSPS bị glyphosate ức chế cây sẽ chết. Tính kháng glyphosate được phân lập từ vi khuẩn đất (*Agrobacterium sp.*) bởi một bản gene EPSPS kháng được glyphosate. Cũng có thể chuyển gene *glyphosate oxidoreductase* (GOX) có nguồn gốc từ vi khuẩn *Achromobacter* có khả năng khử độc glyphosate. Vì thế, gene mã hoá GOX được dùng làm gene chọn lọc. Dùng gene này làm gene chọn lọc không những thuận lợi trong thao tác chuyển gene mà còn làm cho cây chuyển gene có thêm tính trạng kháng được thuốc trừ cỏ trên đồng ruộng.

Gene kháng glufosinate : Glufosinate có tên thương mại là Basta, Liberty, Ignite. Thành phần mang hoạt tính của các thuốc trừ cỏ nhóm glufosinate là *phosphinothricin* có cấu trúc tương tự glutamine. Cấu trúc này cũng giống cấu trúc của bialaphos (PTI), một kháng sinh của *Streptomyces* (hình 6.10). Tác động của phosphinothricin là ức chế glutamine synthase (GS), một enzyme chủ chốt trong chu trình khử độc ammoni hình thành trong quá trình trao đổi chất nitơ trong thực vật. Tính kháng glufosinate được thực hiện thông qua chuyển gene mã hoá phosphinothricin acetyltransferase (PAT) có nguồn gốc từ các loài xạ khuẩn như *Streptomyces hygroscopicus* (Thompson et al., 1987) hay *S. viridochromogenes* (Wohlleben et al., 1988). Gene *bar* (bialaphos resistance) mã hoá phosphinothricin acetyl transferase (PAT) khử độc phosphinothricin làm cho GS không bị ức chế.

Bialaphos



Hình 6.10. Cấu trúc bialaphos.

Gene kháng bromoxynil : Thuốc trừ cỏ có hoạt chất bromoxynil thương mại hoá dưới tên gọi Buctril chủ yếu diệt cây cỏ lá rộng thông qua ánh chế quang hợp. Gene mã hoá bromoxynil nitrilase (BXN) phân lập từ vi khuẩn *Klebsiella pneumoniae* khử độc bromoxynil. Trong trường hợp tế bào, mô nuôi cấy *in vitro* hiệu quả chọn lọc thông qua ánh chế quang hợp rất hạn chế.

6.2.4.2. Gene kháng bệnh cây trồng

Bệnh cây trồng chủ yếu phải kể đến các bệnh do virus, do vi khuẩn và do nấm gây ra. Sau đây là các nhóm gene chuyên dụng để tạo tính kháng bệnh cho cây.

(I) Gene kháng virus

Gene protein vỏ (Coat protein CP) : Có ba kiểu thí nghiệm chuyển gene thành công nhất, đó là tăng tính kháng virus, kháng chất diệt cỏ và kháng côn trùng. Từ nhiều năm, người ta đã biết thực vật có thể kháng đối với các chủng virus gây bệnh nếu trước đó bị nhiễm chung với cùng loại hay với virus tương tự. Hiện tượng này được gọi là bảo vệ chéo (Fulton, 1968). Công nghệ chuyển gene đã thành công trong việc tạo ra tính kháng virus thông qua chuyển gene mã hoá protein vỏ của virus. Đó là công nghệ CP do Monsanto đăng ký bản quyền phát minh. Hình như protein vỏ của thế virus đóng vai trò quyết định đối với tính kháng chéo, bởi vì khi gene mã hoá protein vỏ của nhiều loại virus RNA được thể hiện trong cây thì phản ứng bảo vệ chéo biểu hiện (Powell-Abel et al., 1986 ; Stark và Beachy, 1989 ; Lawson et al., 1990 ; van der Wilk et al., 1991). Cơ chế của bảo vệ chéo còn chưa được giải thích đầy đủ và cũng không nhất thiết phải có sự biểu hiện gene protein vỏ ở mức cao (Hemenway et al., 1988 ; Beachy et al., 1990). Điều lý thú còn chưa được khám phá là liệu protein vỏ của các loại virus DNA có thể tạo nên phản ứng bảo vệ chéo không.

Tính kháng virus còn có thể đạt được khi biến nạp thực vật với DNA có những đoạn mã hoá RNA cực ngắn (microsatellite) (Gerlach et al., 1987 ; Harison et al., 1987). Cũng chưa rõ cơ chế kháng chéo nhưng có thể giả thuyết rằng giữa thế virus gây nhiễm và đoạn RNA vẹt tinh có sự cạnh tranh nhất định đối với các thành phần tế bào chủ.

Gene replicase (Nib) : Phối hợp thêm với gene mã hoá RNA replicase tính kháng virus của cây trồng được chuyển gene tăng lên rõ rệt và mức độ bền vững của tính kháng cũng đảm bảo hơn.

(II) Gene kháng vi khuẩn gây bệnh

Gene Xa21 : Bệnh hại do vi khuẩn gây ra ít có thuốc trừ rét tiên và phổ biến. Nếu dùng các loại kháng sinh như ở người và động vật nuôi thì giá thành rất cao. Vì vậy, việc tạo giống mang tính chống chịu vi khuẩn gây bệnh là hướng nghiên cứu có triển vọng kinh tế cao. Hiện nay có những cố gắng trong việc phân lập các gene kháng bệnh vi khuẩn tự nhiên để dùng cho công nghệ gene. Diễn hình nhất là trường hợp của bệnh bạc lá vi khuẩn lúa do *Xanthomonas oryzae* pv. *oryzae* gây ra. Từ cây lúa đại *Oryza longisminata* có tính kháng tự nhiên do gene *Xa21* quyết định đã được phân lập thành công. Gene *Xa21* mã hoá loại protein kinase có chức năng làm thay đổi hoạt tính của protein (của vi khuẩn gây bệnh). Hiện nay, gene *Xa21* đã được chuyển vào cây lúa thành công và cho thấy cây lúa chuyển gene có khả năng kháng được nhiều nòi *X. oryzae* pv. *oryzae* khác nhau (Wang et al., 1996).

(III) Gene kháng nấm gây bệnh

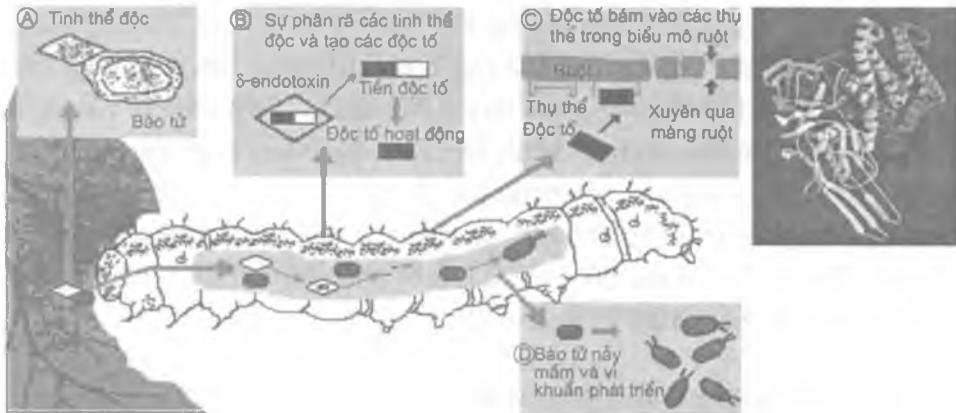
Gene chitinase : Nhiều loài thực vật, động vật và vi sinh vật sinh ra những protein độc đối với nấm. Gene của một số loại protein này được phân lập cho mục đích tăng cường khả năng chống nấm và côn trùng. Ví dụ một loại protein của hạt lúa mạch có khả năng làm mất hoạt tính của ribosome. Khi biểu hiện gene này ở cây thuốc lá thì protein của gene biến nap gây độc đối với nấm bệnh *Rhizoctonia solani* (Tommerup et al., 1991). Nấm này cũng gây nên bệnh bạc lá ở lúa. Ngoài khả năng gây độc cho côn trùng chitinase còn có khả năng khống chế nhiều loại nấm, bởi vì thành tế bào nấm cấu tạo bởi chitin, khi bị phân huỷ các sợi nấm ngừng sinh trưởng (Schlumbaum et al., 1986).

6.2.4.3. Gene kháng côn trùng có hại

Gene mã hoá protein độc tố : Từ những năm 1930, các nhà khoa học Đức đã phát hiện ra loài vi khuẩn *Bacillus thuringiensis* có khả năng tạo ra một loại protein độc dạng tinh thể gọi là δ-endotoxin có thể giết chết nhiều loại côn trùng khi tế bào vi khuẩn tạo bào tử. Tới nay, có khoảng

6000 chủng *B. thuringiensis* đã được phân lập và hàng trăm dạng khác nhau của độc tố được phát hiện. Ký hiệu của một độc tố là hệ Tứ phân, 4 số và chữ, ví dụ như *cryIA(a)* – con số 1 sau tên *cry* đến nay đã là số 51 ; chữ cái A đã là B và C ; chữ cái (a) nay đã là (b), (c) và (d) và con số 1 sau cùng đã lên đến con số 11. Danh sách đầy đủ công bố trong : <http://www.lifesci.sussex.ac.uk/home/Neil-Crickmore/Bt/toxins2.html>.

Vi khuẩn *Bacillus* được dùng trong sản xuất thuốc trừ sâu sinh học trên hầu hết các quốc gia. Bước tiến mới của công nghệ sinh học là chuyển trực tiếp gene *cry* vào cây trồng để chúng tự sản xuất ra độc tố. Khi biểu hiện được gene độc tố của *Bacillus thuringiensis* (Bt) trong cây trồng (Fischhoff et al., 1987 ; Vaeck et al., 1987 ; Perlak et al., 1990 ; 1991) thì có nghĩa là đã chuyển được tính kháng sâu cho chúng. Tuy nhiên, gene *cry* dùng cho cây thường là gene đã được thay đổi mã di truyền để cho chúng hoạt động biểu hiện tốt trong thực vật. Để có hiệu quả các gene này cần được biểu hiện trong bộ phận của cây mà côn trùng (sâu non và bọ trĩ trưởng thành) ưa ăn. Các chất ức chế tác dụng lên bộ phận tiêu hoá. Độc tố Bt gắn vào glycoprotein niêm mạc ruột, đặc biệt là đoạn ruột giữa làm cho ruột bị thủng, rò rỉ dịch tiêu hoá vào đường dẫn máu gây nên hiện tượng bị loãng dịch tiêu hoá và nhiễm trùng máu (Hofte & Whiteley, 1989) ở côn trùng.



Hình 6.11. Cơ chế hoạt động và cấu trúc phân tử của δ-endotoxin do gene cry mã hóa trong vi khuẩn *Bacillus thuringiensis* chủng israelis.

– Ngoài δ-endotoxin, các tế bào vi khuẩn *Bacillus* còn có khả năng tổng hợp những loại protein độc đối với côn trùng ngay trong giai đoạn

sinh trưởng dinh dưỡng của tế bào. Loại protein đó được gọi là protein sinh dưỡng diệt sâu (Vegetative insecticidal protein, Vip). Vip là đối tượng mới mà nhiều công ty sản xuất thuốc diệt sâu đang săn lùng.

– *Gene ức chế enzyme dịch tiêu hoá* : Dịch tiêu hoá của côn trùng chứa amylase để tiêu hoá chất bột và protease để tiêu hoá protein. Có những loại gene mã hoá protein ức chế hoạt tính amylase (AI) và gene mã hoá protein ức chế hoạt tính trypsin (TI). Khi sử dụng những gene này để chuyển vào cây trồng cũng thu được kết quả tạo ra tính kháng côn trùng ở chúng (Hilder et al., 1987 ; Johnson et al., 1989). Chất ức chế amylase có thể rất độc đối với côn trùng do khả năng kèm hãm tiêu hoá tinh bột. Trong tự nhiên, các chất ức chế proteinase và α-amylase thường được tạo ra ở các vết thương như một phản ứng tự vệ của cây chống lại côn trùng (Bowles, 1990).

– *Gene mã hoá lectin (GNA)* : Lectin là một phức hợp nhiều protein khác nhau về cấu trúc phân tử và đặc tính sinh hoá và tính đặc hiệu trong liên kết với carbohydrate. Loại lectin gắn kết mannose (*Galanthus nivalis* agglutinin, GNA) được xác định là độc với nhiều loài côn trùng thuộc các bộ Homoptera, Coleoptera và Lepidoptera. Năm 1991, gene *gna* được phân lập, biến đổi và biến nạp vào nhiều loại cây trồng khác nhau nhằm tạo ra tính ngán ăn đối với côn trùng trích hút. Tác dụng độc đối với côn trùng của lectin là do chúng có tác dụng tiêu huỷ đối với glycoprotein của đường tiêu hoá (Chrispeels và Raikhel, 1991).

– *Gene mã hoá chitinase* : Chitinase có tác dụng độc đối với côn trùng nếu nó phát huy được tác dụng phân huỷ lớp chitin bảo vệ niêm mạc khi kết hợp với những gene khác tính kháng côn trùng cùng được cải tiến.

6.2.4.4. Gene bất dục đực

Ưu thế lai đã chứng tỏ giá trị to lớn của nó ở một số loài cây trồng giao phấn. Việc sản xuất hạt ngô lai đang bán trên thị trường là dự hoàn toàn trên hiện tượng này. Việc khai thác hiện tượng bất dục đực cho những loài cây trồng tự phôi như lúa thì còn tương đối mới. Bất dục đực tế bào chất (CMS) là chìa khóa cho việc sản xuất hạt giống lai thường phẩm và không ít những nghiên cứu đã được tiến hành nhằm tìm kiém

hiện tượng tương tự ở các loài cây trồng khác nhau. Hiện nay, có nhiều hy vọng tạo ra được hệ thống nhân tạo điều khiển tính bất dục đực thuận nghịch do nhân tế bào chỉ phối bằng công nghệ gene.

Một hướng tiếp cận khác là dùng gene mã hoá ribonuclease của vi khuẩn (Barnase) được điều khiển bởi đoạn khởi động đặc trưng cho tế bào nuôi (tapetum) của túi phấn. Kết quả là các tế bào này bị chết và cây đó trở nên bất dục đực (Mariani et al., 1990). Để hoàn thiện hệ thống này cần tìm ra cơ chế để làm ngược tính bất thụ, nghĩa là khôi phục lại được tính hữu thụ. Người ta đang hy vọng có thể đạt được bằng biểu hiện có điều khiển gene tạo chất ức chế đối với Barnase.

6.2.4.5. Gene ngược chống chín nhũn và làm chín chậm

Gene ngược (antisense) là những phân tử tương tác lên sợi bổ sung của acid nucleic làm thay đổi quá trình biểu hiện gene. Một trong hai sợi của phân tử DNA được dùng làm khuôn để tổng hợp ra phân tử RNA bổ sung. Nếu sợi khuôn có trình tự ngược chiều thì sợi RNA sẽ có trình tự thuận chiều. Vì phân tử DNA có hai sợi nên sợi còn lại được gọi là sợi không phiên mã, có trình tự bổ sung với sợi phiên mã, hay thuận chiều với sợi mRNA. Cụ thể như sau :

DNA sợi 1 : sợi thuận chiều (sense strand).

DNA sợi 2 : sợi ngược chiều (antisense strand) phiên mã thành → sợi RNA thuận chiều (sense).

Gene ngược đang mở ra một hướng mới cho việc cải tiến giống cây trồng được ứng dụng nhiều đối với các tính trạng khác nhau, chủ yếu liên quan đến chất lượng chế biến và chất lượng bảo quản sản phẩm. Nguyên lý hoạt động của gene ngược là một đoạn phân tử RNA có trình tự nucleotide ngược chiều với trình tự nucleotide của sợi DNA giữ vai trò sense của gene được tổng hợp thành cDNA rồi chuyển vào thực vật. Tại đó, đoạn antisense này sẽ được phiên mã thành một đoạn mRNA bổ sung với sản phẩm mRNA của gene cần bị ức chế. Hai phân tử mRNA này kết hợp thành sợi đôi thông qua kết cặp các base bổ sung, là cho mRNA không được dịch mã ở bộ máy ribosome và sản phẩm protein của gene không được tạo ra.

Thời gian bảo quản của cà chua phụ thuộc vào tốc độ chín. Những nghiên cứu sinh hoá cho thấy hoạt tính polygalacturonase (PG)

[poly(1,4- α -D-galacturonide)-glucanohydrolase] quyết định quá trình thuỷ phân pectin và galacturonan trong vỏ quả xanh làm cho chúng mềm khi chín. Thông qua tác động của gene ngược dưới dạng phân tử RNA ngược chiều, quá trình sinh tổng hợp PG bị úc chế. Quả vẫn chín nhưng chậm lại và đặc biệt là không bị nhũn. Tuy nhiên, trước khi thuỷ phân, pectin cần được chuyển hoá thành trạng thái demethylester nhờ pectin methylesterase (PME). Tiến hành dùng kỹ thuật chuyển gene ngược úc chế PME làm tăng độ cứng của quả trong quá trình chín.

Quá trình chín của quả thường liên quan đến hoạt động sinh tổng hợp ethylen. Chuyển các antisense của các gene 1-aminocyclopropane-1-carboxylate (ACC) synthase và ACC oxidase (ACO) vào cây trồng có thể úc chế quá trình hình thành ethylen, dẫn đến sự làm chậm quá trình chín của quả (delayed rippening). Công nghệ này góp phần tăng khả năng vận chuyển đi xa của nhiều loại quả nhiệt đới như du dù, xoài, đồng thời cũng làm tăng tuổi hàng hoá khi được bày bán trên giá hoa quả của các siêu thị.

Chalcone synthase (CHS) là một enzyme trong sinh tổng hợp flavonoide và rất cần cho quá trình tạo sắc tố hoa. Gene ngược của CHS được dùng để hạn chế hoạt tính của CHS và làm thay đổi màu hoa ở cây *Petunia* và cây thuốc lá (van der Krol et al., 1988).

6.2.4.6. Gene chống chịu điều kiện sinh thái bất lợi

Có hai xu hướng tìm kiếm các gene liên quan đến tính chống chịu của cây trồng đối với điều kiện sinh thái phi sinh học bất lợi. Xu hướng thứ nhất là tìm những gene liên quan trực tiếp đến khả năng chịu hạn, chịu lạnh, ngập úng hay phèn mặn. Hướng này đã thu được những kết quả nhất định. Tuy vậy, trong những năm gần đây, cùng với sự phát triển những kỹ thuật phân tích mạnh hơn như đọc trình tự gene, phân tích khói phổ, phân tích microarray và các công cụ xử lý số liệu của tin sinh học mạnh hơn cho phép người ta hướng những nghiên cứu về gene chống chịu theo hướng các yếu tố điều khiển, ví dụ như các transcription factor (TF) nhằm tiếp cận cách điều hoà tổng thể hơn của thực vật khi gặp điều kiện bất lợi.

(I) Gene liên quan đến tính chịu lạnh

Khi tiếp xúc với nhiệt độ thấp, một số gene được hoạt hoá sinh tổng hợp ra những protein mới (Ito et al., 1994). Hiện tượng này được phát hiện

trên hàng loạt đối tượng khác nhau, ví dụ ở lá rau chân vịt (Anderson et al., 1994), cây mạ non 3 lá (Binh et al., 1989, 1992 ; Hahn et al., 1989), ở cây *Mesembryanthemum* (Bohnert et al., 1988), ở lúa mạch (Cattivelli et al., 1989, 1990), ở cỏ *Dactylis glomerata* (Yoshida, 1984), ở cải dâu (Jiang et al., 1996), ở tảo lục (Sato, 1994), ở đậu tương (Takahashi, 1996). Tuy nhiên, từ chỗ phát hiện ra hoạt tính gene đến việc phân lập được gene thì còn là một bước đi dài. Cho đến nay mới có những trường hợp sau, gene liên quan đến tính chịu lạnh được phân lập và mô tả : Nhóm gene *cor* từ cây *Arabidopsis* (Gilmour et al., 1992, 1988 ; Hajela et al., 1990 ; Kurkela et al., 1992 ; Li et al., 1994 ; Uemura et al., 1996 ; Webb et al., 1996), nhóm gene *cas* từ cây cỏ ba lá (Mohapatra et al., 1988, 1989 ; Wolfraim et al., 1993 ; Wolfraim et al., 1993). Hầu hết, các gene này đều được chứng minh là liên quan đến tính chịu băng giá. Riêng nhóm gene liên quan đến tính chịu lạnh của vùng nhiệt độ không đóng băng thì hầu như không được nghiên cứu. Những kết quả của chúng tôi liên quan đến gene chịu lạnh của lúa được trình bày trong công trình này là những kết quả mới như kết quả thẩm định của Ngân hàng dữ liệu gene thế giới khẳng định.

(II) Gene chịu hạn

Việc tìm kiếm gene chịu hạn được nhiều phòng thí nghiệm chú ý trên nhiều đối tượng khác nhau. Ví dụ : cỏ ba lá (Laberge et al., 1993), lúa mỳ (Labhilili et al., 1995), lúa mạch (Singh et al., 1989, 1973 ; Xu et al., 1996), ngô (Vilardell et al., 1990), *Arabidopsis* (Welin et al. 1994), lúa nước (Claes et al., 1990 ; Morgan, 1984 ; Mullet, 1988 ; Mundy et al., 1988 ; Yamaguchi-Shinozaki et al., 1989). Hai nhóm gene được mô tả chi tiết là nhóm LEA của cây lúa mạch (Singh et al., 1989 ; Singh et al., 1973 ; Xu et al., 1996) và nhóm RAB của cây lúa nước (Claes et al., 1990 ; Morgan, 1984 ; Mullet, 1988 ; Mundy et al., 1988 ; Yamaguchi-Shinozaki et al., 1989). Nhóm thứ nhất mã hoá những protein xuất hiện vào thời kỳ muộn của quá trình hình thành phôi sau khi thụ phấn và đó là những protein góp phần bảo vệ phôi trong quá trình ngủ và chịu hạn của phôi trong hạt khô. Nhóm thứ hai liên quan đến tác động ức chế sinh trưởng tế bào của ABA khi thực vật gặp điều kiện bất lợi. Các gene RAB phản ứng với ABA ngoại sinh và nội sinh cho ra những protein có chức năng ức chế và bảo vệ. Đây là hai nhóm gene có triển vọng nhất đang được lưu ý sử dụng trong công nghệ gene.

(III) Gene liên quan đến tính chịu mặn

Một dòng cDNA liên quan đến tính chịu mặn của tế bào thuốc lá, các gene liên quan đến sinh tổng hợp proline như GAT đang là mối quan tâm của nhiều phòng thí nghiệm với hy vọng gia tăng sinh tổng hợp proline là gia tăng được tính chịu mặn. Hoạt động của gene *ornithine aminotransferase (OAT)*, một gene chủ chốt trong chu trình sinh tổng hợp proline của cây *Arabidopsis*, tăng lên mạnh mẽ khi cây bị xử lý muối (Tran Thanh Thu, 1996) và gene *OAT* đã được phân lập cho mục đích nghiên cứu chuyển gene.

(IV) Gene liên quan đến tính chịu độc tố nhôm

Đã có những nghiên cứu di truyền khẳng định tính chịu nhôm do một gene kiểm soát. Gene *Alt1* sản sinh ra những protein liên quan đến hiện tượng tế bào đầu rẽ cây chịu độc tố nhôm tiết ra các chất hữu cơ. Các cố gắng bước đầu để phân lập và xác định gene *Alt1* này đang được triển khai (Randall, 1994). Đây là tiền đề rất quan trọng cho kỹ thuật gene nhằm nâng cao tính chịu nhôm.

(V) Gene liên quan đến tính chịu ngập úng

Gene alcoledhydrogenase giúp cho cây có được năng lượng ATP trong điều kiện thiếu oxy khi cây bị ngập nước. Ngoài ra, các gene đặc trưng như gene làm cho lóng thân dài nhanh khi bị ngập nước (rất rõ ở lúa nếp) cũng đang được tiến hành lập bản đồ gene nhằm tiến đến mục tiêu phân lập gene (Vanavichit, 1997).

6.2.4.7. Gene có giá trị y dược

Xu thế nghiên cứu phát triển cơ thể sinh vật thành lò phản ứng sinh học (bioreactor) để sản xuất các hoạt chất có giá trị y dược đang được chú ý nhiều ở các quốc gia có nền Công nghệ sinh học phát triển mạnh. Sau đây là liệt kê những nhóm gene đang được chú ý.

(I) Gene tạo kháng thể

Một khả năng rất hấp dẫn trong công nghệ gene thực vật là khai thác tính đặc hiệu của kháng thể để ức chế hoặc làm gián đoạn một quá trình, chẳng hạn như quá trình virus nhân trong tế bào và truyền giữa các tế bào. Đã có những thành công được trình bày (Hiatt et al., 1989 ; During et al.,

1990). Bởi vì kháng thể bao gồm 2 tiểu phần : Chuỗi nhẹ và chuỗi nặng, vì thế cần phải tạo ra được cả hai chuỗi trong cây để có kháng thể hoàn chỉnh. Hiatt et al. (1989) đã chuyển vào một dòng thuốc lá gene chuỗi nhẹ và một dòng khác là gene chuỗi nặng. Cả hai dòng đều tạo được sản phẩm protein, nhưng chỉ có con lai của chúng mới có biểu hiện có kháng thể với đầy đủ chức năng của cả hai chuỗi. During et al. (1990) đã biến nạp cây thuốc lá với cả hai gene của hai chuỗi, nhưng chịu sự điều khiển của hai đoạn khởi động khác nhau và đã thu được kháng thể tích luỹ trong tế bào. Tương tự Chaudhary et al. (1990) đã dung hợp cả hai gene chuỗi ngắn và chuỗi dài chung vào một đoạn dịch thông trước khi biến nạp và vẫn phát hiện được kháng thể có hoạt tính.

(II) Gene tạo kháng nguyên

Các protein do những đoạn epitope của genom virus khác nhau đang được sử dụng để chuyển vào cây trồng nhằm thu được sản phẩm là các protein kháng nguyên có thể sử dụng để phát triển các loại vaccine thực vật ăn được (plant edible vaccine). Hiện nay, gene protein vỏ (CP, G) của các loại virus như virus dại, virus cúm A, virus viêm não Nhật Bản đang được chuyển vào cây trồng (xem Tổng quan Lê Trần Bình và Cao Huyền Trang, 2005).

6.2.4.8. Biến đổi gene nhằm tăng cường sự biểu hiện

Có nhiều thử nghiệm nhằm tăng cường quá trình biểu hiện gene chuyển trong thực vật chủ. Các thử nghiệm đó liên quan đến tìm kiếm đoạn điều khiển (promoter) mạnh, tìm kiếm các đoạn tăng cường (enhancer) hay gắn thêm các yếu tố liên quan đến quá trình phiên mã và dịch mã. Sau đây là một vài ví dụ chọn lọc.

(I) Điểm gắn ribosome

Gene có thể được biến đổi có thêm điểm gắn ribosome cần thiết trên phân tử mRNA để có hoạt động biểu hiện cao hơn. Ở vi khuẩn, đoạn gắn ribosome (chuỗi Shine-Dalgarno) được mô tả rất rõ nhưng ở thực vật thì khác hơn và thường không thống nhất. Từ 75 đoạn DNA nhân của thực vật bậc cao được công bố (Joshi, 1987) tìm thấy đoạn bảo thủ của codon khởi đầu ATG là TAAACAATGGCT. Hình như nucleotide thứ 3 trước ATG có một vai trò quan trọng nhất định, đó thường là một base dạng

purin (cụ thể là A) đối với những gene mã hoá một protein được biểu hiện mạnh (Heidecker và Messing, 1986 ; Joshi, 1987 ; Kozak, 1989). Hiệu suất dịch mã được tăng cường khi ghép thêm một đoạn dẫn đầu không dịch mã của virus và giữa promoter với đoạn mã hoá của gene (Gallie et al., 1987 ; Jobling và Gehrke, 1987).

(II) Đoạn intron

Một hiện tượng khá hấp dẫn là hiệu quả của các đoạn intron đối với sự biểu hiện gene lạ trong cây một lá mầm. Callis et al. (1987) phát hiện thấy intron làm tăng biểu hiện gene trong nuôi cấy tế bào ngô. Tanaka et al. (1990) thông báo mức độ biểu hiện cao hơn 90 lần của gene *GUS* trong cây lúa chuyển gene khi gắn thêm đoạn intron của gene catalase vào gene *GUS* của cây đậu trắng. Sự gia tăng biểu hiện gắn liền với gia tăng mật độ mRNA và tạo khuyễn đối với đoạn intron. Hiện tượng gia tăng này lại không xảy ra trong cây thuốc lá chuyển gene. Ở cây lúa chuyển gene, biểu hiện gene tăng lên khi ghép đoạn intron thứ nhất của gene *Act1* actin vào giữa promoter CaMV 35S và gene *GUS* (Wu et al., 1991). Maas et al. (1991) nghiên cứu tác động của exon 1 và intron 1 ở gene *Shrunken-1* của cây ngô lên biểu hiện tạm thời của gene *CAT* trong mô ngô, lá lúa chuyển gene. Đoạn exon làm tăng 10 lần, đoạn intron làm tăng 100 lần, còn khi tổ hợp cả hai đoạn exon và intron thì hiệu quả gia tăng đạt 1000 lần khi ghép chúng vào giữa promoter 35S với gene *CAT*. Tuy nhiên, ở tế bào trân của cây thuốc lá, tác động của đoạn exon là kích thích còn của đoạn intron lại là kìm hãm. Hiện tại vẫn chưa rõ vì sao cây một lá mầm và cây hai lá mầm lại phản ứng khác nhau đối với sự có mặt của intron như vậy.

(III) Điểm glycosyl hoá

Nhiều protein khi xâm nhập vào hệ thống lưới nội bào thường chịu tác động glycosyl hoá ở đoạn amino acid ngắn thông qua enzyme glycosyl transferase. Ở những protein không bao giờ xâm nhập lưới nội bào cũng phát hiện thấy chuỗi amino acid tương tự. Khi chuyển gene như vậy vào thực vật có thể xảy ra hiện tượng glycosyl hoá những chỗ trước đây không bị glycosyl hoá và lúc đó sản phẩm protein này bị mất hoạt tính. Ví dụ, trường hợp của gene β -glucuronidase của *E. coli* (Jefferson et al., 1986) được xác định là có hai vị trí bị glycosyl hoá : Đó là vị trí

N358LS và N423IS. Gene *GUS* biểu hiện tương đối mạnh khi enzyme hình thành trong tế bào chất (Jefferson et al., 1987), trong lục lạp (Kavanagh et al., 1988) hoặc trong ty thể (Schmitz và Lonsdale, 1989) nhưng nếu enzyme này được chuyển đến lưỡi nội bào thì bị glycosyl hoá, kết quả là mất hoạt tính và nhuộm kém với cơ chất X-gluc (Iturriaga et al., 1989). Vấn đề này được giải quyết thông qua thay đổi hai điểm N358 bằng đột biến thành serine, threonine hoặc proline và N423 biến mất ngẫu nhiên (Farrell và Beachy, 1990). Lúc này, protein GUS được vận chuyển qua lưỡi nội bào tới không bào mà vẫn duy trì được hoạt tính cao.

(IV) Cặp đôi AT/GC

Gene độc tố BT cần phải biến đổi nhiều (Perlak et al., 1990 ; 1991). Có hai gene *cryIIA(b)* và *cryIIA(c)* thể hiện rất yếu trong cây thuốc lá và cây cà chua chuyển gene. Có 3 vấn đề chính liên quan : (1) Mã cặp đôi ; (2) Chuỗi tín hiệu polyadenyl hoá không đúng chỗ ; (3) mRNA không bền vững. Cả ba vấn đề đều cùng là hậu quả của một nguyên nhân : Tỉ lệ AT cao. Mã giàu AT vẫn thường có ở những gene độc tố nhưng rất hiếm ở thực vật (ví dụ TTA là mã của leucine). Vì thế, khi dịch mã một gene là giàu TTA ở thực vật thường có hiện tượng ngừng trong bộ máy ribosome làm cho sinh tổng hợp protein bị kết thúc sớm. Hiện tượng này xảy ra với gene BT vì trong đó có nhiều chuỗi AATTAA dạng tín hiệu polyadenyl hoá (Dean et al., 1986). Ngoài ra chuỗi ATTTA được xác định là làm giảm độ bền của mRNA. Murray et al. (1991) đã chứng minh được mRNA của gene *cryIA(b)* không bền trong tế bào trân và cây cà rốt chuyển gene. Khi thay 21% các gốc base 1845 bp của gene này bằng gốc phù hợp mà không làm thay đổi trình tự amino acid được mã hoá thông qua con đường hoá tổng hợp thì Perlak et al. (1991) đã thu được 10 lần gia tăng sự thể hiện gene *cryIA(b)* và 100 lần gene *cryIA(c)*, qua đó làm cho cây chuyển gene trở nên độc hơn nhiều đối với côn trùng.

Dịnh vị trong tế bào : Mỗi loại protein thực vật thường có vị trí hoạt động riêng, rất đặc trưng trong tế bào. Phần lớn hoạt động trong tế bào chất. Những loại protein đó thường được sinh tổng hợp bởi ribosome tự do trong tế bào chất và có đoạn tín hiệu đặc biệt để được giữ lại trong tế bào chất. Nhiều loại protein được tổng hợp trong tế bào chất, nhưng lại được chuyển vào lục lạp, ty thể hoặc nhân. Phải có đoạn tín hiệu đặc biệt cho sự định hướng đó (Boutry et al., 1987 ; Von Heijne et al., 1989 ;

Lassner et al., 1991). Ngoài ra, hiện tượng định hướng trong cơ quan tử, ví dụ như trong các khoang của lục lạp cũng đòi hỏi phải có đoạn tín hiệu (Keegstra et al., 1989). Tín hiệu hướng về lục lạp và ty thể thường nằm ở đầu N của chuỗi protein và mất đi khi protein xâm nhập vào trong cơ quan tử. Các loại protein nhân thường quá lớn để có thể khuếch tán tự do qua các lỗ màng nhân cho nên khi có một đoạn tín hiệu ngắn nằm ở giữa phân tử. Chuỗi ngắn gồm 7 amino acid là Pro-Lys-Lys-Lys-Arg-Lys-Val của phân tử kháng nguyên T của virus Simian 40 được chứng minh là tín hiệu đủ cho nhiều loại protein khác nhau có thể xâm nhập nhân tế bào động vật có vú (Kalderon et al., 1984). Cũng chuỗi đó có khả năng chuyển được enzyme RNA polymerase (100 kDa) của thực khuẩn T7 (Lassner et al., 1991) và β-glucuronidase (68 kDa) (van der Krol và Chua, 1991) vào nhân của tế bào thuốc lá chuyển gene. Các gốc kiềm cao liên kết DNA của 3 loại protein nhân thực vật được coi là yếu tố thúc đẩy quá trình chuyển nháp β-glucuronidase vào trong nhân của cây thuốc lá chuyển gene (van der Krol và Chua, 1991). Có thể chúng có liên quan đến tính hướng nhân của protein B-ZIP.

Nhiều loại protein được tổng hợp trong ribosome ở tế bào chất cho nên lúc đầu hoàn toàn tự do, nhưng về sau nhanh chóng liên kết với lưới nội màng ở đầu N của chuỗi peptide đang lớn dần. Chính đầu N của chuỗi peptide tạo nên tín hiệu để xâm nhập vào lưới nội bào (Burr và Burr, 1981 ; von Heijne, 1985 ; Torrent et al., 1986 ; DellaPenna và Bennett, 1988). Đoạn peptide tín hiệu dài 20-25 gốc amino acid bao gồm 3 vùng được đặt tên là n, h và c. Vùng n có 3-5 gốc chứa một gốc diện tích dương duy nhất ; vùng h có 7-15 gốc kị nước và vùng c có 3-5 gốc bé trung tính, nhưng rất quan trọng được nhận biết bởi một loại protease và cắt rời đoạn peptide tín hiệu ngay sau khi phân tử protein xâm nhập vào lưới nội bào. Protein không có chuỗi peptide tín hiệu sẽ được vận chuyển theo một cơ chế chọn trước tới lưới nội bào rồi bị tiết ra ngoài (Denecke et al., 1990). Còn những protein có thêm một đoạn tín hiệu khác nữa thì được tích luỹ lại trong các hạt protein, trong lysosome, trong không bào, trong peroxysome và trong các thể hạt khác có nguồn gốc lưới nội bào (Chrispeels, 1991 ; Dorel et al., 1989 ; Saalbach et al., 1991). Có một số ít protein được mã hoá và tổng hợp ngay trong các loại cơ quan tử như lục lạp và ty thể (Mullet, 1988 ; Newton, 1988).

6.2.5. Phương pháp biến nạp gene vào thực vật

Genom của thực vật có các phần khác nhau, phần chính nằm ở nhân, một phần nhỏ nằm ở lục lạp và một phần nhỏ nữa nằm ở tì thể. Hầu hết các nghiên cứu chuyển gene đều thử nghiệm việc đưa gene lạ vào bộ gene của nhân. Tuy nhiên, gần đây càng có nhiều quan tâm đổi mới việc chuyển gene cơ quan tử vì lý do an toàn sinh học. Gene lục nạp được biến nạp ở tảo lục *Chlamydomonas* (Boynton et al., 1988 ; Blowers et al., 1989). Chuyển gene ty thể thành công ở nấm men (Butow và Fox, 1990).

6.2.5.1. Điều kiện cho chuyển gene ở thực vật

Các điều kiện cần thiết cho việc chuyển gene thực vật là yếu tố mang gene cần chuyển (vector), gene cần chuyển, đoạn khởi động (promoter) và đoạn kết thúc (terminator). Vector cần có những bộ phận sau : Một đoạn khởi đầu cho tái bản ; một gene chỉ thị chọn lọc để thao tác trong *E. coli* và một gene chỉ thị khác được khống chế bởi một promoter mạnh của thực vật phục vụ cho việc chọn lọc ở thực vật. Plasmid T-DNA được dùng trong chuyển gene thông qua *Agrobacterium* được gắn thêm một đoạn khởi đầu, một đôi đoạn nháy lại của T-DNA biên trái và biên phải trợ giúp cho plasmid được ghép nối vào DNA cây chủ (Horsch và Klee, 1986). Đoạn nháy lại biên phải bắt buộc phải có (Wang et al., 1984). Trong trường hợp biến nạp gene lục lạp, người ta dùng đoạn khởi đầu tái bản của DNA lục lạp nhằm giúp cho gene biến nạp tồn tại trong một thể tái bản độc lập trong lục lạp của tế bào chủ.

Có hai loại promoter dùng cho chuyển gene thực vật, đó là promoter thường trực (constructive promoter) như CaMV 35S hoạt động thường xuyên trong mọi bộ phận, mọi thời điểm và mọi hoàn cảnh ở cây một lá mầm và cây hai lá mầm (Jefferson et al., 1987 ; Benfey và Chua 1989 ; Battaraw và Hall, 1990). Đoạn khởi động như vậy rất cần cho những nghiên cứu đầu tiên về cây chuyển gene. Loại thứ hai là promoter chỉ hoạt động ở từng loại mô (specific promoter), ví dụ như ở củ, ở hạt ; ở từng thời điểm (development stage depending promoter) ví dụ như : khi ra hoa, khi ngủ nghỉ và hoàn cảnh nhất định (inducible promoter) (Kuhlenmeier et al., 1989) như sốc nhiệt (Baumann et al., 1987), ánh sáng (Kuhlenmeier et al., 1989), bị vết thương (Keil, 1990), bị nấm (Schmid et al., 1990). Có nhiều loại promoter như vậy đã được biết đến và cũng có khá nhiều phương pháp khác nhau để phát hiện ra những loại

promoter với những tính trạng lý thú (Kerbundit et al. S, 1991). Promoter thường có cấu trúc khá đồng bộ cho phép nhận được những mức độ thể hiện gene khác nhau thông qua hoán vị và tổ hợp các vùng khác nhau của một promoter. Điều đó thể hiện rất rõ ở promoter CaMV 35S (Benfey và Chua, 1989). Hoạt lực của promoter có thể được tăng cường khi bố trí hai hoặc nhiều promoter nối tiếp hoặc tổ hợp với nhau (Comai et al., 1989). Đoạn kết thúc phiên mã của thực vật thường chứa một tín hiệu polyadenyl hoá đó là AATAAA ; ATTAA hoặc AACCAA (Dean et al., 1986 ; Joshi, 1987). Hai đoạn kết thường dùng trong biến nạp thực vật là đoạn kết của CaMV 35S và đoạn kết nos. Đầu polyA có tác dụng bảo vệ các phân tử mRNA bền vững chống lại sự phân huỷ (Gallie et al., 1989). Gene chọn lọc giúp cho việc nhận biết thể biến nạp được dễ dàng (Dekeyser et al., 1989).

6.2.5.2. Các phương pháp chuyển gene vào thực vật

Có nhiều phương pháp biến nạp gene áp dụng cho thực vật, tuy nhiên, những phương pháp có giá trị thực tiễn quan trọng chỉ hạn chế trong hai nhóm phương pháp chính. Đó là (1) biến nạp trực tiếp DNA bằng tác nhân hoá học như polyethylene glycol (PEG), xung điện (electroporation) và súng bắn gene (gene gun) và (2) biến nạp DNA gián tiếp thông qua vi khuẩn *Agrobacterium tumefaciens*.

(I) Chuyển gene trực tiếp bằng hoá chất

Trong nhiều năm, kỹ thuật chuyển gene thông qua tiếp nhận DNA tinh chế vào tế bào trân là giải pháp kỹ thuật quan trọng của công nghệ gene thực vật. Thuốc lá và hoa mào gà là hai đối tượng được sử dụng đầu tiên cho chuyển gene trực tiếp nhờ tác động của poly-Lornithine hoặc polyethylene glycol (PEG). Nguyên lý chung của phương pháp là trong điều kiện dung dịch chứa nồng độ PEG phân tử lượng cao (~8000) thì tế bào trân cạnh tranh với PEG các phân tử nước theo nguyên tắc thẩm thấu. Kết quả là tế bào trân mất nước và nhân nhúm lại. Khi pha loãng và rửa sạch PEG khỏi dung dịch, các phân tử nước được tế bào tiếp nhận trở lại các phân tử DNA qua màng nguyên sinh chất. Phương pháp này nhanh, rẻ vì không cần thiết bị, nhưng không xác định được lượng DNA xâm nhập vào tế bào, vì thế rất hay được áp dụng cho những nghiên cứu biểu hiện tạm thời các cấu trúc gene mới.

(II) Chuyển gene bằng xung điện

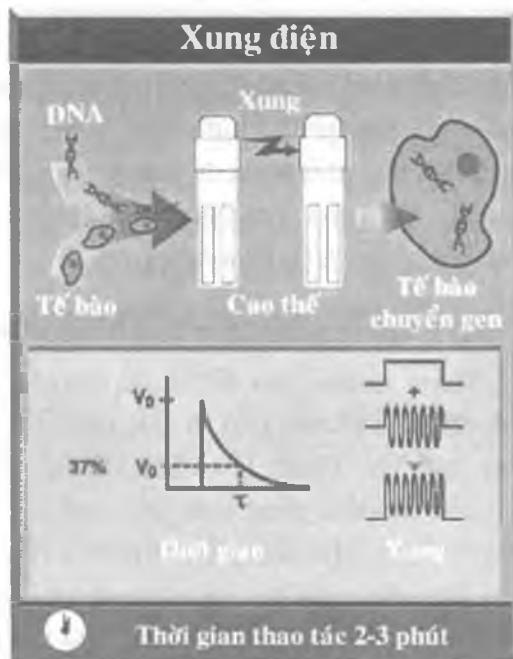
Thiết bị biến nạp gene bằng xung điện được gọi là gene pulser, hoạt động trên nguyên tắc khi cho một dòng điện một chiều chạy qua dung dịch tế bào trắn, các tế bào sẽ sắp xếp thành từng chuỗi từ điện cực dương sang điện cực âm. Khi một dòng điện mạnh chạy qua trong một quãng thời gian vài phần nghìn giây, màng nguyên sinh chất giữa chúng bị đánh thủng thành lỗ (poration tức là khoan lỗ) và làm cho các phân tử DNA mang điện tích âm xâm nhập vào trong tế bào (hình 6.12). Hiệu quả chuyển gene trực tiếp được tăng lên đáng kể nhờ kỹ thuật xung điện này. Tuy nhiên, phương pháp này cũng thường được áp dụng cho những nghiên cứu nhanh đổi với quá trình biểu hiện tạm thời của gene trong tế bào thực vật vì lượng DNA xâm nhập vào tế bào cũng không được xác định chính xác.



A



B



C

Hình 6.12. Thiết bị xung điện Gene Pulser của Hãng BioRad (A) dùng cho chuyển gene ở tế bào trắn và tế bào vi sinh vật với các loại cuvette có dung tích khác nhau (B) và sơ đồ nguyên lý hoạt động (C).

Mặc dù kỹ thuật này chủ yếu được dùng trong nghiên cứu biểu hiện tạm thời các gene trong tế bào trân, đã có nhiều cải tiến đối với kỹ thuật này để ứng dụng cho tế bào còn nguyên vẹn thành cellulose (Fromm et al., 1985). Tuy nhiên, để đạt hiệu quả cao, tế bào và mô trước khi biến nạp phải được xử lý trước như làm bị thương hoặc xử lý với dung dịch có áp suất thẩm thấu cao hoặc chứa enzyme (macerozyme).

Cũng có nhiều loại mô và tế bào được xác định là có khả năng tiếp nhận gene (khả biến) như phôi non của ngô, lúa và lúa mỳ. Các yếu tố chính tác động lên hiệu quả biến nạp bằng xung điện là : Cường độ điện trường, điện dung, tính chất ion và nồng độ đệm, thời gian tiền xử lý. Nếu gây vết thương thì có hạn chế mức độ tiết nuclease khi bị thương. Cũng có thể xử lý nóng đối với mô biến nạp bằng xung điện để nâng cao hiệu quả.

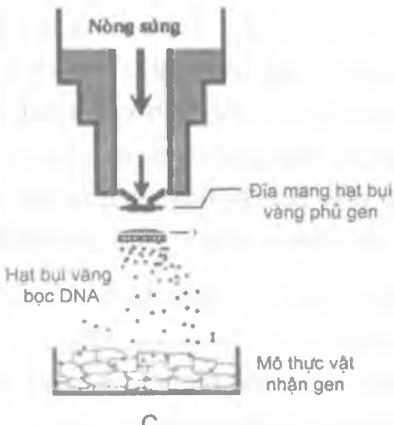
Hiệu quả cuối cùng của biến nạp trực tiếp gene vào tế bào trân phụ thuộc vào hiệu quả chọn lọc dòng mang gene biến nạp và tỷ lệ tái sinh cây từ dòng tế bào chọn được. Những thí nghiệm thời kỳ đầu mới tập trung vào tế bào trân của các cây thuộc họ Cà (Solanaceae) vì chúng dễ tái sinh cây và gene dễ chọn lọc là gene vi khuẩn mã hoá neomycin phosphotransferase (npt II). Sau đó, kỹ thuật nuôi cây và tái sinh cây từ tế bào trân đã thành công trên nhiều loài cây khác nhau, kỹ thuật chuyển gene thông qua tế bào trân cũng đã được triển khai có kết quả trên nhiều đối tượng quan trọng như lúa japonica (Cornejo và Meredith, 1985 ; Hayashimoto et al., 1990) và Indica (Datta et al., 1992 ; Peng et al., 1990), ngô, cỏ chăn nuôi.

(III) Chuyển gene bằng máy bắn gene

Klein (1987) là tác giả đầu tiên tiến hành chuyển gene vào tế bào biểu bì của cây hành bằng kỹ thuật bắn gene. Nguyên lý chính của kỹ thuật bắn gene là DNA được bọc ra ngoài các hạt bụi vàng (hoặc Wolfram) kích thước khoảng 1 µm, hạt vàng bọc gene được bắn vào khối mô đặt phía trước luồng đạn nhờ áp lực cao (3500 psi) của một luồng khí helium (hình 6.13). Đến nay, kỹ thuật bom hạt (particle bombardement) có nhiều tên gọi khác nhau như chuyển nạp sinh học (biolistics) ; bom vi tiêm (microprojectile bombardement) ; gia tốc hạt (particle acceleration) nhưng đều có chung một nội dung là bắn gene (gene gun).



A



B



C

Hình 6.13. Thiết bị súng bắn gene (gene gun) loại đơn giản, gọn nhưng kém hiệu quả (A), loại đồng bộ do BioRad sản xuất (B) hoạt động trên nguyên tắc gia tốc hạt (C).

Kỹ thuật bắn gene có những ưu điểm chính như sau :

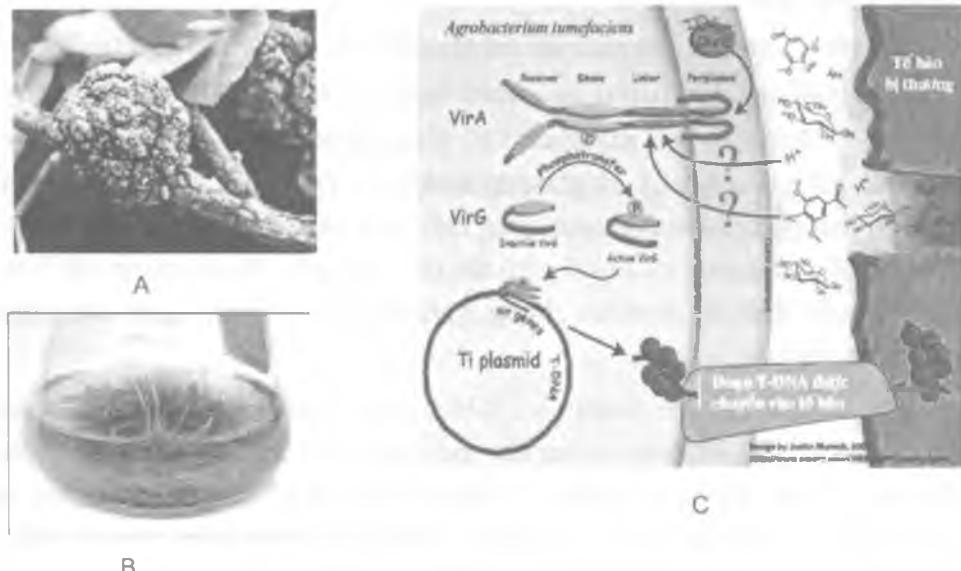
- Chuyển được gene cho các khối mô có tổ chức. Ưu điểm này cho phép chọn những loại mô có triển vọng tái sinh cây tốt để chuyển gene.
- Hệ thống chuyển gene toàn năng : Về nguyên tắc có thể bắn gene cho bất kỳ loại mô, cơ quan nào và kết quả nhuộm tạm thời khẳng định điều đó, nhưng kết quả tái sinh cây còn phụ thuộc vào loài và loại mô chọn.
- Chuyển được gene cho các giống cây trồng có giá trị. Không những áp dụng tốt cho cây mô hình mà ngay cả những giống cây đang phổ biến như : đậu tương, ngô, lúa (Chen et al., 1993 ; Christou et al., 1991 ; 1995) cũng cho phép tiến hành chuyển gene.

(IV) Chuyển gene gián tiếp thông qua *Agrobacterium tumefaciens*

Vi khuẩn đất *Agrobacterium* có hai loài được dùng trong kỹ thuật di truyền, đó là *A. tumefaciens* và *A. rhizogenes*. Hình 6.14 mô tả tác động gây nên khối u crown gall của *A. tumefaciens* và gây nên búi rễ của *A. rhizogenes* đối với nhiều loài thực vật mà bản chất của nó là gene của vi khuẩn được chuyển vào tế bào thực vật.

Hiện tượng gene của vi khuẩn đất được chuyển vào tế bào thực vật được coi là hiện tượng chuyển gene tự nhiên. Đoạn gene được chuyển vào thực vật nằm trong một plasmid khổng lồ kích thước trên 400 kb có tên

gọi là Ti plasmid. Không phải toàn bộ DNA của Ti plasmid xâm nhập vào tế bào thực vật mà chỉ có một đoạn gene nằm giữa biên trái (left border LB) và biên phải (right border RB) của đoạn DNA vận chuyển (transfer DNA, T-DNA) tách khỏi phân tử Ti và xâm nhập vào tế bào. Hơn thế nữa, đoạn T-DNA này còn có khả năng hòa nhập với DNA nhiễm sắc thể và trở thành một bộ phận bền vững trong genom của thực vật.



Hình 6.14. Hiện tượng gây khối u crown gall do vi khuẩn *Agrobacterium tumefaciens* (A) và gây bùi rễ do vi khuẩn *A. rhizogenes* (B) mà bản chất là các gene liên quan đến sinh tổng hợp phytohormon IAA được chuyển nạp từ vi khuẩn sang tế bào thực vật (C).

Khi các nhà nghiên cứu thành công trong việc sử dụng đoạn T-DNA của plasmid Ti (Bevan et al., 1983 ; Chan et al., 1992 ; Horsch và Klee, 1986 ; Skriver và Mundy, 1990 ; Wang et al., 1984) để đưa gene xâm nhập vào bộ genome của tế bào thực vật thì việc chuyển gene thông qua *Agrobacterium* trở thành công cụ rất hữu hiệu trong công nghệ gene thực vật, kể cả cây mít lá mầm (Hiei et al., 1994 ; Le et al., 1994 ; Liu et al., 1992).

Trước hết, cần phân biệt rõ hai họ vector dùng trong biến nạp thông qua *Agrobacterium tumefaciens* :

(1) Vector liên hợp (cointegrate) : Là một vector Ti mang một số đoạn chức năng của vector thông thường khác của *E. coli*. Một vector liên hợp gồm những phần sau : Một vị trí thích hợp cho phép ghép nối các

đoạn gene quan tâm (gene of interest = GOI), thường là có nguồn gốc từ một plasmid của vi khuẩn *E. coli*; Gene chỉ thị có tính kháng kháng sinh để chọn lọc hoạt động được trong *E. coli* và *A. tumefaciens*; Gene chọn lọc hoạt động được ở tế bào thực vật; Một đoạn khởi đầu của *E. coli*, không hoạt động ở *A. tumefaciens*. Như vậy, thực chất vector liên hợp là một vector lai có chỗ cho GOI đi nhờ như một hành khách vào trong tế bào thực vật.

So với *E. coli* thì *Agrobacterium* không được thích hợp lấm trong các thao tác biến nạp trực tiếp tế bào đông lạnh nhờ CaCl_2 . Vì vậy, phải dùng một hệ vi khuẩn *E. coli* mang một số plasmid hỗ trợ như plasmid mob+ có đoạn khởi đầu ColE1 và plasmid tương trợ pHelper có khả năng vận chuyển loại plasmid mang gene quan tâm GOI vào tế bào *Agrobacterium*. Hình 6.14 giới thiệu một sơ đồ phối tác các chủng vi khuẩn mang các loại plasmid cần thiết để có được chủng *Agrobacterium* mang gene cần biến nạp (pVector).

(2) Vector nhị thể (binary) : Khắc phục tình trạng khó nhân bản trong *A. tumefaciens* loại vector nhị thể được thiết kế bao gồm các phần như sau : Vị trí ghép nối đa điểm ; Đoạn khởi đầu tái bản hoạt động có ở *E. coli* và *A. tumefaciens* (ví dụ RK2) ; Gene chọn lọc hoạt động ở cả vi khuẩn và thực vật ; Có khả năng vận chuyển (như : oriV, oriT, trfA), chỉ cần khi chuyển gene thông qua *A. tumefaciens*, không cần khi chuyển gene trực tiếp ; Đoạn T-DNA biên (border), đặc biệt là đoạn bên phải bắt buộc phải có. Ngoài ra còn một số thành phần phụ trợ khác như : Đoạn kích thích vận chuyển T-DNA (T-DNA transfer stimulator sequence, TSS).

Hệ vector nhị thể được thiết kế bao gồm hai plasmid tự tái bản. Loại vector vận tải mang GOI giữa đoạn ngoại biên (border) của T-DNA và một plasmid hỗ trợ có gene vir đảm nhiệm chức năng vận chuyển DNA vào tế bào thực vật. Plasmid hỗ trợ được cải biến từ các Ti plasmid bình thường bằng cách loại bỏ gene kích thích tế bào thực vật phát triển thành khối u, nhưng vẫn duy trì khả năng xâm nhập vào tế bào thực vật.

Cũng đã có các thiết kế plasmid nhị thể dùng cho loài vi khuẩn *Agrobacterium rhizogenes*, loài này không gây nên khối u mà gây ra các cấu trúc như rễ, thành từng chùm như túm lông nên được gọi là rễ lông.

Với các kiểu plasmid như trình bày trên có thể tiến hành biến nạp gián tiếp thông qua *Agrobacterium*, nhưng cũng có thể nuôi cấy vi khuẩn để tách chiết và tinh chế plasmid rồi tiến hành biến nạp theo phương thức trực tiếp.

6.2.5.3. Mục đích chuyển gene vào một số cây trồng ở nước ta

(I) Chuyển gene vào cây bông

Hiện nay, khó khăn lớn nhất đối với nghề trồng bông là giống năng suất thấp và bị sâu hại rất nghiêm trọng. Một vụ sản xuất kéo dài 4 tháng thường phải phun 16-18 lần thuốc hoá học làm cho giá thành bông hạt tăng đáng kể vượt qua cả giá nhập khẩu. Hơn nữa, do phải sử dụng nhiều thuốc hoá học, vấn đề đảm bảo sức khoẻ cho người trồng bông trở thành yếu tố hạn chế mở rộng diện tích cây bông. Giống năng suất cao có thể tạo ra bằng con đường lai tạo, nhưng giống kháng sâu thì duy nhất phải chọn con đường chuyển gene. Các gene thuộc nhóm *cry* đang được sử dụng rộng rãi. Viện Công nghệ sinh học đã tập hợp được các gene trên, thông qua kết hợp với Viện Nghiên cứu và Phát triển cây bông để tiến hành chuyển gene kháng sâu, gene chịu hạn bằng các phương pháp khác nhau.

(II) Chuyển gene vào cây hoa

Nghề trồng hoa cắt đang góp phần thay đổi cơ cấu cây trồng và tăng thu nhập cho người nông dân. Tuổi thọ của hoa cắt là yếu tố quyết định việc mở rộng thị trường trong và ngoài nước. Tham gia vào quá trình già hóa của hoa làm nở và rụng cánh nhanh là các gene điều khiển sinh tổng hợp ethylene. Các gene ngược của chúng hoạt động theo cơ chế tạo RNA sợi kép không dịch mã thành protein được. Hiện nay, các gene ACO antisense ức chế tạo ethylene đã được phân lập. Cơ chế thứ hai là làm tăng sinh tổng hợp các loại phytohormone thuộc nhóm cytokinin làm cho rau xanh và hoa tươi lâu. Cơ chế sử dụng các kích thích sinh tổng hợp cytokinin đã được Gan (1995) quan tâm.

Hiện nay, Viện Công nghệ sinh học đang có trong tay nhóm gene *ACO antisense*, gene *chitinase* và gene *bar*. Các nhóm gene này sẽ được chúng tôi sử dụng để chuyển vào một số loại cây hoa với mục đích làm hoa lâu tàn.

(III) Chuyển gene vào cây trồng rừng

Cây trồng rừng thường bị sâu phá hoại, mọt khát, các đối tượng cây thân gỗ vào loại khó nuôi cấy mô vì vậy, cần sớm có những nghiên cứu, dù chỉ là bước đầu để tiếp cận nội dung này. Đối tượng chính được chọn là cây *Paulownia* vốn là cây đã được nuôi cấy thành công ở nhiều phòng thí nghiệm trong nước. Các gene lựa chọn là gene chỉ thị (GUS, kháng kanamycin) để tối ưu hệ thống chuyển gene và tiếp đến là các gene kháng côn trùng (*cry*) và kháng bệnh. Mục đích lớn nhất là tạo được phương pháp chuyển gene ở cây gỗ trồng rừng và tạo được một số dòng chuyển gene có giá trị.

(IV) Chuyển gene vào cây lúa

Hiện nay, khá nhiều gene kháng bệnh đã được đưa vào cây trồng lương thực và thực phẩm như ngô, khoai tây và cải dầu. Lúa là cây trồng lương thực quan trọng số một của nước ta, các giống lúa thuộc 2 nhóm quan trọng nhất là lúa hàng hóa và lúa đặc sản chất lượng cao đều bị nhiễm bệnh nặng, nhất là khi chuyển chúng từ phía nam ra hoặc từ Trung Quốc về : vụ xuân bị đạo ôn và vụ mùa bị bạc lá. Gene *Xa21* kháng bệnh bạc lá vi khuẩn sẽ được chuyển vào 2-3 giống chọn lọc cho mỗi nhóm nhằm tạo ra tính kháng rộng các nòi *Xanthomonas oryzae* (Zhanget al., 1998). Gene GNA chuyển vào giống phía nam để tạo tính kháng rầy nâu (Sudhakar et al., 1998). Các gene *CryIA(b,c)* sẽ được chuyển vào lúa để tạo tính kháng sâu. Tính chịu hạn và chịu muối được tăng cường thông qua gene *P5CS* điều khiển sinh tổng hợp prolin.

(V) Chuyển gene kháng virus đốm vòng vào cây đu đủ

Cây đu đủ là loại cây trồng ăn quả khá phổ biến của nhiều quốc gia, tuy nhiên, đây cũng là loại cây trồng bị bệnh virus một cách nặng nề. Hầu như 100% các vườn đu đủ đều bị bệnh, sau 2 năm phải huỷ và trồng lại sau khi đổi cây trồng 1-2 vụ. Bệnh đốm vòng đu đủ do PSRV (papaya ring spot virus) gây nên. Tính kháng PSRV được tạo nên bởi quá trình chuyển gene CP hoặc và gene NiB của PSRV vào cây đu đủ và đã thành công trên nhiều quốc gia như Hoa Kỳ, Thái Lan, Đài Loan, Malaysia. Đáng lưu ý là cấu trúc những gene này của chủng virus thu thập ở vùng địa lý nào thì mới có khả năng kháng lại virus gây bệnh ở vùng đó, tức là không thể dùng gene của virus thu nhận từ Hawaii để tạo tính kháng đối

với chủng PSRV của Thái Lan. Vì vậy, các nhóm nghiên cứu phải tiến hành điều tra chủng giống, thiết kế và chuyển gene theo một yêu cầu nghiêm ngặt thì mới thu được kết quả. Gần đây, có những cố gắng mới nhằm thiết kế cấu trúc gene CP da đoạn (multiple gene fragment approach) có các đoạn gene của nhiều chủng virus có nguồn gốc khác nhau với hy vọng thu được tính kháng phổ rộng. Chúng tôi cũng đang tiếp cận cả hai cách thức này.

6.2.6. Phương pháp phân tích thực vật chuyển gene

Điều quan trọng sau khi chuyển gene là cần theo dõi gene chuyển vào cây sẽ được tế bào, đặc biệt là bộ genome của tế bào chủ, tiếp nhận như thế nào. Hiện nay có nhiều nhóm phương pháp cần được tiến hành trong quá trình tạo cây trồng chuyển gene. Các nhóm phương pháp đó bao gồm :

1. Phân tích tính kháng thuốc
2. Phân tích hoá tế bào
3. Phân tích bằng kỹ thuật sinh học phân tử
4. Phân tích hoạt động của gene chuyển (transgene)

Mỗi nhóm phương pháp hoạt động trên những nguyên lý khác nhau được trình bày dưới đây, các bước kỹ thuật của phương pháp sẽ được trình bày chi tiết trong chương này.

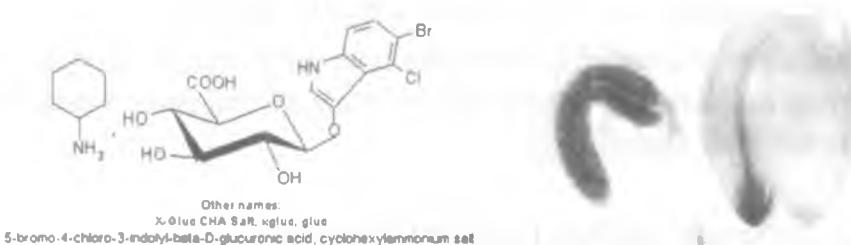
6.2.6.1. Phương pháp phân tích tính kháng thuốc

Thực chất phân tích tính kháng thuốc là chọn dòng tế bào, mô và cây mang tính kháng khuỷc thông qua hoạt động của gene chuyển. Thuốc được hiểu ở đây là thuốc kháng sinh như kanamycin khi dùng gene *nptII* hoặc hygromycin khi dùng gene *hpt* hoặc thuốc diệt cỏ basta khi dùng gene *bar*. Bước phân tích này có thể áp dụng cho mọi giai đoạn sau khi biến nạp từ trạng thái tế bào đến mô sẹo, chồi tái sinh, cây hoàn chỉnh hay cây non gieo từ hạt. Cách thức tiến hành đơn giản : chỉ bổ sung thuốc theo nồng độ nhất định vào môi trường nuôi cấy rồi quan sát ảnh hưởng của thuốc lên mô hoặc cây. Thông thường, mô và cây mất khả năng tạo lục lạp rồi sau đó chết tạo màu nâu đen.

Phương pháp thử tính kháng thuốc còn được pháp triển thành phương pháp tạo vết tẩy trên lá (leaf bleach method). Thuốc kháng sinh được pha ở nồng độ cao (2-5 lần so với thử *in vitro*), bổ sung Tween để tăng khả năng bám dính. Sau đó, dùng bút lông quét lên mặt lá xanh của cây cần thử. Cũng có thể nhúng sợi chỉ bằng bông thấm dịch thuốc rồi vắt sợi chỉ qua nhiều lá của nhiều cây. Sau 3-5 ngày có thể quan sát thấy nơi bôi thuốc hay nơi sợi chỉ vắt qua hình thành các vết trắng hoặc nâu trên mặt lá do bị mất diệp lục. Ở những cây có tính kháng, các vết đó không hình thành. Để bảo đảm độ chính xác cần tiến hành thử lại những cá thể được chọn 2-3 lần nhắc lại.

6.2.6.2. Phương pháp phân tích hoá tế bào

Có thể nhuộm hoá tế bào để khẳng định sự có mặt của sản phẩm GUS thông qua hoạt động biểu hiện của gene *gus*. Nhuộm GUS được thực hiện trong thí nghiệm xác định biểu hiện tạm thời, mô sẹo, chồi mầm và rễ mới tái sinh. Nhuộm GUS cũng được thực hiện để khẳng định gene chuyển được di truyền sang thế hệ sau khi nhuộm phôi của hạt hình thành từ cây chuyển gene.



Hình 6.15. Cấu trúc phân tử X-Gluc dùng làm cơ sở cho phương pháp nhuộm GUS đối với mô thực vật chuyển gene. Khi X-Gluc bị GUS khử tạo ra sản phẩm sơ cấp không màu, nhưng sau đó sản phẩm này bị ôxy hoá và dimer hoá thành sản phẩm cuối cùng có màu xanh indogo đặc trưng.

Sản phẩm sơ cấp 5-Br-4Cl-3-indolyl vẫn còn là một chất dễ tan và không màu. Nhưng ngay sau đó sản phẩm này bị ôxy hoá và dimer hoá thành một phức hệ không tan có màu xanh (indigo). Sự có mặt của ferri và ferrocyanide trong dung dịch nhuộm gia tăng quá trình hình thành sản phẩm màu cuối cùng, đồng thời cũng góp phần hạn chế việc khuyếch tán sản phẩm sơ cấp đến những vùng không có mặt của GUS.

6.2.6.3. Phương pháp phân tích sinh học phân tử

Sau bước chọn lọc trên môi trường chứa các tác nhân chọn lọc như kháng sinh, thuốc trừ cỏ, vẫn đề cốt yếu còn lại là phải phân tích được số phận của gene biến nạp trong tế bào nhận bằng các kỹ thuật sinh học phân tử như sau :

- PCR.
- Lai Southern.
- Phân tích biến hiện gene.

Trên nguyên tắc gene ngoại lai khi được biến nạp vào tế bào có thể tồn tại trong tế bào chủ ở 3 trạng thái : (1) Tạm thời ở dạng DNA tự do ; (2) Lâu dài dưới dạng một thể plasmid độc lập tự nhân ; (3) Ổn định như một đoạn DNA của một genome trong tế bào chủ và được nhân lên theo dạng tương hợp hay không tương hợp. đương nhiên, trạng thái thứ ba được ưa chuộng nhất vì tính bền vững của thể biến nạp, nhưng nó hoàn toàn phụ thuộc vào quá trình tái tổ hợp. Hầu hết các nghiên cứu biến nạp gene thuộc nhân vẫn có thể phát hiện thấy hiện tượng nhân không tương hợp đối với đoạn gene lạ khi hoà đồng vào DNA nhân. Bởi vì vị trí đoạn gene lạ được ghép nối ảnh hưởng đến biểu hiện của chính nó hay gene chủ gần đó, cho nên người ta đang quan tâm nhiều đến tái tổ hợp tương hợp hoặc là ghép nối định hướng (Yoder và Kmiec, 1991). Nếu kỹ thuật này thành công thì khả năng thay thế gene trở thành hiện thực và sẽ có giá trị lớn về mặt thực tiễn và lý luận.

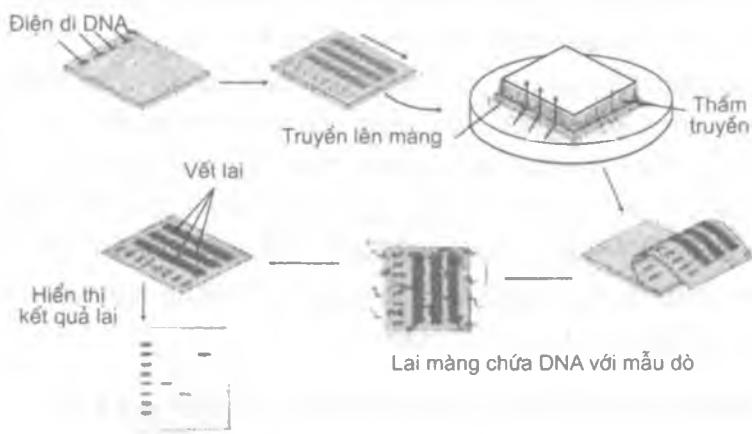
(I) Phương pháp phân tích cây chuyển gene bằng PCR

Đây là phương pháp đơn giản, dễ thực hiện với các phòng thí nghiệm chuyên về công nghệ gene thực vật. Gene chuyển là gene đã được biết vì vậy chỉ cần sử dụng cặp mồi đặc hiệu cho gene đó và tách chiết DNA từ mẫu thực vật cần xác định làm khuôn là có thể tiến hành PCR. Nếu sản phẩm PCR có kích thước đúng như dự kiến thì mẫu phân tích được coi là dương tính. Tuy nhiên, PCR dương tính không đồng nghĩa với chuyển gene thành công vì có nhiều cách để giải thích sự có mặt của DNA trong mẫu phân tích : (1) Vi khuẩn *Agrobacterium* mang gene chuyển còn tồn tại trong khối mô hay trong các gian bào của mẫu phân tích. Hiện tượng này xuất hiện ở nhiều loại đối tượng. Nếu là mẫu phân tích của thực vật

dã qua nhân giống hữu tính, tức là qua thế hệ T1, T2 dưới dạng hạt rồi thì hoàn toàn loại trừ khả năng này ; (2) Gene chuyển tồn tại tự do trong tế bào chất, có thể biến mất qua sinh sản hữu tính ; (3) Gene chuyển không hoạt động, tức là không được biểu hiện thành protein có chức năng sinh học. Vì thế, kết quả phân tích bằng PCR mới chỉ có giá trị định hướng ban đầu (hầu hết các tạp chí chuyên ngành không chấp nhận những công bố chuyển gene khi tác giả mới đạt đến kết quả phân tích PCR).

(II) Lai Southern

Lai Southern là việc khẳng định sự tồn tại của gene chuyển trong genome của cây nhận, thông qua lai DNA tổng số của mẫu phân tích với mẫu DNA dò. Mẫu dò chính là một đoạn của gene chuyển có kích thước từ 100-300 bp được đánh dấu phóng xạ hay huỳnh quang. Chi tiết các bước kỹ thuật lai Southern nêu kỹ trong phần 3. Có thể tóm tắt các bước như sau (hình 6.16) :



Hình 6.16. Nguyên lý lai Southern để khẳng định sự có mặt của gene chuyển trong genome cây nhận : Sau khi xử lý cắt hạn chế DNA mẫu phân tích được điện di, thẩm truyền lên màng và lai với mẫu dò đánh dấu và hiện thị lên phim. Gene chuyển hiện lên thành băng vạch kích thước khác nhau tùy thuộc vị trí cắt hạn chế, từ đó có thể kết luận về số bản copy của gene chuyển trong genome cây nhận.

1. Tách chiết DNA tổng số của mẫu cần phân tích.
2. Xử lý cắt hạn chế với 1-2 enzyme.
3. Điện di DNA trên gel agarose.
4. Thẩm truyền bản gel lên màng nitrosocellulose (còn gọi là blotting).

5. Tổng hợp DNA mẫu dò có đánh dấu bằng phóng xạ hay huỳnh quang bằng PCR với mỗi ngẫu nhiên (6-mer oligo).
6. Lai mẫu dò với màng mang DNA.
7. Hiển thị trên phim X quang.
8. Phân tích kết quả.

Lai Southern góp phần tiến xa hơn một bước trong việc đánh giá kết quả chuyển gen : Một mặt kết quả dương tính khi lai khẳng định gene chuyển tồn tại trong genome cây nhện, đồng thời số lượng băng vạch dương tính cho biết số lượng copy của bản gene chuyển có mặt trong genome cây nhện. Trong nhiều trường hợp số lượng copy đó không phải chỉ là 1 bản, mà là 2 hay nhiều bản.

(III) Phân tích biểu hiện gene

Khái niệm biểu hiện gene được hiểu trong sinh học phân tử là quá trình hoạt động của gene để tạo ra sản phẩm cuối cùng là protein. Quá trình biểu hiện gene trải qua hai bước : (1) Bước phiên mã từ DNA cho ra mRNA. Muốn đánh giá sự có mặt của sản phẩm phiên mã, người ta tiến hành kỹ thuật lai Northern ; (2) Bước dịch mã từ mRNA thông qua bộ máy ribosome, phân tử polypeptide được tổng hợp với trình tự amino acid chính xác do trình tự nucleotide trong đoạn DNA mã hoá của gene quyết định. Muốn đánh giá sự có mặt của protein do gene chuyển tạo ra phải tiến hành kỹ thuật lai Western.

Kỹ thuật lai Northern về nguyên lý cũng dựa trên tính cặp đôi base bổ sung của hai sợi acid nucleic, trong trường hợp mẫu là mRNA có trong RNA tổng số tách từ mẫu cần phân tích. Sau khi điện di và thám truyền RNA tổng số lên màng nitrosocellulose, màng được lai với mẫu dò là các oligonucleotide đánh dấu, được tổng hợp bằng PCR từ gene cần tìm với mỗi ngẫu nhiên dài 6 nucleotide. Trong những trường hợp dương tính trên màng nitrosocellulose xuất hiện những băng sản phẩm lai. Gene được phiên mã càng mạnh thì tín hiệu lai càng mạnh.

Kỹ thuật lai Western dựa trên nguyên lý của phản ứng miễn dịch giữa kháng nguyên (là sản phẩm protein cần tìm) và kháng thể đặc hiệu đối với kháng nguyên đó. Trước hết protein toàn phần được tách chiết từ mẫu cần phân tích, sau điện di trên gel polyacrylamide

(PAGE) protein được thảm truyền lên màng nitrosocellulose rồi lai với kháng thể (kháng thể được sản xuất nhờ kháng nguyên gốc tinh sạch hay kháng nguyên tái tổ hợp thông qua kỹ thuật sản xuất kháng thể đơn dòng hay đa dòng). Quá trình hiển thị sản phẩm lai có thể thực hiện thông qua phản ứng tạo màu với phosphatase kiềm hay hiển thị thông qua phản ứng dạ quang ECL.

6.2.6.4. Phương pháp phân tích chức năng sinh học

Chỉ làm sau khi đã xác định nguyên liệu tạo được là cây chuyển gene thực sự thông qua các kỹ thuật sinh học phân tử ở mức độ phòng thí nghiệm. Xác định sự có mặt của gene trong tế bào cây chuyển gene được tiến hành bằng kỹ thuật PCR, với cặp mồi đặc hiệu của gene chuyển và DNA của dòng cây cần kiểm tra. Xác định gene chuyển được gắn vào genome cây tái sinh hay không phải thực hiện phép lai phân tử Southern mà mẫu dò là đoạn gene chuyển được đánh dấu huỳnh quang, còn DNA nhân cao phân tử được tách từ mô của cây chuyển gene. Xác định gene chuyển có được phiên mã thành mRNA hay không, tức là có được bộ máy sinh tổng hợp protein chấp nhận hay không được thực hiện bằng kỹ thuật lai Northern giữa mẫu dò là đoạn DNA của gene đánh dấu huỳnh quang và RNA toàn phần của mô cây chuyển gene. Bước cuối cùng là xác định gene chuyển có biểu hiện và tạo ra sản phẩm cuối cùng của nó là protein bằng kỹ thuật lai Western giữa protein của cây chuyển gene và kháng thể đặc hiệu với protein đó được tinh sạch từ nguồn khác. Mặt khác, hoạt tính gene chuyển còn được thử bằng các phép thử chỉ thị như thử tính kháng kháng sinh, thử tính diệt sâu nhưng ở quy mô phòng thí nghiệm và quy mô nhà lưới.

Điểm tiếp theo là phải xác định phương thức và đặc điểm di truyền của gene chuyển và tính trạng được gene chuyển mã hoá, đồng thời kiểm tra mức độ đồng hợp tử của thế hệ con của chúng bằng cách lai chéo hay tự phôi.

Biến nạp gene với mục đích cuối cùng là đưa được tính trạng mong muốn vào cây trồng. Vì thế, khâu phân tích trực tiếp tính trạng đó là khâu có giá trị quyết định cuối cùng. Gene chuyển thuộc các nhóm khác nhau, vì thế phương pháp phân tích cũng khác nhau, phụ thuộc vào tính trạng đang được quan tâm. Có thể nhóm các phương pháp thành các nhóm như sau :

(I) Phân tích tính kháng bệnh

Tính kháng virus, kháng vi khuẩn hay kháng nấm được thử bằng cách lây nhiễm nhân tạo tại các phòng thí nghiệm về bệnh học thực vật với các chủng virus, vi khuẩn hoặc nấm. Sau khâu khẳng định tính kháng bệnh trong phòng thí nghiệm, cần thực hiện bước tiếp theo là kiểm tra tính kháng trên đồng ruộng có lây nhiễm nhân tạo. Bước cuối cùng là khảo nghiệm trên đồng ruộng ở nhiều địa điểm và qua ít nhất hai vụ liên tục.

(II) Phân tích tính chống chịu

Tính chống chịu thường được kiểm tra bằng thí nghiệm gây điều kiện ngoại cảnh bất lợi nhân tạo, tuy nhiên, đây là việc làm không dễ. Đối với khô hạn có nhiều cách bố trí thí nghiệm khác nhau : (1) Xử lý hạn bằng cách xử lý PEG hay đường sucrose. Các chất này cạnh tranh nước với bộ rễ của cây ; (2) Bố trí thí nghiệm gây hạn nhân tạo trong nhà hoặc buồng trồng cây bách cách không tưới nước hoặc tưới phục hồi rất hạn chế. Loại thí nghiệm này được tiến hành với cây trồng trên cát sạch, đất trồng cây trong chậu hoặc các khối đất nén ; (3) Tiến hành thí nghiệm trên hiện trường tại những vùng khô hạn. Israel là nơi lý tưởng được nhắc đến đối với việc tiến hành các thí nghiệm về hạn vì ở đây có những vùng và thời kỳ hàng nửa năm không có mưa tự nhiên. Tương tự như vậy, khi nghiên cứu về tính chịu mặn, tính chịu phèn cần có những cách bố trí thí nghiệm tự nhiên và hợp lý, nhưng không được quá khác biệt với điều kiện bất lợi của tự nhiên.

(III) Phân tích gene cải tiến chất lượng sản phẩm

Đây là loại phân tích dễ tiến hành nhất. Khi quan tâm đến tính trạng chất lượng nào và chuyển những gene phù hợp thì sẽ tiến hành các phân tích về chỉ tiêu chất lượng đó. Ví dụ, tăng hàm lượng β -carotene thì phân tích trực tiếp hàm lượng chất này bằng khối phổ. Tăng chất lượng tinh bột, tăng độ dẻo của hạt gạo thì xác định hàm lượng amylose, khi giảm được hàm lượng amylose xuống dưới 18% là đạt yêu cầu. Những tính trạng liên quan đến chất lượng bảo quản, chất lượng chế biến cần phải được đánh giá theo cách thức riêng phù hợp với từng loại sản phẩm.

(IV) Đánh giá an toàn sinh học cây chuyển gene

Cây chuyển gene, sau bước đánh giá ở mức độ phòng thí nghiệm cần phải kiểm tra mức độ an toàn trên đồng ruộng, mức độ an toàn đối với sức khoẻ vật nuôi, con người và khả năng ảnh hưởng lên môi trường. Quy định chung khi khảo nghiệm trồng cây chuyển gene trên đồng ruộng cần phải thu được các thông tin cơ bản sau :

1. Gene được chuyển có nguồn gốc từ đâu ? Gene được chuyển vào cây trồng thông qua vector nào ? Gene được chuyển kèm những gene gì ?
2. Gene chuyển hoạt động trong điều kiện nào ? Sản phẩm tạo ra là gì ?
3. Gene được chuyển và gene chuyển kèm có thể bị phát tán (chuyển sang cây khác) của hệ sinh thái trong khu vực thông qua hạt phấn của cây chuyển gene hay không ? Nếu có thì nguy cơ gì có thể xảy ra ?
4. Sản phẩm của gene chuyển có gây những thay đổi cơ bản về chất đối với sản phẩm nông nghiệp so với sản phẩm đó được sản xuất ra bằng giống cây bình thường hay không ?
5. Chất (protein) do gene chuyển tạo ra có gây dị ứng cho người và vật nuôi sử dụng chúng hay không ? Cần tiến hành kiểm tra độc hại theo quy định hiện hành của ngành y tế về an toàn thực phẩm.

Phương pháp thử nghiệm trên đồng ruộng và đánh giá an toàn sinh học cây chuyển gene cần được tiến hành bảo đảm các nội dung sau :

1. Trồng cách ly và phân tích di truyền mức độ phát tán phấn hoa nếu có đối với cây trồng một năm.
2. Đánh giá hoạt tính gene chuyển bằng phép thử sinh học về tính kháng sâu, kháng bệnh.
3. Thủ độc tính và tính gây dị ứng của protein gene chuyển theo quy định của Bộ Y tế.

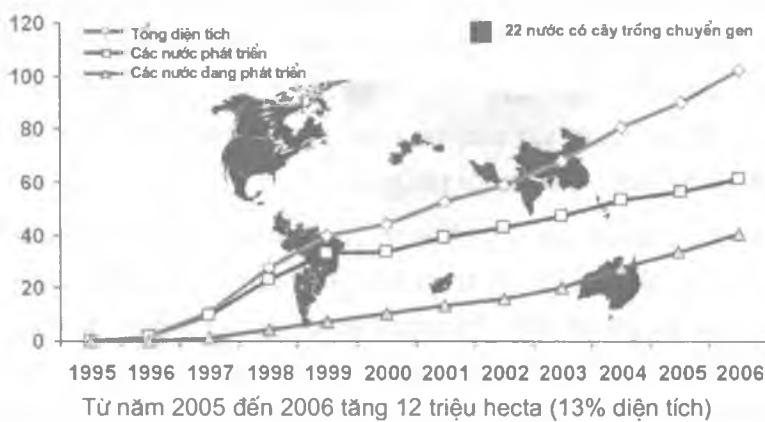
6.2.7. Ý nghĩa kinh tế – xã hội

6.2.7.1. Thành tựu trên thế giới

Năm 1990, lần đầu tiên cây thuốc lá chuyển gene được đưa vào thử nghiệm trong sản xuất (Cohen, 2000). Tính đến năm 1996 đã có 7 loại

cây trồng chuyển gene đó là : bông, cà chua, thuốc lá, ngô, khoai tây, đậu tương và cải dầu với tổng diện tích là 1,7 triệu ha (de la Cruz, Reynaldo, 1999). Tiếp theo đó, đối tượng cây chuyển gene được mở rộng sang nhiều đối tượng cây trồng khác như lúa, chuối, du đủ và diện tích trồng cây đã được chuyển nạp gene cũng không ngừng tăng lên. Đến năm 2006, diện tích cây chuyển gene đã lên tới xấp xỉ trên 100 triệu ha, được trồng nhiều ở Mỹ, Canada, Australia, Argentina và Trung Quốc (hình 6.17, theo James, 2006). Việc ứng dụng công nghệ sinh học nói chung và phát triển kỹ thuật chuyển gene vào cây trồng nói riêng đã đem lại nhiều cơ hội to lớn, cho phép tạo ra hàng loạt cây trồng mang các gene ngoại lai quy định những tính trạng quan trọng như kháng sâu, cải tiến chất lượng sản phẩm, kháng nấm, kháng virus.

Tình hình cây trồng chuyển gen trên toàn thế giới
Triệu hecta (1996 đến 2006)



Hình 6.17. Tình hình triển khai sản xuất cây trồng chuyển gene trên qui mô toàn thế giới.
: tổng diện tích trên cả thế giới, ■ : khu vực các nước phát triển, ▲ : các nước đang phát triển (James, 2006).

Cây chuyển gene bao gồm nhiều loại cây trồng khác nhau như : cỏ ba lá, bí ngô, hoa cúc, ngô, bông, dưa chuột, khoai tây, cải dầu, lúa gạo, đậu tương, hoa hướng dương và cà chua. Có 10 nhóm tính trạng được khảo nghiệm chủ yếu liên quan đến việc tăng cường tính chống chịu đói với sâu bệnh và điều kiện ngoại cảnh bất lợi. Đó là tính kháng virus, kháng vi khuẩn, kháng nấm, kháng côn trùng, kháng chất diệt cỏ, kháng kim loại nặng và chống băng giá. Các tính trạng như chất lượng sản phẩm, trong

đó có tính chín chậm, thay đổi hàm lượng protein hạt, tăng hàm lượng bột của củ, tăng hàm lượng sterol, giảm mùi hôi trong dầu ăn thực vật, biến đổi màu sắc hoa. Trong tất cả các trường hợp, tính trạng được thử phải là kết quả chuyển nạp một gene ngoại lai vào cây trồng.

6.2.7.2. Thành tựu trong nước

Ở Việt Nam, việc nghiên cứu cây trồng chuyển gene cũng bắt đầu được tiến hành ở một số phòng thí nghiệm. Đồng thời, chúng ta cũng đang mở rộng quan hệ hợp tác với các tổ chức nghiên cứu quốc tế và các công ty nước ngoài nên không thể tránh khỏi việc du nhập của các cây và sản phẩm chuyển gene. Chính vì thế, để quản lý và sử dụng cây trồng chuyển gene một cách hiệu quả, việc xác định phương pháp nhận biết cây chuyển gene là rất có ý nghĩa đối với tình hình thực tiễn hiện nay.

Tạo được những quy trình kỹ thuật nền về chuyển gene ở cây trồng, các quy phạm cơ bản để đánh giá cây trồng chuyển gene góp phần phát triển lĩnh vực khoa học tạo giống và năng lực quản lý nhà nước về an toàn sinh học đối với sinh vật chuyển gene là một bước chuẩn bị quan trọng trong quá trình nắm bắt cơ hội phát triển.

Đưa công nghệ tạo cây chuyển gene thành một công cụ hỗ trợ đắc lực cho công tác tạo giống truyền thống, đồng thời nâng cao năng lực tiếp cận các kỹ thuật hiện đại trong lĩnh vực tạo giống và khoa học cây trồng nói chung là góp phần đưa nền khoa học tạo giống nước ta tiến thêm một bước theo hướng khoa học hiện đại, góp phần đưa nền khoa học nước ta từng bước hội nhập khu vực và quốc tế trên các lĩnh vực về công nghệ sinh học thực vật nói chung và công nghệ sinh học cây trồng nói riêng.

Rút ngắn thời gian và nâng cao hiệu quả công tác tạo giống cây trồng, đặc biệt đối với những cây trồng đòi hỏi thời gian dài trong lai tạo giống, giảm chi phí sản xuất khi không phải dùng nhiều thuốc trừ sâu, hoá chất bảo vệ thực vật, đồng thời đảm bảo sức khoẻ cho con người và bảo vệ môi trường, giảm chi phí bảo vệ thực vật, hạ giá thành nông sản, tăng thu nhập cho người nông dân là những lợi ích về kinh tế và môi trường mà cây trồng chuyển gene có thể mang lại.

Đa dạng hoá sản phẩm, kéo dài tuổi thọ sản phẩm, mở rộng thị trường (hoa tươi, quả nhiệt đới), tạo thêm công việc cho nông dân, góp phần chuyển dịch cơ cấu trong sản xuất nông nghiệp có ý nghĩa kinh tế xã hội quan trọng trong giai đoạn hiện nay.

Đưa nước ta hội nhập bình đẳng vào trào lưu phát triển nông nghiệp công nghệ cao của khu vực và thế giới thông qua giảm sử dụng hoá chất và tăng cường chất lượng sản phẩm.

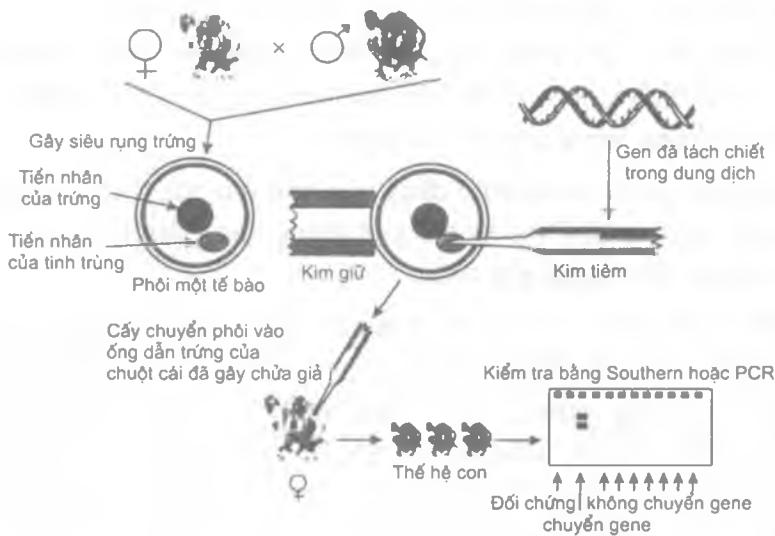
Trước mắt, nước ta tập trung chuyển gene có tính kháng côn trùng, kháng bệnh virus vào những đối tượng cây trồng không làm thực phẩm gồm loại cây công nghiệp như : bông vải, hoa cảnh ; cây lâm nghiệp như : bạch đàn, thông, hông, keo, bước đầu thử nghiệm các công nghệ tiên tiến như công nghệ mRNA vào việc tạo tính kháng bệnh virus. Một số phòng thí nghiệm cũng đã nghiên cứu chuyển gene kháng nguyên của virus vào cây trồng, thử nghiệm phương thức sản xuất vaccine dùng đường miệng.

6.3. TẠO GIỐNG ĐỘNG VẬT CHUYỂN GENE

6.3.1. Kỹ thuật chuyển gene ở động vật

Động vật chuyển gene là động vật tiếp nhận gene ngoại lai xen vào DNA genome và được di truyền qua các thế hệ. Có nhiều phương pháp chuyển gene vào động vật, nhưng có thể chia thành ba nhóm chính sau đây :

1. Chuyển gene bằng cách dung hợp các loại tế bào với nhau hay chuyển cả nhân tế bào cho vào tế bào nhận. Chuyển gene bằng tế bào gốc phôi cũng có thể được xếp chung vào nhóm phương pháp này mặc dù mới ra đời trong những năm gần đây.
2. Chuyển gene gián tiếp thông qua vector là virus hay plasmid. Vấn đề đặt ra là phải chèn được gene chuyển vào vector và phải chọn được loại vector cho phép gene chuyển chèn ổn định vào nhiễm sắc thể của tế bào nhận và không tác động lên hoạt động của gene sau này.
3. Chuyển gene trực tiếp thông qua phương pháp chuyển các đoạn nhiễm sắc thể nguyên hay vi tiêm gene dưới dạng DNA tinh sạch vào tinh trùng hay tiền nhân của hợp tử (hình 6.18).



Hình 6.18. Kỹ thuật chuyển gene trực tiếp bằng vi tiêm vào tiền nhân của hợp tử.

6.3.2. Kỹ thuật hỗ trợ chuyển gene ở động vật

Phòng thí nghiệm tiến hành chuyển gene động vật cần làm chủ các kỹ thuật phụ trợ, nhưng có tính quyết định sau :

1. Phân lập và thiết kế gene : Kỹ thuật phân lập gene được trình bày ở mục 6.2.3. Khi thiết kế vector để chuyển gene ở động vật chú ý tới mục đích nghiên cứu để sử dụng loại promoter thích hợp. Các promoter ở tế bào động vật có nguồn gốc từ động vật như metallothionein (MT), thymidine kinase, β -actin, amylase, insulin, β -lactoglobulin, adiposite P2 hoặc từ virus động vật như Simian virus (SV40), Rous sarcoma virus (RSV).
2. Công nghệ tế bào và phôi động vật bao gồm các kỹ thuật cụ thể như nuôi cấy tế bào, thụ tinh và nuôi cấy phôi *in vitro* và cấy chuyển phôi là những kỹ thuật bắt buộc phải làm chủ trước khi triển khai chuyển gene đối với động vật có vú. Ở các đối tượng như cá hay lưỡng cư, bò sát hay chim có thể bỏ qua khâu cấy chuyển phôi vì trứng hay hợp tử sau khi nhận gene chuyển được nuôi dưỡng hay ấp trực tiếp thành con non. Tuy nhiên, ở động vật có vú, kỹ thuật

xử lý gây siêu rụng trứng thực hiện được nhờ sử dụng các hormone cần được áp dụng thành thục.

3. Kỹ thuật đánh giá động vật chuyển gene, bao gồm các phương pháp sinh học phân tử như lai Southern để xác định sự có mặt của gene chuyển, lai Northern để chứng minh gene chuyển được hoạt hoá dịch mã và lai Western nhằm khẳng định sự biểu hiện thành công gene chuyển thành sản phẩm protein. Gene có thể được xác định thông qua PCR với mỗi đặc hiệu, sản phẩm phiên mã có thể được đánh giá bằng phương pháp RT-PCR, sản phẩm dịch mã được đánh giá bằng phương pháp ELISA hoặc kỹ thuật miễn dịch phóng xạ (RIA).

Theo dõi các thế hệ sau của động vật chuyển gene (F_1 , F_2 , F_3) để xác định gene chuyển có di truyền hay không và số lượng bản sao của gene chuyển là bao nhiêu trong genome của vật chủ là những nội dung cần tiếp tục sau khi chứng minh chuyển gene thành công.

6.3.3. Những gene được sử dụng chuyển vào động vật

6.3.3.1. Gene sinh trưởng

Hiện nay, các nghiên cứu tập trung chủ yếu vào việc đưa tổ hợp bao gồm gene cấu trúc của hormone sinh trưởng và promoter metallothionein vào các loài động vật như : chuột, cá, thỏ, lợn, cừu và bò. Hiệu ứng chuyển gene mang lại không giống nhau, lý do chính là cấu trúc gene chuyển trong thiết kế còn chưa thực sự thích hợp với từng loại động vật, ví dụ như promoter.

6.3.3.2. Gene mã hoá protein có giá trị làm thuốc

Các gene mã hoá hormone được coi là gene đích để sản xuất các loại protein có giá trị y dược cao. Khả năng này trở thành hiện thực nếu sử dụng tuyến sữa của động vật bậc cao để sản xuất ra các protein dạng hormone thông qua việc gắn gene đích với promoter của gene β -lactoglobulin điều khiển sự biểu hiện của gene ở tuyến sữa (hình 6.19).



Hình 6.19. Nguyên lý kỹ thuật của quy trình sản xuất protein tuyến sữa ở động vật có vú thông qua chuyển gene kiểm soát bởi promoter của gene β -lactoglobulin.

Bảng 6.17. Biểu hiện protein dược phẩm trong sữa động vật chuyển gene (Wall, 1997).

Loài chuyển gene	Gene chuyển (nguồn)	Promoter (nguồn)	Tài liệu tham khảo
Chuột	Albumin huyết thanh (người)	β -LG (cừu)	Shani, 1992
Chuột	CFTR (người)	β -CN (nhím)	DiTullio, 1992
Dê, chuột, chuột, thỏ	Chất hoạt hoá plasminogen mő (người)	WAP (chuột), WAP (chuột), α S1-CN (bò), α S1-CN (bò)	Ebert, 1991 Riego, 1993
Thỏ	Interleukin-2 (người)	β -CN (nhím)	Buhler, 1995
Chuột, lợn	Protein C (người)	WAP (chuột)	Valender, 1992
Chuột	Superoxide dismutase (người)	β -LG (cừu), WAP (chuột)	Hansson, 1994
Chuột	Trophoblastin (cừu)	α -LA (bò)	Stinnakre, 1991
Chuột	Urokinase (người)	α S1-CN (bò)	Meade, 1990
Chuột, cừu	Yếu tố đông máu IX (người)	β -LG (cừu), β -LG (cừu)	Yull, 1995 Simons, 1988
Chuột, chuột, cừu	α 1-antitrypsin (chuột)	WAP (thỏ), β -LG (cừu), β -LG (cừu)	Bischoff, 1992 Archibald, 1990 Wright, 1991
Chuột	β -interferon (người)	WAP (chuột)	Schellander, 1992
Chuột	γ -interferon (người)	β -LG (cừu)	Dobrovolsky, 1993

Những protein dược phẩm đang được nghiên cứu để sản xuất qua tuyến sữa của động vật được trình bày trên bảng 6.17. Có thể một vài ví dụ điển hình như : (1) α 1-antitrypsin và yếu tố làm đông máu IX

của người đã được tiết ra trong sữa chuột, sữa cừu với nồng độ tương ứng là 5 mg/ml và 25 mg/ml. (2) Chất hoạt hoá plasminogen mō người làm tăng đông máu đã được tiết ra ở sữa dê và sữa chuột. (3) Gene *urokinase* người đã được đưa thành công vào lợn và tiết ra ở tuyến sữa nhờ gene khởi động α -casein bò. (4) Protein C người được tạo ra từ sữa chuột và sữa lợn chuyển gene. (5) Riêng hormone sinh trưởng người nếu được sản xuất qua sữa thỏ thì giá thành sẽ giảm đáng kể so với hiện tại sản xuất nhờ vi sinh.

Để tiến hành tinh chế sản phẩm protein được thuận tiện có những thử nghiệm theo hướng phát triển động vật chuyển gene sản xuất ra được phẩm ở trong bàng quang và tiết vào nước tiểu của chúng. Các gene này mã hoá cho protein uroplakin (một thành phần tham gia hình thành nên màng bàng quang). Kerr (1998) đã nghiên cứu tạo ra chuột chuyển gene sản xuất hormone sinh trưởng người từ nước tiểu. Gene hormone sinh trưởng người được nối với promoter của gene uroplakin. Promoter này kiểm soát vị trí và thời gian hoạt động của gene. Chuột mang gene ngoại lai đã tạo ra 500 ng hormone sinh trưởng người trong 1 ml lít nước tiểu thải ra.

6.3.3.3. Gene tăng cường chống chịu bệnh tật và ngoại cảnh bất lợi

Đến nay, người ta đã biết được một số gene có khả năng kháng bệnh và chống chịu được sự thay đổi điều kiện môi trường của vật nuôi. Đưa gene *Mx* vào lợn để tạo ra được giống lợn miễn dịch với bệnh cúm, hay đưa gene *IgA* vào lợn, cừu, mở ra khả năng tạo được các giống vật nuôi miễn dịch được với bệnh.

Ở cá, chuyển gene chống lạnh AFP (antifreeze protein) đã tạo ra được các giống cá có khả năng bảo vệ cơ thể chống lại sự lạnh giá (cá hồi, cá vàng), cho phép mở rộng khả năng sống của các loài cá nuôi vào mùa đông. Đây là một thuận lợi lớn cho việc nuôi trồng các nguồn thuỷ sản quan trọng.

6.3.3.4. Gene năng suất và chất lượng vật nuôi

Nhiều phương pháp đã được đề xuất để nâng cao chất lượng dinh dưỡng và để cải tiến hiệu quả của các sản phẩm được sản xuất từ sữa như phomát, kem và sữa chua (bảng 6.18).

Bảng 6.18. Một số thay đổi thành phần của sữa được để xuất (Wall, 1997).

Sự thay đổi	Kết quả
Tăng α -CN và β -CN	Tăng khả năng bền vững của sữa đồng cho việc làm phomát, cải tiến tính bền vững đối với nhiệt và tăng hàm lượng calci
Tăng vị trí phosphoryl hoá trong casein	Tăng hàm lượng calci và cải tiến sự hoà nhũ tương
Đưa các trình tự phân giải protein vào casein	Tăng tốc độ phát triển kết cấu (cải tiến sự chín của phomát)
Tăng nồng độ κ -CN	Tăng tính ổn định của sự kết tụ casein, giảm kích thước micelle, giảm đông keo và đông tụ
Tiết β -LG	Giảm đông keo ở nhiệt độ cao, cải tiến tính tiêu hóa, giảm dị ứng và giảm nguồn cysteine sơ cấp trong sữa
Tăng α -CN và β -CN	Giảm lactose, tăng khả năng thương mại của sữa, giảm sự hình thành các tinh thể nước đá và sự điều khiển tính thấm của tuyến sữa
Thêm lactoferrin người	Tăng cường sự hấp thu sắt và bảo vệ chống lại sự nhiễm trùng ruột
Thêm các trình tự phân giải protein vào β -CN	Tăng tốc độ chín của phomát
Giảm sự biểu hiện của acetyl-CoA cacboxylase	Giảm hàm lượng mỡ, cải tiến chất lượng dinh dưỡng và giảm giá thành sản xuất sữa
Biểu hiện gene Ig	Bảo vệ chống lại các bệnh như <i>Salmonella</i> và <i>Listeria</i>
Thay thế các gene protein sữa bò bằng các gene protein sữa người	Tạo ra sữa giống như sữa người

Trong hướng này nổi bật là những nghiên cứu nâng cao chất lượng sữa bò, sữa cừu bằng cách chuyển gene lactose vào các đối tượng quan tâm. Sự biểu hiện của gene này được điều khiển bởi promoter của tuyến sữa. Trong sữa của những động vật chuyển gene này, đường lactose bị thủy phân thành đường galactose và đường glucose. Do vậy, những người không quen uống sữa cũng có thể sử dụng được sữa này mà không cần quá trình lên men. Mới đây, các nhà khoa học (Brigid Brophy, 2003) đã chuyển thêm các gene mã hoá β -casein (CSN2) và κ -casein (CSN3) bò vào các nguyên bào sợi của bò và tạo ra bò chuyển gene cho sữa có mức β -casein và κ -casein cao hơn bình thường : hàm lượng β -casein tăng lên 8-20%, hàm lượng κ -casein tăng gấp 2 lần và tỉ lệ κ -casein so với casein tổng số thay đổi một cách đáng kể. Hai loại casein là protein chủ yếu ở

trong sữa và là thành phần chính của sữa đông, chìa khoá của sự sản xuất phomát và sữa chua. Các protein này rất quan trọng, chúng làm cho sữa có hàm lượng protein cao nhưng chứa nhiều nước.

Hiện tại, người ta chú ý tới việc đưa một số gene của vi sinh vật vào cơ thể động vật. Tiến bộ nổi bật nhất trong hướng này là đưa gene mã hoá enzyme chịu trách nhiệm tổng hợp cysteine vào cừu. Cysteine là amino acid được tổng hợp từ serine nhờ hai enzyme serine trans-acetylase và O-acetylserine sulfhydrylase. Hai gene chịu trách nhiệm tổng hợp hai enzyme này là *cysE* và *cysK*. Cysteine là amino acid cơ bản rất quan trọng trong sự phát triển của lông. Những cố gắng để bổ sung amino acid này vào thức ăn đều không đạt kết quả do chúng bị phân huỷ trong ống tiêu hoá của động vật. Bởi vậy, nếu đưa được gene tổng hợp cysteine vào cơ thể động vật sẽ làm tăng năng suất lông lên rất nhiều.

Các nhà khoa học Australia đã dùng gene tổng hợp cysteine có nguồn gốc từ vi sinh vật (*E. coli* và *S. typhimurium*) để đưa vào cừu. Hai gene *cysE* và *cysK* phân lập từ hai chủng vi sinh được gắn với nhau thành một đoạn DNA, sau đó được gắn tiếp với promoter methallothionein. Ở chuột được chuyển tổ hợp gene này thì ở hầu hết các cơ quan đều có mặt tổ hợp và lượng lông tăng lên rất nhiều. Sự biểu hiện ở chuột và cừu giống nhau nên trong tương lai không xa sẽ tạo ra được giống cừu chuyển gene cysteine với năng suất lông tăng lên rất nhiều lần.

Tương tự, việc đưa gene tổng hợp amino acid cơ bản như threonine và lysine có nguồn gốc vi sinh vật vào cơ thể động vật để làm tăng hiệu quả sử dụng thức ăn của vật nuôi là có triển vọng trong tầm tay.

Việc nghiên cứu để tạo ra vật nuôi chuyển gene hormone sinh trưởng có tốc độ lớn nhanh, hiệu suất sử dụng thức ăn cao và hướng sản xuất các protein quý dùng chữa bệnh cho con người nhờ tuyển sữa động vật là có nhiều triển vọng hơn cả. Lý do là hormone sinh trưởng cũng như nhiều protein quý chỉ do một gene chịu trách nhiệm tổng hợp nên dễ biểu hiện hơn là những tính trạng do nhiều gene tham gia. Các protein được cơ thể động vật sản xuất phân lớn tiết ra qua tuyến sữa nhờ promoter β -lactoglobulin nên không gây độc hại cho cơ thể vật nuôi và có thể khai thác nhiều lần.

6.3.3.5. Gene ức chế thải ghép tạng

Gene *Fas* được phát hiện như một tác nhân làm chết các tế bào miễn dịch đã hoạt hoá làm chậm quá trình đào thải đối với tạng ghép. Người ta

kỳ vọng dùng lợn chuyển gene *Fas* làm nguồn cung cấp tạng ghép do tạng có kích thước tương tự của người, không nhiều bệnh lây sang người và có thể nuôi dưới điều kiện kiểm soát nghiêm ngặt. Ngoài ra, cũng còn có sự khác nhau rất lớn trong sự phù hợp của tạng lợn đối với người : theo thứ tự càng ít thích nghi : tim, tuy còn tạm được, rồi đến thận và cuối cùng là gan.

Có những thử nghiệm dùng tế bào ngoài cơ thể, như tế bào gan chẳng hạn để giải độc cho máu người. Nhưng đó cũng chỉ là giải pháp tạm thời giúp cho bệnh nhân có thể sống được cho đến khi có được cơ quan người để thay thế.

6.3.3.6. Gene m恁 để nghiên cứu bệnh ở người

Bảng 6.19. Một số bệnh được điều trị bằng mô hình chuột chuyển gene (Hardin, 1994).

Bệnh	Bệnh	Tài liệu tham khảo
Cystic fibrosis	CFTR	O'Neal, 1993 ; Dorin, 1992 ; Snowaert, 1992
Atherosclerosis	<i>Apo E, Ape(a), Apo A-II</i>	Havel, 1989 ; Zhang, 1992 ; Shimano, 1992 ; Fabry, 1993
Liệu pháp gene kháng Atherosclerosis	<i>Apo Al, Ape E, LDLR</i>	Plump, 1992 ; Goldstein, 1989 ; Warden, 1993
β-Thalassemia	β-globin	Yokode, 1990
Thiếu máu hồng cầu hình liềm	βs (và các dạng đột biến)	Yokode, 1990
Bệnh viêm ruột	Interleukine-2, Interleukine-10, T-cell receptor, β, MHC II	Podoisky, 1991 ; Mombaerts, 1993 ; Kuhn, 1993 ; Sadlack, 1993
Thiếu hụt miễn dịch kết hợp năng	Yokode, 1990	Rubin, 1991 ; Mombaerts, 1992
Liệu pháp gene loạn dưỡng cơ	Dystrophin	CFTR
Bệnh Alzheimer	β-amyloid	Cox, 1993 ; Sandhu, 1991
ALS (amyotrophic lateral sclerosis)	Neurofilament heavy chain	Yokode, 1990
Bệnh tiểu đường phụ thuộc insulin	Dystrophin	Stewart, 1993
CFTR	Các gene ung thư và các gene ức chế ung thư	Dystrophin

Có thể nói thành công lớn nhất trong lĩnh vực chuyển gene ở động vật là việc tạo ra các dòng chuột chuyển gene nhằm giúp nhân loại tìm ra chức năng sinh học của gene và đặc biệt là khả năng điều trị các bệnh về gene thông qua liệu pháp gene (genetherapy). Việc nghiên cứu các nguyên nhân chủ yếu của hơn 3000 bệnh di truyền của người đã được tiến hành với mục đích để phát minh các liệu pháp gene tế bào sinh dưỡng và tìm ra các phương pháp điều trị hiệu quả. Các dòng chuột nội phổi đặc biệt di truyền các kiểu hình mong muốn một cách tự phát đã cung cấp các mô hình hữu ích cho việc nghiên cứu sự phát sinh bệnh của người. Tuy nhiên một số hạn chế cần khắc phục như : (1) Khó gặp được các dòng động vật mang các triệu chứng bệnh đặc biệt và nếu có thì việc duy trì cũng rất tốn kém ; (2) Biểu hiện bệnh di truyền ở người và động vật không hoàn toàn giống nhau, nhất là những bệnh về thần kinh nên rất khó nhận ra và mô tả chính xác.

Trong thập kỷ qua, nhiều dòng chuột chuyển gene đã được tạo ra như các mô hình nghiên cứu bệnh tâm thần, tim mạch, phổi, ung thư, viêm nhiễm và miễn dịch cũng như để nghiên cứu cơ chế và sự rối loạn của chuyển hoá, sự sinh sản và sự phát triển sớm ở người (bảng 6.19).

6.3.3.7. Gene phục vụ nghiên cứu độc học

Việc thử nghiệm các chất gây đột biến và các chất gây ung thư bằng động vật chuyển gene có khả năng chỉ thị là một ứng dụng trong xét nghiệm phát hiện chất gây đột biến *in vivo* trong một loạt các loại mô khác nhau, kể cả các tế bào mầm.

Xét nghiệm các chất gây đột biến : Các dòng chuột chuyển gene như Mutamouse và Big Blue chứa gene chuyển *lacZ* và *lacI* của *E. coli* đã tích hợp vào trong genome của chuột. Thông qua hoạt động của các gene chuyển này trên đĩa nhuộm màu mà biết được genome của chuột thử có chịu tác động đột biến hay không.

Xét nghiệm các chất gây ung thư : Nguyên lý sử dụng động vật chuyển gene cho xét nghiệm các chất gây ung thư là sự có mặt của một gene chuyển thích hợp sẽ không trực tiếp gây ra khối u nhưng sẽ cho thấy tính mẫn cảm cao với chất gây ung thư. Rút ngắn thời gian xét nghiệm. Có các dòng chuột mang gene chuyển được dùng để làm chuột xét nghiệm chất gây ung thư là (1) Chuột Eμ-pim-1 mang gene ung thư đã hoạt hoá pim-1,

gene này có tốc độ gây ra khối u tự phát thấp, nhưng rất mẫn cảm với chất gây ung thư. (2) Chuột chuyển gene mang một gene ung thư đã hoạt hoá (v-H-ras, c-H-ras) hoặc một gene ức chế khối u bất hoạt (p53). (3) Chuột chuyển gene với gene sửa đổi DNA đã bất hoạt (XPA).

6.3.4. Một số động vật chuyển gene và ứng dụng

6.3.4.1. Thành tựu chuyển gene động vật

Cho đến nay, kỹ thuật vi tiêm DNA vào tiền nhân được triển khai thành công nhất trong việc tạo ra các động vật chuyển gene. Hình 6.20 sưu tập một số hình ảnh về động vật chuyển gene như chuột, thỏ, lợn, cá và khỉ. Tuy nhiên, tỷ lệ hành công của phương pháp này ở các loài động vật khác nhau là rất khác nhau, cho thấy nhu cầu cải tiến phương pháp đang là một thách thức lớn. Ngoài ra, mục đích chuyển gene cũng là một nội dung rất đáng bàn vì kinh phí cho những thí nghiệm này rất lớn.

Số lượng các loài động vật được chuyển gene ngày càng phong phú, nhưng có thể tổng kết thành tựu chuyển gene trên động vật tập trung thành công nhất trên các đối tượng sau :

1. Chuột : Chuột chuyển gene trở thành động vật mô hình nghiên cứu thuốc và bệnh di truyền. Cuộc chạy đua này thu hút những khoản đầu tư khổng lồ của các công ty y dược và của nhiều chương trình nghiên cứu quốc gia.
2. Cá : Thành công ở cá chuyển gene đã được thương mại hóa. Ví dụ điển hình việc chuyển các loại gene sinh trưởng của người, của bò và của cá vào các loại cá như cá hồi cầu vồng, cá hồi, cá rô phi, cá vàng, cá mú vằn, cá medaka, cá chép, cá nheo Mỹ, cá trê châu Phi, cá vền biển (seabream), cá hồi chấm hồng Bắc cực (Arctic charr), trong đó thành công nhất là cá chép, cá hồi, cá trê phi. Phương pháp vi tiêm đem lại hiệu quả cao nhất.
3. Gà : Nghiên cứu gà chuyển gene đang được sử dụng cho rất nhiều mục đích khác nhau như : (1) Chế tạo vaccine. (2) Sản xuất kháng thể trong trứng : chống nhiễm khuẩn (như *E. coli*) ; chống ung thư ở người ; immoglobulin ở chim hoặc IgY một cách đặc biệt, thay thế cho việc sử dụng các động vật thí nghiệm. (3) Sản xuất trứng mang

dược chất : hormone sinh trưởng người để chữa bệnh lùn, trứng cholesterol thấp, isoflavon đậu nành trong trứng. (4) Tạo dòng gà thịt giảm mỡ.



A



B



C



D



E



F

Hình 6.20. Một số động vật chuyển gene thành công. A. Chuột chuyển gene hormone sinh trưởng (bên phải) và chuột đối chứng (bên trái) (Palmiter và Brinster, 1982). B. Elba, thỏ chuyển gene protein huỳnh quang màu xanh lá cây (Kac, 2001). C. Lợn chuyển gene siêu nạc (Hammer et al, 1985). D. Cá hồi (Chinook salmon) chuyển gene hormone sinh trưởng. E. Vi tiêm DNA gene phospholipase A2 (PLA2) vào phổi giai đoạn sớm của muỗi gây sốt vàng da nhằm hạn chế ký sinh trùng phát triển trong ruột muỗi. F. ANDi – chú khỉ rhesus chuyển gene GP chào đời ngày 2/10/2000.

6.3.4.2. Ứng dụng của động vật chuyển gene

Mặc dù đã trình bày ở trên, có thể tóm tắt khía cạnh ứng dụng của nghiên cứu và phát triển sinh vật chuyển gene là nhằm các mục tiêu ứng dụng như sau : (1) Nghiên cứu cơ bản khám phá chức năng sinh học, khả năng điều hoà, mối tương tác của gene. (2) Phục vụ phát triển nông lâm ngư nghiệp và công nghiệp thực phẩm trên phương diện cải tiến giống vật

nuôi theo hướng năng suất, chất lượng, khả năng chống chịu bệnh và gia tăng giá trị thương phẩm. (3) Phục vụ y dược bảo vệ sức khoẻ con người thông qua nghiên cứu thuốc mới, vaccine, bệnh di truyền và liệu pháp gene, thông qua cung cấp tạng ghép. (4) Phát triển công nghiệp thông qua các vật liệu đặc biệt như tơ nhện trong sữa cừu chuyển gene, xét nghiệm độc học công nghiệp.

Tuy nhiên, việc kiểm soát của luật pháp đối với các giá trị của động vật chuyển gene cần hết sức chặt chẽ. Trước khi được sử dụng làm thực phẩm và lưu hành trên thị trường chúng phải vượt qua được các thử nghiệm rất ngặt nghèo về mặt an toàn thực phẩm. Khả năng rủi ro của chuyển gene đối với môi trường và hệ sinh thái là tồn tại khi nuôi trồng động vật chuyển gene được đề cập thông qua phát tán gene thiểu kiểm soát. Mặt khác, sự phát triển lan tràn của chúng làm mất tính cân bằng của hệ sinh thái, làm giảm tính đa dạng sinh học của quần thể. Do vậy, hiện nay, động vật chuyển gene được nuôi ở những khu vực được giám sát hết sức chặt chẽ để giảm thiểu tối đa khả năng lây lan vào môi trường.

TÀI LIỆU THAM KHẢO

(Mục 2 – Sao chép, phiên mã và dịch mã)

1. Adams RPL, Knowler JT, Leader DP (1992) *Biochemistry of the Nucleic Acids*, 11th edn. Chapman & Hall, London.
2. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994) *Molecular Biology of the Cell*, 3rd edn. Garland Publishing, New York.
3. Carradine CR, Drew HR (1997) *Understanding DNA. The molecule and how it works*. Academic Press, London.
4. Kozak M (1996) Interpreting cDNA sequences : some insights from studies on translation *Mamm. Genome* 7. 563-574.
5. Lewin B (1997) *Genes VI*. Oxford University Press, Oxford.
6. Lodish H, Baltimore D, Berk A, Zipursky L, Matsudaira P, Darnell J (1995) *Molecular Cell Biology*, 3rd edn. Scientific American Books, New York.
7. Stanley JP, Guthrie C (1998) Mechanical devices of the spliceosome : motors, clocks, springs and things. *Cell* 92. 315-326.
8. Tarn WY, Steitz J (1997) Pre-mRNA splicing : the discovery of a new spliceosome doubles the challenge. *Trends Biochem Sci* 22. 132-137.
9. Wickens M, Anderson P, Jackson RJ (1997) Life and death in the cytoplasm : messages from the 3' end. *Curr Opin Genet Dev* 7. 220-232

TÀI LIỆU THAM KHẢO

(Mục 4 – Lập bản đồ gene)

1. Alberts et al. (2002) The Molecular Biology of the Cell, 4th Edition Garland Publishing.
2. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA Struhl K (Eds) (1992) Short protocols in molecular biology 3rd Edition. John Wiley & Sons, Inc.
3. Baulcombe D (2002) RNA silencing. Current Biology 12 : 82-84.
4. Benjamin, Lewin (2000) Genes VII. Oxford University Press.
5. Berg, Tymoczlo, Stryer (2002) Biochemistry 5th Edition Freeman.
6. Bloom MV, Freyer GA, Mickios DA (1996) Laboratory DNA science : an introduction to recombinant DNA techniques and methods of genome analysis.
7. Brown TA Essential Molecular Biology – A Practical Approach Volumes I and II by IRL.
8. Gelvin SB, Schilperoort RA (Eds.) (1994) Plant molecular biology manual. Kluwer Academic Publisher, Dordrecht, Boston, London.
9. Glover DM (1985) DNA Cloning Volume 1 GLO LONG Hall, Introduction to Genetic Engineering 5th Edition Blackwell Scientific Publishers, Oxford, UK.
10. Jorgensen, R. (1993). The germinal inheritance of epigenetic information in plants. Phil Trans R Soc Lond B 339 :173-181.
11. Kakutani T, Jeddeloh, JA, Flowers SK, Munakata K, Richards E (1996) Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci USA 93 :12406-12411.
12. Latchman (1995) Gene Regulation – A Eukaryotic Perspective 2nd Edition Chapman and Hall.
13. Lê Duy Thành (1995) Di truyền học. NXB Nông nghiệp.
14. Lê Trần Bình, Hồ Hữu Nhị, Lê Thị Muội (1997) Công nghệ sinh học thực vật trong cải tiến giống cây trồng. NXB Nông nghiệp.

15. Lindsey, K. (1998). Transgenic Plant Research : Harwood Academic Publishers.
16. Lodish H, Baltimore D, Berk A, Zipursky SL, Matsudaira P, Darnell J (1995). Molecular Cell Biology. New York : Scientific American Books.
17. Lodish, et al. (2001) Molecular Cell Biology. 4th Edition Freeman.
18. Micklos DA, Freyer GA (1990) DNA Science..a first course in recombinant DNA technology Carolina Biological Supply Co. Cold Spring Harbor : Cold Spring Harbor Laboratory Press.
19. Nicholl DST (2002) An Introduction to Genetic Engineering 2nd Edition, Cambridge.
20. Prejls F, Meyer P (1992). The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in Petunia hybrida. *Plant J* 2 : 465 – 475.
21. Sambrook J, Russell DW (Eds.) Molecular Cloning. A laboratory Manual. 3rd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
22. Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM (1987). Molecular Biology of the Gene. Menlo Park, California : Benjamin/Cummings.
23. Williams J, Ceccarelli A, Wallace A (2001) Genetic Engineering 2nd Edition BIOS Scientific Publishers Ltd.

MỤC LỤC

1. VẬT CHẤT DI TRUYỀN	3
1.1. Cấu trúc và chức năng nucleic acid	3
1.1.1. Thành phần nucleic acid	4
1.1.2. Cấu trúc deoxyribonucleic acid	7
1.1.3. Cấu trúc ribonucleic acid	11
1.1.4. Chức năng của deoxyribonucleic acid	12
1.1.5. Chức năng của ribonucleic acid	16
1.2. Cấu trúc và chức năng protein	22
1.2.1. Thành phần của protein	23
1.2.2. Cấu trúc protein	26
1.2.3. Sinh tổng hợp protein	32
1.2.4. Xếp gấp, biến đổi và phân hủy protein	43
1.2.5. Chức năng xúc tác sinh học của enzyme	51
1.2.6. Chức năng vận động cơ học của protein	62
1.2.7. Cơ chế chung điều hòa chức năng protein	70
2. SAO CHÉP, PHIÊN MÃ VÀ DỊCH MÃ	79
2.1. Thành phần cấu tạo và liên kết hóa học	79
2.1.1. Cấu trúc polymer mạch thẳng	79
2.1.2. Liên kết trong các đại phân tử	84
2.1.3. Tầm quan trọng của liên kết hydrogen	86
2.2. Cấu trúc và sao chép DNA	87
2.2.1. Cấu trúc của DNA : một helix kép đối song	87
2.2.2. Tinh bản bảo thủ và bản liên tục	91

2.3. Phiên mã RNA và biểu hiện gene	95
2.3.1. Dòng chảy thông tin di truyền	95
2.3.2. Một DNA biểu hiện một protein hoặc RNA	97
2.3.3. Phiên mã DNA (gene) chỉ định tổng hợp RNA	98
2.3.4. Phiên mã các gene eukaryote	100
2.3.5. Biểu hiện gene đặc hiệu mô	102
2.4. Quá trình chế biến RNA	103
2.4.1. Cắt dán RNA	103
2.4.2. Hoạt động của thẻ phiên mã RNA polymerase II	107
2.5. Dịch mã	111
2.5.1. Quá trình tổng hợp polypeptide ở ribosome	111
2.5.2. Code di truyền là suy biến	114
2.5.3. Cải biến hậu dịch mã	116
2.5.4. Sự tiết protein	120
2.5.5. Cấu trúc protein biến đổi và phức tạp	123
3. ĐỌC TRÌNH TỰ VÀ NGHIÊN CỨU GENOME	126
3.1. PCR	126
3.2. Tối ưu phản ứng PCR	135
3.2.1. Khuôn DNA	135
3.2.2. Mồi	137
3.2.3 dNTP	139
3.2.4. Polymerase	139
3.2.5. Nồng độ Mg ²⁺	140
3.2.6. Phụ gia	141
3.2.7. Đêm PCR	142
3.2.8. Lớp chống bốc hơi	142
3.2.9. Chuẩn bị hỗn hợp PCR	143

3.3. Xử lý sau PCR	155
4. LẬP BẢN ĐỒ GENE	157
4.1. Giải trình tự nucleotide	157
4.1.1. Các phương pháp xác định trình tự nucleotide	157
4.1.2. Khung đọc của gene	159
4.2. Enzyme hạn chế và ứng dụng trong lập bản đồ gene	160
4.2.1. Sự phát hiện enzyme hạn chế	160
4.3. Lập bản đồ cắt hạn chế	163
4.3.1. Sự cần thiết của việc lập bản đồ hạn chế	163
4.3.2. Lập bản đồ hạn chế một loại phân tử DNA	164
4.3.3. Sử dụng điểm cắt hạn chế làm chỉ thị di truyền	167
4.4. Lai phân tử trong phân tích Southern, Northern và Western	168
4.4.1. Nguyên lý chung của lai phân tử	168
4.4.1.1. Thám truyền đại phân tử lên màng lọc	168
4.4.1.2. Đánh dấu và phát hiện phân tử	169
4.4.2. Lai Southern	170
4.4.3. Lai Northern	171
4.4.4. Lai Western	171
4.4.5. Kỹ thuật PCR	172
4.5. Lập bản đồ RFLP	173
4.5.1. Giới thiệu chung về RFLP	173
4.5.2. Thư viện chỉ thị phân tử	176
4.5.3. Xây dựng bản đồ di truyền RFLP	177
4.5.4. So sánh chỉ thị di truyền truyền thống và RFLP	180
4.5.5. Lập bản đồ di truyền bằng chỉ thị RFLP	181
4.5.5.1. Chọn cây bố mẹ	181
4.5.5.2. Tạo quần thể lập bản đồ	182

4.5.5.3. Đánh giá RFLP trong quản thể lập bản đồ	182
4.5.5.4. Ứng dụng RFLP trong chọn, tạo giống cây trồng	184
4.6. Đánh giá tính đa dạng di truyền bằng kỹ thuật RAPD	184
4.7. Lấy dấu DNA bằng kỹ thuật AFLP	186
5. KỸ THUẬT DNA TÁI TỔ HỢP	187
5.1. Các khái niệm cơ bản	187
5.2. Tạo DNA tái tổ hợp	189
5.2.1. Tách chiết DNA	192
5.2.2. Cắt DNA	194
5.2.3. Gắn dinh DNA	197
5.2.4. Khuếch đại DNA tái tổ hợp	198
5.3. Nhận dòng một gene đặc hiệu	199
5.3.1. Lựa chọn vector nhận dòng	199
5.3.1.1. <i>Plasmid</i>	199
5.3.1.2. <i>Phage lambda</i>	202
5.3.1.3. <i>Cosmid</i>	203
5.3.1.4. Các <i>phage sợi đơn</i>	204
5.3.1.5. Các vector biểu hiện	205
5.3.2. Tạo thư viện DNA	205
5.3.3. Tìm các dòng đặc trưng sử dụng mẫu dò	208
5.3.3.1. Mẫu dò để tìm DNA	208
5.3.3.2. Mẫu dò tìm các protein	212
5.3.4. Tìm các dòng đặc trưng bằng bổ trợ chúc năng	214
5.3.5. Nhận dòng vị trí	217
5.3.6. Nhận dòng gene bằng gắn đuôi	218

5.4. Sử dụng DNA nhân dòng	220
5.4.1. Sử dụng DNA nhân dòng làm mẫu để	220
5.4.2. Đò tìm nucleic acid đặc trưng trong hỗn hợp	220
5.5. Xác định trình tự DNA	224
5.6. Phát hiện và khuếch đại các trình tự bằng PCR	228
5.7. Định vị gene trên bản đồ hạn chế	230
5.8. Một thế kỷ truy tìm di truyền alkaptonuria	234
6. ỨNG DỤNG KỸ THUẬT DI TRUYỀN	238
6.1. Sản xuất protein tái tổ hợp	238
6.1.1. Lựa chọn tế bào chủ	238
6.1.2. Lựa chọn vector	240
6.1.3. Lựa chọn tag – dung hợp	242
6.1.4. Lập kế hoạch sản xuất protein tái tổ hợp	243
6.1.5. Protein dung hợp GST	246
6.1.5.1. Nhận gene và sản xuất protein	246
6.1.5.2. Tế bào chủ	247
6.1.5.3. Vector	247
6.1.5.4. Tinh chế	248
6.1.5.5. Định lượng protein	249
6.1.5.6. Cắt bỏ đoạn tag bằng enzyme	250
6.1.6. Protein dung hợp (His) ₆	251
6.1.6.1. Nhận gene và sản xuất protein	251
6.1.6.2. Tinh chế	251
6.1.6.3. Phát hiện protein	252
6.1.7. Xử lý thể protein vùi	253
6.1.8. Làm tan các dạng protein hạt	254
6.1.9. Tái thiết cấu trúc 3D	255
	321

6.1.10. Tái tạo cấu trúc 3D trên cột	255
6.1.11. Thu hoạch và chiết suất protein tái tổ hợp	256
6.1.12. Sản xuất insulin tái tổ hợp	257
6.2. Tạo giống cây trồng chuyển gene	259
6.2.1. Lựa chọn đối tượng cây trồng	259
6.2.1.1. Các đối tượng cây trồng chuyển gene	260
6.2.1.2. Các loại sinh phẩm, vật tư và hoá chất	260
6.2.2. Sưu tập nguồn gene khác nhau	261
6.2.3. Phân lập gene	262
6.2.3.1. Phân lập gene bằng PCR	262
6.2.3.2. Phân lập gene từ thư viện cDNA	263
6.2.3.3. Phân lập gene thông qua phân tích proteomic	265
6.2.4. Các nhóm gene có giá trị đang được sử dụng	266
6.2.4.1. Gene công cụ	266
6.2.4.2. Gene kháng bệnh cây trồng	270
6.2.4.3. Gene kháng côn trùng có hại	271
6.2.4.4. Gene bất dục đực	273
6.2.4.5. Gene ngược chống chín nhũn và làm chín chậm	274
6.2.4.6. Gene chống chịu điều kiện sinh thái bất lợi	275
6.2.4.7. Gene có giá trị y được	277
6.2.4.8. Biến đổi gene nhằm tăng cường sự biểu hiện	278
6.2.5. Phương pháp biến nạp gene vào thực vật	282
6.2.5.1. Điều kiện cho chuyển gene ở thực vật	282
6.2.5.2. Các phương pháp chuyển gene vào thực vật	283
6.2.5.3. Mục đích chuyển gene vào một số cây trồng ở nước ta	289
6.2.6. Phương pháp phân tích thực vật chuyển gene	291
6.2.6.1. Phương pháp phân tích tính kháng thuốc	291
6.2.6.2. Phương pháp phân tích hoá tế bào	292

<i>6.2.6.3. Phương pháp phân tích sinh học phân tử</i>	293
<i>6.2.6.4. Phương pháp phân tích chức năng sinh học</i>	296
6.2.7. Ý nghĩa kinh tế - xã hội	298
<i>6.2.7.1. Thành tựu trên thế giới</i>	298
<i>6.2.7.2. Thành tựu trong nước</i>	300
6.3. Tạo giống động vật chuyển gene	301
<i>6.3.1. Kỹ thuật chuyển gene ở động vật</i>	301
<i>6.3.2. Kỹ thuật hỗ trợ chuyển gene ở động vật</i>	302
<i>6.3.3. Những gene được sử dụng chuyển vào động vật</i>	303
<i>6.3.3.1. Gene sinh trưởng</i>	303
<i>6.3.3.2. Gene mã hoá protein có giá trị làm thuốc</i>	363
<i>6.3.3.3. Gene tăng cường chống chịu bệnh tật và ngoại cảnh bất lợi</i>	306
<i>6.3.3.4. Gene năng suất và chất lượng vật nuôi</i>	306
<i>6.3.3.5. Gene ức chế thải ghép tạng</i>	308
<i>6.3.3.6. Gene mẫu để nghiên cứu bệnh ở người</i>	309
<i>6.3.3.7. Gene phục vụ nghiên cứu độc học</i>	310
<i>6.3.4. Một số động vật chuyển gene và ứng dụng</i>	311
<i>6.3.4.1. Thành tựu chuyển gene động vật</i>	311
<i>6.3.4.2. Ứng dụng của động vật chuyển gene</i>	312
Tài liệu tham khảo	314
Mục 2 – Sao chép phiên mã và dịch mã	314
Mục 4 – Lập bản đồ gene	315

Chịu trách nhiệm xuất bản :

Chủ tịch HĐQT kiêm Tổng Giám đốc NGÔ TRẦN ÁI

Phó Tổng Giám đốc kiêm Tổng biên tập NGUYỄN QUÝ THAO

Tổ chức bìa thảo và chịu trách nhiệm nội dung :

Phó Tổng biên tập PHAN XUÂN THÀNH

Giám đốc CTCP Dịch vụ xuất bản Giáo dục Hà Nội PHAN KẾ THÁI

Biên tập nội dung :

NGUYỄN THU HUYỀN – LÊ THỊ PHƯỢNG

Biên tập kỹ thuật :

TRẦN THANH HẰNG

Trình bày bìa :

NGUYỄN THANH LONG

Sửa bản in :

LÊ NHƯ HÀ – BÍCH THUÝ

Chế bản :

CÔNG TY CỔ PHẦN THIẾT KẾ VÀ PHÁT HÀNH SÁCH GIÁO DỤC

Tổng phát hành

NHÀ XUẤT BẢN GIÁO DỤC TẠI TP. ĐÀ NẴNG

CÔNG NGHỆ SINH HỌC

TẬP MỘT : CÔNG NGHỆ GEN

Mã số : 7K623H9-NĐN

In 1.000 cuốn, khổ 16x24cm tại Công ty Cổ phần In và Dịch vụ Quảng Nam - 260 Hùng Vương, Thành phố Tam Kỳ. Số xuất bản : 90-2009/CXB/4-82/GD.
In xong và nộp lưu chiểu tháng 9 năm 2009.



VƯƠNG MIỆN KIM CƯỜNG
CHẤT LƯỢNG QUỐC TẾ

1979-2009



Barcode area containing the number 8934980973370

Cơ sở công nghệ sinh học 11



70,000



Giá : 70.000đ