

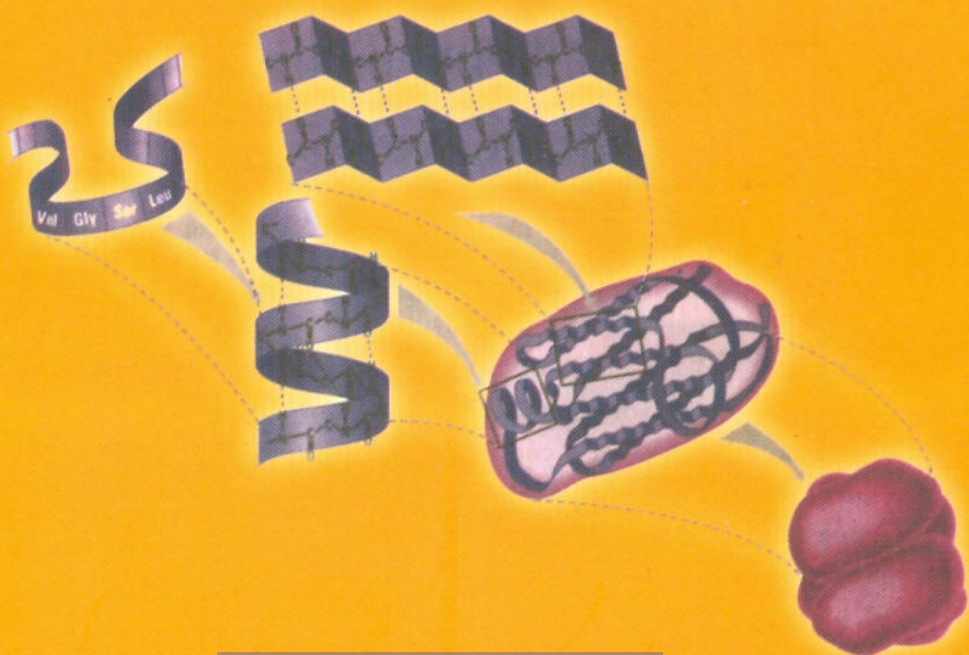
DH2.56

BỘ Y TẾ

SINH HỌC PHÂN TỬ

(DÙNG CHO ĐÀO TẠO DƯỢC SĨ ĐẠI HỌC)

GS. TS. NGUYỄN VĂN THANH (Chủ biên)



**THƯ VIỆN
HUBT**
NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM
TÀI LIỆU PHỤC VỤ THAM KHẢO NỘI BỘ

7

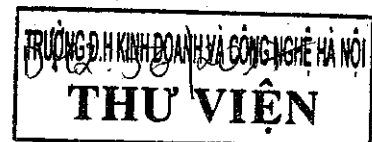
BỘ Y TẾ

SINH HỌC PHÂN TỬ

(DÙNG CHO ĐÀO TẠO DƯỢC SĨ ĐẠI HỌC)

Mã số: Đ.20.X.06

(Tái bản lần thứ ba)



NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM



Chỉ đạo biên soạn:

CỤC KHOA HỌC CÔNG NGHỆ VÀ ĐÀO TẠO – BỘ Y TẾ

Chủ biên:

GS.TS. NGUYỄN VĂN THANH

Ban biên soạn:

GS.TS. NGUYỄN VĂN THANH

PGS.TS. TRẦN CÁT ĐÔNG

PGS.TS. Trần Thu Hoa

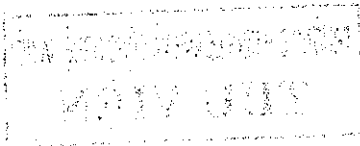
TS. HUỲNH THỊ NGỌC LAN

TS. NGUYỄN TÚ ANH

Tham gia tổ chức bản thảo:

TS. NGUYỄN MẠNH PHA

ThS. PHÍ VĂN THÂM



Lời giới thiệu

Thực hiện Nghị định 43/2000/NĐ-CP ngày 30/08/2000 của Chính phủ quy định chi tiết và hướng dẫn triển khai Luật Giáo dục, Bộ Giáo dục & Đào tạo và Bộ Y tế đã phê duyệt, ban hành các chương trình khung cho đào tạo Dược sĩ Đại học. Bộ Y tế đã tổ chức thẩm định sách và tài liệu giảng dạy - học các môn học cơ sở và chuyên môn theo chương trình mới nhằm từng bước xây dựng bộ sách chuẩn trong công tác đào tạo Dược sĩ Đại học ngành Y tế.

Sách được khoa Dược Đại học Y Dược thành phố Hồ Chí Minh biên soạn dựa trên chương trình khung đã được Bộ Giáo dục & Đào tạo và Bộ Y tế ban hành. Sách gồm 09 bài, mỗi bài được trình bày nổi bật các nội dung: mục tiêu, nội dung chuyên môn, câu hỏi lượng giá và tài liệu đọc thêm; đảm bảo yêu cầu cơ bản về kiến thức, tính chính xác và khoa học, cập nhật tiến bộ khoa học kỹ thuật vận dụng thực tiễn.

Nội dung tài liệu chủ yếu cung cấp những kiến thức cơ bản về sinh học phân tử, cụ thể là các quá trình sao chép, phiên mã, dịch mã, điều hoà hoạt động gen ở tế bào nhân nguyên thủy và tế bào nhân thật, các nguyên nhân gây đột biến và các quá trình sửa sai. Ngoài ra, còn cung cấp các kiến thức về bộ gen của tế bào nhân thật và các kỹ thuật phân tích ADN hiện đại.

Đối tượng sử dụng sách, chủ yếu là sinh viên các trường đại học Dược. Ngoài ra, sinh viên các trường đại học khác cũng có thể sử dụng như tài liệu tham khảo. Xin trân trọng giới thiệu cùng bạn đọc.

Năm 2006 cuốn sách đã được Hội đồng chuyên môn thẩm định sách giáo khoa và tài liệu dạy - học chuyên ngành Dược của Bộ Y tế thẩm định và được Bộ Y tế ban hành làm tài liệu dạy - học chính thức của Ngành Y tế trong giai đoạn hiện nay. Trong thời gian từ 3 đến 5 năm, sách cần được chỉnh lý, bổ sung, cập nhật.

Cục Khoa học Công nghệ và Đào tạo xin chân thành cảm ơn GS.TS. Nguyễn Văn Thanh, PGS.TS. Trần Cát Đông và các giảng viên Khoa Dược Đại Học Y Dược Thành phố Hồ Chí Minh đã bỏ nhiều công sức để biên soạn cuốn sách này. Sách giáo khoa không thể tránh khỏi những thiếu sót, chúng tôi rất mong nhận được ý kiến đóng góp của đồng nghiệp và sinh viên để cuốn sách có chất lượng tốt hơn.

CỤC KHOA HỌC CÔNG NGHỆ VÀ ĐÀO TẠO – BỘ Y TẾ





**THƯ VIỆN
HUBT**

TÀI LIỆU PHỤC VỤ THAM KHẢO NỘI BỘ

LỜI NÓI ĐẦU

Năm 1909, Johansen W. xuất bản chuyên khảo “Các yếu tố của học thuyết đúng đắn về biến dị và di truyền” (*Element de exakten Erblchkeitslehre*), trong đó lần đầu tiên xuất hiện từ gen. Năm 1953, Watson và Crick khám phá ra mô hình xoắn kép ADN đã thúc đẩy nhanh chóng sự phát triển của di truyền học ở mức độ phân tử. Năm 1965, J. Watson xuất bản sách “Sinh học phân tử của gen” dày 494 trang và đến năm 1976, trong lần tái bản lần thứ ba đã dày lên 739 trang. Tiếp sau đó, hàng loạt tài liệu về sinh học phân tử ra đời.

Cho đến ngày 26/06/2000, tại Washington D.C, công ty tư nhân Celera Genomics (Anh) và Dự án Bộ gen Người (*Human Genome Project*) của Viện nghiên cứu Quốc gia về Sức khỏe của Hoa Kỳ (*National Institute of Health*) đã phác thảo bản đồ bộ gen người. Theo đó, bộ gen người có 3,12 tỉ nucleotid và 97% tổng nucleotid đã được xác định trình tự, trong đó có 85% số trình tự đã đặt đúng vị trí. Khoảng 3% ADN có chứa gen, 97% còn lại là ADN “không chức năng”. Trong tổng số 3% ADN này có khoảng 30-50 nghìn gen.

Ngày 14/4/2003, Tổ chức Quốc tế Định trình tự Bộ gen Người (*International Human Genome Sequencing Consortium*) tuyên bố đã hoàn thành những công đoạn cuối cùng của bản đồ gen người.

Kế bản đồ gen, các phương pháp tìm gen có tính chất trị liệu sẽ được thực hiện ở quy mô lớn và Dược lý bộ gen (*Pharmacogenomics*) sẽ khám phá sâu hơn bộ gen người để ứng dụng trong ngành Dược.

Sách giáo khoa “Sinh học phân tử” nhằm giúp cho sinh viên Dược khoa hiểu được cấu trúc cơ bản và chức năng của gen. Sách giáo khoa gồm các chương: Nhập môn, Sao chép ADN, Các loại ARN, Sự phiên mã và mã di truyền, Sự tổng hợp protein, Điều hòa hoạt động gen, Nhiễm sắc thể nhân thật, Đột biến gen và Các phương pháp cơ bản trong nghiên cứu sinh học phân tử.

Các tác giả rất mong nhận được sự góp ý của độc giả.

CÁC TÁC GIẢ





**THƯ VIỆN
HUBT**

TÀI LIỆU PHỤC VỤ THAM KHẢO NỘI BỘ

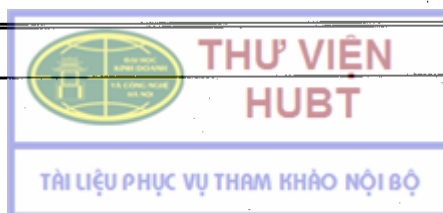
MỤC LỤC

Lời giới thiệu	3
Lời nói đầu	5
Các từ viết tắt	11
Bảng đối chiếu từ khóa Việt-Anh	13
Bảng đối chiếu từ khóa Anh-Việt	19
Bài 1. Nhập môn sinh học phân tử	25
1.1. Lược sử	25
1.1.1. Giai đoạn hình thành các tiền đề.....	25
1.1.2. Giai đoạn Sinh học phân tử ra đời.....	26
1.1.3. Sinh học phân tử hiện đại.....	26
1.2. Đối tượng và phương pháp nghiên cứu	30
1.2.1. Định nghĩa.....	30
1.2.2. Học thuyết trung tâm.....	31
1.2.3. Các phương pháp nghiên cứu.....	31
1.3. Những đóng góp lớn của sinh học phân tử hiện nay	32
1.3.1. Genomics: giải mã bộ gen và ngành hệ gen học.....	32
1.3.2. Proteomics: phân tích biến động protein và ngành hệ protein học.....	34
1.3.3. Genomics, proteomics và sự phát triển thuốc.....	35
1.3.4. Sản xuất và sử dụng chip ADN.....	38
1.3.5. Chuyển gen vào cây trồng.....	38
1.3.6. Tin sinh học.....	39
1.3.7. Công nghệ nano sinh học.....	39
Bài 2. Sao chép ADN	41
2.1. Khái niệm	41
2.2. Sự sao chép của ADN	42
2.2.1. Thí nghiệm của Meselson và Stahl.....	43
2.2.2. Các yếu tố cần thiết cho sự sao chép ADN.....	43
2.2.3. Các ADN polymerase.....	44
2.2.4. Quá trình sao chép ADN ở <i>E. coli</i>	46
2.2.5. Sao chép ADN ở tế bào nhân thật.....	51
2.2.6. Sự sao chép ở virus và phage.....	53
2.3. Sửa sai trong sao chép và khi không sao chép	57
2.3.1. Sửa sai trong sao chép.....	58



2.3.2. Sửa sai khi không sao chép.....	58
Bài 3. Các loại ARN	61
3.1. Khái niệm	61
3.2. Các ARN và vai trò của chúng	63
3.2.1. ARN ribosom.....	63
3.2.2. Các ARN vận chuyển.....	68
3.2.3. ARN thông tin.....	71
3.2.4. Các ARN nhân nhỏ và các ARN tế bào chất nhỏ	73
3.2.5. Cắt nối ở ARN tế bào nhân thật.....	74
Bài 4. Sự phiên mã và mã di truyền	82
4.1. Mở đầu	82
4.2. Nguyên tắc chung	83
4.3. Sự phiên mã ở tế bào nhân nguyên thủy	84
4.3.1. ARN polymerase của <i>E. coli</i>	85
4.3.2. Promoter.....	85
4.3.3. Khởi đầu và nối dài chuỗi ARN.....	89
4.3.4. Kết thúc quá trình tổng hợp ARN.....	90
4.4. Sự phiên mã ở tế bào nhân thật	91
4.4.1. Các gen gián đoạn.....	92
4.4.2. ARN polymerase của tế bào nhân thật.....	93
4.4.3. Sự phiên mã do ARN polymerase I.....	93
4.4.4. Sự phiên mã bởi polymerase II.....	95
4.4.5. Sự phiên mã bởi ARN polymerase III	97
4.5. Phiên mã ngược ở retrovirus	99
4.6. Mã di truyền	100
Bài 5. Quá trình dịch mã	104
5.1. Mở đầu	104
5.2. Các yếu tố cần thiết cho quá trình dịch mã	104
5.2.1. Ribosom	104
5.2.2. ARN thông tin.....	106
5.2.3. ARN vận chuyển và các enzym tổng hợp aminoacyl- tARN	108
5.2.4. Các yếu tố dịch mã.....	111
5.3. Diễn biến dịch mã ở ribosom (chu trình ribosom)	111
5.3.1. Khởi đầu.....	112
5.3.2. Nối dài.....	117
5.3.3. Kết thúc.....	120
5.4. Nhu cầu năng lượng cho quá trình sinh tổng hợp protein	120

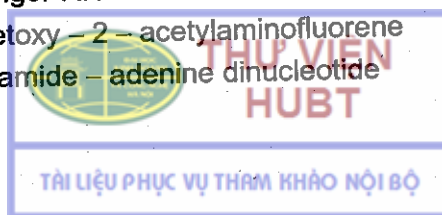
5.5. Độ chính xác của quá trình dịch mã	121
5.5.1. Aminoacyl hóa tARN	121
5.5.2. Tương tác codon - anticodon	122
5.5.3. Kết thúc sớm, kết thúc muộn	122
5.6. Các yếu tố ức chế quá trình dịch mã	122
Bài 6. Điều hòa hoạt động gen	126
6.1. Mở đầu	126
6.2. Điều hòa quá trình sao chép	127
6.3. Điều hòa quá trình phiên mã	128
6.3.1. Kiểm soát cảm ứng âm	130
6.3.2. Kiểm soát ức chế âm	134
6.3.3. Cơ chế điều hòa suy giảm	136
6.3.4. Kiểm soát cảm ứng dương	137
6.3.5. Đa kiểm soát	138
6.4. Kiểm soát sau dịch mã	138
Bài 7. Bộ gen tế bào nhân thật	143
7.1. Mở đầu	143
7.2. Tổ chức bộ gen ở tế bào nhân thật	143
7.2.1. Kích thước của bộ gen tế bào nhân thật - giá trị C	143
7.2.2. Sơ đồ khái quát về các loại trình tự ADN	144
7.3. Các mức độ điều hòa biểu hiện gen	153
7.3.1. Mức độ chất nhiễm sắc	153
7.3.2. Mức độ phiên mã	154
7.3.3. Mức độ sau phiên mã	155
7.3.4. Mức độ dịch mã	155
7.3.5. Mức độ sau dịch mã	155
7.4. Điều hòa hoạt tính gen của tế bào nhân thật	156
7.4.1. Các promoter	157
7.4.2. Các yếu tố tăng cường (enhancer)	158
7.4.3. Các yếu tố protein tham gia vào quá trình điều hòa biểu hiện gen, các yếu tố <i>trans</i>	159
7.4.4. Hormon	161
7.5. Sự kiểm soát các chất thường gặp trong nhân	161
7.5.1. Sự dôi dào ARN	161
7.5.2. Sự khuếch đại gen	162
Bài 8. Đột biến gen	165
8.1. Mở đầu	165



8.2. Các loại đột biến	168
8.2.1. Đột biến điểm.....	168
8.2.2. Đột biến đa điểm.....	169
8.3. Nguyên nhân đột biến	170
8.3.1. Đột biến tự nhiên	170
8.3.2. Đột biến cảm ứng.....	171
8.4. Các cơ chế chống lại đột biến	178
8.4.1. Đảo nghịch sai hỏng.....	178
8.4.2. Loại bỏ hư hỏng.....	179
8.4.3. Dung nạp ADN sai hỏng.....	181
8.4.4. Sửa sai nhờ hệ thống SOS (cấp cứu).....	183
8.5. Các tính trạng đột biến và protein đột biến	183
8.5.1. Đột biến sinh dưỡng ở vi sinh vật.....	184
8.5.2. Đột biến và bệnh ở người.....	185
8.6. Đột biến gen và ung thư	186
8.6.1. Các nhóm gen liên quan tới ung thư.....	187
8.6.2. Gen ung thư và gen tiền ung thư.....	187
8.6.3. Các đột biến ở gen ức chế khối u.....	188
8.7. Các hệ thống chọn lọc đột biến ở vi sinh vật	188
8.7.1. Phương pháp đề kháng.....	189
8.7.2. Phương pháp làm giàu chậm.....	189
8.7.3. Phương pháp làm giàu hạn chế.....	189
8.7.4. Phương pháp làm giàu nhờ penicillin	189
8.7.5. Phương pháp lọc	190
8.7.6. Phương pháp in	190
Bài 9. Một số phương pháp và kỹ thuật cơ bản trong sinh học phân tử	194
9.1. Công cụ cơ bản	194
9.1.1. Chủng vi sinh vật.....	194
9.1.2. Vector tạo dòng.....	195
9.1.3. Các enzym làm biến đổi acid nucleic.....	199
9.2. Một số phương pháp và kỹ thuật cơ bản	205
9.2.1. Chiết tách và tinh chế vật liệu di truyền.....	205
9.2.2. Kỹ thuật lai acid nucleic	209
9.2.3. Kỹ thuật PCR (polymerase chain reaction).....	212
9.2.4. Kỹ thuật tái tổ hợp ADN và tạo dòng gen	218
9.2.5. Xác định trình tự của acid nucleic	223
Đáp án	229
Tài liệu tham khảo	230

CÁC TỪ VIẾT TẮT

A	Adenine
ADN	Acid desoxyribonucleic
AIDS	Acquired immune – deficiency syndrome
AP	Apurinic or apyrimidinic
ARN	Acid ribonucleic
bp	Base pair
C	Cytosine
cADN	Complementary DNA
cds	Coding sequence
CpG	C phosphat G
CTF	CCAAT binding Trascription Factor
DMD	Duchenne muscular dystrophy
DNA	Desoxyribonucleic acid
Dnase	Deoxyribonuclease
dNTP	Desoxyribonucleozid triphosphat
dsDNA	Double – stranded DNA
eEF	Eukaryote elongation factor
EF	Elongation factor
eIF	Eukaryote initiation factor
eRF	Eukaryote release factor
EGF	Epidermal growth factor
G	Guanine
GMO	Genetic modified organism
Hfr	High frequency recombination
Hft	High frequency transduction
HIV	Human immunodeficiency virus
hnRNA	Heterogenous nuclear RNA
lcr	Insensitive to catabolite repression
IF	Initiation factor
IGS	Internal guide sequence
Kb	Kilobase
LINE	Long interspersed repetitive element
mRNA	Messenger RNA
NAAAF	N – acetoxy – 2 – acetylaminofluorene
NAD	Nicotinamide – adenine dinucleotide



NER	Nucleotide excision repair
NF	Neurofibromatosis
NF1	Neurofibromatosis type 1
NMN	Nicotinamide mononucleotide
NMP	Nucleotide monophosphat
Ori	Origin of replication
PCR	Polymerase chain reaction
PDGF	Platelet – derived growth factor
PFGE	Pulse field gel electrophoresis
Pi	Inefficient promoter
PKU	Phenylketonuria
PR	Photoreactivation
Rad	Radiation absorbed dose
rARN	Ribosome RNA
RE	Restriction enzyme, restriction endonuclease
Rem	Roentgen equivalent man
RF	Release factor
RFI	Replicative form i
RFLP	Restriction fragments length polymorphism
Rnase	Ribonuclease
RNP	Ribonucleoprotein
RRF	Ribosome release factor
scRNA	Small cytoplasmic RNA
SINE	Short interspersed repetitive element
snRNA	Small nuclear RNA
SRP	Signal recognition particle
SSB	Single – stranded binding
ssADN	Single – stranded DNA
T	Thymine
tRNA	Transfer RNA
THE	Transposable human element
THR	Thyroid Hormon Receptor
tRNA	Transfer RNA
Tm	Melting temperature
U	Uracile
UTR	Untranslated region

BẢNG ĐỐI CHIẾU THUẬT NGỮ VIỆT – ANH

Tiếng Việt	Tiếng Anh
ADN bổ sung	Complementary DNA
ADN sợi đôi	Double – stranded DNA
ADN sợi đơn	Single – stranded DNA
ADN vi vệ tinh	Microsatellite DNA
Alkapton niệu	Alkaptonuria
ARN nhân không đồng nhất, hnARN	Heterogenous nuclear RNA
ARN nhân nhỏ, snARN	Small nuclear RNA
ARN polymerase phụ thuộc ADN	DNA – dependent RNA polymerase
ARN ribosom	Ribosomal RNA
ARN tế bào chất nhỏ, scARN	Small cytoplasmic RNA
ARN thông tin	Messenger RNA
ARN vận chuyển	Transfer RNA
Base đồng đẳng	Base analogue
Bệnh hồng ban lupus	Systemic lupus erythematosus
Bệnh loạn dưỡng cơ Duchenne	Duchenne muscular dystrophy
Biến nạp	Transformation
Bộ ba kết thúc	Stop codon
Bộ gen	Genome
Bộ suy giảm	Attenuator
Bong bóng sao chép	Replication bubble
Chạc ba sao chép	Replicating fork
Chất cảm ứng	Inducer
Chất dị nhiễm sắc	Heterochromatin
Chất dị nhiễm sắc cơ cấu	Constitutive heterochromatin
Chất hoạt hoá	Activator
Chất nguyên nhiễm sắc	Euchromatin
Chất nhiễm sắc	Chromatin



Tiếng Việt	Tiếng Anh
Chất siêu ức chế	Superrepressor
Chất ức chế	Inhibitor, repressor
Chất ức chế gốc	Aporepressor
Chip ADN	DNA microarray
Chuỗi xoắn kép	Double helix
Chuyển vị nghịch	Retrotransposition
Cơ chế bán bảo tồn	Semiconservative mechanism
Cơ chế bảo tồn	Conservative mechanism
Cơ chế lăn vòng	Rolling circle mechanism
Công nghệ di truyền	Genetic engineering
Công nghệ sinh học	Biotechnology
Công nghệ nano sinh học	Bionanotechnology
Dấu ấn ADN	DNA fingerprint
Dị xúc tác	Heterocatalysis
Dự án bộ gen người	Human genome project
Đa cistron	Polycistron
Đa kiểm soát	Multiple control
Đầu dính	Cohesive end
Đầu tù	Blunt end
Điểm gốc sao chép	Origin of replication
Điện di gel	Gel electrophoresis
Điện di trường xung	Pulse field gel electrophoresis
Điều hoà âm	Negative control
Điều hoà dương	Positive control
Định kiểu gen	Genotyping
Đoạn chèn	Insertion sequence
Đoạn dò	Probe
Đoạn Okazaki	Okazaki fragment
Đối mã	Anticodon
Đơn cistron	Monocistron
Đơn vị phiên mã	Operon
Đơn vị sao chép	Replicon
Đột biến cảm ứng	Induced mutation
Đột biến cùng nghĩa	Sense mutation
Đột biến lặn	Silent mutation

Tiếng Việt**Tiếng Anh**

Đột biến lệch nghĩa	Missense mutation
Đột biến tự nhiên	Spontaneous mutation
Đột biến vô nghĩa	Nonsense mutation
Enzym cắt giới hạn	Restriction enzyme
Enzym phiên mã ngược	Reverse transcriptase
Gây đột biến điểm định hướng	Site – directed mutagenesis
Gen có thể cảm ứng	Inducible gene
Gen có thể ức chế	Repressible gene
Gen điều hoà	Regulatory gene
Gen giả	Pseudogene
Gen nhảy	Transposon
Gen tiền ung thư	Proto – oncogene
Gen ung thư	Oncogene
Hạch nhân	Nucleolus
Hạt F	F – prime
Bộ gen học	Genomics
Hệ protein	Proteome
Hệ protein học	Proteonomics
Heterogenous nuclear RNA	hnRNA
Hoá miễn dịch	Immunochemistry
Học thuyết trung tâm	Central dogma
Hộp Pribnow	Pribnow box
Hộp TATA	TATA box
Hợp tử	Zygote
Hợp tử không hoàn toàn	Merozygote
Khả nạp	Competence
Khuôn mẫu	Template
Kiểm soát âm	Negative control
Kiểm soát cảm ứng âm	Negative inducible control
Kiểm soát cảm ứng dương	Positive inducible control
Kiểm soát dương	Positive control
Kiểm soát sau dịch mã	Post – translation control
Kiểm soát ức chế âm	Negative repressible control
Kiểu gen	Genotype
Kiểu hình	Phenotype



Tiếng Việt	Tiếng Anh
Kỹ thuật di truyền	Genetic engineering
Lai tại chỗ	In situ hybridization
Lai vết ADN	Southern blot
Lai vết ARN	Northern blot
Lai vết protein	Western blot
Lồng ghép	Catena
Lục lạp	Chloroplast
Lực tiếp hợp	Matting force
Mã di truyền	Genetic code
Màng tế bào chất	Cell membrane
Mồi	Primer
Môi trường tối thiểu	Minimal media
Ngoài nhiễm sắc thể	Extrachromosome
Ngón tay kẽm	Zinc finger
Nhân con	Nucleolus
Nhiễm sắc thể	Chromosome
Nhiễm sắc thể đa sợi	Polytene chromosome
Nicotinamide mononucleotide	NMN
Nicotinamide – adenine dinucleotide	NAD
Nút	Loop
Nút kẹp tóc	Hairpin loop
Operon có thể cảm ứng	Inducible operon
Phage ôn hoà	Temperate phage
Phần của gen cấu trúc không mang mã	Intron
Phần của gen cấu trúc mang mã	Exon
Phần thừa ở đầu	Terminal redundancy
Phản ứng chuỗi polymerase	Polymerase chain reaction
Phenyl ceton niệu	Phenylketonuria
Phức hợp khởi động đóng	Closed promoter complex
Phức hợp khởi động mở	Open promoter complex
Phức hợp nhận diện tín hiệu xuất	Signal recognition particle
Promoter vệ tinh	Spacer promoter
Protein căng mạch	Single – stranded binding protein
Protein điều hoà	Regulatory protein
Protein gắn chóp	Cap – binding protein

Tiếng Việt	Tiếng Anh
Protein hiệu ứng	Effector protein
Protein ức chế	Repressor protein
Quang phục hồi	Photoreactivation
Sản phẩm khuếch đại	Amplicon
Sau dịch mã	Post – translation
Sinh học phân tử	Molecular biology
Sinh vật chuyển gen	Genetic modified organism
Sợi sớm	Leading strand
Sợi khuôn sớm	Leading strand template
Sợi khuôn muộn	Lagging strand template
Sợi muộn	Lagging strand
Sự biến hình dị lập thể	Allosteric transformation
Sự cắt nối	Splicing
Sự đa hình độ dài các đoạn cắt giới hạn	Restriction fragments length polymorphism
Sự dị biệt codon	Codon bias
Sự hồ biến	Tautomerization
Sự phiên mã	Transcription
Sự sao chép	Replication
Sự tự cắt nối	Self – splicing
Sự tự điều hoà	Autoregulation
Sự ức chế ngược	Feedback inhibition
Sửa chữa bằng cách cắt nucleotid	Nucleotide excision repair
Tác nhân gây đột biến	Mutagen
Tải nạp	Transduction
Tế bào cho	Donor
Tế bào nhận	Recipient
Tế bào nhân nguyên thủy	Prokaryote
Tế bào nhân thật	Eukaryote
Thành tế bào	Cell wall
Thể cắt nối	Spliceosome
Thể đa hình	Polymorphism
Thể khuyết dưỡng	Auxotroph
Thể nguyên dưỡng	Prototroph
Thể nhân	Nucleoid

TRƯỜNG D. H. KINH DOANH VÀ CÔNG NGHỆ HÀ NỘI
THU VIỆN HUBT

TRƯỜNG D. H. KINH DOANH VÀ CÔNG NGHỆ HÀ NỘI
THU VIỆN

Tiếng Việt	Tiếng Anh
Thể sao chép	Replicative form
Thể sinh dưỡng	Soma
Thụ thể	Receptor
Thực khuẩn thể	Bacteriophage
Tiền thực khuẩn	Prophage
Tiền virus	Provirus
Tiếp hợp	Conjugation
Tin sinh học	Bioformatics
Tổ chức hạch nhân	Nucleolus organizer
Trình tự chung	Consensus sequence
Trình tự điều hoà	Operator
Trình tự hướng dẫn nội tại	Internal guide sequence
Trình tự kết thúc	Terminator
Trình tự mã hoá	Coding sequence
Trình tự tăng cường	Enhancer sequence
Ung thư giác mạc di truyền	Retinoblastoma
Vị trí AP	AP site
Vị trí điều hoà	Regulatory site
Vị trí khởi đầu	Initiator site
Vỏ virus	Capsid
Vùng khởi động	Promoter
Vùng không dịch mã	Untranslated region
Vùng không mã hoá	Noncoding region
Vùng mã hoá	Coding region
Yếu tố di truyền động nghịch	Retroposon, retrotransposon
Yếu tố di truyền vận động	Transposable genetic element
Yếu tố giới tính	Sex – factor
Yếu tố kéo dài	Elongation factor
Yếu tố kết thúc	Release factor
Yếu tố khởi đầu	Initiation factor
Yếu tố phóng thích ribosom	Ribosome release factor
Yếu tố tăng cường	Enhancer
Yếu tố tăng trưởng thượng bì	Epidermal growth factor
Yếu tố tăng trưởng từ tiểu cầu	Platelet – derived growth factor



BẢNG ĐỐI CHIẾU THUẬT NGỮ ANH – VIỆT

Tiếng Anh	Tiếng Việt
Activator	Chất hoạt hoá
Alkaptonuria	Alkapton niệu
Allosteric transformation	Sự biến hình dị lập thể
Amplicon	Sản phẩm khuếch đại
Anticodon	Đối mã
AP site	Vị trí AP
Aporepressor	Chất ức chế gốc
Attenuator	Bộ suy giảm
Autoregulation	Sự tự điều hoà
Auxotroph	Thể khuyết dưỡng
Bacteriophage	Thực khuẩn thể
Base analogue	Base đồng đẳng
Bioinformatics	Tin sinh học
Bionanotechnology	Công nghệ nano sinh học
Biotechnology	Công nghệ sinh học
Blunt end	Đầu tù
Cap	Chóp
Cap – binding protein	Protein gắn chóp
Capsid	Vỏ virus
Catena	Lông ghép
Cell membrane	Màng tế bào chất
Cell wall	Thành tế bào
Central dogma	Học thuyết trung tâm
Chloroplast	Lục lạp
Chromatin	Chất nhiễm sắc
Chromosome	Nhiễm sắc thể
Closed promoter complex	Phức hợp khởi động đóng
Coding region	Vùng mã hoá



Tiếng Anh	Tiếng Việt
Coding sequence	Trình tự mã hoá
Codon bias	Sự dị biệt codon
Cohesive	Dính
Cohesive end	Đầu dính
Competence	Khả nạp
Complementary DNA	ADN bổ sung
Conjugation	Tiếp hợp
Consensus sequence	Trình tự chung
Conservative mechanism	Cơ chế bảo tồn
Constitutive heterochromatin	Chất dị nhiễm sắc cơ cấu
Corepressor	Chất đồng ức chế
De novo	Mới
DNA fingerprint	Dấu ấn ADN
DNA microarray	Chip ADN
DNA – dependent RNA polymerase	ARN polymerase phụ thuộc ADN
Donor	Tế bào cho
Double helix	Chuỗi xoắn kép
Double – stranded DNA	ADN sợi đôi
Duchenne muscular dystrophy	Bệnh loạn dưỡng cơ Duchenne
Effector protein	Protein hiệu ứng
Elongation factor	Yếu tố kéo dài
Enhancer	Yếu tố tăng cường
Enhancer sequence	Trình tự tăng cường
Epidermal growth factor	Yếu tố tăng trưởng thượng bì
Euchromatin	Chất nguyên nhiễm sắc
Eukaryote	Tế bào nhân thật
Exon	Phần của gen cấu trúc mang mã
Extrachromosome	Ngoài nhiễm sắc thể
Feedback inhibition	Sự ức chế ngược
F – prime	Hạt F
Gap	Khoảng trống
Gel electrophoresis	Điện di gel
Genetic code	Mã di truyền
Genetic engineering	Công nghệ di truyền, kỹ thuật di truyền
Genetic modified organism	Sinh vật chuyển gen



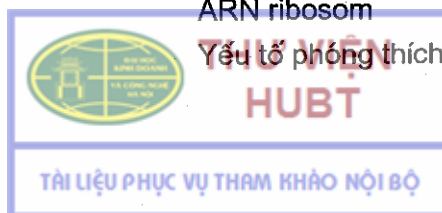
Tiếng Anh	Tiếng Việt
Genome	Bộ gen
Genomics	Bộ gen học
Genotype	Kiểu gen
Genotyping	Định kiểu gen
Hairpin loop	Nút kẹp tóc
Heterocatalysis	Đị xúc tác
Heterochromatin	Chất dị nhiễm sắc
Heterochromatin	Chất dị nhiễm sắc
Heterogenous nuclear RNA	ARN nhân không đồng nhất, hnARN
Human genome project	Dự án bộ gen người
Immunochemistry	Hoá miễn dịch
In silico	Trên máy tính
In situ	Tại chỗ
In situ hybridization	Lai tại chỗ
In vitro	Trong ống nghiệm
In vivo	Trên sinh vật
Induced mutation	Đột biến cảm ứng
Inducer	Chất cảm ứng
Inducible gene	Gen có thể cảm ứng
Inducible operon	Operon có thể cảm ứng
Inhibitor	Chất ức chế
Initiation factor	Yếu tố khởi đầu
Initiator site	Vị trí khởi đầu
Insertion sequence	Đoạn chèn
Internal guide sequence	Trình tự hướng dẫn nội tại
Intron	Phần của gen cấu trúc không mang mã
Lagging strand	Sợi muộn
Lagging strand template	Sợi khuôn muộn
Leading strand	Sợi sớm
Leading strand template	Sợi khuôn sớm
Loop	Nút
Matting force	Lực tiếp hợp
Merozygote	Hợp tử không hoàn toàn
Messenger RNA	ARN thông tin
Microsatellite DNA	ADN vi vệ tinh



Tiếng Anh**Tiếng Việt**

Minimal media	Môi trường tối thiểu
Missense mutation	Đột biến lệch nghĩa
Molecular biology	Sinh học phân tử
Monocistron	Đơn cistron
Multiple control	Đa kiểm soát
Mutagen	Tác nhân gây đột biến
Negative control	Điều hoà âm, kiểm soát âm
Negative inducible control	Kiểm soát cảm ứng âm
Negative repressible control	Kiểm soát ức chế âm
Noncoding region	Vùng không mã hoá
Nonsense mutation	Đột biến vô nghĩa
Northern blot	Lai vết ARN
Nucleoid	Thể nhân
Nucleolus	Hạch nhân
Nucleolus organizer	Tổ chức hạch nhân
Nucleotide excision repair	Sửa chữa bằng cách cắt nucleotid
Okazaki fragment	Đoạn Okazaki
Oncogene	Gen ung thư
Open promoter complex	Phức hợp khởi động mở
Operator	Trình tự điều hoà
Operon	Đơn vị phiên mã
Phage	Thực khuẩn
Phenotype	Kiểu hình
Phenylketonuria	Phenyl ceton niệu
Photoreactivation	Quang phục hồi
Platelet – derived growth factor	Yếu tố tăng trưởng từ tiểu cầu
Polycistron	Đa cistron
Polymerase chain reaction	Phản ứng chuỗi polymerase
Polymorphism	Thể đa hình
Polytene chromosome	Nhiễm sắc thể đa sợi
Positive control	Điều hoà dương, kiểm soát dương
Positive inducible control	Kiểm soát cảm ứng dương
Post – translation	Sau dịch mã
Post – translation control	Kiểm soát sau dịch mã
Pribnow box	Hộp Pribnow

Tiếng Anh	Tiếng Việt
Primer	Mồi
Probe	Đoạn dò
Prokaryote	Tế bào nhân nguyên thủy
Promoter	Vùng khởi động
Prophage	Tiềm thực khuẩn
Proteome	Hệ protein
Proteomics	Hệ protein học
Proto – oncogene	Gen tiền ung thư
Prototroph	Thể nguyên dưỡng
Provirus	Tiền virus
Pseudogene	Gen giả
Pulse field gel electrophoresis	Điện di trường xung
Receptor	Thụ thể
Recipient	Tế bào nhận
Regulatory gene	Gen điều hoà
Regulatory protein	Protein điều hoà
Regulatory site	Vị trí điều hoà
Release factor	Yếu tố kết thúc
Replicating fork	Chạc ba sao chép
Replication	Sự sao chép
Replication bubble	Bong bóng sao chép
Replicative form	Thể sao chép
Replicon	Đơn vị sao chép
Repressible gene	Gen có thể ức chế
Repressor	Chất ức chế
Repressor protein	Protein ức chế
Restriction enzyme	Enzym cắt giới hạn
Restriction fragments length polymorphism	Sự đa hình độ dài các đoạn cắt giới hạn
Retinoblastoma	Ung thư võng mạc
Retroposon	Yếu tố di truyền động nghịch
Retrotransposition	Chuyển vị nghịch
Retrotransposon	Yếu tố di truyền động nghịch
Reverse transcriptase	Enzym phiên mã ngược
Ribosomal RNA	ARN ribosom
Ribosome release factor	Yếu tố phóng thích ribosom



Tiếng Anh	Tiếng Việt
Rolling circle mechanism	Cơ chế lăn vòng
Self – splicing	Sự tự cắt nối
Semiconservative mechanism	Cơ chế bán bảo tồn
Sex – factor	Yếu tố giới tính
Signal recognition particle	Phức hợp nhận diện tín hiệu xuất
Silent mutation	Đột biến cùng nghĩa, đột biến lặn
Single – stranded binding protein	Protein gắn sợi đơn
Single – stranded DNA	ADN sợi đơn
Site – directed mutagenesis	Gây đột biến điểm định hướng
Small cytoplasmic RNA	ARN tế bào chất nhỏ, scARN
Small nuclear RNA	ARN nhân nhỏ, snARN
Soma	Thể sinh dưỡng
Southern blot	Lai vết ADN
Spacer promoter	Promoter vệ tinh
Spliceosome	Thể cắt nối
Splicing	Sự cắt nối
Spontaneous mutation	Đột biến tự nhiên
Stop codon	Bộ ba kết thúc
Superrepressor	Chất siêu ức chế
Systemic lupus erythematosus	Bệnh hồng ban lupus
TATA box	Hộp TATA
Tautomerization	Sự hỗ biến
Temperate phage	Phage ôn hoà
Template	Khuôn mẫu
Terminator	Trình tự kết thúc
Transcription	Sự phiên mã
Transduction	Tải nạp
Transfer RNA	ARN vận chuyển
Transformation	Biến nạp
Transposable genetic element	Yếu tố di truyền vận động
Transposon	Gen nhảy
Untranslated region	Vùng không dịch mã
Western blot	Lai vết protein
Zinc finger	Ngón tay kẽm
Zygote	Hợp tử

Bài 1

NHẬP MÔN SINH HỌC PHÂN TỬ

MỤC TIÊU

- Trình bày được lịch sử và phương pháp nghiên cứu sinh học phân tử.
- Nêu được mục tiêu và đối tượng môn học.
- Trình bày được các thành tựu do sinh học phân tử đem lại.

1.1. LƯỢC SỬ

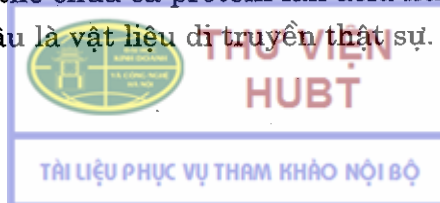
Quá trình phát triển của sinh học phân tử (SHPT) có thể chia thành ba giai đoạn chính.

1.1.1. Giai đoạn hình thành các tiền đề

Friedrich Miescher khám phá acid nucleic vào năm 1869, khi đó ông chỉ là một bác sĩ 22 tuổi. Khi thủy phân mủ bệnh nhân bằng pepsin và acid hydrochloric rồi chiết với ether, ông ta thu được nhân tế bào. Khi xử lý nhân tế bào với kiềm, ông thu được một chất tủa mà ông gọi là nuclein và do tính chất acid của nó nên được gọi là acid nucleic. Về sau Miescher sử dụng tinh trùng của cá hồi để chiết acid nucleic cho nghiên cứu của mình.

Gregor Mendel khám phá quy luật di truyền năm 1865, khi nghiên cứu sự truyền tính trạng trên cây đậu Hà Lan.

Vào đầu thế kỷ XX, **Walter Sutton** và **Theodor Boveri** đã thiết lập mối quan hệ giữa Di truyền học và Sinh học khi quan sát sự phân ly của nhiễm sắc thể ở tế bào. Nhưng vì nhiễm sắc thể chứa cả protein lẫn acid nucleic nên người ta bắt đầu nghiên cứu để xác định đâu là vật liệu di truyền thật sự.



Frederick Griffith là một nhân viên y tế ở London với nhiệm vụ chính là nghiên cứu các cơn dịch. Năm 1928 đã nghiên cứu sự nhiễm phế cầu khuẩn trong đại dịch cúm sau Thế chiến I, trong đó ghi nhận sự chuyển đổi kiểu hình từ R sang S của phế cầu có thể xảy ra ngay cả khi dùng tế bào chết.

Năm 1944, **Oswald Avery, Colin MacLeod, và Maclyn McCarty** đã công bố bằng chứng thực nghiệm đầu tiên rằng chính ADN là vật liệu di truyền. Avery và các cộng sự đọc bài báo của Griffith và họ cố lặp lại thí nghiệm với ý định tìm kiếm chất chịu trách nhiệm cho sự chuyển đổi nhưng không thành công. Sau đó nhờ một phương pháp tách loại protein ra khỏi dung dịch của **M. G. Sevag** họ đã thành công trong việc chứng minh ADN chính là tác nhân truyền tính trạng.

Năm 1951, **Erwin Chargaff**, đã chứng minh trong ADN tỷ lệ nucleotid Adenin luôn bằng Thymin; Cytosin bằng Guanin hay nói cách khác tỷ lệ base purin bằng tỷ lệ base pyrimidin; trong khi tỷ lệ $A + T \neq G + C$ và thay đổi theo loài.

Nhưng đến năm 1952, vai trò di truyền của ADN mới được xác nhận. Khi **Hershey và Chase** sử dụng kỹ thuật đánh dấu phóng xạ để chứng minh ADN chứ không phải protein là chất mang thông tin di truyền.

1.1.2. Giai đoạn Sinh học phân tử ra đời

Năm 1953, mô hình cấu trúc ADN của **Watson – Crick** ra đời, được xem là khám phá lớn nhất trong Sinh học của thế kỷ. Mô hình xoắn kép ADN do James D. Watson và Francis H. C. Crick xây dựng là sự kết hợp của công trình về tỷ lệ nucleotid do Edwin Chargaff đề xướng và công trình nghiên cứu sợi ADN bằng tán xạ tia X của **Rosalind Elsie Franklin và Maurice Wilkins**.

Năm 1961, **M. Nirenberg và J. Matthei** tìm ra bộ mã di truyền nhân tạo đầu tiên. Đến giữa những năm 1961, toàn bộ 64 codon (bộ ba mã hoá) đã được xác định.

Đặc biệt năm 1961, **F. Jacob và J. Monod** tìm ra cơ chế di truyền điều hoà tổng hợp protein.

Năm 1962, Warner Arber phát hiện ra enzym cắt giới hạn trong tế bào vi khuẩn.

Năm 1967, enzym ADN ligase được chiết xuất. Enzym này được xem là keo dán phân tử, có thể nối hai sợi ADN lại với nhau.

Năm 1970, **Hamilton Smith** chiết được enzym cắt giới hạn.

1.1.3. Sinh học phân tử hiện đại

Trong thập niên 70 của thế kỷ XX, kỹ thuật di truyền ra đời tạo nên cuộc cách mạng mới trong di truyền và sinh học phân tử. Việc xác định trình tự nucleotid

(ADN sequencing) của gen đã nhanh chóng dẫn đến kỹ thuật mới: gây đột biến điểm định hướng cho phép tạo các biến đổi tùy ý trên gen.

Năm 1973, A.C. Chang và Herbert Boyer, ... đã tạo được những plasmid tái tổ hợp đầu tiên. Phương pháp này được mở rộng vào năm 1973 và người ta đã nối nhiều đoạn ADN vào plasmid SC101 tách ra từ *E. coli*. Các phân tử tái tổ hợp này có thể tự sao chép khi đưa vào tế bào *E. coli* khác. Như vậy, Sinh học phân tử đã thúc đẩy sự ra đời của một ngành công nghệ mới là công nghệ di truyền.

Năm 1983, Randy Saiki và Kary Mullis..., đã công bố việc sử dụng PCR để khuếch đại các đoạn gen β -globin *in vitro*. Từ đây thao tác ADN (ADN manipulation), xác định trình tự gen đã phát triển rộng rãi và ứng dụng trong nhiều lĩnh vực khác nhau như chẩn đoán, biến đổi di truyền, xác định phả hệ,...

Có thể nói SHPT trong giai đoạn hiện đại được áp dụng trong ba lĩnh vực chủ yếu:

- Nghiên cứu cơ bản về cấu trúc và chức năng của từng gen.
- Sản xuất protein hữu ích bằng phương pháp mới.
- Tạo ra các sinh vật biến đổi gen (GMO).

Việc hoàn thành bản đồ gen người đã giúp giải quyết những vấn đề lớn như điều trị ung thư, phát triển phôi, biệt hoá tế bào.

Vào đầu năm 1990, sự kết hợp giữa Sinh học phân tử và Tin học đưa đến một phương pháp mới là *in silico* (nghiên cứu sinh học trên máy điện toán).

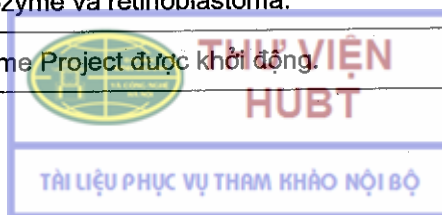
Bảng 1.1. Các cột mốc quan trọng trong sự hình thành và phát triển của SHPT

Thời gian	Sự kiện
1830	Phát hiện protein.
1833	Phát hiện và phân lập enzym đầu tiên.
1865	Di truyền học ra đời: Gregor Mendel khám phá quy luật di truyền các tính trạng.
1869	Friedrich Miescher khám phá acid nucleic
1879	Fleming phát hiện chất nhuộm sắc.
1900	Drosophila (ruồi giấm) được sử dụng trong các nghiên cứu gen.
1906	Thuật ngữ "Di truyền học" lần đầu tiên được sử dụng.
1911	Rous phát hiện virus gây ung thư đầu tiên.
1915	Phát hiện thực khuẩn (phage).



Thời gian	Sự kiện
1919	Lần đầu tiên từ "Công nghệ sinh học" (<i>biotechnology</i>) được dùng.
1938	Thuật ngữ "Sinh học phân tử" (<i>molecular biology</i>) được đưa ra bởi Warren Weaver.
1941	Thuật ngữ "Công nghệ di truyền" (<i>genetic engineering</i>) được dùng lần đầu, bởi nhà vi sinh vật học người Hà Lan A. Jost trong bài giảng tại học viện kỹ thuật Lwow, Ba Lan.
1944	Oswald Avery và cộng sự chứng minh ADN mang thông tin di truyền.
1946	Phát hiện rằng vật liệu di truyền từ các virus khác nhau có thể phối hợp để tạo ra virus mới, một thí dụ về tái tổ hợp di truyền.
1947	McClintock khám phá yếu tố di truyền vận động "gen nhảy" ở bắp.
1949	Pauling chứng minh bệnh hồng cầu lưỡi liềm là một "bệnh phân tử" do đột biến trong protein hemoglobin.
1953	James Watson và Francis Crick công bố cấu trúc xoắn kép ADN, mở ra thời đại mới của sinh học.
1955	Phân lập được enzym chịu trách nhiệm tổng hợp acid nucleic.
1956	Kornberg khám phá ADN polymerase I, dẫn đến việc làm sáng tỏ cơ chế sao chép ADN.
1958	Chứng minh bệnh hồng cầu lưỡi liềm chỉ do đột biến 1 acid amin.
	ADN được tách ra trong ống nghiệm lần đầu tiên.
1959	Các bước trong quá trình sinh tổng hợp protein được phác họa.
1960	Sử dụng tính chất bổ sung cặp base để tạo phân tử ARN – ADN lai.
	Tim thấy ARN thông tin.
1961 – 1966	Khám phá mã di truyền gồm các bộ ba (codon).
1967	Máy định trình tự protein tự động đầu tiên hoàn chỉnh.
1969	Tổng hợp enzym <i>in vitro</i> lần đầu tiên.
1970	Phát hiện enzym giới hạn, mở đường cho việc tạo dòng gen.
1971	Tổng hợp được gen hoàn chỉnh đầu tiên.
1973	Stanley Cohen và Herbert Boyer cắt và nối ADN thành công (dùng enzym giới hạn và ligase) và tạo ra ADN mới ở vi khuẩn.
1976	Các công cụ tái tổ hợp ADN lần đầu tiên được áp dụng cho người rối loạn di truyền.

Thời gian	Sự kiện
	Lai phân tử được áp dụng để chuẩn đoán tiền sinh bệnh alpha thalassemia.
	Gen của nấm men được biểu hiện ở <i>E. coli</i> .
	Trình tự của một gen được xác định.
1977	Lần đầu tiên biểu hiện gen người ở vi khuẩn.
	Phát triển kỹ thuật định trình tự nhanh các ADN dài nhờ điện di.
1978	Xác định các cấu trúc chi tiết của virus.
	Chứng minh khả năng đưa đột biến điểm vào vị trí đặc hiệu trên ADN.
Thập niên 70	Phát hiện các polymerase.
	Hoàn chỉnh kỹ thuật định trình tự acid nucleic.
	Định hướng gen.
	Cắt nối ARN.
1980	Trao bằng sáng chế kỹ thuật tạo dòng cho Cohen và Boyer.
	Phát triển máy tổng hợp gen đầu tiên.
	Giải Nobel Hoá học cho kỹ thuật tái tổ hợp phân tử: Berg, Gilbert, Sanger.
1982	Applied Biosystems, Inc., giới thiệu máy định trình tự protein pha khí đầu tiên, giảm thiểu lượng protein cần để định trình tự.
1983	Tổng hợp nhiễm sắc thể nhân tạo đầu tiên.
	Các định vị di truyền đầu tiên cho các bệnh di truyền được phát hiện.
1984	Phát triển kỹ thuật dấu ấn ADN.
	Toàn bộ bộ gen của HIV được tạo dòng và định trình tự.
1985	Tìm thấy các định vị di truyền của bệnh thận và xơ nang.
	PCR được phát minh. Và trở thành công cụ chính trong nghiên cứu và phát triển sản phẩm khắp thế giới.
	Kết quả dấu ấn di truyền đầu tiên được sử dụng để tranh tụng tại toà.
1988	Quốc hội Mỹ thông qua Human Genome Project.
1989	Bắt đầu Plant Genome Project.
Thập niên 80	Các nghiên cứu ADN được dùng để xác định lịch sử tiến hoá.
	Phát hiện ribozyme và retinoblastoma.
1990	Human Genome Project được khởi động.



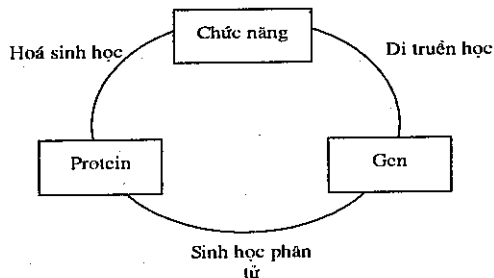
Thời gian	Sự kiện
1994	Gen ung thư vú đầu tiên được phát hiện.
1995	Trình tự bộ gen hoàn chỉnh của một sinh vật không phải virus được xác định: <i>Hemophilus influenzae</i> .
1996	Phát hiện gen liên quan đến bệnh Parkinson.
1998	Bộ gen của động vật đầu tiên, giun <i>C. elegans</i> , được định trình tự.
	Bản đồ sơ bộ của bộ gen người được công bố cho thấy có hơn 30000 gen.
Thập niên 90	Bản án đầu tiên được tuyên dựa vào bằng chứng dấu ấn di truyền ở Anh
2000	Bản đồ gen hoàn chỉnh của thực vật đầu tiên: <i>Arabidopsis thaliana</i> .
	Công bố bản đồ sơ bộ của bộ gen người.
2001	Hoàn chỉnh bản đồ gen cây lương thực đầu tiên: cây lúa.
	Hoàn chỉnh bản đồ gen của vi khuẩn nông nghiệp: <i>Sinorhizobium meliloti</i> , vi khuẩn cố định đạm và <i>Agrobacterium tumefaciens</i> , thuốc diệt sâu rầy thực vật.
2002	Bản đồ proteom sơ bộ đầu tiên của nấm men.
	Thành lập tổ hợp quốc tế định trình tự các ký sinh trùng gây sốt rét và muỗi sốt rét.
	Các nhà khoa học làm sáng tỏ các yếu tố kiểm soát sự biệt hoá tế bào gốc, xác định trên 200 gen liên quan.
2003	Công bố phiên bản thô của bộ gen người hoàn chỉnh.
2004	FDA lần đầu tiên công nhận hệ thống ADN microarray, Ampli Chip Cytochrome P450 Genotyping Test, để giúp chẩn đoán bệnh.
	Giải mã Bộ gen Gà.
20/10/2004	Công bố bản mô tả khoa học hoàn chỉnh của Bộ gen Người cho thấy có khoảng 20000 – 25000 gen mã hoá cho protein.

1.2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

1.2.1. Định nghĩa

Sinh học phân tử là một bộ phận của Sinh học, khoa học về sự sống, với đối tượng nghiên cứu là sự sống ở cấp độ phân tử, tập trung vào các khía cạnh về cấu trúc, sự sao chép và biểu hiện của gen; sự tương tác và chức năng sinh lý của các

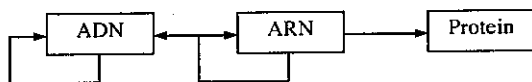
sản phẩm của gen. Các khoa học sinh học khác như Di truyền học hay Hoá sinh cũng nghiên cứu sinh học ở mức độ phân tử, tuy nhiên thiên về chức năng sinh học của gen hay sản phẩm của gen hơn là chính các phân tử này (Hình 1.1).



Hình 1.1. Quan hệ giữa sinh học phân tử với các khoa học sinh học khác

1.2.2. Học thuyết trung tâm

Học thuyết trung tâm (Central Dogma) của Sinh học phân tử được phát biểu lần đầu tiên bởi Francis Crick năm 1958 về sự luân chuyển thông tin của sinh vật: "Thông tin khi đã chuyển sang protein thì không thể lấy ra lại được". Nghĩa là thông tin không thể chuyển ngược từ protein đến acid nucleic. Thông tin ở đây là trình tự chính xác các nucleotid của acid nucleic quy định trình tự acid amin của protein.



Hình 1.2. Sự luân chuyển thông tin của sinh vật

Theo đó thông tin có thể được chuyển giữa các phân tử ADN thông qua sự sao chép hoặc giữa phân tử ADN và ARN thông qua sự phiên mã hoặc từ acid nucleic đến protein nhờ sự dịch mã. Ngoài ra, thông tin còn có thể di chuyển giữa ARN đến ARN trong quá trình sinh sản của các ARN virus hay quá trình xử lý cắt nối của ARN hoặc ngược lại từ ARN đến ADN trong quá trình phiên mã ngược của retrovirus. Các quá trình này nghiên cứu chi tiết trong Sinh học phân tử.

1.2.3. Các phương pháp nghiên cứu

Từ cuối những năm 50 đến đầu những năm 60 của thế kỷ XX, các nhà Sinh học phân tử đã có được các kỹ thuật để phân lập, xác định đặc tính và thao tác trên các phân tử sinh học như ADN, ARN và protein. Các kỹ thuật này bao gồm:

1.2.3.1. Tạo dòng biểu hiện

Tạo dòng biểu hiện được sử dụng để nghiên cứu chức năng protein. Đoạn ADN cần quan tâm được tạo dòng vào một plasmid và đưa vào tế bào để biểu hiện. Bằng cách thay đổi cơ cấu biểu hiện trên plasmid hoặc trình chế protein người ta có thể

hiểu được cách thức điều hoà hoạt động của gen, quan hệ của nó với hoạt động của các gen hay protein khác cũng như cấu trúc và chức năng của nó.

1.2.3.2. PCR

PCR là một kỹ thuật cho phép khuếch đại một đoạn gen đặc hiệu để thu được nhiều bản sao phục vụ nghiên cứu. Bên cạnh đó, PCR cũng cho phép dễ dàng đưa các thay đổi, đột biến vào ADN để nghiên cứu.

1.2.3.3. Điện di gel

Đây là một công cụ phân tích chính trong Sinh học phân tử. Nó dựa vào việc phân tách các phân tử sinh học như ADN, ARN, protein trong điện trường, trong đó các phân tử này sẽ di chuyển dưới tác động của điện trường xuyên qua một hệ gel, do đó chúng có thể được tách ra theo kích thước hay điện tích.

1.2.3.4. Lai vết Southern và Northern

Kỹ thuật lai vết Southern được sử dụng để phát hiện sự hiện diện của một trình tự ADN đặc hiệu trong mẫu bằng cách chuyển các phân tử ADN đã được phân tách bằng điện di gel sang một màng rắn và cho lai để phát hiện nhờ một đoạn dò đặc hiệu có đánh dấu. Kỹ thuật lai vết Northern với nguyên lý tương tự nhưng được áp dụng cho đối tượng ARN, được sử dụng để nghiên cứu sự biểu hiện của các ARN.

1.2.3.5. Lai vết Western và hoá miễn dịch

Protein cũng có thể được phân tách trên gel và lai vết trên màng rắn (lai vết Western) và phát hiện bằng kỹ thuật hoá miễn dịch. Nhờ đó có thể nghiên cứu sự hiện diện của protein trong tế bào và liên hệ với sự biểu hiện của gen tương ứng.

1.2.3.6. Kỹ thuật array

Kỹ thuật array cho phép đặt rất nhiều mẫu dò acid nucleic vào các chấm nhỏ trên chip. Các mẫu dò được thiết kế có trình tự bổ sung với trình tự acid nucleic (ADN hay ARN) đích. Array được dùng để nghiên cứu sự biểu hiện gen, phát hiện các biến động về trình tự acid nucleic trong tế bào cần nghiên cứu. Array cũng có thể được thiết kế sử dụng kháng thể để nghiên cứu sự biến động của protein (protein array).

1.3. NHỮNG ĐÓNG GÓP LỚN CỦA SINH HỌC PHÂN TỬ HIỆN NAY

1.3.1. Genomics: giải mã bộ gen và ngành hệ gen học

Việc giải mã gen, trước đây, theo phương pháp thủ công, mỗi tuần, mỗi người chỉ thực hiện được một vài phản ứng giải trình tự với năng suất 300 bp/phản ứng. Ngày nay với hệ thống máy mao mạch có thể xác định tự động đồng thời 96 phản ứng với độ dài trên 1000



bp/phân ứng. Nhờ đó đề án giải trình tự bộ gen người dài 3,2 tỷ nucleotid đã hoàn thành vào tháng 4/2003.

99% bộ gen người đã được giải trình tự với nội dung chính xác 99,99%. Thành tựu về giải mã bộ gen người và nhiều sinh vật khác như ruồi giấm, giun tròn, chuột và các vi sinh vật khác như *E. coli*, *Saccharomyces cerevisiae*, *C. briggsae*, *Drosophila pseudobsura* và gần đây của cây lúa nước, đã mở ra một ngành khoa học mới là “Bộ gen học - Genomics” chuyên nghiên cứu về cấu trúc và chức năng của toàn bộ bộ gen sinh vật.

Genomics (bộ gen học) là khoa học nghiên cứu một cách hệ thống các trình tự ADN đầy đủ (genome) của sinh vật. Genomics thiết lập thông tin về sự sống bằng các tiếp cận một cách hệ thống và có thể tiến hành ở quy mô công nghiệp. Genomics khác với genetics (di truyền học). Genetics nhắm vào các gen đơn lẻ, tại một thời điểm, như là chụp ảnh nhanh. Genomics cố gắng nhắm vào tất cả các gen như một hệ động học, qua thời gian, để xác định chúng tương tác nhau và ảnh hưởng như thế nào đến các con đường sinh học, các mạng lưới và sinh lý trong ngữ cảnh tổng quát hơn. Như vậy, genomics chính là một nhánh khoa học giải mã thông tin được lưu trong ADN và khám phá các thông tin như số lượng gen, tổ chức và nội dung của bộ gen. Thông tin này được ứng dụng nhiều trong y tế, nông nghiệp, sinh học tiến hóa và các lĩnh vực khoa học khác.

Có nhiều loại genomics tùy theo khía cạnh và lĩnh vực mà nó nghiên cứu áp dụng, ví dụ như plant genomics, pharmacogenomics,...

Khả năng giải trình tự đầy đủ bộ gen sinh vật đã đẩy mạnh việc xác định chức năng các gen ở mức độ nghĩa rộng của bộ gen. Vì các gen có chức năng liên quan được điều hoà cùng nhau, các kỹ thuật để đánh giá sự biểu hiện toàn bộ gen sẽ giúp nhận định cơ chế khởi đầu và cụm các trình tự gen mới có chức năng liên quan. Xếp các trình tự gen thành các nhóm chức năng theo sự biểu hiện của gen sẽ cho hướng dẫn cơ bản các thực nghiệm bổ sung, nhằm xác định đặc điểm chức năng chính xác của sản phẩm cuối cùng của gen. Trong hai thập kỷ vừa qua, các kỹ thuật để đánh giá sự biểu hiện gen đã có những tiến bộ, từ các phương pháp trên các gen đơn lẻ đặc hiệu (ví dụ lai vết Northern, slot và dot; phiên mã ngược và PCR bán định lượng và định lượng; phương pháp bảo vệ nuclease) đến các kỹ thuật tập trung vào nhận định tất cả các gen có biểu hiện khác nhau giữa hoặc trong các mẫu thử nghiệm.

Các phương pháp để nhận định các khác biệt của biểu hiện gen gồm có lai loại trừ (subtractive hybridization), trình diện khác biệt (differential display), phân tích hàng loạt biểu hiện gen (SAGE, serial analysis of gene expression) và lai vi bản trái (microarray hybridization).

1.3.2. Proteomics: phân tích biến động protein và ngành hệ protein học

Proteomics là khoa học nghiên cứu ở qui mô lớn về protein nhằm thiết lập các nhận dạng, số lượng, cấu trúc và chức năng sinh hóa và tế bào của tất cả protein trong cơ thể, cơ quan hoặc tiểu cơ quan cũng như sự thay đổi các đặc tính này theo không gian, thời gian và trạng thái sinh lý. Proteome (hệ protein) là toàn bộ protein được biểu hiện bởi một loại tế bào hay sinh vật nào đó tại một thời điểm trong một điều kiện nhất định.

Proteomics thường được xem là bước tiếp theo sau genomics trong nghiên cứu các hệ thống sinh học. Nó phức tạp hơn genomics nhiều, chủ yếu vì bộ gen của sinh vật ổn định hơn, trong khi proteome khác nhau giữa các tế bào và thay đổi liên tục do các tương tác của nó với bộ gen và môi trường. Bên cạnh đó, một sinh vật có sự biểu hiện protein khác nhau hoàn toàn ở các phần khác nhau trong cơ thể. Một khó khăn chính nữa là tính phức tạp về cấu trúc của protein.

Kiến thức về proteomics sẽ giúp hiểu sinh vật tốt hơn nhiều so với genomics. Các phương pháp như phosphoproteomics và glycoproteomics được dùng để nghiên cứu các biến đổi sau dịch mã (post – translational modifications). Các nghiên cứu proteomics bao gồm nghiên cứu tương tác của các protein và nhận diện protein.

1.3.2.1. Tương tác của các protein

Hầu hết các protein hoạt động hợp tác với các protein khác và một trong những đích của proteomics là nhận diện tương tác của các protein. Đây thường là đầu mối quan trọng về chức năng của các protein mới tìm ra. Đã có một vài phương pháp để dò các tương tác protein – protein. Phương pháp kinh điển là phân tích lai đôi trên nấm men (yeast two – hybrid analysis). Các phương pháp mới gồm có vi bản trải protein (protein microarrays), sắc kế ái lực miễn dịch (immunoaffinity chromatography), tiếp theo là khối phổ (mass spectrometry), và kết hợp các phương pháp thực nghiệm như trình diện phage và các phương pháp máy tính.

1.3.2.2. Nhận diện protein

Để nghiên cứu proteomics, trước hết protein phải được phân giải, đôi khi ở qui mô lớn. Có thể tách protein bằng phương pháp điện di hai chiều trên gel (two-dimensional gel electrophoresis). Trong kỹ thuật này, đầu tiên protein được tách theo điểm đẳng điện và sau đó theo trọng lượng phân tử. Các vết protein trên gel có thể được quan sát nhờ các chất nhuộm hóa học khác nhau hoặc các dấu hiệu huỳnh quang. Thường thì các protein có thể được định lượng dựa vào cường độ nhuộm màu của chúng. Các protein

được nhận diện sau khi đã được tách riêng và định lượng. Các vết riêng rẽ được cắt ra khỏi gel và cắt thành các peptid bằng các enzym thủy phân protein. Sau đó các peptid này được nhận diện bằng khối phổ, đặc biệt là khối phổ MALDI-TOF (Matrix-Assisted Laser Desorption-Ionization Time-Of-Flight).

Các hỗn hợp protein cũng có thể được phân tích mà không cần tách trước. Quy trình này bắt đầu bằng phân giải các protein trong hỗn hợp. Các peptid tạo thành thường được bơm vào cột sắc ký lỏng hiệu năng cao (HPLC) để tách các peptid dựa vào tính không ưa nước. HPLC có thể được ghép trực tiếp với MALDI-TOF. Các peptid tách ra từ cột có thể được nhận diện bằng khối phổ song song (MS/MS). Trong một số trường hợp, khối phổ có thể được dùng để xác định trình tự acid amin của protein.

1.3.2.3. Các kỹ thuật được sử dụng trong proteomics

– Điện di một chiều và hai chiều trên gel (One – and two – dimensional gel electrophoresis).

– Tán xạ tia X và cộng hưởng từ hạt nhân (X – ray crystallography and nuclear magnetic resonance).

– Khối phổ song song kết hợp với sắc ký đảo pha hoặc điện di hai chiều.

– Khối phổ (không song song), thường là MALDI-TOF.

– Sắc ký ái lực, kỹ thuật lai đôi trên nấm men, FRET (Fluorescence Resonance Energy Transfer), và SPR (Surface Plasmon Resonance) được dùng để nhận định tương tác protein-protein và protein-DNA.

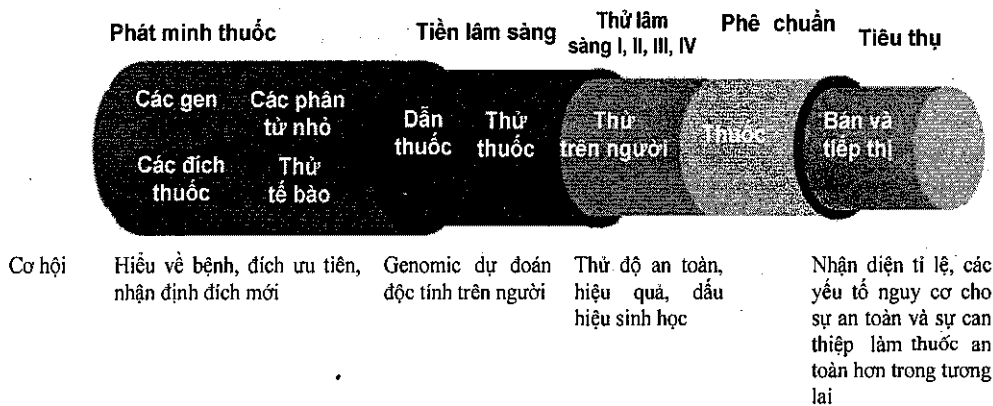
– Công cụ tin sinh học.

Với các kỹ thuật sắc ký, điện di, trước đây người ta chỉ nghiên cứu được từng loại protein riêng rẽ. Ngày nay, khi phối hợp sắc ký đa chiều và khối phổ người ta có thể phân tích cùng lúc 5000 loại protein và kết quả cho phép chẩn đoán sớm những bệnh hiểm nghèo như ung thư máu. Proteomics nhằm tiến tới giải mã chức năng sinh học của hệ gen, song trước mắt chỉ mới nghiên cứu được những biến đổi hoạt động của các nhóm gen trong điều kiện bệnh lý, cung cấp thông tin cho việc chẩn đoán sớm, phòng trừ và điều trị nhiều loại bệnh khác nhau.

1.3.3. Genomics, proteomics và sự phát triển thuốc

Phát minh và phát triển thuốc là một quá trình dài, để sản phẩm ra thị trường cần khoảng 15 năm từ khi có ý tưởng cho thuốc có đích tiềm năng. Các giai đoạn cơ bản quá trình được minh họa trong Hình 1.3.





Hình 1.3. Quá trình phát minh và phát triển thuốc và các cơ hội tham gia của genomics, proteomics

1.3.3.1. Pharmacogenomics và chiến lược phát triển thuốc

Pharmacogenomics (dược lý bộ gen) nghiên cứu sự ảnh hưởng của di truyền lên đáp ứng của cơ thể với thuốc. Nó làm thay đổi cơ bản cách sử dụng thuốc và quá trình phát triển thuốc. Nguyên nhân của sự thay đổi là những thành tựu trong lĩnh vực bộ gen người (human genome, bản đồ đa hình đơn nucleotid (single-nucleotide polymorphism, SNP), định tít gen (genotyping) và tin sinh học (bioinformatics).

Các lợi ích mà pharmacogenomics đem lại gồm có:

- Cho các thuốc công hiệu hơn. Các thuốc mới dựa trên các phân tử protein, enzym, và RNA liên quan đến các gen và bệnh được chế tạo. Nghiên cứu các đích tác động phân tử cho phép đưa ra liệu pháp có tính chuyên biệt. Phân tử thuốc có thể được tối ưu hóa cho các đích chuyên biệt này do đó có tác động mạnh và chuyên biệt hơn.

- Cho các thuốc tốt hơn, an toàn hơn ngay từ đầu. Thay vì phương pháp thử-và-sai (mò mẫm) ở các bệnh nhân phù hợp với các thuốc đúng, từ ban đầu các bác sĩ có thể phân tích lược đồ di truyền của bệnh nhân và kê toa thuốc tốt nhất đã có. Đây không chỉ là công việc tạm thời để tìm thuốc đúng, mà thời gian bình phục nhanh và tăng độ an toàn vì các phản ứng phụ có thể xảy ra sẽ được hạn chế.

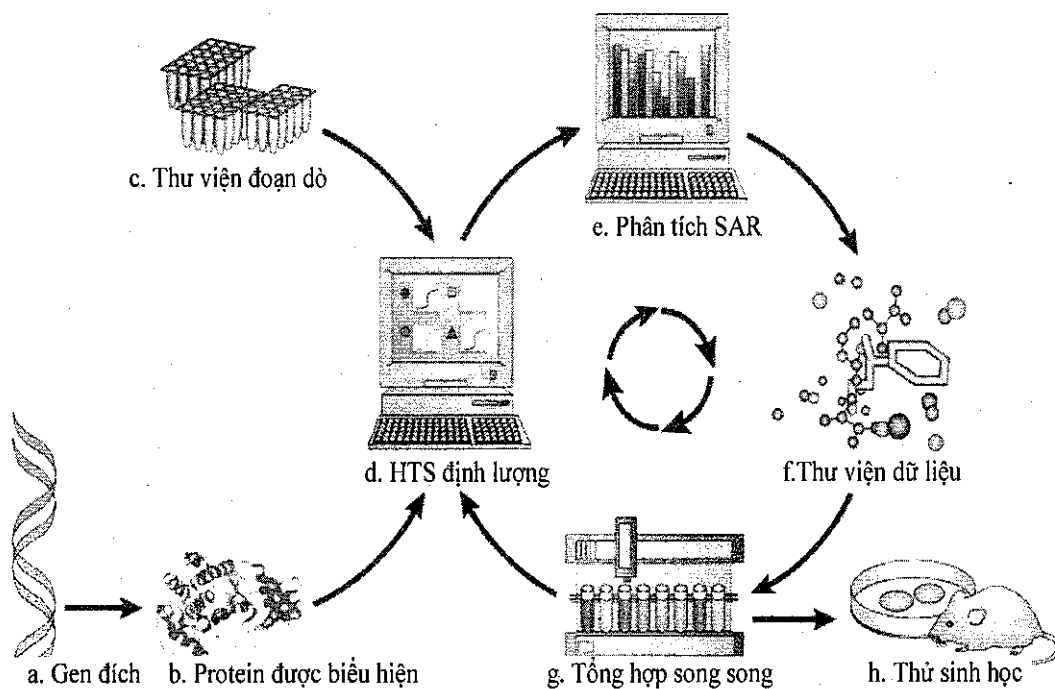
- Xác định liều thuốc thích hợp bằng phương pháp chính xác hơn. Các phương pháp hiện nay để định liều cơ sở theo trọng lượng và tuổi sẽ được thay thế bằng định liều cơ sở theo di truyền của con người - cơ thể xử lý thuốc và thời gian chuyển hóa nó như thế nào. Điều này sẽ tối đa hóa giá trị của liệu pháp và còn giảm khả năng quá liều xảy ra.

- Tiến xa trong sàng lọc bệnh. Kiến thức mã di truyền người sẽ cho phép con người tạo lối sống thích hợp và các thay đổi môi trường sớm để tránh hoặc giảm mức

độ trầm trọng của bệnh di truyền. Tương tự như vậy, tiến bộ trong hiểu biết tính nhạy cảm của bệnh đặc biệt sẽ cho phép đưa ra biện pháp theo dõi và điều trị cẩn thận ở những giai đoạn thích hợp nhất để tối đa hóa liệu pháp cho chúng.

– Cải tiến trong phát minh, phát triển thuốc. Việc thử thuốc được tiến hành trên nhóm dân số có di truyền chuyên biệt - cho mức độ thành công lớn hơn. Giá thành và rủi ro của các thử nghiệm lâm sàng giảm do chỉ nhắm vào những người có thể đáp ứng với thuốc.

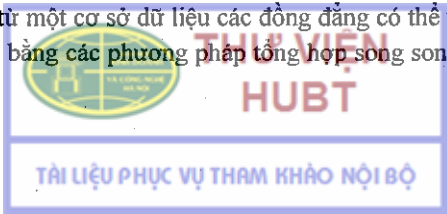
– Giảm chi phí trong chăm sóc sức khỏe. Giảm số lượng các phản ứng phụ của thuốc, số lượng thử thuốc thất bại, thời gian để thuốc được phê duyệt, thời gian bệnh nhân cần dùng thuốc, số lượng thuốc mà bệnh nhân phải dùng để điều trị hiệu quả, ảnh hưởng của bệnh lên cơ thể (do phát hiện sớm) giúp giảm chi phí chăm sóc sức khỏe.



HTS: High Throughput Screening - Sàng lọc hiệu năng cao
 SAR: Structural Activity Relationship - Liên quan cấu trúc tác dụng

Hình 1.4. Một ví dụ thực tiễn và kinh tế của chiến lược chemogenomics

Trình tự gen của các đích đã được xác định bằng các tiếp cận genomics được tạo dòng và biểu hiện thành các protein đích (a,b) và các protein này thích hợp cho việc sàng lọc với một thư viện mẫu dò gồm các hợp chất giống thuốc (c). Các hợp chất này được sàng lọc để tìm các phân tử có hoạt tính bằng cách sử dụng một thử nghiệm liên kết có tính phổ quát và định lượng (d). Các phân tử có hoạt tính ban đầu hay các dữ liệu định lượng về hoạt tính-cấu trúc phát sinh từ thử nghiệm liên kết được phân tích (e) và sử dụng để thiết lập chiến lược chọn lọc cho sự tổng hợp mới các hợp chất với các đặc tính được cải tiến. Các hợp chất này được chọn từ một cơ sở dữ liệu các đồng đẳng có thể tổng hợp được của thư viện mẫu dò ban đầu (f) được tổng hợp bằng các phương pháp tổng hợp song song (g) và được thử nghiệm (d) để



thiết lập profile hoạt tính-cấu trúc của đích đang nghiên cứu và để tinh chỉnh các tiêu chí sàng lọc ở các vòng tiếp theo. Trong mỗi chu kỳ, ưu thế được gán cho các ứng viên tổng hợp bằng cách sử dụng các quá trình tối ưu hóa đa biến được thiết kế nhằm đảm bảo các hợp chất không chỉ được tối ưu hóa về ái lực gắn với đích mà còn có các đặc tính giống thuốc để cho phép chúng được sử dụng trực tiếp như là các hợp chất mẫu trong các mô hình sinh học hay tế bào thích hợp (h).

1.3.3.2. Chiến lược chemogenomics để phát minh thuốc

Các thành tựu trong genomics và proteomic để nhận diện các đích mới cho sự can thiệp của thuốc là các cơ hội chưa từng có cho phát minh các tác nhân mới có các kiểu liệu pháp tác động mới.

Sự tiếp cận của "chemogenomics" để đánh giá đích dùng thông tin cơ bản cung cấp bởi trình tự đích để làm protein và rồi sau đó phát minh "hợp chất công cụ" phân tử nhỏ tương tác với đích. Hợp chất công cụ có thể được đánh giá trong mô hình bệnh để thử trực tiếp giả thuyết trị liệu. Sự tiếp cận này có thể được thực hiện như một quá trình song song và đặc biệt thích hợp để phát minh thuốc trong những họ lớn. Một ví dụ của chiến lược chemogenomics để phân lập các đích phân tử được phác thảo trong Hình 1.4.

1.3.4. Sản xuất và sử dụng chip ADN

Chip ADN có hình dạng giống như một con chip với những chấm ADN thay cho transistor. Là một mảnh màng liên kết cao có kích thước 20 × 40 mm được in trên đó các đoạn ADN. Ví dụ chip bộ gen người được in 20000 – 50000 gen ở những điểm chấm vuông cực nhỏ.

Khi lai chip ADN này với sản phẩm phiên mã của bộ gen cơ thể, các chấm ADN đổi màu tương ứng với mức độ hoạt động của những gen tương ứng trong cơ thể ở trạng thái và thời điểm nghiên cứu. Từ đó kết luận được tình trạng bệnh lý của bệnh nhân.

Chip ADN đang dần dần trở thành công cụ chẩn đoán trong công nghiệp lên men vi sinh, trong y học dự phòng, trong kiểm dịch và an toàn thực phẩm và trong kiểm soát môi trường.

1.3.5. Chuyển gen vào cây trồng

Người ta thực hiện chuyển vào cây trồng các loại gen tăng cường khả năng kháng sâu bệnh như gen kháng sâu nhóm oxy/VIP, gen kháng virus nhóm CP/Nbi, gen kháng thuốc diệt cỏ nhóm bar. Những công trình chuyển gen lạ vào thuốc lá cho thấy thực vật có biểu hiện gen lạ giống như *E.coli*. Vào những năm 1996, các nhà công nghiệp (Monsanto) đã đưa ra thị trường những sản phẩm cây chuyển gen đầu tiên như bắp (ngô), đậu nành và bông vải. Năm 1996 trên thế giới mới có



0,5 ha cây trồng chuyển gen được đưa vào sản xuất. Đến nay (2003) diện tích trồng cây chuyển gen trên quy mô toàn cầu đã lên tới 67 triệu ha. Ngày nay, người ta đang tìm cách đưa gen sản xuất vaccin, gen sản xuất dược chất (các pepetid nhỏ), các chất dinh dưỡng (vitamin A),... để cây trồng từ lương thực thực phẩm trở thành cây sản xuất dược liệu có giá trị kinh tế cao.

Trong số 50 loài cây trồng mang gen lạ đang được thử nghiệm thì các cây bông vải kháng sâu, cây đậu tương, cây ngô kháng sâu, kháng thuốc diệt cỏ chiếm tổng số trên 90% diện tích.

1.3.6. Tin sinh học

Công nghệ tin sinh học bao gồm các phương pháp khai thác nhanh ngân hàng dữ liệu, phân tích trình tự và cấu trúc ADN và protein. Tin sinh học đang cải tiến phương pháp xử lý phân tích số liệu, cải thiện khả năng dự đoán vùng hoạt động, vùng ngưng nghỉ của bộ gen, cải tiến khả năng phỏng đoán phản ứng tế bào đối với tác nhân ngoại sinh, thiết lập nên các cấu trúc phân tử có hoạt lực cao và định hướng phân hoá tế bào một cách hiệu quả. Trên thế giới hiện có ba ngân hàng dữ liệu gen lớn nhất là GenBank (NCBI Entrez Nucleotide – Mỹ), EMBL (European Molecular Biology Laboratory – Châu Âu) và DDBJ (DNA Data Bank of Japan – Nhật Bản) lưu trữ trên 9 tỷ dữ liệu về gen.

Một số thành tựu nổi bật của tin sinh học là: chương trình NMR đa chiều thiết lập cấu trúc không gian protein. Chương trình FASTA so sánh trình tự gen và protein ra đời trước 1990 cho phép so sánh tự động, miễn phí trình tự một đoạn gen dài khoảng 1000 bp với trình tự đã công bố trong vài phút. Chương trình BLAST, trung tâm NCGR, chip ADN thiết lập trước năm 2000 và gần đây là đề án IBM Blue Gene được bắt đầu, hệ chương trình trọn gói EMBOSS được bán, chip ADN bộ gen người được đưa ra thị trường. Đến thời điểm này có trên 60 công ty lớn chuyên dịch vụ và kinh doanh trên lĩnh vực sinh tin học đang hoạt động. DNASTar và GCG là hai trong những công ty thành công nhất trong lĩnh vực cung ứng phần mềm phân tích gen.

1.3.7. Công nghệ nano sinh học

Đây là lĩnh vực đa ngành, tập trung khai thác vật liệu, thiết bị hoặc phương pháp của các chất ở phạm vi kích thước tới hạn nằm giữa chiều dài phân tử và bước sóng ánh sáng khả kiến từ 0,1 đến 500 nm. Công nghệ nano sinh học là:

– Phương hướng mới cho phép thu nhận những thông tin về hệ thống sinh học ở mức chấm lượng tử, tạo đầu dò nano với kích thước phân tử dùng trong chẩn đoán bệnh.



- Phương pháp *in situ* mới cung cấp thông tin tốt hơn về chức năng tế bào.
- Công nghệ thao tác cải biến 2 chiều và 3 chiều đối với mô và tế bào.
- Vận chuyển và phân phối thuốc hoặc gen vào mô và tế bào thông qua việc khống chế kích thước hạt, hoạt hoá và giải phóng hoạt chất thuốc thông qua cơ chế và thiết bị bơm kích thước nano, van tế bào và cơ quan nhân tạo.

Trong Y học, người ta hy vọng phẫu thuật gen, phẫu thuật tế bào, trị liệu tế bào, trị liệu gen, tổng hợp gen chẩn đoán đại tế bào, các hệ thống cơ khí điện tử nano y học – đó là các "công cụ nano" thông minh, robot mô kích thước nano,... sẽ được đưa vào thực tiễn trong vài chục năm tới.

CÂU HỎI

1. Ai là người xác nhận vai trò di truyền của ADN?
 - a) Frederick Griffith
 - b) Oswald Avery
 - c) Hershey và Chase
 - d) Erwin Chargaff
 - e) Watson và Crick
2. Ai là người đưa ra mô hình xoắn kép của ADN?
 - a) Frederick Griffith
 - b) Oswald Avery
 - c) Hershey và Chase
 - d) Erwin Chargaff
 - e) Watson và Crick
3. Bản đồ gen người khi hoàn chỉnh cho thấy có bao nhiêu gen mã hoá cho protein?
 - a) 10 000 – 15 000
 - b) 15 000 – 20 000
 - c) 20 000 – 25 000
 - d) 25 000 – 30 000
 - e) 30 000 – 35 000
4. Sinh học phân tử là khoa học sinh học nghiên cứu về
 - a) Hoá học của các phân tử sinh học.
 - b) Ảnh hưởng của các đột biến di truyền.
 - c) Chức năng của protein.
 - d) Chức năng của gen.
 - e) Quan hệ giữa gen và sản phẩm của nó.
5. Nội dung chính của học thuyết trung tâm của Sinh học phân tử:
 - a) Thông tin khi đã chuyển sang protein thì không thể lấy ra lại được.
 - b) Thông tin được lưu trữ trên ARN không thể chuyển sang ADN.
 - c) Thông tin chỉ luân chuyển giữa các dạng acid nucleic khác nhau.
 - d) Thông tin chỉ được lưu trữ trên ADN.
 - e) Protein không mang thông tin di truyền.

Bài 2

SAO CHÉP ADN

MỤC TIÊU

- Trình bày được quá trình sao chép ADN.
- Tóm tắt được cách thức sao chép của acid nucleic của virus.
- Mô tả được quá trình sửa chữa ADN.

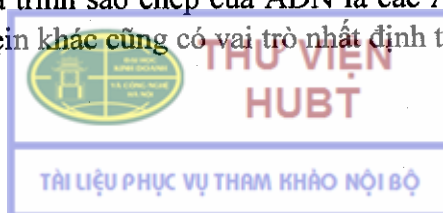
2.1. KHÁI NIỆM

Một tế bào vi khuẩn điển hình có thể chứa khoảng 2000 loại protein, các tế bào nhân thật chứa khoảng 50000. Thông tin về các loại protein này được mã hoá trong các phân tử acid nucleic. Ở đa số loài, vật liệu di truyền chủ yếu là ADN, ở một số virus là ARN.

Các phân tử ADN có thể khác nhau về thành phần, trật tự sắp xếp các nucleotid, nhưng đều có cấu trúc cơ bản giống nhau. Cấu trúc xoắn kép gồm hai chuỗi đường – phosphat ở mặt ngoài của phân tử, còn các base nitơ ở bên trong được nối với nhau bằng các liên kết hydro (Hình 2.1).

Mặc dù các base A, T, G, và C có kích thước khác nhau nhưng phân tử ADN vẫn có thể tự điều chỉnh để có được cấu trúc xoắn kép đối xứng qua trục. Trật tự của các base nitơ tạo nên thông tin di truyền. Sự khác nhau trong trật tự này tạo nên sự đa dạng của sinh vật. Do đó, nếu xét nghiệm dấu ấn hồng cầu thì có thể biết được 70% bí mật di truyền, nghĩa là hàng triệu người mới có hai người giống nhau, nhưng nếu xét nghiệm ADN thì biết được 99% bí mật di truyền, nghĩa là trên 70 tỉ người mới có hai người giống nhau.

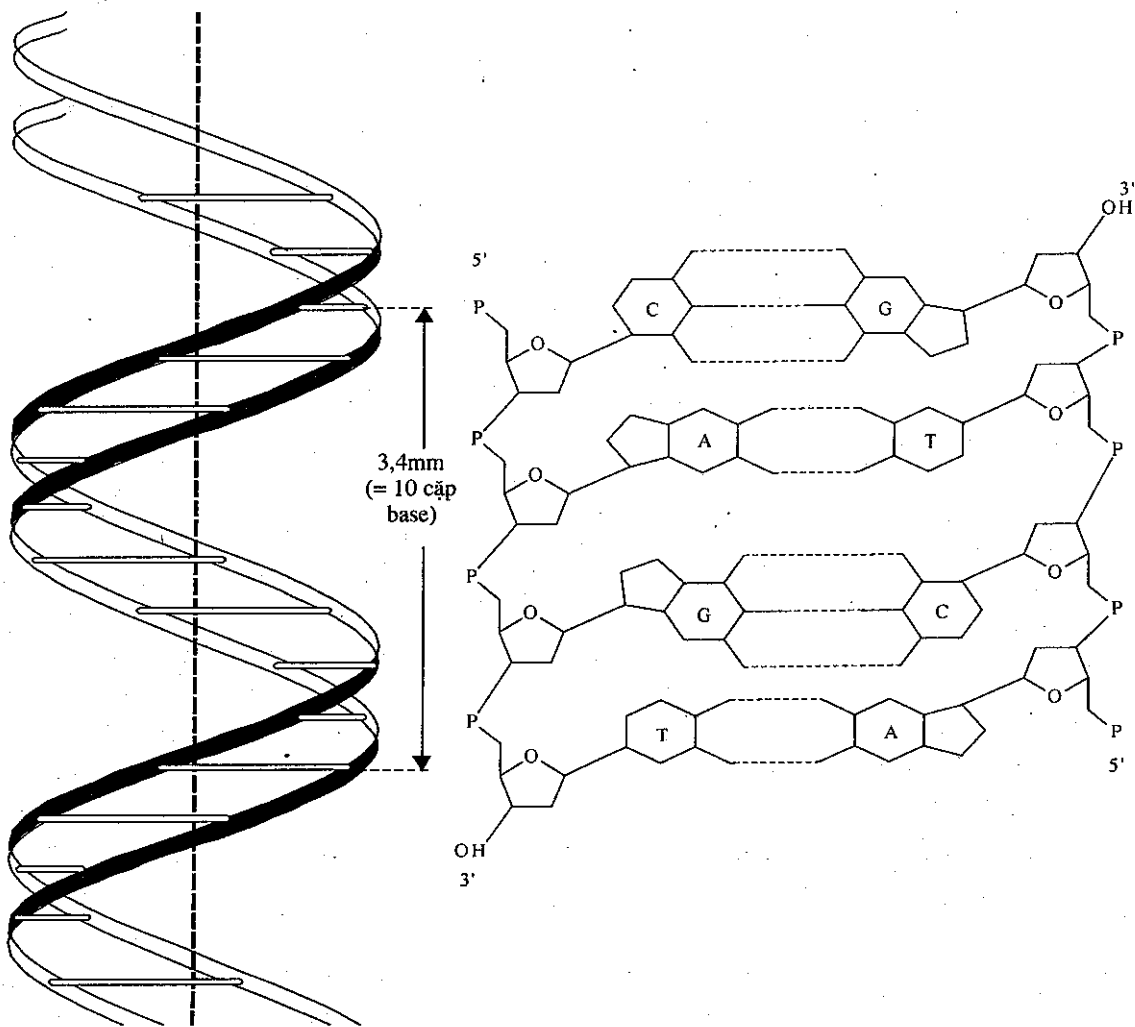
Thông tin di truyền của mỗi sinh vật được lưu trữ giống nhau ở mức tế bào. Khi các tế bào phân chia thông tin di truyền được chuyển nguyên vẹn sang tế bào mới nhờ quá trình sao chép. Enzym xúc tác quá trình sao chép của ADN là các ADN-polymerase phụ thuộc ADN. Một số enzym và protein khác cũng có vai trò nhất định trong tiến trình này.



Các yếu tố vật lý, hóa học... có thể tác động lên ADN, làm phá hủy ADN hay gây đột biến, sự hư hại này có thể được sửa chữa để giữ bộ gen sinh vật được ổn định. Tuy nhiên, trong một số trường hợp các yếu tố này gây ra các biến đổi không sửa chữa được, dẫn đến biến đổi bộ gen (genome). Tất cả những thay đổi này có thể di truyền được qua con đường sao chép ADN.

2.2. SỰ SAO CHÉP CỦA ADN

Người ta đưa ra hai cơ chế cơ bản của sự sao chép ADN: một cơ chế bảo tồn (conservative mechanism) trong đó phân tử ADN con gồm hai chuỗi hoàn toàn mới hoặc cơ chế bán bảo tồn (semiconservative mechanism), trong đó phân tử ADN con gồm một chuỗi mẹ kết hợp với một chuỗi mới được tổng hợp. Cơ chế sau được chứng minh bằng thí nghiệm Meselson và Stahl (1958).

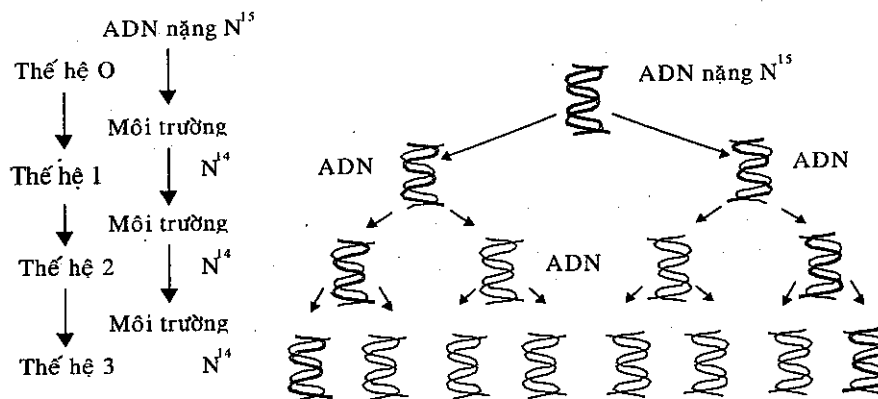


Hình 2.1. Cấu trúc phân tử xoắn kép của ADN

2.2.1. Thí nghiệm của Meselson và Stahl

Nếu tế bào *Escherichia coli* được cung cấp một nguồn nitơ duy nhất là amoni, chúng có thể tạo nên tất cả các nucleotid thuộc nhóm purin và pyrimidin cần cho sự tổng hợp ADN. Nếu nitơ được sử dụng là đồng vị nitơ nặng (N^{15}) ADN được tạo ra sẽ nặng hơn ADN được tổng hợp bằng cách dùng nitơ nhẹ (N^{14}). Hai loại ADN nặng (N^{15}) hay nhẹ (N^{14}) có thể được tách ra bằng siêu ly tâm trong thang nồng độ cesium clorid.

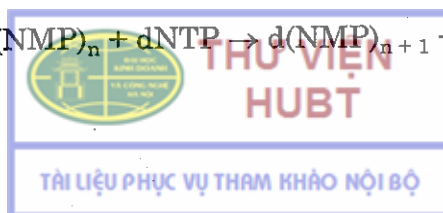
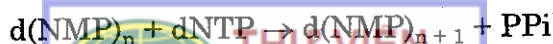
Trong thí nghiệm này, vi khuẩn được cho phát triển vài thế hệ trong môi trường N^{15} để đảm bảo cho tất cả ADN của vi khuẩn đều là loại nặng. Bây giờ mới chuyển tế bào vào môi trường có chứa N^{14} , như là nguồn cung cấp nitơ duy nhất và để cho phân chia trong môi trường N^{14} , khi khối lượng ADN tăng lên gấp đôi, ADN được chiết ra khỏi tế bào và được ly tâm trên thang nồng độ cesium clorid. Nếu cơ chế bảo tồn được thực hiện, hai băng ADN phải được thấy rõ sau khi ly tâm: một băng nguyên thủy (N^{15}) và một băng chứa N^{14} (Hình 2.2). Nếu theo cơ chế bán bảo tồn thì chỉ có một băng có tỷ trọng trung gian chứa $N^{14.5}$ xuất hiện. Nếu ADN được đun nóng, rồi làm nguội nhanh để tách hai mạch, ly tâm trên thang nồng độ cesium clorid sẽ thấy hai băng tương ứng với ADN N^{15} và ADN N^{14} .



Hình 2.2. Thí nghiệm Meselson và Stahl

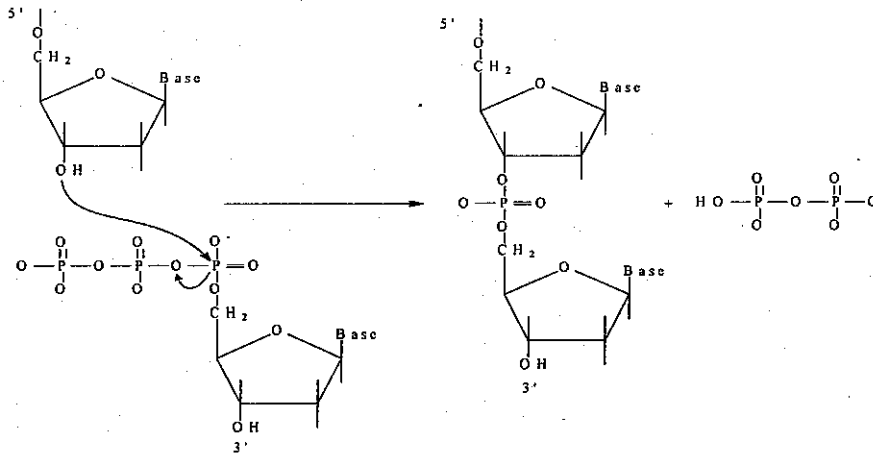
2.2.2. Các yếu tố cần thiết cho sự sao chép ADN

Tất cả mọi sinh vật đều đòi hỏi những yếu tố sau đây cho quá trình sao chép: khuôn mẫu (phân tử ADN ban đầu), 4 loại desoxyribonucleotid triphosphat (dNTP), enzym ADN polymerase và $in\ Mg^{2+}$, (đồng yếu tố của AND polymerase). Tiến trình này có thể được tóm tắt bằng phương trình:



Trong đó, dNTP là một desoxyribonucleotid triphosphat và d(NMP)_n là một polymer có n desoxyribonucleotid.

Việc thêm một desoxyribonucleotid mới vào chuỗi ADN được trình bày ở Hình 2.3:



Hình 2.3. Sự hình thành liên kết phosphodiester giữa desoxy – ribonucleotid và 3' – OH của chuỗi ADN đang sao chép

Pyrophosphat tạo ra sẽ được thủy phân thành phosphat vô cơ, làm phản ứng xảy ra theo chiều phải, kéo dài chuỗi ADN. Cơ chế tương tự cũng xảy ra trong việc hoạt hoá acid amin, sinh tổng hợp glycogen và hoạt hoá acid béo để tăng hiệu suất của sản phẩm.

2.2.3. Các ADN polymerase

Enzym tổng hợp mạch ADN mới trong quá trình sao chép là ADN polymerase, đây là nhóm enzym polymerase phụ thuộc ADN, tức sử dụng ADN làm khuôn mẫu. Hoạt tính polymerase thêm các nucleotid vào mạch ADN theo hướng 5'-3' theo nguyên tắc bổ sung với mạch khuôn. ADN polymerase không tự tổng hợp mạch *de novo* mà nó cần nhóm 3'-OH tự do trên mạch cần tổng hợp để bắt đầu.

Ở *E. coli*, người ta tìm thấy 5 loại ADN polymerase gồm: I, II, III, IV và V. Đặc tính của các ADN polymerase khác nhau được tóm tắt trong Bảng 2.1.

Bảng 2.1. Đặc điểm của ADN polymerase từ *E. coli* và retrovirus

Polymerase	<i>E. coli</i>			Virus
	I	II	III	
Khối lượng phân tử (Mr)	109 000	120 000	180 000	160 000
Cấu trúc, đơn vị nhỏ (Mr)	1	1	α 140 000 ε 25 000 θ 10 000	α 65 000 β 95 000

Polymerase	<i>E. coli</i>			Virus
	I	II	III	
Hoạt tính polymer hoá 5' → 3'	+	+	+	+
Hoạt tính exonuclease 5' → 3'	+	-	+	-
3' → 5'	+	+	+	-

Tất cả ADN polymerase của tế bào nhân nguyên thủy đều có hoạt tính exonuclease, vì chúng thủy phân mạch đơn ADN của chuỗi từ đầu 3' theo hướng 3' → 5' (hoạt tính exonuclease 3' → 5'); hoặc chúng có thể thủy phân chuỗi ADN từ đầu 5' (hoạt tính exonuclease 5' → 3').

ADN polymerase I của *E. coli* là một chuỗi polypeptid lớn có trọng lượng phân tử 109.000. Trong mỗi phân tử có chứa một nguyên tử kẽm. Enzym polymerase chỉ có thể xúc tác sự polymer hóa theo hướng 5' → 3' do thêm một nucleotid từ deoxyribonucleotid-5'-triphosphat vào nhóm 3'-hydroxy của chuỗi ADN (Hình 2.3).

ADN polymerase I chịu trách nhiệm sửa chữa các ADN hư hỏng và có vai trò phụ trong sao chép.

ADN polymerase II của *E. coli* rất giống ADN polymerase I nhưng không có hoạt tính exonuclease 5' → 3'. ADN polymerase II cần cho sự tái khởi động chạc ba sao chép khi nó bị khóa bởi tổn thương trong ADN.

ADN polymerase III, còn gọi là replicase, là enzym sao chép chính ở tế bào nhân nguyên thủy, nó là phần xúc tác của ADN polymerase III holoenzym. Đây là một phức hợp có phân tử lượng khoảng 900 kDa chứa 10 loại protein được tổ chức thành 4 loại phức hợp con như sau:

- Hai bản sao của lõi xúc tác được cấu tạo từ các tiểu đơn vị α (mang hoạt tính ADN polymerase), ϵ (mang hoạt tính sửa lỗi 3'-5' exonuclease) và θ (có chức năng kích thích hoạt tính exonuclease).

- Hai bản sao của tiểu đơn vị nhị trùng τ , có chức năng liên kết hai lõi xúc tác với nhau.

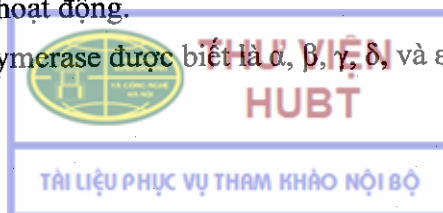
- Hai bản sao của "phức kẹp", chịu trách nhiệm giữ các lõi xúc tác trên sợi khuôn. Mỗi kẹp được cấu tạo bởi 2 tiểu đơn vị β .

- Phức γ bao gồm 5 protein, gọi là "phức tải kẹp" có vai trò đặt kẹp lên sợi ADN.

ADN polymerase IV và V liên quan đến việc cho phép quá trình sao chép bỏ qua một số loại tổn thương trong ADN.

ADN polymerase của thực khuẩn thường kết hợp với các protein khác của nó hay của tế bào chủ để tạo thành dạng hoạt động.

Ở nhân thật có 5 loại polymerase được biết là α , β , γ , δ , và ϵ (Bảng 2.2).



Các enzym này có cơ chế tác động tương tự với các enzym của tế bào nhân nguyên thủy. Polymerase ϵ và δ cần cho sự sao chép của ADN nhân. Polymerase ϵ cũng có vai trò trong sửa chữa ADN nhân. Polymerase α liên quan đến sự mồi (khởi đầu) sao chép. Polymerase β cấu tạo từ một chuỗi polypeptid đơn, đóng vai trò sửa chữa. Polymerase γ được tìm thấy trong ty thể và chịu trách nhiệm sao chép ADN ty thể.

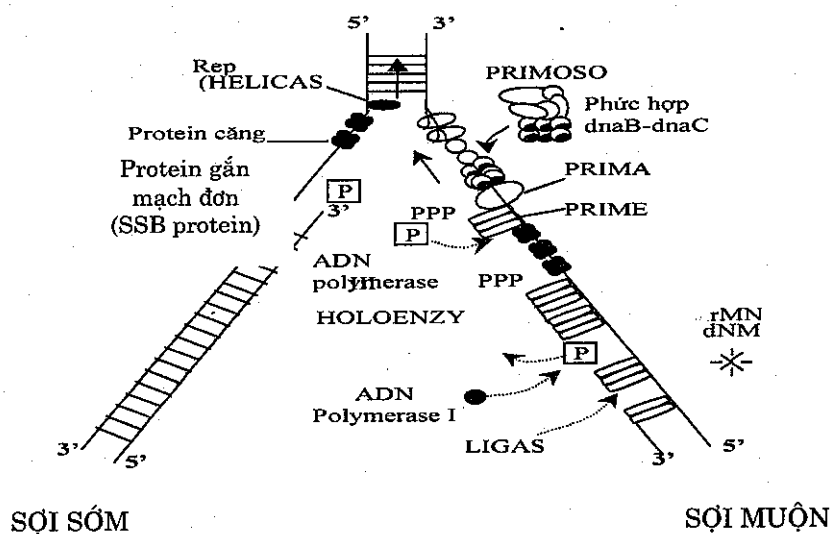
Bảng 2.2. ADN polymerase của nhân thật

Polymerase	α	β	γ	δ	ϵ
Mr (kDa)	350	39	200	250	350
Đơn vị nhỏ	4	1	2	4	4
Vị trí	Nhân	Nhân	Ty thể	Nhân	Nhân
Chức năng	Khởi đầu sao chép AND nhân	Sửa chữa	Sao chép ADN ty thể	Sao chép ADN nhân	Sao chép ADN nhân

2.2.4. Quá trình sao chép ADN ở *E. coli*

2.2.4.1. Chạc ba sao chép

ADN của tế bào nhân nguyên thủy có dạng vòng, xoắn kép. Quá trình sao chép ADN thường bắt đầu từ một bong bóng sao chép, đây là chỗ phình khởi đầu sự tổng hợp ADN. Các nút sao chép chính là các chạc ba sao chép (Hình 2.4).



Hình 2.4. Chạc ba sao chép trong quá trình sao chép ADN

ADN ty thể có dạng vòng và có hai chạc ba sao chép. Sự tổng hợp ADN của nhiễm sắc thể vi khuẩn luôn luôn bắt đầu ở cùng một điểm. Điểm này được gọi là vị trí Origin hay vị trí Ori, là một vùng gồm 254 cặp base. Dừng các chủng vi

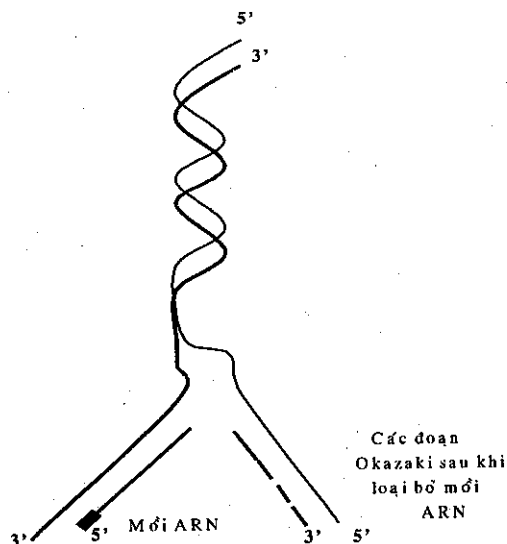
khuẩn đột biến, người ta đã chứng minh được rằng chỉ có một số nhỏ trong số các base này là cần cho sự khởi đầu sao chép. Ở *E. coli*, quá trình sao chép bắt đầu khi protein B nhận biết được điểm khởi sự sao chép này (OriC).

Các sợi ADN được tách ra và một sợi ADN có hướng 3' → 5' sẽ được sao chép trực tiếp bởi ADN polymerase III. Sợi con bổ sung vừa tạo ra được gọi là sợi sớm (leading strand). Sợi gốc còn lại có hướng 5' → 3', không được sao chép cho đến khi một phần của sợi gốc được tháo xoắn. Bản sao này được gọi là sợi muộn (lagging strand). Sợi muộn được tổng hợp bằng một quá trình sao chép không liên tục. Sự sinh tổng hợp sợi muộn và sợi sớm được điều hoà bởi sự tạo nút thắt (loop) của sợi muộn, nhờ đó sợi muộn cũng được tổng hợp cùng hướng với sợi sớm.

Việc sao chép không liên tục tạo ra các đoạn ADN ngắn vào khoảng 1000 – 2000 nucleotid gọi là đoạn Okazaki (Hình 2.5). Các đoạn này sau đó được nối với nhau bởi ADN ligase để tạo ra sợi ADN liên tục.

Đoạn mỗi ARN luôn luôn cần cho sự tổng hợp ADN. ADN polymerase đòi hỏi một đoạn ARN ngắn có đầu 3' – OH tự do để bắt đầu sự tổng hợp và gắn được desoxyribonucleotid vào mạch con, bổ sung được với ADN khuôn.

Mỗi ARN được tổng hợp bởi một enzym đặc biệt có tên gọi là primase. Chúng có thể kéo dài sự khởi động chuỗi *de novo*. Primase nhận ra trình tự đặc hiệu trên chuỗi ADN đơn. Tự bản thân primase không hoạt động được, nó phải tạo phức hợp với vài chuỗi polypeptid để tạo thành primosome chức năng. Các trình tự ARN được tạo ra bởi primosome là những đoạn ngắn khoảng 5 – 10 base đặc hiệu. Các protein nhận diện được gọi là N – protein, chọn các điểm (origin) trên ADN, tại đó primase có thể hoạt động. Các đoạn mỗi ARN sẽ bị ADN polymerase I loại khỏi chuỗi ADN. Enzym này đồng thời làm nhiệm vụ lấp đầy các khoảng trống bằng các desoxyribonucleotid thích hợp do việc loại môi. Các đoạn ADN được nối lại với nhau bởi ADN ligase tạo thành sợi ADN liên tục.

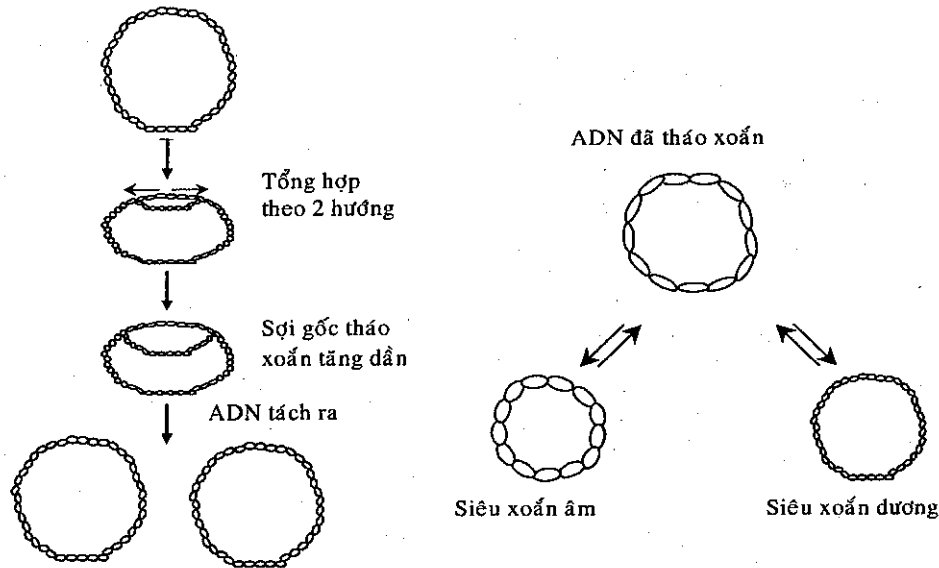


Hình 2.5. Sự hình thành các đoạn Okazaki của sợi chậm ADN do sự sao chép không liên tục của sợi gốc

2.2.4.2. Cấu trúc theta (θ)

Sự sao chép ADN vòng tạo ra cấu trúc theta, hình thành bởi hai chạc ba sao chép xuất phát từ một vị trí Origin. Sự tổng hợp ADN được tiến hành theo cả hai chiều thuận và ngược kim đồng hồ cùng một lúc. Để tạo ra các sợi đơn ADN, hai sợi ADN gốc phải vặn xoắn (quay). Cứ mỗi 10 base được sao chép thì phân tử ADN phải vặn xoắn 1 vòng.

Chạc ba sao chép của *E. coli* cần phải di chuyển với tốc độ 800 base/giây, đòi hỏi ADN gốc phải tháo xoắn với tốc độ 80 vòng/giây. Sự tháo xoắn ADN gây ra hiện tượng siêu xoắn (supercoiling). Việc tháo xoắn sợi kép dẫn đến siêu xoắn âm (xoắn ngược chiều và làm giảm số vòng), trong khi đó, sự xoắn thêm sẽ dẫn đến siêu xoắn dương (xoắn kép cùng chiều, số vòng tăng) (Hình 2.6).

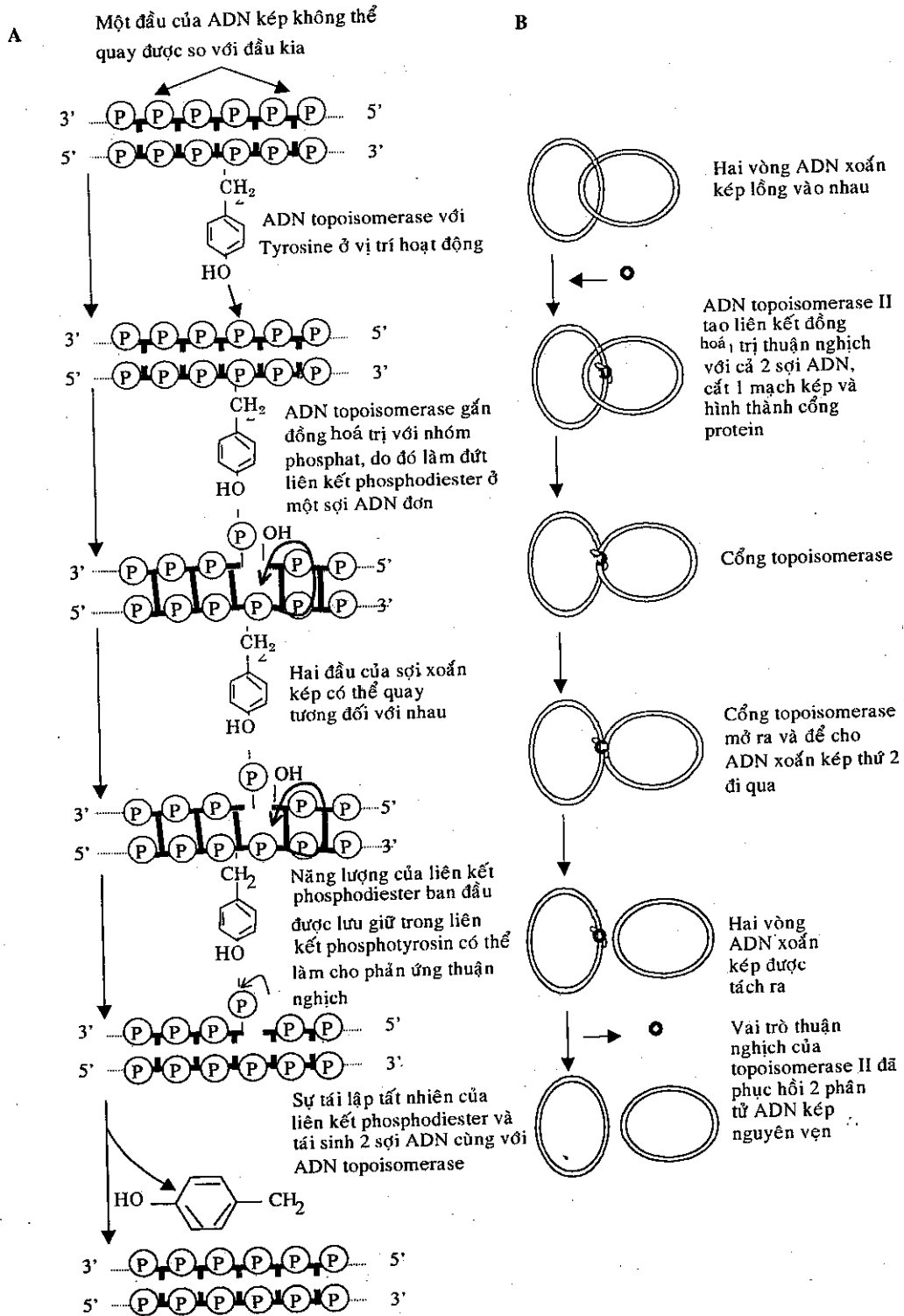


Hình 2.6. Sự hình thành cấu trúc theta do sự sao chép hai hướng của ADN kép

Cơ chế sao chép bán bảo tồn của ADN đòi hỏi các điểm cắt ở một sợi hay ở cả hai sợi ADN để tách chuỗi. Một nhóm enzym, gọi là topoisomerase sẽ chuyển trạng thái topo của ADN sang trạng thái khác. Topoisomerase II còn gọi là gyrase tạo ra các dạng siêu xoắn âm (phải sang trái). Topoisomerase I gắn với ADN, cắt một sợi, cho phép ADN xoắn kép quay quanh một điểm làm mất đi sự xoắn (Hình 2.7 – A).

Phần Tyrosin của topoisomerase gắn với gốc phosphat tự do trên ADN gốc và phức hợp sẽ quay. Lúc bấy giờ, enzym tách khỏi phức hợp và sợi ADN được tạo thành. Loại topoisomerase I tham gia vào quá trình tạo chạc ba ở vi khuẩn cũng được tìm thấy ở tế bào động vật.

Hai sợi ADN lồng vào nhau được gọi là vòng lồng ghép (catena) hay ADN lồng ghép và thường được tạo ra khi sao chép ADN vòng. Loại topoisomerase II thường làm mất đi vòng catena, cắt một sợi ADN kép và để phân tử ADN sợi kép còn lại đi xuyên qua điểm cắt, sau đó sợi kép ban đầu được nối lại (Hình 2.7 – B).



Hình 2.7. Vai trò của topoisomerase

A – Topoisomerase I trong tháo xoắn ADN (trái).

B – Topoisomerase II (Gyrase) trong việc tách hai phần tử ADN mạch kép lồng vào nhau (catena)

2.2.4.3. Enzym helicase

Các enzym helicase cần thiết để hỗ trợ cho sự tháo xoắn ADN gốc vì hai sợi đơn ADN không thể tự nhiên tách ra được (Hình 2.4).

Helicase cần ATP như một cofactor. Helicase dùng năng lượng tự do khi thủy phân ATP để đi dọc theo sợi ADN, làm tăng tốc độ tách hai sợi ADN.

Có hai loại helicase khác nhau: một là Rep – protein gắn với ADN gốc, trực tiếp tổng hợp sợi sớm và di chuyển theo hướng 3' → 5'. Helicase thứ hai gắn với sợi khuôn kia để tổng hợp sợi chậm, enzym này tạo phức hợp với primase để tạo mỗi ARN.

2.2.4.4. Các protein khác tham gia sao chép

Các SSB – protein (protein gắn mạch đơn) giữ cho các sợi đơn ADN do helicase tạo ra không chập lại với nhau. Nhờ sự hiện diện của SSB – protein mà ADN ở dạng tháo xoắn rời nhau và ổn định, lý tưởng cho việc sao chép. Nếu các SSB – protein bị tách ra sẽ làm xuất hiện các nút kẹp tóc (hairpin loop) trong ADN gốc và như vậy sẽ ngăn chặn sự sao chép.

Ba protein khác hỗ trợ cho việc gắn ADN polymerase và hiệp đồng tác động như ATPase, thủy phân ATP để tạo năng lượng tự do cho helicase hoạt động. Tác động hiệp đồng của các protein này thể hiện trong Hình 2.4 và Hình 2.9 và chức năng của chúng được tóm tắt trong Bảng 2.3.

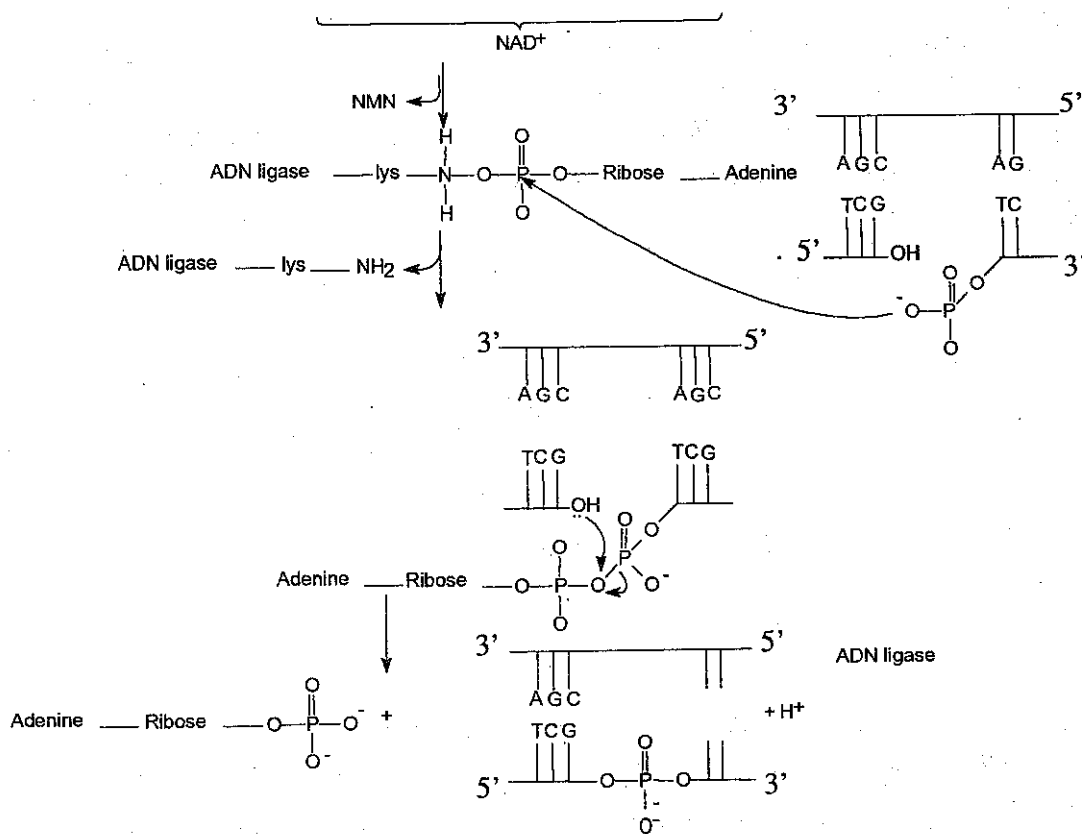
Bảng 2.3. Chức năng của các protein sao chép quan trọng của E. coli

Protein	Chức năng
Protein B	Nhận biết điểm Ori
N – protein	Cho phép primase hoạt động
Primase	Tổng hợp mỗi ARN
Rep protein	Tháo xoắn sợi sớm
Helicase	Tháo xoắn sợi chậm
SSB – protein	Ổn định sợi ADN đơn
ADN polymerase III	Tổng hợp ADN
ADN topoisomerase I	Cắt khóa một sợi đơn ADN
ADN topoisomerase II	Cắt khóa cả hai sợi đơn ADN
ADN ligase	Nối các đầu của polynucleotid đã hình thành

2.2.4.5. ADN ligase

ADN ligase gồm các enzym xúc tác hình thành liên kết phosphodiester giữa các polynucleotid được hình thành.

ADN ligase liên quan tới việc tổng hợp ADN không liên tục bằng cách nối các đoạn Okazaki sau khi thay thế các đoạn mỗi ARN. Chúng cũng cần thiết cho việc sửa chữa các ADN hư hỏng. Cơ chế tác động của các enzym này khác với các ADN polymerase là do liên kết pyrophosphat của NAD^+ tham gia (Hình 2.8).



Hình 2.8. Cơ chế hoạt động của ADN ligase
(NMN = Nicotinamide mononucleotide)

2.2.5. Sao chép ADN ở tế bào nhân thật

2.2.5.1. Cơ chế sao chép

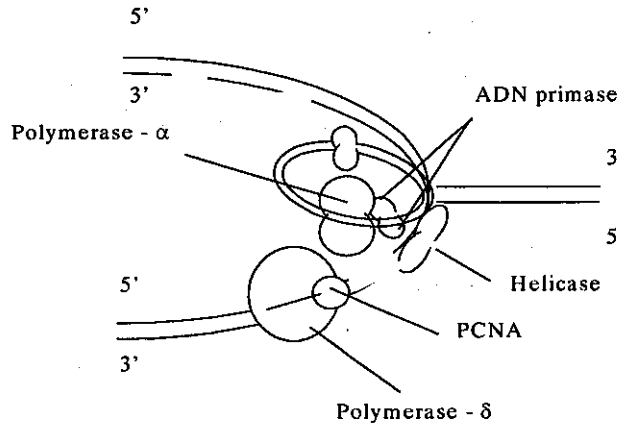
Cơ chế tổng hợp ADN ở tế bào nhân thật cũng tương tự như ở tế bào nhân nguyên thủy. Điểm khác biệt chủ yếu là ADN ở tế bào nhân thật đóng cuộn trong nhiều nhiễm sắc thể và dài hơn. Tốc độ di chuyển ADN polymerase ở tế bào nhân thật chậm hơn rất nhiều so với tế bào nhân nguyên thủy (khoảng 50 nucleotid/giây). Nhưng tế bào nhân thật chứa tới 20 000 phân tử enzym. Do vậy, ở nhiễm sắc thể tế bào nhân thật hình thành một lượng lớn chạc ba sao chép, khoảng 2000 hay nhiều hơn. Các đoạn Okazaki nhỏ hơn, dài khoảng 40 - 300 base. Do đó, tốc độ sao chép ADN ở tế bào nhân thật nhanh hơn rất nhiều so với

E. coli. Điểm khác biệt cơ bản là ADN tế bào nhân thật có nhiều replicon, ví dụ ở *Saccharomyces cerevisiae* có tới 500 replicon tức là có 500 điểm Ori.

Cơ chế và sự điều hoà sao chép ở tế bào nhân thật đã được nghiên cứu bằng một mô hình đơn giản, đó là nghiên cứu sự sao chép ADN virus trong tế bào động vật.

Hai loại virus, là Adenovirus và SV40 (virus linh trưởng) được dùng làm các mô hình này. Cơ chế sao chép của nhiễm sắc thể vòng của SV40 cũng tương tự với cơ chế sao chép từ một vị trí khởi đầu (Ori) của nhiễm sắc thể tế bào nhân thật.

Mô hình tạo chạc ba sao chép ở tế bào nhân thật dựa trên các nghiên cứu này được trình bày ở Hình 2.9.



Hình 2.9. Mô hình sao chép ở tế bào nhân thật

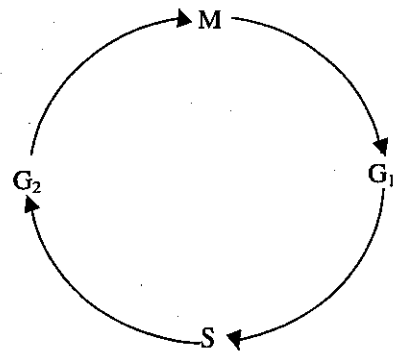
(PCNA = proliferating cell nuclear antigen = kháng nguyên tăng sinh nhân tế bào)

Có hai loại polymerase tham gia vào quá trình sao chép: Polymerase δ tham gia tổng hợp sợi sớm và polymerase α tổng hợp sợi muộn. Sợi muộn được thắt nút xung quanh polymerase δ , cho phép enzym di chuyển theo hướng chạc ba sao chép, giống như sự sao chép ở tế bào nhân nguyên thủy.

2.2.5.2. Chu kỳ tế bào

Sự sao chép ADN diễn ra như một phần của tiến trình phối hợp của sự phân chia tế bào. Chu kỳ tế bào nhân thật gồm 4 pha riêng biệt: pha M, G_1 , S và G_2 (Hình 2.10).

Pha G_1 là thời kỳ trước khi ADN bắt đầu tổng hợp, độ dài của G_1 thay đổi từ phút, giờ, tuần thậm chí đến hàng năm. Tế bào không bao giờ phân chia là tế bào trong đó G_1 bị ngưng trệ, thường được coi như là giai đoạn G_0 .

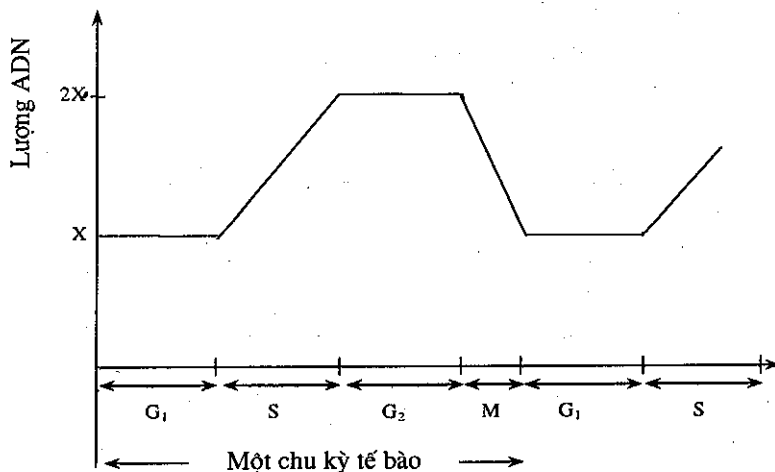


Hình 2.10. Chu kỳ tế bào nhân thật

Trong pha S, ADN được sao chép và lượng ADN tăng gấp đôi (Hình 2.11). Các histone mới cũng được tổng hợp trong pha này, tạo thành 2 bộ nhiễm sắc thể. Tuy nhiên, các nhiễm sắc tử vẫn dính chung cho đến khi phân chia.

Pha G_2 ngắn hơn pha G_1 , trong pha này người ta biết được ít những điều gì xảy ra ở tế bào. Pha G_2 kết thúc với các dấu hiệu đầu tiên của sự phân bào.

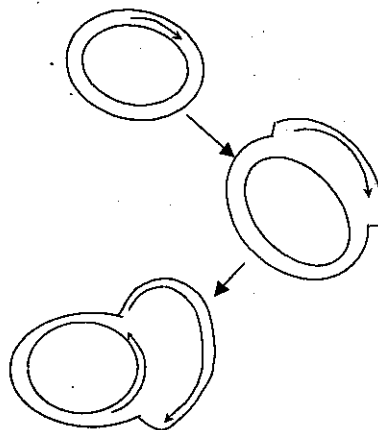
Pha M hay pha phân bào dẫn đến sự hoà tan màng nhân, sự tách nhiễm sắc thể và sự phân chia tế bào.



Hình 2.11. Lượng ADN được tạo ra trong chu kỳ tế bào

2.2.5.3. Sự sao chép của ADN ty thể và lạp thể

Ty thể và có khi cả lạp thể, chứa ADN polymerase. Cụ thể là ở ty thể có ADN polymerase γ . Sự sao chép bắt đầu từ việc sao chép một sợi của ADN gốc. Sự sao chép này tiến hành chưa được nửa đường thì sự sao chép của sợi kia bắt đầu. Kết quả tạo nên cấu trúc D. Kết quả của quá trình sao chép là hình thành sự lồng ghép của 2 vòng ADN và topoisomerase II cần cho việc tách đôi 2 vòng kép ADN ty thể (Hình 2.12).



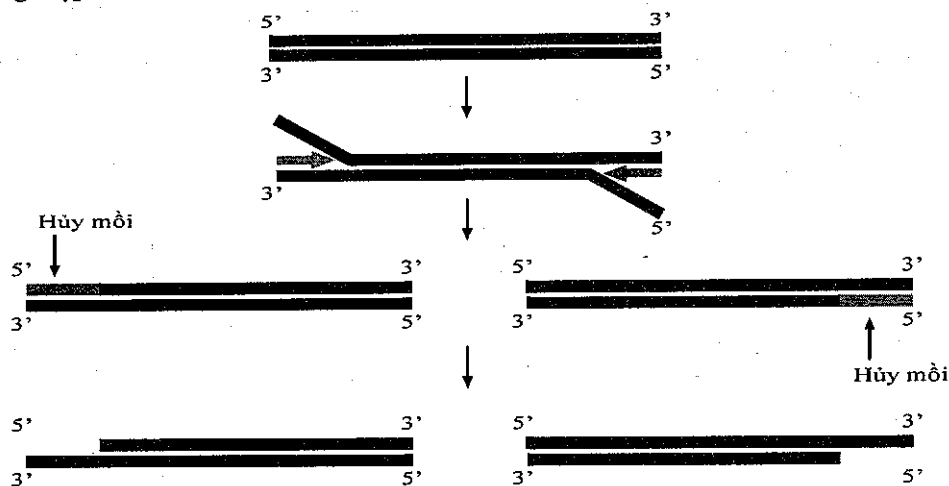
Hình 2.12. Sự hình thành cấu trúc D trong quá trình sao chép ADN ty thể

2.2.6. Sự sao chép ở virus và phage

Sự sao chép vật liệu di truyền của virus hay phage được thực hiện bình thường trong trường hợp bộ gen là ADN dạng vòng đôi. Tuy nhiên, nếu bộ gen virus là ADN dạng thẳng hay ADN sợi đơn hay ARN thì cần có cơ chế sao chép đặc biệt.

2.2.6.1. Sao chép theo kiểu dạng thẳng

Nếu có dạng thẳng, bộ gen virus sẽ trở nên ngắn hơn sau mỗi lần được sao chép, trừ khi virus có cơ chế đặc biệt để khắc phục điều này. Do ADN polymerase cần có mỗi ARN và chỉ hoạt động tổng hợp ADN theo chiều 5' → 3' nên nó không thể tổng hợp chính xác các ADN thẳng (Hình 2.13).

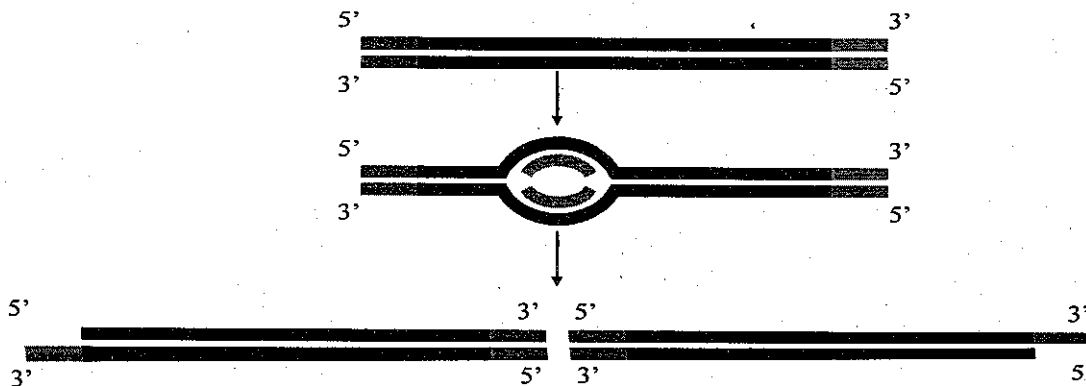


Hình 2.13. Sao chép ADN theo kiểu dạng thẳng

Để tránh hiện tượng này, các phage sử dụng một trong hai cách sau: phage λ vòng hoá bộ gen của nó nhờ các trình tự *cos*, còn phage T7 thành lập các phức nối.

Sao chép của phage T7:

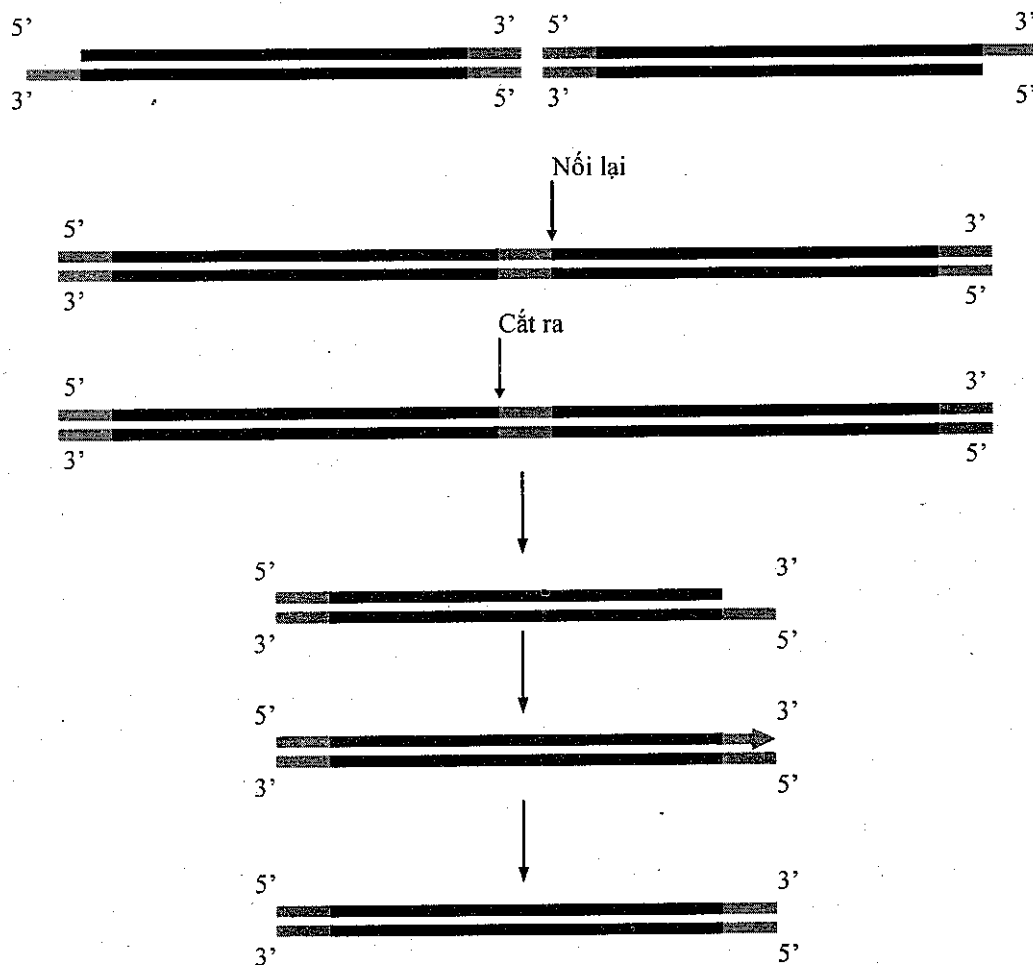
Phage T7 có bộ gen là dsADN thẳng, dài 39 937 bp. Sự sao chép bắt đầu tại một điểm cách đầu tận 5' khoảng 5900 bp và diễn ra theo cả hai hướng. Việc sao chép này gặp trở ngại của sự sao chép ADN thẳng nói trên. Để tránh trở ngại này, bộ gen của nó có 160 bp đầu tiên ở bên trái giống hệt 160 bp cuối cùng ở bên phải, gọi là các phần thừa ở đầu. Sau vòng sao chép đầu tiên hai phân tử con thu được có một đầu dính với trình tự của phần thừa (Hình 2.14).



Hình 2.14. Bước sao chép đầu tiên ở phage T7

Phần ssADN ở phía phải (3') của một phân tử con có thể gắn bổ sung với phần phía trái (cũng là 3') của phân tử khác. Khoảng trống được lấp nucleotid bởi ADN polymerase và được nối lại bởi ADN ligase. Phân tử nhị trùng thu được có thể được cắt ra trở lại, nhưng khi đó sẽ tạo ra đầu dính 5' thay vì 3', đầu dính 5' có thể dùng làm khuôn mẫu để hoàn chỉnh ADN sợi đôi (Hình 2.15).

T7 có gen mã hoá cho ADN ligase (gen 1.3), SSB (gen 2.5) và ADN polymerase (gen 5) của riêng nó.

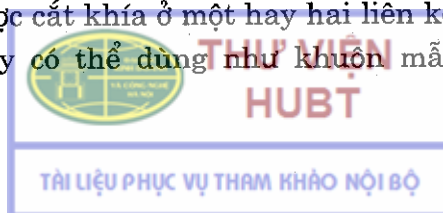


Hình 2.15. Phức nối được cắt ra và tái tạo bộ gen của phage T7

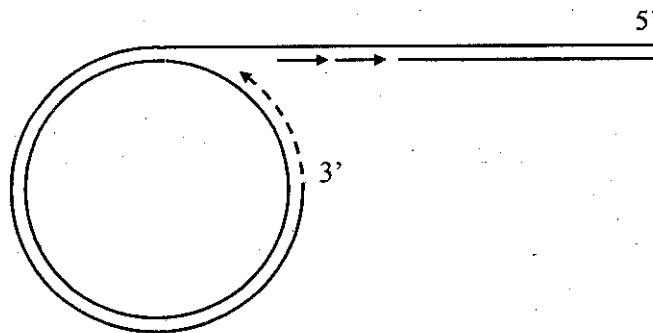
2.2.6.2. Sao chép theo kiểu lăn vòng

Sự sao chép theo kiểu theta không phải là cách duy nhất để sao chép các phân tử ADN vòng. Một cơ chế khác sao chép ADN theo kiểu lăn vòng mặc dù sản phẩm tạo ra không phải là ADN vòng mà là ADN dạng thẳng.

Phân tử ADN vòng được cắt khía ở một hay hai liên kết phosphodiester, tạo ra đầu 3' OH tự do, đầu này có thể dùng như khuôn mẫu để sao chép bởi ADN



polymerase. Khi đầu 3' OH được kéo dài, đầu 5' bị đẩy ra khỏi vòng dưới dạng sợi đơn, trong khi vòng ADN lăn để tiếp tục kéo dài đầu 3'. Kiểu tổng hợp trên vòng tương tự như tổng hợp sợi sớm. Trong lúc đó đầu 5' bị đẩy ra có thể dùng làm khuôn mẫu để tổng hợp sợi đôi theo kiểu tương tự sợi muộn nếu có sự hiện diện của đoạn môi thích hợp (Hình 2.16).



Hình 2.16. Sao chép theo kiểu lăn vòng

Nếu kiểu tổng hợp này tiếp tục, sự tổng hợp trên sợi 5' bị đẩy ra sẽ tạo thành một sợi ADN phức chứa nhiều bản sao liên tiếp của phân tử ADN vòng. Nghĩa là, nhiều bản sao của bộ gen được tạo ra.

Kiểu này gặp ở phage lambda và tạo ra nhiều bản sao nối tiếp, trong khi ở phage M13 chỉ tạo một bản sao.

Sao chép của phage M13:

Phage M13 (và các phage dạng sợi khác) có bộ gen là một phân tử ADN sợi đơn vòng. Khi nhiễm tế bào *E. coli*, phân tử này được bơm vào và được bao bởi các SSB protein. Vì phage M13 không có gen mã hoá cho ADN polymerase riêng nên phải lệ thuộc hoàn toàn vào bộ máy của tế bào chủ để sao chép. Mặc dù bộ gen là ADN sợi đơn của nó rất thích hợp để làm khuôn mẫu tổng hợp ADN, nhưng lại không thích hợp với việc tổng hợp ARN bởi ARN polymerase hay primase.

Tuy nhiên, một phần của bộ gen có dạng kẹp tóc sợi đôi. Vùng này có thể đóng vai trò promoter cho ARN polymerase của tế bào chủ để phiên mã một đoạn môi ARN ngắn. Sự phiên mã cũng làm mất cấu trúc kẹp tóc. ADN Polymerase III bấy giờ có thể tham gia và tổng hợp phân tử ADN sợi đôi, gọi là thể sao chép I (RFI – replicative form I).

Việc sao chép tiếp của RFI không thực hiện theo kiểu theta mà là lăn vòng. gp2 endonuclease, được mã hoá bởi gene 2 của phage, cắt khía ADN của RFI tại vị trí đặc hiệu (điểm gốc trên sợi dương) và sự sao chép lăn vòng diễn ra nhưng không tạo phân tử phức mà gp2 endonuclease cắt phân tử một lần nữa để hoàn tất bản sao. Sản phẩm thu được gồm một phân tử ADN vòng đơn (sợi bị thể ra được nối lại thành vòng) và một phân tử ADN sợi đôi là RFI. Phân tử ssADN vòng lúc này có thể được sao chép bằng cách lặp lại từ đầu quy trình trên.

Để có được ssADN như ban đầu để đóng gói vào capsid, các phân tử ssADN bị thế ra cần được bao bởi các SSB.

Sao chép ở phage Lambda:

Phage lambda có bộ gen là ADN sợi đôi thẳng. Tuy nhiên, ở mỗi tận cùng của bộ gen là một đoạn sợi đơn dài 12 nucleotid và được gọi là vị trí *cos*, hai đoạn đơn ở vị trí *cos* có thể gắn bổ sung với nhau để biến bộ gen thẳng thành dạng vòng đôi.

Sau khi phage nhiễm tế bào *E. coli* và tiêm ADN vào, nhiễm sắc thể của nó vòng hoá nhờ các vị trí *cos*. Nhờ đó nó tránh bị cắt bởi các exonuclease của vi khuẩn và có thể được sao chép theo chu trình tiêu giải tiềm ẩn. Quá trình sao chép của phage lambda gồm hai giai đoạn.

Giai đoạn sớm:

Phage lambda khởi đầu sao chép theo mô hình *theta*. Điểm gốc sao chép (*ori*) nằm ở gen O, có sản phẩm tương tự DnaA (yếu tố khởi đầu sao chép). Phân tử này gắn trình tự lặp ở điểm gốc và khởi đầu việc tách hai sợi ADN tại chỗ. Một gen khác, gen P, có sản phẩm với chức năng tương tự DnaC (yếu tố nạp helicase). Sản phẩm này giúp DnaB (helicase) gắn vào ADN đã tách mạch. Sau đó, các thành phần khác của bộ máy sao chép của vi khuẩn gắn vào và bắt đầu việc sao chép. Kiểu sao chép này kéo dài khoảng 5 - 15 phút.

Giai đoạn muộn:

Sau 15 phút, phage lambda chuyển sang kiểu sao chép lăn vòng. Hiện chưa biết nguyên nhân và cơ chế của sự chuyển đổi này.

Hậu quả của cơ chế lăn vòng là ADN của phage được sao chép thành một ADN kép dài chứa nhiều bản sao liên tiếp của bộ gen phage, cách nhau bằng các trình tự *cos*. Phage có gen mã hoá cho một enzym là Terminase, enzym này nhận diện vị trí *cos* (ở dạng sợi kép) và cắt tại đó để biến phân tử ADN dài thành các bản sao đơn của bộ gen phage với các vị trí *cos* ở hai đầu.

In vivo, quá trình xử lý bản sao lớn thành bộ gen đơn có sự tham gia của các protein capsid. Các protein này nhận diện vị trí *cos* thứ nhất và nếu khoảng cách đến vị trí *cos* thứ hai phải nằm trong giới hạn 75 - 105% so với chiều dài nguyên thủy của bộ gen phage thì capsid sẽ đóng gói ADN tạo thành phage hoàn chỉnh.

2.3. SỬA SAI TRONG SAO CHÉP VÀ KHI KHÔNG SAO CHÉP

ADN được sao chép bởi ADN polymerase với độ chính xác cao (cứ khoảng $10^8 - 10^{12}$ base thì có một base sai được gắn vào mỗi lần sao chép). Sự sao chép chính xác của phân tử ADN có tầm quan trọng sống còn đối với hoạt động bình thường của tế bào.



2.3.1. Sửa sai trong sao chép

Người ta đã dùng các nucleotid và ADN polymerase để tổng hợp ADN *in vitro*. Sai sót trong trường hợp này là 10^{-5} tức là trong 100 000 nucleotid có một sai sót. Ở *E. coli* có $3 \cdot 10^6$ cặp base, vậy mỗi lần sao chép theo tính toán như trong thực nghiệm phải có tới 30 sai sót. Tuy nhiên, trên thực tế sai sót trong tự nhiên thấp hơn nhiều.

Bằng cách đánh giá tần số các đột biến mới xuất hiện trong quần thể lớn và theo dõi biến đổi enzym nào đó trong nuôi cấy mô tế bào, người ta tính được sai sót trong cơ thể sinh vật khi sao chép *in vivo* là 10^{-9} tức là một sai sót trên một tỉ base. Như vậy tế bào có $3 \cdot 10^9$ cặp base mỗi lần sao chép chỉ có 3 sai sót.

Mức chính xác cao của sao chép trong cơ thể sinh vật có được nhờ các cơ chế sửa sai:

Hướng sao chép bao giờ cũng từ đầu 5' → 3' để việc sửa sai chính xác.

Ở tế bào nhân nguyên thủy ADN polymerase I và III vừa polymer hoá, vừa có hoạt tính exonuclease 5' → 3' và 3' → 5'. Nếu trên đường di chuyển để polymer hoá gặp base sai, ADN – polymerase sẽ lùi lại cắt bỏ theo hướng 3' → 5' (hoạt tính exonuclease 3' → 5').

Ở tế bào nhân thật, hoạt tính exonuclease chỉ có được ở polymerase δ và ε. Không tìm thấy hoạt tính exonuclease ở polymerase α, có thể do một tiểu đơn vị của nó bị mất đi trong quá trình chiết tách enzym này.

ADN ty thể dễ bị đột biến, nhưng polymerase γ của nó lại không có hoạt tính exonuclease.

2.3.2. Sửa sai khi không sao chép

Phân tử ADN có thể bị biến đổi ngay cả khi không sao chép. Các biến đổi đột biến này xảy ra với tần số khá cao. Nhờ cơ chế sửa sai nên tần số đột biến được duy trì ở mức thấp.

Các enzym nhận biết gắn vào các trình tự sai và cắt rời đoạn sai, rồi dùng mạch đơn đúng làm khuôn mẫu để tổng hợp lại chỗ bị sai cho đúng.

Hàng loạt enzym đặc hiệu làm nhiệm vụ dò tìm và sửa sai. Có khoảng 20 enzym rà soát đọc ADN để dò tìm các base bị biến đổi hoá học, mỗi enzym có một chức năng chuyên biệt cho một loại sai hỏng. Khoảng 5 enzym khác đặc hiệu cho các liên kết cộng hoá trị sai giữa base với các chất hoá học khác hoặc giữa các base kề nhau trên một mạch. Một số enzym khác phát hiện sự bất cặp sai, như trường hợp mất purin. Tổng cộng có khoảng 50 enzym chuyên biệt phát hiện và sửa các sai hỏng trên ADN.

Tóm tắt:

Sự sao chép chính xác của ADN nhiễm sắc thể là một sự khởi đầu cần thiết cho sự sinh sản. Sự sao chép theo cơ chế bán bảo tồn và mỗi phân tử ADN mới bao gồm một mạch ADN gốc và một mạch mới được tổng hợp. Sự tổng hợp ADN được thực hiện bởi các ADN polymerase chuyên biệt, chúng dùng ADN có sẵn làm khuôn mẫu. Chỉ có loại ADN polymerase III là cần thiết cho sự sao chép ADN của tế bào nhân nguyên thủy. Tuy nhiên, ở các tế bào nhân thật có một số polymerase tham gia vào quá trình sao chép. Một số protein khác cũng tham gia vào quá trình tổng hợp ADN ở cả tế bào nhân thật lẫn tế bào nhân nguyên thủy.

Các phân tử ADN có thể bị hư hại bởi các tác nhân khác nhau và sự hư hại này có thể được sửa chữa để hạn chế các đột biến có hại. Quá trình sửa chữa được điều khiển bởi các hệ thống có sự tham gia của enzym.

CÂU HỎI

1. ADN sao chép theo cơ chế bán bảo toàn vì từ một gen ban đầu tạo ra
 - a) 2 gen con chứa các nucleotid cũ và mới xen kẽ.
 - b) 2 mạch đơn ADN chứa các nucleotid cũ và mới xen kẽ.
 - c) 1 gen con hoàn toàn mới, 1 gen con hoàn toàn cũ.
 - d) 2 gen con, mỗi gen chứa một mạch mới, một mạch cũ.
 - e) 1 gen ban đầu đồng thời tồn tại với 2 gen con vừa mới vừa cũ.
2. Chọn tổ hợp sai
 - a) ADN polymerase α – Nhân – Khởi đầu sao chép.
 - b) ADN polymerase β – Nhân – Sao chép sợi sớm.
 - c) ADN polymerase γ – Ty thể – Sao chép ADN.
 - d) ADN polymerase δ – Nhân – Sao chép sợi sớm.
 - e) ADN polymerase ϵ – Nhân – Sao chép ADN.
3. ADN polymerase đóng vai trò sửa chữa của tế bào nhân thật
 - a) I
 - b) II
 - c) α
 - d) β
 - e) b và d
4. Điểm khởi đầu sao chép là
 - a) Nút sao chép
 - b) Chạc ba sao chép
 - c) Vị trí Origin
 - d) Vị trí Okazaki
 - e) Bong bóng sao chép



5. Ý nào đúng với đoạn Okazaki ở tế bào nhân nguyên thủy?
- Gồm khoảng 1000 – 2000 nucleotid
 - Được nối lại bằng ADN ligase
 - Nối với nhau tạo thành sợi sớm
 - Còn gọi là loop
 - a và b
6. Vị trí Origin
- Điểm khởi đầu sự sao chép
 - Được nhận diện bởi protein B
 - Gồm 254 cặp base
 - a và b
 - a, b và c
7. Primase là enzym
- Tự bản thân không hoạt động được
 - Gồm nhiều N – protein
 - Còn gọi là primosome
 - Lấp đầy các GAP bằng dNTP
 - a và c
8. Primase bắt đầu hoạt động khi
- N – protein được nhận diện
 - N – protein nhận diện được Ori
 - Protein – B nhận diện được Ori
 - Tạo phức hợp với các chuỗi polypeptid
 - Tạo phức hợp với N – protein
9. Bản sao ADN vòng được tổng hợp, tách ra khỏi bản gốc nhờ
- Topoisomerase I
 - Topoisomerase II
 - Việc tháo xoắn âm
 - Phần Tyrosin của Topoisomerase
 - Gốc phosphat tự do của Topoisomerase
10. Phage lambda sao chép bộ gen của nó theo kiểu
- Theta
 - Theta và lăn vòng
 - Sao chép ADN thẳng
 - Sao chép ADN vòng
 - Sao chép ngược

Bài 3

CÁC LOẠI ARN

MỤC TIÊU

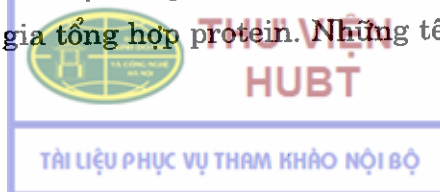
- Mô tả được cấu trúc của từng loại ARN.
- Giải thích được chức năng và vai trò của các loại ARN.
- Trình bày được các cách thức biến đổi các bản sao nguyên thủy ARN và các chức năng của sự biến đổi này.

3.1. KHÁI NIỆM

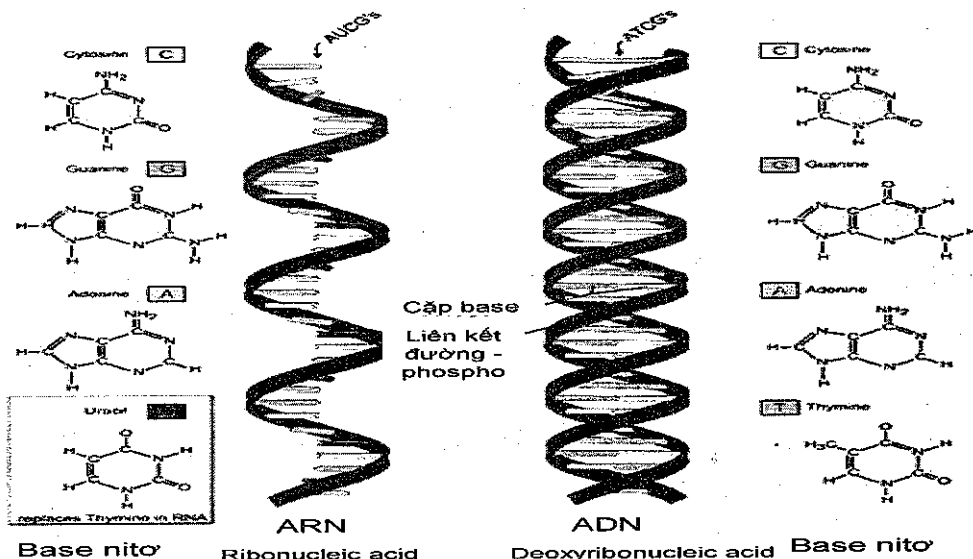
ADN mang thông tin di truyền của tế bào nhưng bản thân nó không trực tiếp chỉ huy quá trình tổng hợp protein, vì ADN nhiễm sắc thể chỉ mang một bản sao duy nhất cho mỗi gen, trong khi tế bào có hàng nghìn phân tử protein cùng loại tồn tại. Mặt khác, không phải tất cả các gen đều được biểu hiện thành protein cùng lúc mà các gen khác nhau sẽ biểu hiện khác nhau tại mỗi thời điểm trong vòng đời của tế bào và tùy theo điều kiện môi trường. Do vậy, cần có cơ chế trung gian để khuếch đại thông tin di truyền từ ADN đến protein, đồng thời kiểm soát sự biểu hiện của gen theo nhu cầu của tế bào. ARN đóng vai trò trung gian quan trọng này. Từ mỗi gen có thể có nhiều bản sao ARN để chỉ huy quá trình tổng hợp protein, đồng thời vòng đời của ARN ngắn nên sự biểu hiện của gen qua trung gian ARN dễ dàng được kiểm soát bởi nhu cầu của tế bào và điều kiện môi trường.

Ngoài vai trò trung gian trong quá trình biểu hiện gen, nhiều ARN còn có vai trò cấu trúc hay xúc tác.

Ở tế bào nhân thật, ARN được tổng hợp ngay trong nhân (chứa ADN), sau đó đi vào tế bào chất để tham gia tổng hợp protein. Những tế bào giàu ARN tổng hợp



protein nhiều hơn. Ví dụ: các tế bào tổng hợp nhiều protein như ở gan, lá lách, tuyến tụy của tằm chứa ARN nhiều hơn so với tế bào ít tổng hợp protein như ở thận, tim, phổi.



Hình 3.1. So sánh cấu trúc ARN và ADN

Về cấu trúc, ARN giống ADN: mạch polyribonucleotid thẳng của ARN cũng chứa 4 loại ribonucleotid: Adenin (A), Guanin (G), Cytosin (C) và Uracil (U) thay cho Thymin (T) (Hình 3.1). Tất cả ARN đều được phiên mã từ các gen tương ứng trên ADN.

ARN được chia thành 2 nhóm chính: ARN mã hóa và ARN không mã hóa (Hình 3.2). Các loại ARN cụ thể gồm các loại:

- mARN - ARN thông tin.
- rARN - ARN ribosom.
- tARN - ARN vận chuyển.

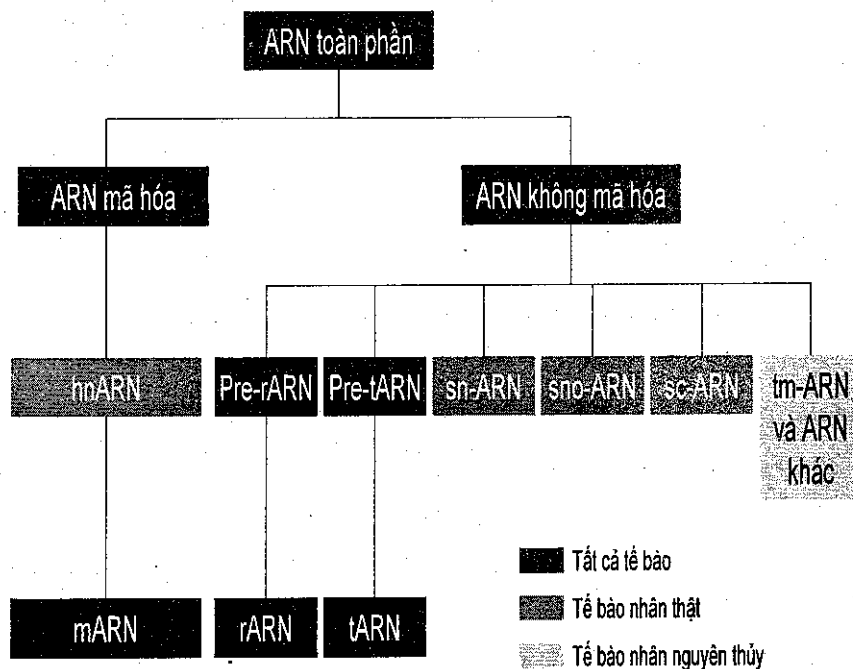
- pre - rARN (tiền rARN) là ARN được tổng hợp từ ADN, sau khi chúng được cắt nối sẽ trở thành rARN.

- pre - tARN (tiền tARN) là ARN tổng hợp từ ADN, sau khi được cắt nối sẽ trở thành tARN.

- hnARN - ARN nhân không đồng nhất, sẽ tạo mARN sau khi được cắt nối.
- snARN - ARN nhân nhỏ.
- scARN - ARN tế bào chất nhỏ.

ARN có thể ở dạng tự do hoặc gắn với protein thành các phức hợp nucleoprotein giữ nhiều vai trò quan trọng trong hoạt động sống của tế bào.

Có tới 7 loại ARN nhỏ khác nhau, trong đó có một số snARN có liên quan đến việc xử lý hnARN thành mARN chức năng. Gần đây người ta mới tìm thấy một vài loại ARN có cả ở tế bào nhân nguyên thủy và tế bào nhân thật là thành phần của enzym có chức năng xúc tác trực tiếp.



Hình 3.2. Các loại ARN

3.2. CÁC ARN VÀ VAI TRÒ CỦA CHÚNG

3.2.1. ARN ribosom

Cấu tạo ribosom phức tạp, gồm các phân tử rARN và khoảng 50 protein. rARN chiếm phân nửa khối lượng của ribosom.

Trong dịch nuôi cấy *Escherichia coli* đang ở giai đoạn phân chia nhanh, ribosom chiếm tới 1/3 khối lượng tế bào và rARN chiếm đến 75% của tổng số ARN.

Mỗi một ribosom có $Mr = 2.5-4.5 \times 10^6$. Có 3 loại ribosom khác nhau, thường được phân biệt dựa vào hệ số lắng S (Svedberg) của chúng khi siêu ly tâm. Ribosom của Eubacteria và lục lạp có hệ số lắng khoảng 70S. Ribosom của tế bào nhân thật là 80S, của ty thể động vật có vú khoảng 50S.

Tất cả ribosom đều được cấu tạo bởi 2 tiểu đơn vị (subunit) có chức năng và cấu trúc riêng biệt, các tiểu đơn vị này có thể tách ra và gắn lại với nhau một cách thuận nghịch tùy theo nồng độ của Mg^{2+} (Hình 3.3, Bảng 3.1).

Bảng 3.1. Thành phần cấu tạo ribosom của vi khuẩn và tế bào nhân thật

Tế bào	Kích thước ribosom	Tiểu đơn vị	rARN	Protein
Vi khuẩn	70S	Lớn (50S)	23S (2.900 nucleotid) 5S (120 nucleotid)	31
		Nhỏ (30S)	16S (1.500 nucleotid)	21
Nhân thật	80S	Lớn (60S)	28S (4.700 nucleotid) 5,8S (160 nucleotid) 5S (120 nucleotid)	49
		Nhỏ (40S)	18S (1.900 nucleotid)	33

3.2.1.1. Ribosom *E. coli*

Ribosom ở *E. coli* được nghiên cứu khá đầy đủ, các chuỗi polynucleotid của 3 loại rARN cùng với cấu trúc bậc 1 của 55 protein của chúng đã được khám phá.

Các protein ribosom (r-protein) của đơn vị nhỏ 30S được đặt tên từ S₁-S₂₁. Các protein của đơn vị lớn 50 S được đặt tên từ L₁-L₃₁. Hầu hết các r-protein có khối lượng tương đối nhỏ (trung bình Mr 17.000) và chứa nhiều acid amin base.

Hầu hết r-protein của ribosom *E. coli* có các trình tự bậc 1 khác nhau, ngoại trừ L₇, L₁₂ và L₈. L₇ là một dạng aminoacetyl hóa L₁₂ (sự acetyl hóa cần thiết cho hoạt tính GTPase trong quá trình dịch mã). L₈ là phức hợp của một phân tử L₁₀ liên kết với L₇ và L₁₂. Đôi khi trong quá trình tinh chế, phức hợp này không tách được.

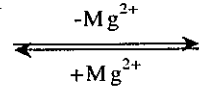
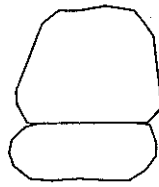
3.2.1.2. Sự lắp ráp ribosom

Các thông tin cho việc tự lắp ráp của ribosom được chứa đựng trong các phân tử thành phần rARN và r-protein. Các tiểu đơn vị ribosom có thể tái hợp *in vitro* đã được sử dụng rộng rãi cho việc nghiên cứu quá trình lắp ráp của ribosom. Các tiểu đơn vị được treo trong một dung dịch có nồng độ muối cao (ví dụ LiCl 4 M và urea 8 M) ngăn cản các tương tác không hóa trị giữa protein và ARN do đó chúng tách ra khỏi nhau. Dung dịch sau đó được thẩm phân để loại bỏ muối và urea, sau đó ủ với nhau. Dưới những điều kiện thích hợp, các tiểu đơn vị được tổ hợp lại theo chức năng của chúng và không thể phân biệt được với sự tổ hợp các tiểu đơn vị tự nhiên.

Khi dùng sự tổ hợp khác nhau giữa các thành phần riêng lẻ của ribosom, người ta có thể xác định được các trình tự liên kết trong quá trình lắp ráp.

Ví dụ, 6 r-protein của 30S của *E. coli* có vị trí kết nối độc lập và chuyên biệt với rARN 16S.

Ribosom 70S của tế bào nguyên thủy (Procarvote)



Tiểu đơn vị 50S



≈ 34 protein



23S rARN
 (≈ 3000 base)



5S rARN
 (≈ 120 base)

Tiểu đơn vị 30S

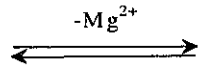
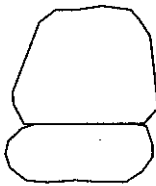


≈ 21 protein



16S rARN
 (≈ 1500 base)

Ribosom 80S của tế bào nhân thật (Eucaryote) của eucaryote



Tiểu đơn vị 60S



≈ 45 protein



28S rARN
 (≈ 5000 base)



5,8S rARN
 (160 base)



5S rARN
 (120 base)

Tiểu đơn vị 40S



≈ 33 protein



18S rARN
 (≈ 2000 base)

Hình 3.3. Thành phần của ribosom 70S (tế bào nhân nguyên thủy) và 80S (tế bào nhân thật)

3.2.1.3. Cấu trúc của rARN

Nhờ tiến bộ của công nghệ di truyền, người ta có thể tạo dòng và xác định trình tự của các rARN từ sinh vật khác nhau, từ ty thể và lục thể. Mặc dù chúng rất khác nhau về kích thước (Bảng 3.1) nhưng đều có thể xác lập được bản đồ cấu trúc bậc 2.



Bảng 3.2. Kích thước theo nucleotid của các rARN khác nhau

<i>Nguồn gốc ribosom</i>	<i>rARN của tiểu đơn vị lớn</i>	<i>rARN của tiểu đơn vị nhỏ</i>
Tế bào chất chuột	4950	1859
Tế bào chất nấm men (<i>Saccharomyces cerevisiae</i>)	3392	1787
<i>Escherichia coli</i>	2904	1541
Lạp thể thuốc lá	2950	1846
Ty thể linh trưởng	1559	854

Cấu trúc bậc 2 chính xác của rARN có tầm quan trọng trong việc lắp ráp ribosom và chúng được lưu trữ tốt trong quá trình tiến hoá. Tất cả các protein của *E. coli* được gắn kết vào rARN cũng có cấu hình bậc hai như vậy và nếu rARN bị biến tính thì protein

Một đặc tính trong cấu trúc bậc 1 của rARN là sự methyl hoá nucleotid. Sự methyl hoá này xảy ra sau khi phiên mã và được thực hiện bởi ARN – methylase trong đó có sử dụng S – Adenosyl Methionin (SAM) như là cơ chất cung cấp methyl. rARN của tế bào nhân nguyên thuỷ và của các bào quan thường methyl hoá ở base khởi đầu (5 – methylcytosin), còn rARN của tế bào chất tế bào nhân thật thường methyl hoá ở vị trí 2' của đường ribose (2' – 0 – methyladenosyl). Các vị trí của phần methyl hoá trong chuỗi rARN được lưu trữ tốt trong các loài khác nhau. Người ta cho rằng sự methyl hoá này đóng vai trò biến đổi bản sao sơ cấp của rARN.

3.2.1.4. Sự biến đổi của rARN

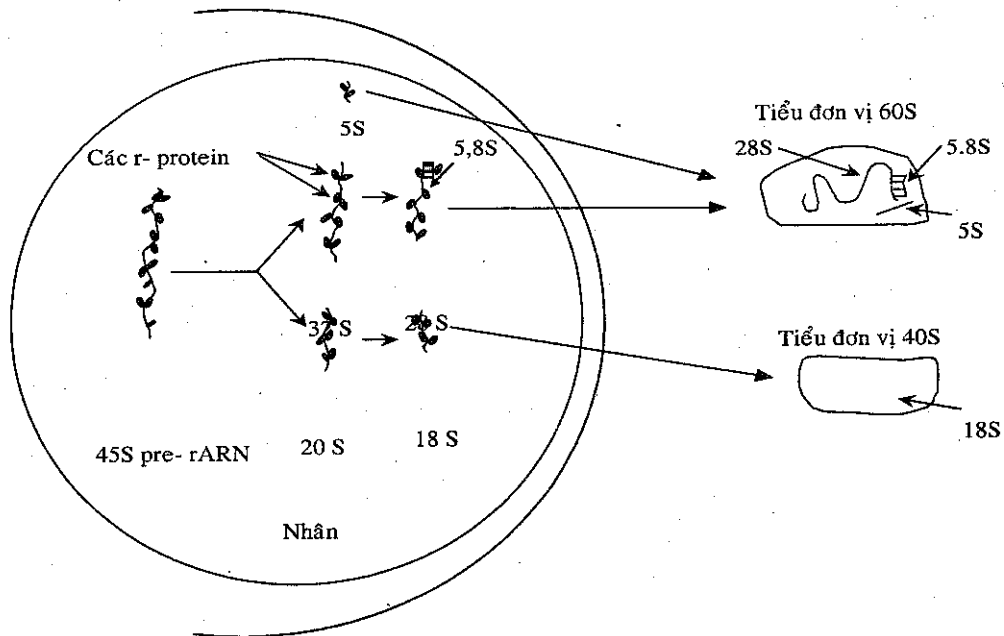
Hầu như ở tất cả các cơ quan, những rARN lớn của 2 tiểu đơn vị ribosom đều được tạo thành từ sự biến đổi và xử lý của các ARN lớn hơn, đó là các pre – rARN. Sự sửa đổi xảy ra trong nhân tế bào, đảm bảo cho 2 ARN lớn được biến đổi thành các ARN trưởng thành có khối lượng tương ứng với từng tiểu đơn vị.

Cách biến đổi của rARN ở tế bào HeLa (người) được minh họa ở Hình 3.4.

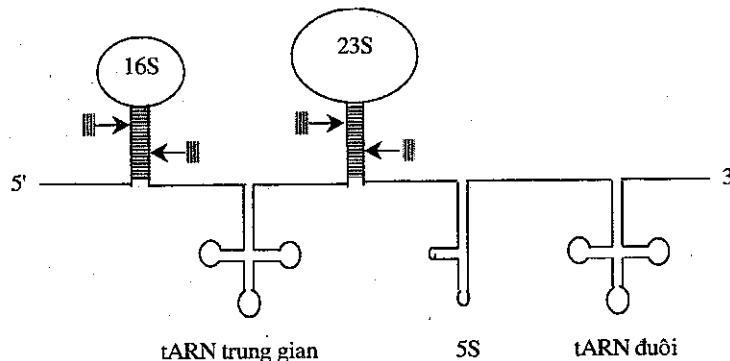
Pre – rARN 45S (gồm khoảng 12000 nucleotid) là bản sao của các gen ARN – ribosom. Các pre – ARN có vài trăm bản sao và tập hợp lại tạo thành phần đầu trong một vùng nhiễm sắc thể và được xem như vùng thiết lập nhân. rARN

45S được biến đổi do đưa vào khoảng 110 nhóm methyl ở các nucleotid đặc hiệu. Hầu hết những phân được methyl hoá đều được bảo tồn ở rARN trưởng thành. Hàng loạt diễn biến trong và ngoài nhân đã tạo ra được các tiền thể trung gian 32S và 20S từ pre - rARN 45S. Sau đó, pre rARN 32S cắt nối thành 28S và 5,8S, còn 20S thành 18S.

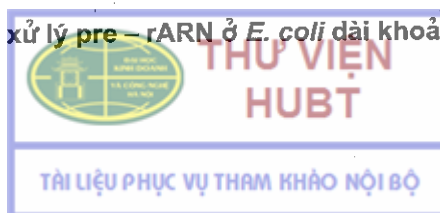
Quá trình sửa đổi của pre - ARN khác nhau về chi tiết giữa tế bào nhân thật và tế bào nhân nguyên thủy. Ở tế bào nhân nguyên thủy, chuỗi rARN 5S là thành phần của đơn vị phiên mã của cả rARN 16S và rARN 23S. Hơn nữa, gen rARN ở tế bào nhân nguyên thủy lại có chứa 2 chuỗi tARN hay nhiều hơn. Những sự phân cắt đầu tiên của pre - rARN được xúc tác bởi ARNse III đặc hiệu cho ARN mạch kép (Hình 3.5).



Hình 3.4. Quá trình sinh tổng hợp các tiểu đơn vị ribosom ở các tế bào HeLa



Hình 3.5. Quá trình xử lý pre - rARN ở *E. coli* dài khoảng 5000 nucleotid.

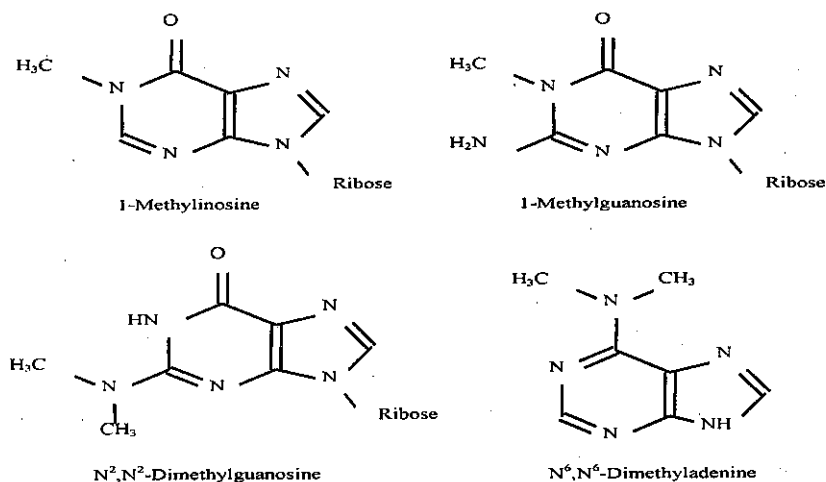


3.2.2. Các ARN vận chuyển

3.2.2.1. Cấu trúc

Năm 1950, F. Crick nêu ra giả thiết về chất nối. Ông cho rằng trước khi gắn thành polypeptid, các amino acid phải gắn qua chất nối trung gian, chất này sẽ bắt cặp đặc hiệu với các base trên mRNA. Vào năm 1957, M. Hoagland và các cộng sự đã tìm ra tARN và chứng minh được rằng mỗi phân tử tARN gắn với một phân tử amino acid và mang chúng đến ribosom.

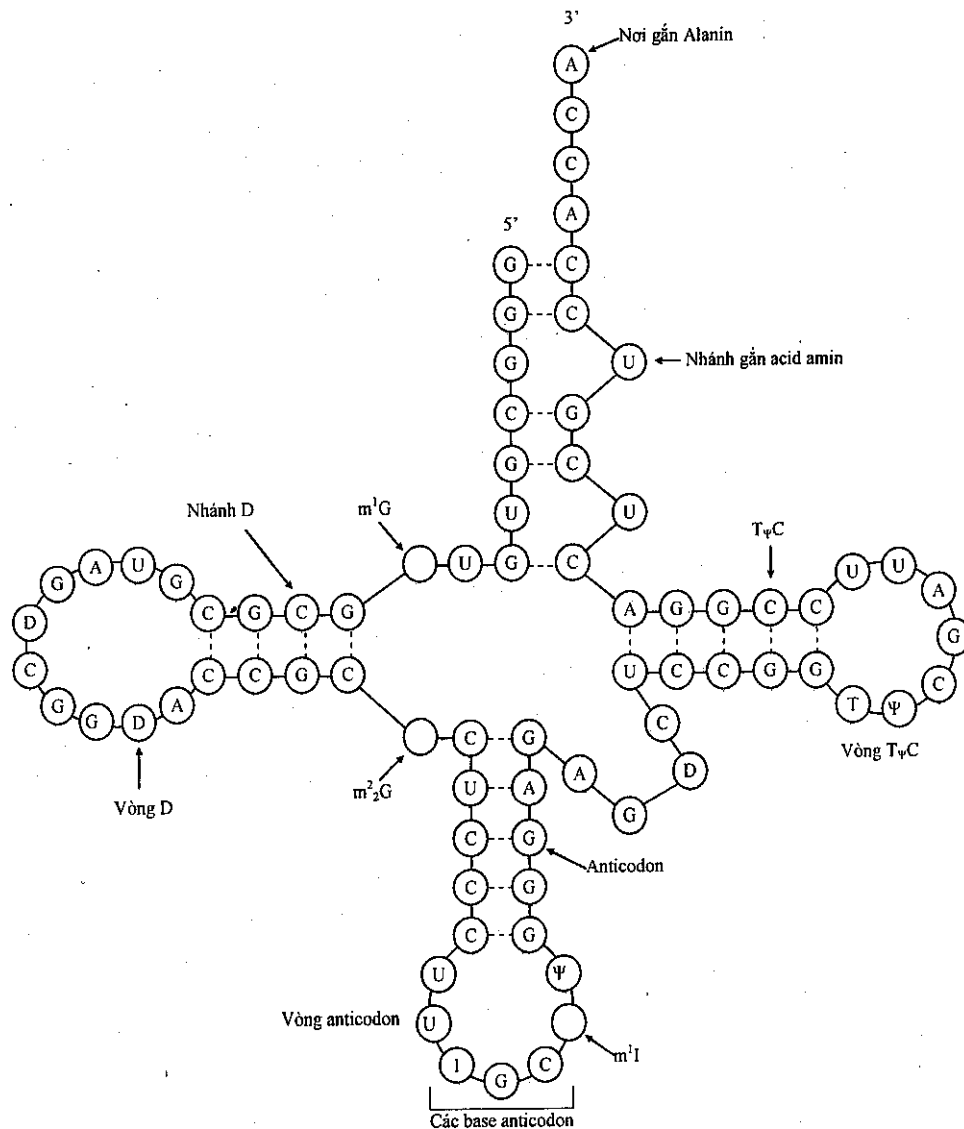
Điểm đặc biệt ở tARN là chứa các base bị biến đổi như: I = Inosin, T = Ribothymidin, Ψ = Pseudouridin, m^1G = Methylguanosin, m^2G = Dimethylguanosin, m^1I = Methylinosin, D = Dihydrouridin (Hình 3.6). Các base biến đổi được tạo thành từ 4 base cơ bản bị thay đổi công thức hóa học sau quá trình phiên mã, những thay đổi này do các enzym tARN-modifying xúc tác. Ở tARN, số base biến đổi chiếm khoảng 10%.



Hình 3.6. Các ribonucleotid bị biến đổi trong ARN

Ít nhất, mỗi loại tARN đặc hiệu cho một loại amino acid. Ví dụ: Arginine chỉ gắn đặc hiệu với ARN vận chuyển arginine - viết tắt là tARN^{Arg}. Tuy nhiên, tất cả các tARN có một số đặc tính cấu trúc chung: chiều dài khoảng 73 - 93 nucleotid, cấu trúc gồm một mạch cuộn lại như hình lá chẻ ba (Hình 3.7). Đầu mút 3' có trình tự kết thúc CCA. Amino acid luôn luôn gắn vào đầu CCA ở vị trí OH 3'.

Các enzym đặc hiệu là aminoacyl tARN synthetase gắn acid amin với tARN tương ứng nhờ năng lượng ATP tạo ra aminoacyl-tARN. Phức hợp này đến ribosom gắn với mRNA. tARN gắn với mRNA bằng sự bắt cặp bổ sung nhờ đối mã.



Hình 3.7. Cấu trúc alanin – tARN

Cấu trúc hình lá chẻ ba của tARN – Alanin của nấm men cho thấy trình tự nucleotid hoàn chỉnh. Ở thú, một vài tARN thiếu vòng xoắn dihydrouracil (vòng D).

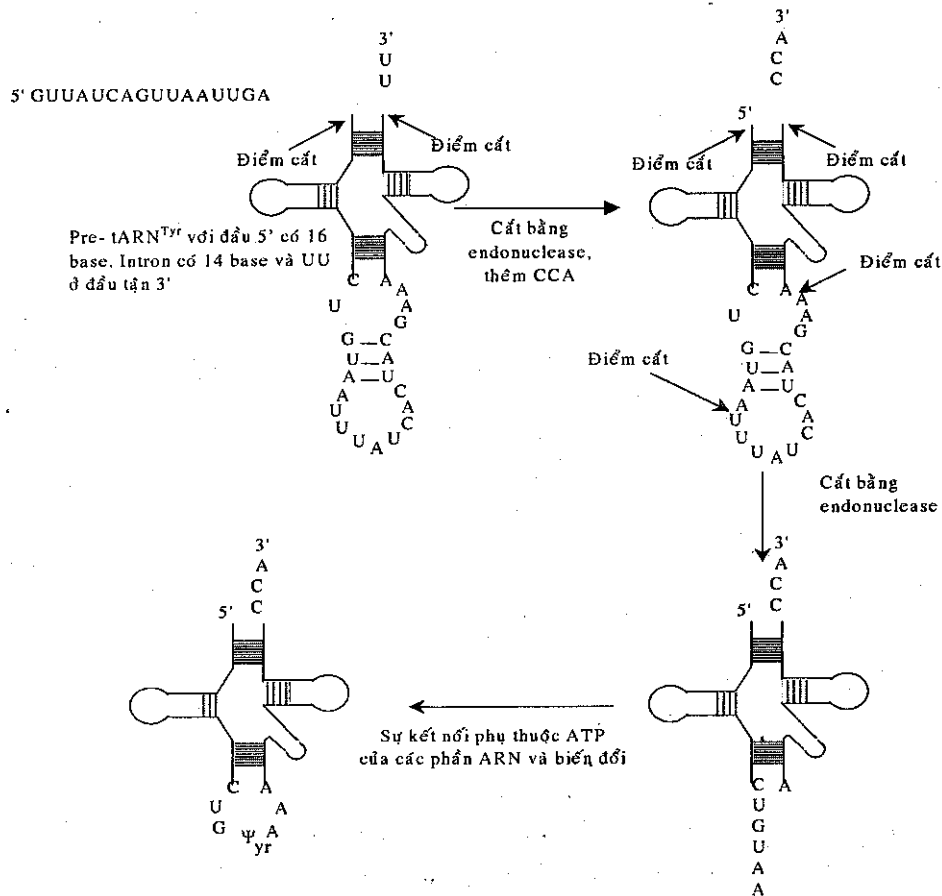
3.2.2.2. Biến đổi tARN

Ở tế bào nhân nguyên thủy và tế bào nhân thật, các tARN đều được tách ra từ những phân tử ban đầu dài hơn nhờ một số phản ứng xử lý. ở *E. coli*, các gen của tARN thường được phát hiện trong các operon hỗn hợp, cùng lúc mã hoá cho 4 tARN (tARN^{Tyr}, tARN^{Thr}, tARN^{Gly}, tARN^{Thr}) và mARN cho sự tổng hợp protein – yếu tố nối dài Tu.

Quá trình cắt nối thủy giải đầu tiên của bản sao sơ cấp được xúc tác bởi RNase, tạo các đầu tận 5' của phân tử tARN.

RNse P là một enzym, gồm 1 ARN có 375 nucleotid và một protein có trọng lượng phân tử 20 000. Trong các điều kiện sinh lý, cả 2 thành phần protein và ARN đều cần những yếu tố để hoạt động, nhưng khi nồng độ Mg^{++} cao thì chỉ có thành phần ARN có hoạt tính xúc tác. Theo sau hoạt động của RNse P, một số phản ứng trong nhân và biến đổi base sẽ tạo nên tARN hoàn chỉnh.

Ở *E. coli*, tARN có trình tự cuối là CCA – OH, phiên mã từ các base bổ sung có trong ADN của nhiễm sắc thể. Trong khi đó ở tế bào nhân thật những base này được thêm vào sau khi phiên mã (Hình 3.8).



Hình 3.8. Quá trình biến đổi của tARN^{Tyr} ở nấm men

3.2.2.3. Các phản ứng trong quá trình xử lý tARN

Một số phản ứng chủ yếu của ARN vận chuyển trong quá trình sinh tổng hợp protein:

- Aminoacyl hoá
- Formyl hoá tARN mở đầu (tế bào nhân nguyên thủy)
- Gắn những yếu tố nối dài

- Gắn ribosom
- Nhận diện codon - anticodon.

tARN có nhiều điểm nhận diện enzym. Một số điểm nhận diện đặc thù cho từng tARN riêng biệt hay cho một số loại đồng đẳng (như aminoacyl hoá, formyl hoá, nhận diện codon - anticodon). Trong khi đó, có những điểm nhận diện đặc hiệu cho tất cả tARN (đó là, sự tương tác với những yếu tố nối dài I, II và sự gắn kết ribosom).

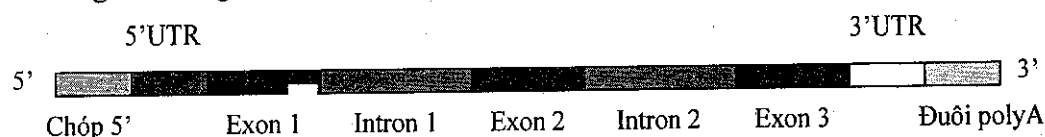
3.2.3. ARN thông tin

mARN mang thông tin qui định trình tự protein đến ribosom, nó được tổng hợp khi enzym ARN polymerase tiếp xúc với trình tự promoter của gen mã hóa cho protein trên phân tử ADN.

3.2.3.1. Cấu trúc

Phân tử mARN gồm 3 phần (Hình 3.9):

1. Vùng 5' không mã hóa (5' UTR - untranslated region).
2. Vùng mã hóa chứa trình tự mã hóa cho protein (ở tế bào nhân thật vùng mã hóa có các trình tự không mã hóa (intron) nằm xen kẽ với các trình tự mã hóa (exon)). Vùng này bắt đầu bằng codon khởi đầu và tận cùng bằng codon kết thúc.
3. Vùng 3' không mã hóa (3' UTR).

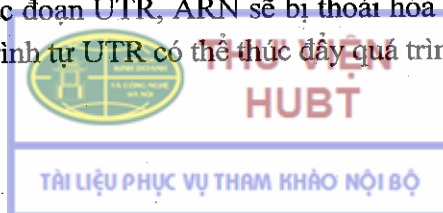


Hình 3.9. Cấu trúc tiền mARN của tế bào nhân thật

ARN polymerase bắt đầu phiên mã từ một vị trí nằm trong trình tự trước vùng mã hóa cho phân tử protein, gọi là đoạn 5' không mã hóa. Do đó, mARN có đoạn đầu mang các tín hiệu cho ribosom nhận biết để gắn vào khi dịch mã.

Ở đầu 3', sau sau codon kết thúc có đoạn 3' không mã hóa mang trình tự của dấu kết thúc (terminater); ở tế bào nhân thật, đây là nơi gắn đuôi polyA.

Do vậy, mARN nguyên vẹn của tế bào vi khuẩn và tế bào nhân thật chứa trình tự nucleotid nhiều hơn số được dùng để mã hóa protein. Các trình tự này - 5'UTR và 3'UTR - được phiên mã từ sợi ADN. Tuy nhiên, do ái lực không ổn định với enzym RNase, đoạn 5' hay 3' có thể làm gia tăng hay giảm độ ổn định của phân tử ARN. Các đoạn UTR giúp phân tử ARN tồn tại lâu hơn trong tế bào trước khi bị thoái hóa, do vậy sẽ tổng hợp được nhiều protein hơn. Nếu không có các đoạn UTR, ARN sẽ bị thoái hóa nhanh hơn và lượng protein sản xuất sẽ ít hơn. Hơn nữa, trình tự UTR có thể thúc đẩy quá trình dịch mã hiệu quả hơn.



Các mARN của tế bào nhân nguyên thủy có thời gian bán hủy ngắn: trung bình 2 phút, của tế bào nhân thật từ 30 phút đến 24 giờ.

3.2.3.2. mARN ở tế bào nhân thật

Ở tế bào nhân thật, sau khi mARN được tổng hợp tại nhân, chúng phải trải qua 2 giai đoạn "hậu phiên mã":

1. Giai đoạn gắn chóp (capping): đầu tiên, nhóm triphosphat tại đầu 5' của phân tử ARN mới tổng hợp sẽ bị cắt nhờ enzym phosphohydrolase cắt liên kết γ - phosphodiester, loại bỏ α và β phosphat. Tiếp theo, enzym guanyltransferase gắn guanin và α phosphat vào β phosphat của đầu 5' bằng liên kết 5' - 5' triphosphat. Sau đó, vị trí N7 của guanin sẽ bị methyl hoá do enzym guanin - 7 - methyltransferase. Cuối cùng, enzym 2' - O - methyltransferase sẽ methyl hoá vị trí 2' của đường ribose.

Phản ứng methyl hoá chóp có thể xảy ra theo 1 trong 3 cách (Hình 3.10):

a) Chóp 0: nhóm methyl gắn vào vị trí G7 do enzym guanin - 7 - methyltransferase xúc tác. Phản ứng này xảy ra ở tất cả mARN tế bào nhân thật.

b) Chóp 1: nhóm methyl gắn vào vị trí 2' - OH đường ribose của nucleotid đầu tiên. Phản ứng do enzym 2' - O - methyltransferase xúc tác, xảy ra ở hầu hết mARN tế bào nhân thật.

c) Chóp 2: phản ứng methyl hoá tương tự xảy ra ở nucleotid thứ 2 với tỉ lệ 10 - 15%.

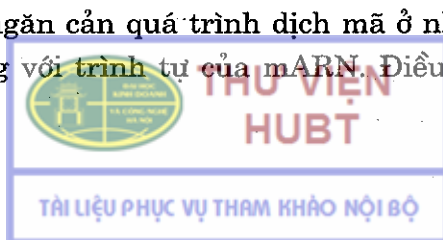
Ngoài ra, còn có quá trình methyl hoá nội phân tử xảy ra tại vị trí N6 của adenin, với tần suất 0,1%.

Giai đoạn này bảo vệ phân tử ARN không bị cắt liên kết phosphoester do các phân tử hydro lân cận và chóp 5' là dấu hiệu nhận biết vị trí gắn của mARN trên ribosom khi dịch mã.

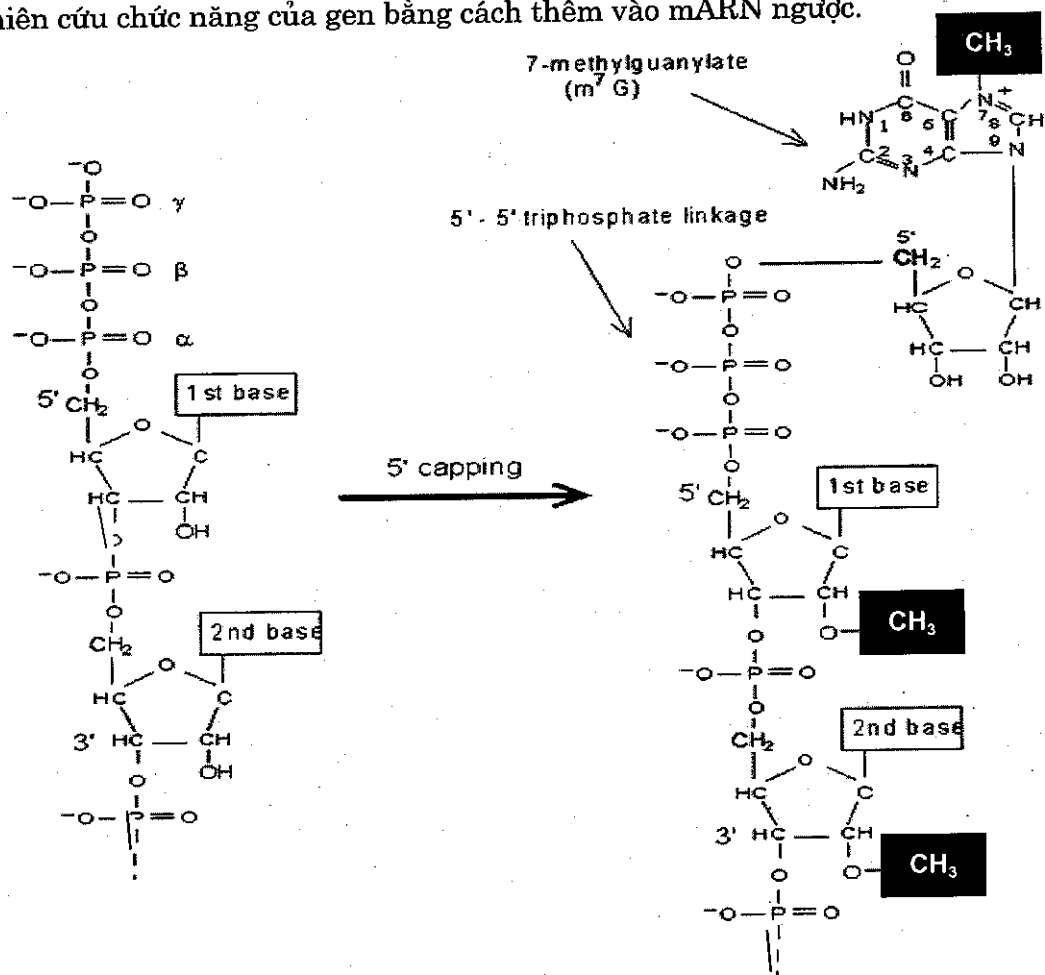
2. Giai đoạn gắn đuôi poly A vào đầu 3': ở động vật có vú khoảng 50 - 250, còn ở nấm men khoảng 100 nucleotid Adenin được gắn vào đầu 3'. Đuôi poly A gắn vào pre - mARN có trình tự nhận diện AAUAAA. Quá trình này do enzym poly A polymerase xúc tác. Chức năng đuôi poly A tương tự chóp 5' trong việc bảo vệ phân tử ARN và gắn với ribosom. Do vậy, thời gian bán hủy được kéo dài và lượng protein tổng hợp cũng nhiều hơn. Ngoài ra đầu 3' cũng có vai trò trong việc vận chuyển mARN từ nhân ra tế bào chất.

3.2.3.3. mARN ngược

mARN ngược có thể ngăn cản quá trình dịch mã ở nhiều tế bào nhân thật khi chúng có trình tự bổ sung với trình tự của mARN. Điều này có nghĩa là một gen



không biểu hiện được thành protein nếu trong tế bào có mARN ngược. Đây có thể là một cơ chế chống lại hiện tượng gen nhảy ngược hay virus; do cả 2 trường hợp đều dùng sợi đôi mARN làm trung gian. Trong sinh hoá, hiện tượng này được dùng để nghiên cứu chức năng của gen bằng cách thêm vào mARN ngược.



Hình 3.10. Gắn chóp ở mARN của tế bào nhân thật

1st base: base ở cacbon vị trí số 1; 2nd base: base ở cacbon vị trí số 2; 5'capping: gắn chóp ở đầu 5'; 5' - 5' triphosphaste linkage: liên kết triphosphat 5' - 5'.

3.2.4. Các ARN nhân nhỏ và các ARN tế bào chất nhỏ

Tế bào nhân thật có chứa nhiều phân tử ARN nhỏ khác ngoài tARN và 5S và 5,8S rARN. *In vivo*, những phân tử nhỏ này phức hợp với các protein đặc hiệu tạo nên các hạt ribonucleoprotein (RNP) được gọi là snRNP hoặc scRNP tương ứng cho nhân hay cho tế bào chất. Chúng dài khoảng 90 - 300 nucleotid.

Việc biến đổi tiền ARN thành ARN được thực hiện với sự tham gia của những phân tử nhỏ là các spliceosom, đó chính là các snRNP. Ngày nay tìm thấy được 7 loại snRNP từ U₁ đến U₇. U₁ và U₂ có vai trò trong việc sửa đổi hnARN thành

mARN hoàn chỉnh, còn U_3 có vai trò trong sửa đổi pre - rARN thành rARN tế bào chất. Các ARN có khả năng xúc tác được gọi là ribozyme. Người bị bệnh tự miễn hồng ban Lupus (systemic lupus erythematosus), do bệnh nhân tạo ra kháng thể gắn với một trong các protein $U_1 - U_7$ snRNP.

scRNP có vai trò biết rõ nhất là 7SL ARN, một thành phần của phức hợp nhận diện tín hiệu xuất SRP. SRP cấu tạo bởi một scARN có chứa 294 nucleotid và 6 polypeptid có Mr từ 9000 - 72000. Khi mARN của protein xuất bắt đầu được dịch mã, SRP tương tác đặc hiệu với đoạn peptid tín hiệu xuất của protein mới sinh và với ribosom, ngăn chặn sự dịch mã cho đến khi ribosom tiếp xúc với lưới nội chất, khi đó nó rời khỏi ribosom và tín hiệu xuất của protein mới sinh gắn vào bộ máy chuyển vị protein trên màng lưới nội chất và sự tổng hợp protein được tiếp tục, protein sinh ra được chuyển vào trong lưới nội chất.

3.2.5. Cắt nối ở ARN tế bào nhân thật

3.2.5.1. Khái niệm

Bộ gen tế bào nhân thật là bộ gen gián đoạn gồm các intron xen kẽ với các exon (Hình 3.9). Exon là trình tự biểu hiện, có ở mARN trưởng thành. Đầu tận cùng 5' và 3' có thể không dịch mã thành trình tự protein, do vậy trình tự exon không hoàn toàn là trình tự mã hóa. Intron là trình tự không chứa mã và không tham gia quá trình biểu hiện, thường có kích thước lớn hơn các exon. Intron có rất ít ở tế bào nhân thật bậc thấp, nằm và hoàn toàn không có ở vi khuẩn trừ một vài ngoại lệ. Bản phiên mã nguyên thủy pre-ARN chứa cả intron lẫn exon và không dùng trực tiếp cho quá trình dịch mã được. Quá trình loại bỏ các intron và nối các exon lại thành bản ARN trưởng thành gọi là quá trình cắt nối ARN.

3.2.5.2. Cắt nối ARN trong nhân tế bào nhân thật

Quá trình cắt nối xảy ra ở các phức hợp giữa bản sao nguyên thủy pre-ARN hay hnARN với protein nhân gọi là hnRNP. Quá trình này diễn ra trong nhân tế bào trước khi mARN di chuyển vào tế bào chất để tham gia dịch mã.

Các yếu tố cần cho sự cắt nối:

Có vài trình tự chung rất ngắn của intron cần thiết cho sự cắt nối, trong khi phần còn lại không có liên quan, cụ thể:

- Nguyên tắc GU-AG (xác định các đầu của intron):

+ GU luôn tận cùng đầu 5' của intron: vị trí cho.

+ AG luôn tận cùng đầu 3' của intron: vị trí nhận.

Bất kỳ vị trí cho nào cũng có thể gắn với vị trí nhận.

- Vị trí nhánh UACUAAC ở năm men hay một trình tự chung ít bảo tồn hơn ở các intron của động vật có vú cũng là yếu tố cần thiết.

Nguyên tắc GU-AG có ở >98% intron của bộ gen người. Ngoài ra, có các kiểu intron khác được xác định bởi kiểu GC-AG, chiếm khoảng 1% và AU-AC chiếm 0,1%.

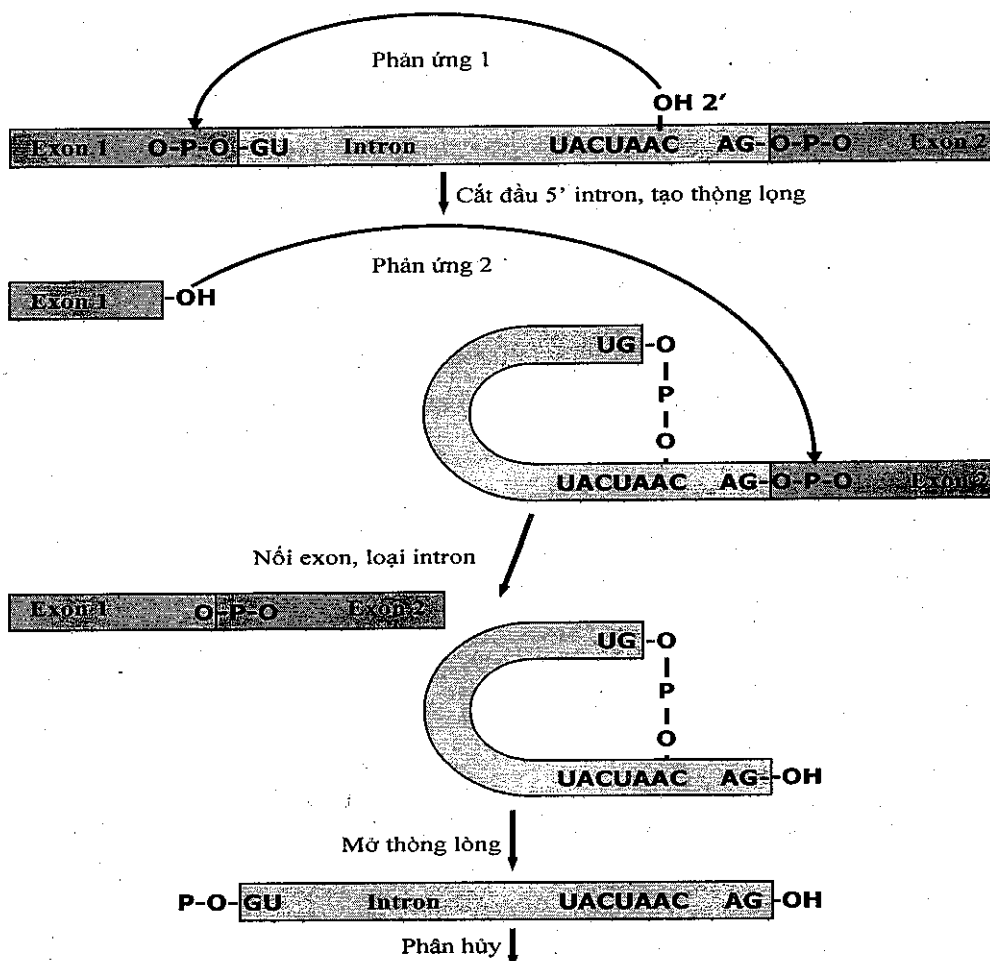
Cơ chế cắt nối thông thường: gồm 2 phản ứng chuyển nhóm ester (Hình 3.11):

– Phản ứng 1: Phản ứng ở đầu 5' của intron, liên quan đến sự hình thành một thòng lọng nối giữa đầu GU của intron qua liên kết 5'-2' với Adenin ở thứ 6 của vị trí nhánh.

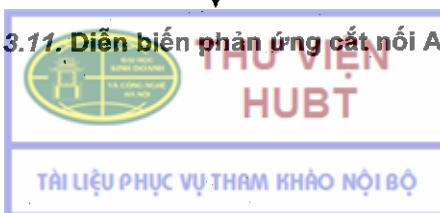
– Phản ứng 2: Đầu 3'-OH của exon lúc này sẽ tấn công vị trí cắt nối 3' (vị trí nhận), sao cho các exon được nối lại và intron được phóng thích dưới dạng thòng lọng.

Thòng lọng sau đó được mở và intron bị phân hủy.

Một số bước phản ứng cần năng lượng của ATP, hầu hết để thúc đẩy các biến đổi về cấu dạng trong ARN và/hay protein. Các vị trí cho (5') và nhận (3') hầu như tương đương, do đó sự hình thành thòng lọng chịu trách nhiệm trong việc lựa chọn vị trí cắt nối 3'. Các kiểu cắt nối khác nhau, tức sự tổ hợp khác nhau của các exon, đối với cùng một phân tử ARN là do các yếu tố protein kích thích việc sử dụng vị trí mới hay ngăn cản việc sử dụng vị trí mặc nhiên.



Hình 3.11. Diễn biến phản ứng cắt nối ARN



Spliceosom (thể cắt nối):

Spliceosom là một cấu trúc phức tạp gồm protein và ARN với khối lượng phân tử lên đến 12 MDa. Nó được lắp ghép từ 5 snARN trong phức hợp với khoảng 40 protein tạo thành 5 snRNP, bên cạnh đó nó cũng chứa khoảng 30 protein khác và 70 yếu tố cắt nối. Các snRNP gồm có U1, U2, U5 và U4/U6 chứa snARN U1, U2, U4, U5, U6. Cấu trúc này ổn định ở các tế bào nhân thật.

Vai trò của các snRNP và sự lắp ghép spliceosom:

– snRNP U1 khởi động quá trình cắt nối bằng cách gắn vào vị trí cho (5') nhờ phản ứng bắt cặp ARN-ARN.

– Phức E chứa snRNP U1 gắn ở đầu 5' của intron, protein U2AF (AF: yếu tố trợ giúp, auxiliary factor) gắn vào rãnh pyrimidin giữa vị trí nhánh và vị trí nhận (3') và các protein SR (yếu tố điều hòa) nối snRNP U1 với U2AF, kiểu gắn khởi đầu này được gọi là “nhận diện intron”. Mặt khác phức E có thể hình thành giữa U2AF ở rãnh pyrimidin và snRNP U1 ở vị trí cắt nối 5' của intron kế tiếp (xuôi dòng), kiểu này gọi là “nhận diện exon”.

– Phức E được chuyển thành phức A khi snRNP U2 gắn vào vị trí nhánh. snRNP U1 tương tác với snRNP U2, giúp vị trí cho và vị trí nhận tiến lại gần nhau.

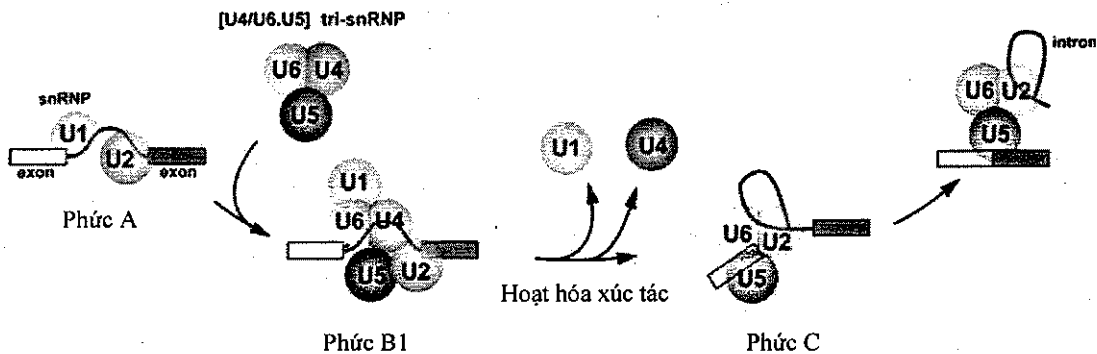
– Nếu phức E được tạo thành theo kiểu “nhận diện exon”, vị trí cắt nối 5' của intron xuôi dòng sẽ bị thay bằng vị trí thích hợp ngược dòng khi phức E được chuyển thành phức A.

– Các vị trí nhận yếu có thể cần có thêm trình tự tăng cường cắt nối (splicing enhancer) nằm ở exon xuôi dòng để gắn các protein SR.

– Khi snRNP U5 gắn với vị trí cho của intron và U4/U6 gắn với U2 (U4 gắn với U6 bằng phản ứng bắt cặp ARN-ARN nhằm giữ U6 không bắt cặp với U2) phức A chuyển thành spliceosom B1, chứa tất cả các thành phần cần thiết cho sự cắt nối.

– Việc phóng thích snRNP U1 cho phép snRNA U6 tương tác với vị trí cắt nối 5', và chuyển spliceosom B1 thành spliceosom B2.

– Khi snRNP U4 tách khỏi U6, snRNA U6 có thể bắt cặp với snRNA U2 để tạo trung tâm xúc tác (phức C).



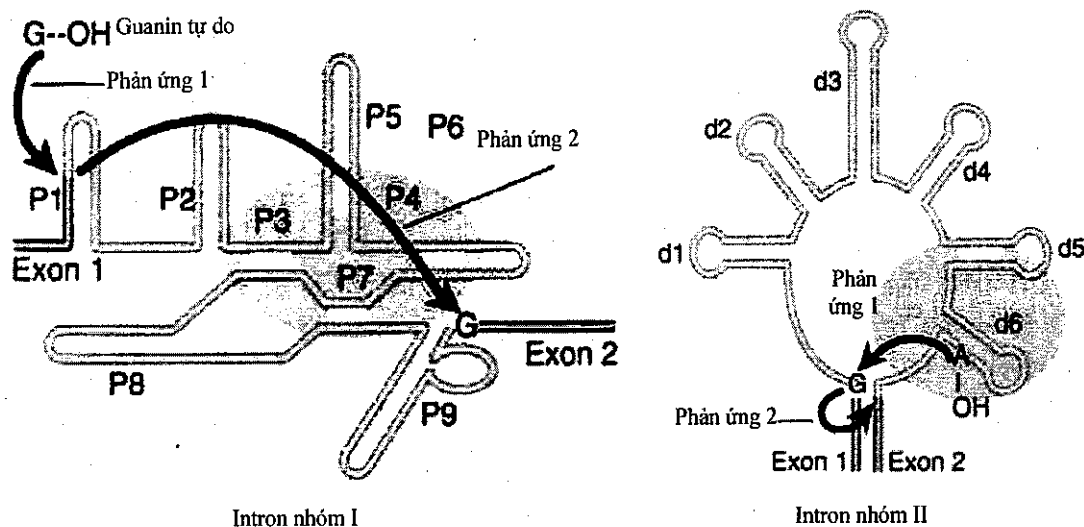
Hình 3.12. Tiến trình lắp ghép và hoạt hóa spliceosom

Kiểu spliceosom dựa trên hoạt tính xúc tác của U2 được gọi là spliceosom U2, chiếm đa số trường hợp cắt nối ở bộ gen người. Tuy nhiên, một số trường hợp cắt nối sử dụng một spliceosom với trung tâm xúc tác dựa trên U12 snRNP, gọi là spliceosom U12. Các intron được cắt nối bằng spliceosom này được gọi là các intron phụ thuộc U12. Spliceosom này có khác biệt về cấu tạo, bao gồm U11 và U12 (có vai trò tương tự U1 và U2), một biến thể của U5, U4_{atac} và U6_{atac}, nhưng nói chung quá trình cắt nối cũng diễn biến tương tự như spliceosom U2.

3.2.5.3. Cắt nối các ARN khác

Tự cắt nối:

Ngoài kiểu cắt nối các intron của ARN trong nhân tế bào nhân thật được xác định theo nguyên tắc GU-AG trên, các intron nhóm I và II được tìm thấy ở vi khuẩn và các bào quan có khả năng tự cắt nối, nghĩa là hoạt tính xúc tác được quy định sẵn trong trình tự của chính intron, do đó không cần đến các snRNP. Intron nhóm I và II sẽ gấp lại (có thể có sự hỗ trợ của protein hay không) tạo cấu trúc bậc hai có hoạt tính xúc tác (ribozym) (Hình 3.13) và nhìn chung phản ứng cắt nối vẫn diễn ra theo hai bước chuyển nhóm ester như trên. Ở intron nhóm I, nhóm OH ái nhân của phản ứng chuyển nhóm ester thứ nhất được cung cấp bởi Guanin tự do. Intron nhóm II cũng sử dụng thông lọng làm trung gian như intron nhân, nhưng có thể thực hiện phản ứng như là đặc tính tự xúc tác của ARN. Các intron này tuân theo nguyên tắc GT-AG, nhưng tạo thành một cấu trúc thứ cấp để giữ vị trí cắt nối đang phản ứng trong sự ghép nối thích hợp.

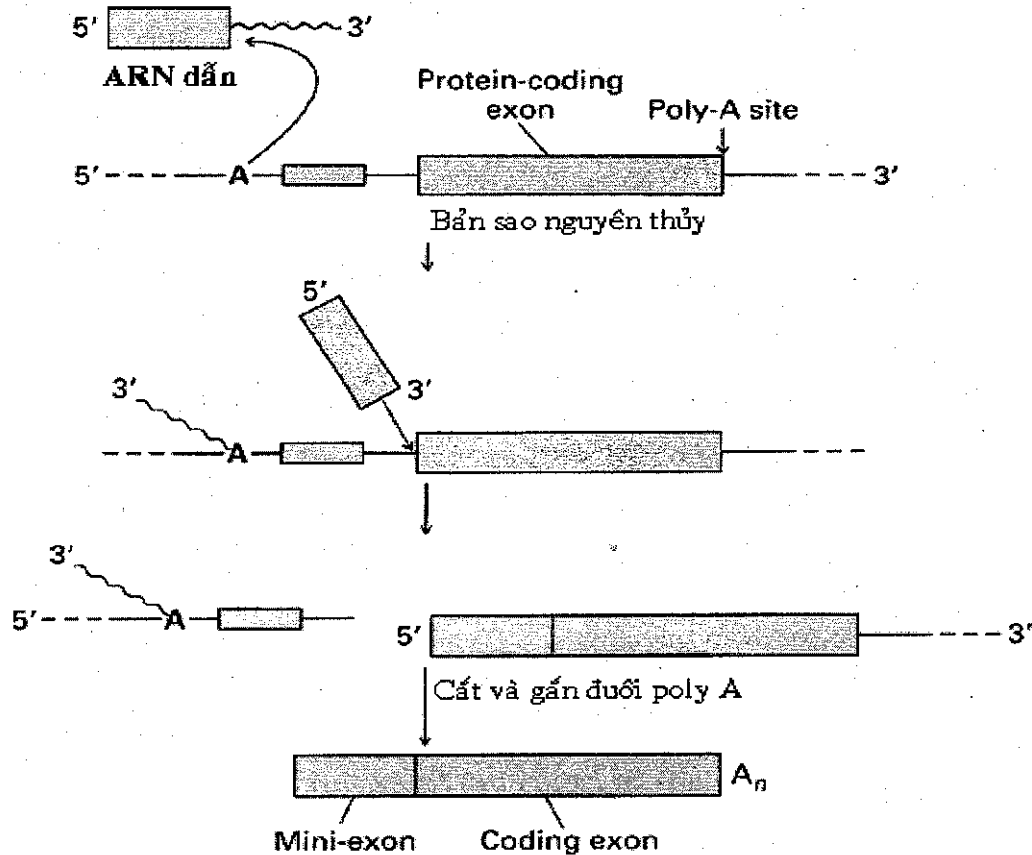


Hình 3.13. Phản ứng tự cắt nối

Cắt nối chéo:

Hầu hết mRNA được tạo ra từ một pre-mRNA duy nhất do các exon nối với nhau. Tuy nhiên, mRNA ở một số sinh vật là sản phẩm do nối các exon từ 2 hay nhiều pre-ARN

khác nhau, gọi là cắt nối chéo (trans-splicing). Cắt nối trans liên quan đến phản ứng giữa SL ARN nhỏ và pre- mARN. SL ARN tương tự với U1 snARN và có thể kết hợp vai trò cung cấp exon và chức năng U1.



Hình 3.14. Cắt nối chéo của ARN dẫn với exon mã hóa protein ở bản sao nguyên thủy polycistron loài Trypanosoma nhờ hai phản ứng chuyển vị ester

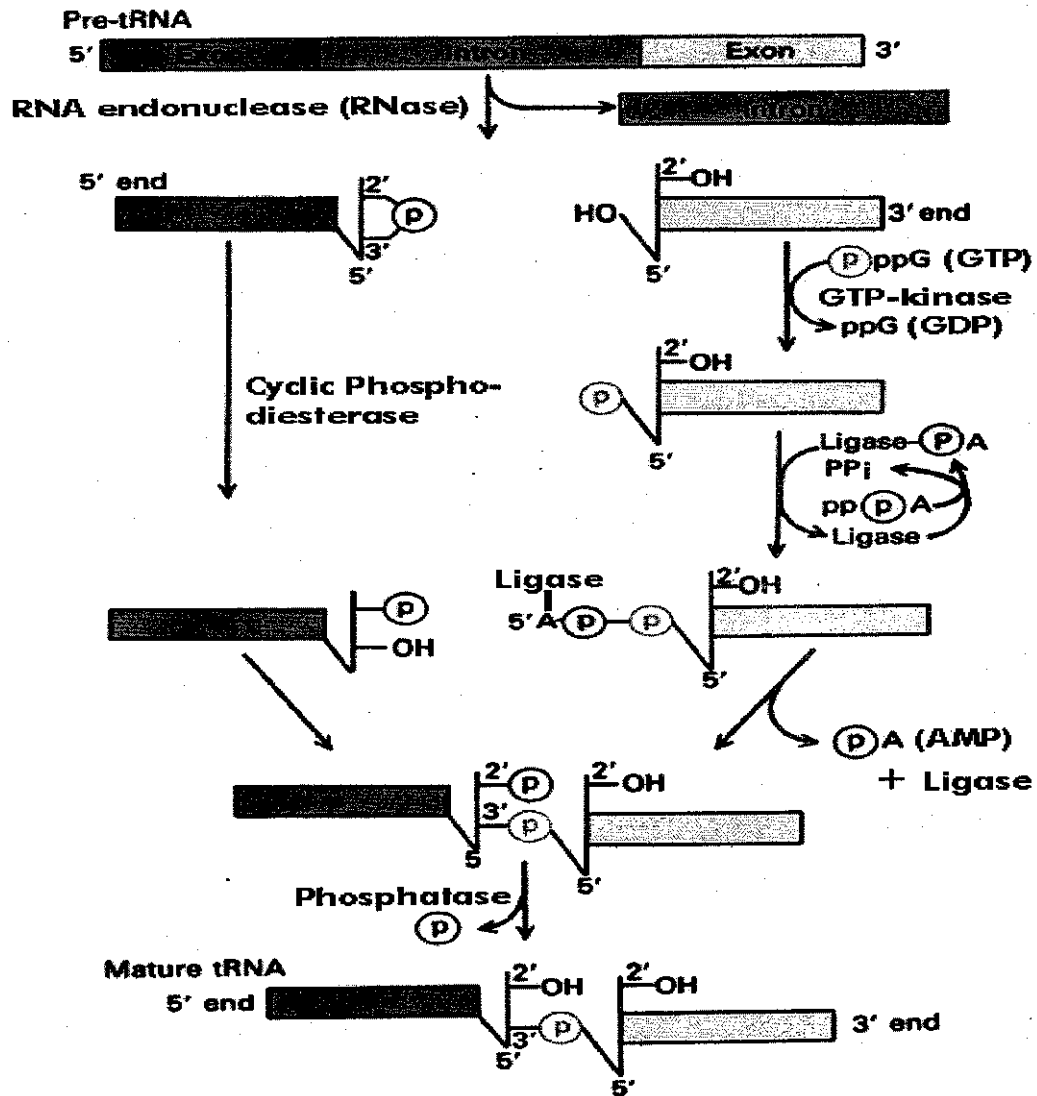
Đầu tiên, vị trí nhánh A ở sợi lên tại vùng giàu pyrimidin (màu xanh) liên kết với vùng 3' của ARN dẫn, giải phóng “tiểu exon” (màu vàng). Sau đó, đầu 3' của tiểu exon tự do liên kết với đầu 5' của exon mã hóa (màu hồng), cắt bỏ đoạn nhánh chứa vùng 3' của ARN dẫn và vùng 5' mã hóa của bản sao nguyên thủy. Các bước này tương ứng với phản ứng chuyển vị ester hóa I, II trong cắt nối pre-mARN ở tế bào nhân thật bậc cao.

Xử lý tARN:

Pre-tARN được cắt và loại bỏ intron để trở thành tARN trưởng thành.

Xử lý đầu 5' và 3':

- Xử lý đầu 5': enzym xúc tác là ribonuclease (ARNseP).
- Xử lý đầu 3': kết hợp endonuclease và exonuclease. Nếu đầu tận cùng 3' chưa có trình tự CCA thì tARN nucleotidyl transferase sẽ bổ sung trình tự này.



Hình 3.15. Cắt nối tARN ở nấm men.

Cắt nối tARN ở nấm men:

- ARN endonuclease (ARNse) cắt tiền tARN tại 2 đầu của intron, tạo nhóm 5'-OH và 2'3'P vòng.
- 2 nửa phân tử tARN cùng tồn tại nhờ liên kết hydro.
- Pi từ GTP gắn vào nhóm 5'-OH của đầu 3' exon nhờ enzym GTP kinase.
- 2 nửa phân tử tARN sau đó liên kết với nhau nhờ enzym ARN ligase.
- Vòng 2'3'P được mở nhờ enzym cyclic phosphodiesterase (là một phần của phản ứng kết nối), giải phóng 2'Pi.
- Phosphatase loại bỏ 2'Pi.

Tóm tắt:

ARN gồm một chuỗi đơn của các ribonucleotid nối với nhau bởi các liên kết phosphodiester. Trình tự nucleotid của ARN bổ sung với trình tự một mạch đơn (mạch khuôn) của ADN xoắn kép.

Có 3 loại ARN chính: mARN, rARN và tARN. Tất cả đều tham gia vào quá trình dịch mã, nhưng chỉ có mARN mang thông tin mã hoá cho cấu trúc cơ bản của phân tử protein. rARN có vai trò trong việc cung cấp nơi tổng hợp protein, ribosom. tARN làm nhiệm vụ vận chuyển acid amin đến ribosom, tương tác với aminoacyl – tARN synthetase đặc hiệu và kết hợp với ribosom.

Tất cả ARN đều được tạo ra từ sự sửa đổi và biến đổi của những tiền ARN lớn. Hơn nữa, tế bào nhân thật còn có một số phân tử ARN nhỏ với những chức năng khác nhau bao gồm thành phần của enzym ribonuclease, di chuyển intron, kết nối exon và một yếu tố kiểm soát việc tổng hợp của protein bài tiết.

CÂU HỎI

1. Tính chất nào **không** phải của tất cả ARN?
 - a) Mạch đơn polynucleotid
 - b) Đường pentose (5C) là ribose
 - c) Ngoài A, G, C thì uracil thay cho thymin
 - d) Được tổng hợp từ trong nhân
 - e) Có liên kết hydro giữa A = T
2. Cấu tạo từ 31 phân tử protein, 1 phân tử rARN 23S, 1 phân tử rARN 5S là tiểu đơn vị:
 - a) 50S
 - b) 30S
 - c) 60S
 - d) 40S
 - e) 70S
3. Cấu tạo từ 49 phân tử protein, 1 rARN 28S, 1 phân tử rARN 5.8S, 1 phân tử rARN 5S là tiểu đơn vị:
 - a) 60S
 - b) 40S
 - c) 50S
 - d) 30S
 - e) 70S
4. Tiểu đơn vị 40S của tế bào nhân thật cấu tạo từ:
 - a) 31 phân tử protein + 1 rARN 23S, 1 rARN 5S

- b) 21 phân tử protein + 1 rARN 16S
 c) 49 phân tử protein + 1 rARN 28S, rARN 5,8S, rARN 5S
 d) 33 phân tử protein + 1 rARN 18S
 e) 49 phân tử protein + 1 rARN 23S + 1 rARN 5S
5. Tiểu đơn vị 30S của tế bào nhân nguyên thủy cấu tạo từ:
 a) 31 phân tử protein + 1 rARN 23S, 1 rARN 5S
 b) 21 phân tử protein + 1 rARN 16S
 c) 49 phân tử protein + 1 rARN 28S, rARN 5.8S, rARN 5S
 d) 33 phân tử protein + 1 rARN 18S
 e) 49 phân tử protein + 1 rARN 23S + 1 rARN 5S
6. Quá trình methyl hoá nhờ ARN – methylase chỉ xảy ra ở:
 a) mRNA
 b) Pre – rARN
 c) tARN
 d) scARN
 e) snARN
7. Tính chất nào **không** đặc hiệu cho tARN?
 a) Chiều dài khoảng 73 – 93 nucleotid
 b) Mạch đơn cuộn hình lá chẻ ba
 c) Đầu mút 3' kết thúc CCA gắn acid amin
 d) Đầu mút 5' kết thúc G
 e) Một loại tARN có thể mang nhiều loại acid amin khác nhau
8. Phản ứng nào **không** phải của tARN trong quá trình sinh tổng hợp protein?
 a) Aminoacyl hoá
 b) Formyl hoá tARN mở đầu
 c) Gắn những yếu tố kết thúc
 d) Gắn ribosom
 e) Nhận diện codon – anticodon
9. Loại snRNP nào tham gia vào việc sửa đổi hnARN thành mRNA hoàn chỉnh?
 a) U1, U3 b) U4, U5 c) U6, U7 d) U1, U2 e) U1, U4
10. Ở tế bào nhân thật mRNA sau khi được phiên mã phải trải qua:
 a) Gắn cap
 b) Gắn đuôi polyA
 c) Cắt nối để loại intron
 d) a, b
 e) a, b và c.

Bài 4

SỰ PHIÊN MÃ VÀ MÃ DI TRUYỀN

MỤC TIÊU

- Trình bày được quá trình phiên mã ở tế bào nhân nguyên thủy và tế bào nhân thật.
- Mô tả được sự phiên mã có điều hoà.
- Trình bày được sự phiên mã ngược, nêu được ý nghĩa của nó trong chu trình sống của Retrovirus.
- Nêu được bản chất của mã di truyền và cách nhận biết.
- Phân biệt được các loại ARN polymerase về cấu trúc và chức năng.

4.1. MỞ ĐẦU

ADN trong nhiễm sắc thể đóng vai trò lưu trữ thông tin di truyền của tế bào. Nó quy định trình tự acid amin của hàng ngàn protein khác nhau. Tính đồng nhất và số lượng các protein này quyết định tính chất sinh hoá căn bản của tế bào. ADN không trực tiếp sắp xếp các acid amin cho sự polymer hoá thành protein mà là mARN, là bản sao bổ sung với một mạch polynucleotid của gen. Quá trình tạo ra các ARN được gọi là sự phiên mã. Tín hiệu được mã hoá trong ADN hướng dẫn enzym thực hiện phiên mã, nơi mở đầu, nơi kết thúc.

Sự điều hoà ở mức độ phiên mã là sự kiểm soát quan trọng hàng đầu trong biểu hiện gen. ARN được phiên mã từ gen phụ thuộc vào một vùng nhất định được gọi là promoter (vùng khởi động). Bên cạnh vùng khởi động, còn có một vài trình tự khác trên ADN cũng có vai trò quan trọng trong việc điều hoà biểu hiện gen. Vấn đề này được nghiên cứu kỹ và biểu hiện rõ ở tế bào tế bào nhân nguyên thủy, còn ở tế bào nhân thật chỉ mới được nghiên cứu từ khi có kỹ thuật ADN tái tổ hợp nhưng cũng thấy được một số khác biệt với tế bào nhân nguyên thủy.

Sự nhận diện các acid amin để liên kết chúng với nhau thành chuỗi polypeptid được xác định bởi các trình tự ribonucleotid trong mARN. Đó là mARN được dịch mã từ "ngôn ngữ" nucleotid của ADN. Mã di truyền mô tả mối quan hệ giữa 2 ngôn ngữ này.

4.2. NGUYÊN TẮC CHUNG

Quá trình truyền thông tin di truyền từ ADN sang ARN được gọi là sự phiên mã. Tham gia vào quá trình phiên mã có hệ thống enzym ARN polymerase, đúng ra, tên chính xác là ARN polymerase phụ thuộc ADN. ADN thể hiện khả năng dị xúc tác tức là làm khuôn để tổng hợp nên một phân tử khác.

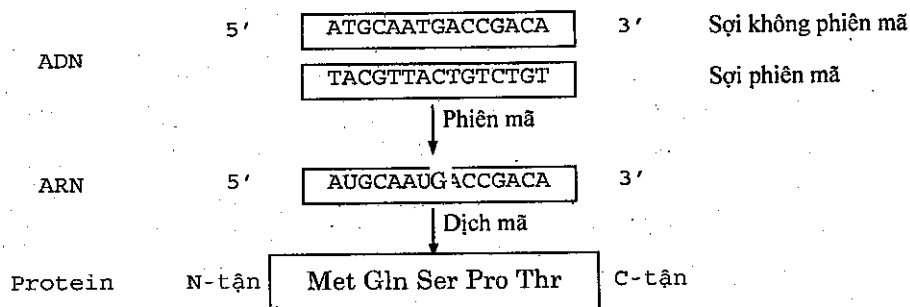
Sự phiên mã thực hiện theo các nguyên tắc:

– Chỉ 1 trong 2 mạch của phân tử ADN dùng làm khuôn để tổng hợp ARN (Hình 4.1)

– ARN polymerase bám vào ADN làm tách mạch và di chuyển theo hướng $3' \rightarrow 5'$ trên ADN để cho "ARN" được tổng hợp theo hướng $5' \rightarrow 3'$.

– ARN polymerase gắn vào ADN gây ra hiện tượng "chảy" tại chỗ bên trong chuỗi xoắn kép ADN. Kết quả làm đứt các liên kết hydro giữa các đôi base bổ sung. Cơ chất cho ARN polymerase là các ribonucleotid 5' – triphosphat ATP, GTP, CTP và UTP. Một ion kim loại hoá trị 2 hoặc là Mg^{2+} hoặc là Mn^{2+} đóng vai trò co – factor.

Cấu trúc của phân tử ARN thường bắt đầu bằng 1 nucleotid purin (ATP hay GTP) bổ sung cho nucleotid ở điểm bắt đầu phiên mã của mạch khuôn mẫu ADN. NTP thứ 2 được đưa vào khuôn mẫu ADN (lại theo nguyên tắc bổ sung cặp base) và liên kết 3' – 5' phosphodiester được hình thành giữa nhóm 3' – OH của NTP đầu tiên và nhóm phosphat (α) của NTP thứ 2.



Hình 4.1. Quá trình chuyển thông tin từ ADN đến protein

Phản ứng xảy ra theo phía thủy phân pyrophosphat vừa được giải phóng. Bằng cách này ARN được tiếp tục kéo dài và một tổ hợp lại ADN-ARN được hình thành. Chuỗi ARN

kéo dài theo hướng 5'-3' và ARN polymerase cùng với vùng sợi kép ADN duỗi xoắn (còn gọi là bong bóng phiên mã) chuyển động dọc theo sợi khuôn mẫu theo hướng 3' → 5'. Tổ hợp lai ADN-ARN trong bong bóng phiên mã có chiều dài khoảng 9 bp. Khi chuỗi ARN mới sinh đạt chiều dài 12 bp, đầu tận 5' của chuỗi ARN mới hình thành sẽ ló ra khỏi bề mặt enzym. Khi bong bóng phiên mã di chuyển, ADN sợi kép phía trước được tháo xoắn và 2 mạch đơn phía sau nó phục hồi trạng thái sợi kép ban đầu. ARN polymerase tiếp tục di chuyển dọc theo mạch ADN khuôn mẫu đến khi gặp điểm kết thúc, đó là tín hiệu được mã hóa trong một trình tự ADN quy định chấm dứt sự phiên mã. ARN polymerase và chuỗi ARN hoàn chỉnh được tách ra khỏi khuôn mẫu ADN.

ARN polymerase thực hiện sự liên kết cho 4 loại nucleotid triphosphat giống nhau, không phân biệt loại này, loại kia. Khác với sự sao chép ADN, mỗi không cần thiết cho sự phiên mã.

Trong đa số trường hợp, chỉ có một trong 2 mạch của ADN được phiên mã và sự phiên mã này được gọi là phiên mã bất đối xứng. Cũng có những trường hợp trong đó sự phiên mã xảy ra ở cả 2 mạch ADN, đó là sự phiên mã đối xứng, ví dụ, sự phiên mã ở ADN ty thể tế bào thú, trong đó cả 2 mạch của bộ gen vòng 16 kb đều được phiên mã ra ARN. Tuy nhiên, sau đó 1 trong 2 bản sao này bị thoái hoá.

Khác với ADN polymerase, ARN polymerase không có một hoạt tính sửa sai nào, vì những nucleotid nào kết hợp không chính xác thường được thay thế ngay bằng những nucleotid phù hợp. Hơn nữa, nếu có lỗi hiếm hoi nào xảy ra thì cũng không di truyền được vì ARN không phải là nơi lưu trữ thông tin di truyền.

Chuỗi ADN được phiên mã tạo ARN được gọi là 1 đơn vị phiên mã và sản phẩm là bản sao nguyên nguyên thuỷ. Bản sao nguyên thuỷ này được biến đổi trước khi đảm nhận vai trò chức năng của nó. Tất cả các loại ARN tế bào (mARN, tARN, rARN và các ARN nhỏ) đều được phiên mã từ ADN.

4.3. SỰ PHIÊN MÃ Ở TẾ BÀO NHÂN NGUYÊN THỦY

Sự phiên mã ở tế bào nhân nguyên thuỷ có những đặc điểm sau:

- Chỉ một loại ARN polymerase chịu trách nhiệm tổng hợp tất cả các loại ARN.
- mARN thường chứa thông tin nhiều gen nối tiếp nhau (polycistron mARN).
- Quá trình tổng hợp mARN được tiến hành khi ARN polymerase bám vào vùng khởi động (promoter). Khi gắn vào, polymerase không tổng hợp mARN ngay. Sự tổng hợp chỉ bắt đầu từ dấu xuất phát, thường là TAC nằm cách 7 - 8 base phía sau chỗ bám về phía đầu 5' của mạch khuôn.

- Ở phần lớn tế bào nhân nguyên thuỷ, quá trình tổng hợp tiếp tục đến khi đọc qua dấu kết thúc. Kết thúc phiên mã, ARN polymerase và mARN tách rời khỏi ADN.

- Quá trình phiên mã và dịch mã xảy ra đồng thời với nhau.

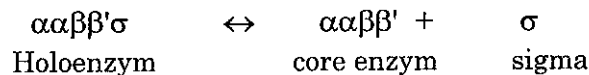
Hầu hết sự hiểu biết về quá trình phiên mã ở tế bào nhân nguyên thủy xuất phát từ sự nghiên cứu trên *E. coli* và thực khuẩn của nó. Cơ chế của sự phiên mã ở *E. coli* tương tự ở hầu hết tế bào nhân nguyên thủy, ngoại trừ *Archaeobacteria* (có intron trong ADN).

Nhiễm sắc thể đơn bội của *E. coli* là một vòng khép kín chứa 4×10^6 bp (cặp base). ADN này có cấu hình tháo xoắn cực đại (negatively supercoiled conformation). *In vivo*, nhiễm sắc thể là một cấu trúc lỏng lẻo gọi là nucleoide và được gắn vào màng tế bào chất. Bộ gen của *E. coli* chứa khoảng 3000 gen cấu trúc khác nhau, hầu hết có mặt trong một bản sao duy nhất, sắp xếp thành 100 đơn vị phiên mã.

4.3.1. ARN polymerase của *E. coli*

Là một protein phức hợp chứa tới 5 tiểu đơn vị gồm 4 loại và có trọng lượng phân tử khoảng 450 000 Da. Toàn bộ enzym được gọi là holoenzym (holozym) chứa các tiểu đơn vị $\alpha_2\beta\beta'\sigma$. Trong đó tiểu đơn vị β và β' chịu trách nhiệm xúc tác, tiểu đơn vị α liên quan đến việc lắp ghép enzym, nhận diện promoter và gắn các protein hoạt hóa, tiểu đơn vị σ (còn gọi là yếu tố σ) liên quan đến tính đặc hiệu với promoter.

Tiểu đơn vị σ (sigma) có thể tách ra thuận nghịch từ holoenzym và được phân tách bằng sắc ký trên cột phosphocellulose:



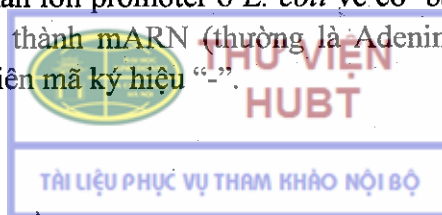
Enzym lõi cấu tạo bởi các tiểu đơn vị $\alpha_2\beta\beta'$ và hầu như giống nhau giữa các loài vi khuẩn. Lõi có khả năng xúc tác và kéo dài chuỗi ARN, nhưng không có khả năng khởi động sự phiên mã. Bình thường nó gắn với ADN tại một số trình tự gọi là "vị trí gắn lõi" với thời gian bán hủy lên đến 60 phút, nhưng không gây ra sự biến đổi nào trên ADN.

Trong mỗi tế bào có nhiều loại yếu tố σ khác nhau, do đó có nhiều holozym khác nhau. Khi yếu tố σ gắn vào enzym lõi, nó làm giảm ái lực của enzym đối với các vị trí gắn lõi hàng vạn lần và thời gian bán hủy chỉ còn < 1 giây. Tuy nhiên, ái lực với promoter lại tăng lên và thời gian bán hủy của phức holozym với promoter lên đến vài giờ. Sự gắn yếu tố σ cũng giúp enzym có khả năng khởi động sự phiên mã. Ái lực và tần suất khởi động phiên mã khác nhau tùy theo mỗi loại yếu tố σ và trình tự promoter cụ thể.

4.3.2. Promoter

Cấu trúc chung:

Promoter là trình tự khởi động sự phiên mã, mỗi đơn vị phiên mã đều có một trình tự promoter ở thượng nguồn. Phần lớn promoter ở *E. coli* về cơ bản đều giống nhau. Nếu ký hiệu base đầu tiên phiên mã thành mRNA (thường là Adenin) là +1, thì các base theo hướng ngược lại với chiều phiên mã ký hiệu "-".



Promoter gồm 2 trình tự. Đó là trình tự TATAAT, cách điểm xuất phát 6-8 base và có base trung tâm ở vị trí -10. Trình tự này còn được gọi là hộp TATA hay hộp Pribnow. Và một trình tự thứ 2 TTGACA có base trung tâm ở vị trí -35, do đó được gọi là vùng -35 (Hình 4.2), (Bảng 4.1). Hai trình tự này là nơi gắn khởi động phiên mã của ARN polymerase. Không phải promoter nào cũng có cấu tạo giống hệt các trình tự này, mà có một ít sai lệch nhỏ, do đó người ta gọi trình tự này là trình tự chung.

Nghiên cứu các trình tự liên quan đến khởi động phiên mã ở *E. coli* cho thấy tín hiệu khởi động bao gồm (Bảng 4.1):

- Vị trí bắt đầu phiên mã là một base purine (chiếm 90% trường hợp), nằm trong trung tâm của trình tự CAT.

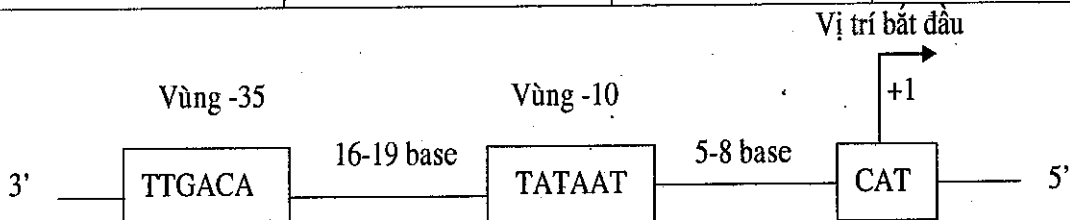
- Vùng -10, có trình tự chung là TATAAT, nằm cách vị trí bắt đầu 7 nucleotid về phía thượng nguồn (thay đổi trong khoảng 5-9 nucleotid). Xác suất xuất hiện của các nucleotid trong vùng -10 là: T₈₀A₉₅T₄₅A₆₀A₅₀A₉₆.

- Vùng -35, có trình tự chung là TTGACA, khoảng cách giữa vùng -10 và -35 là 17 base (thay đổi trong khoảng 16-19 nucleotid). Xác suất xuất hiện của các nucleotid trong vùng -35 là: T₈₂T₈₄G₇₈A₆₅C₅₄A₄₅.

- Trình tự và khoảng cách giữa chúng quyết định việc khởi đầu phiên mã.

Bảng 4.1. Các trình tự chung -10 và -35 của một số promoter của *E. coli*

Promoter	-35	-10	Điểm xuất phát
Trình tự chung (consensus)	TTGACA ← 17 bp →	TATAAT	← 7 bp → A
<i>lac P</i>	TTTACA ← 18 bp →	TATGTT	← 6 bp → A
<i>AraC</i>	CTGACA 17	TGTCAT	6 G
<i>AraBAD</i>	CTGACG 18	TACTGT	8 A
<i>Trp</i>	TTGACA 17	TTAACT	7 A



Hình 4.2. Cấu hình vùng -10 và -35 trên ADN

Chức năng quan trọng của hộp -10 và hộp -35 được phát hiện từ việc nghiên cứu ảnh hưởng của các đột biến thay thế base trong vùng này đối với quá trình phiên mã. Nhiều đột biến như vậy ức chế hay làm tăng sự phiên mã của gen. Trình tự của promoter có tính quyết định đối với cường độ hoặc tần suất khởi động phiên mã. Ví dụ, tế bào *E. coli* chứa khoảng 12 phân tử của *lac repressor* (chất ức chế *lac*-operon) và gen của nó được phiên

mã chỉ một lần mỗi 30 phút. Mặt khác, ARN ribosom, do đòi hỏi số lượng lớn, vì khoảng 10.000 ribosom phải được hình thành sau mỗi thế hệ tế bào (khoảng 20 phút), nên mỗi gen rADN được phiên mã khoảng 2 giây một lần. Sự khác biệt này do cường độ vùng khởi động.

Các nghiên cứu đột biến cho thấy vai trò của vùng -35 liên quan đến việc nhận diện khởi đầu của enzym đối với promoter để gắn vào tạo phức hợp đóng và vùng -10 liên quan đến việc tháo xoắn ADN để chuyển phức hợp đóng thành phức hợp mở.

Yếu tố sigma:

Tế bào có nhiều loại yếu tố sigma để nhận diện các họ promoter khác nhau. Những nghiên cứu trên sự đáp ứng shock nhiệt ở *E. coli* cho thấy rằng khả năng phiên mã các gen khác nhau có thể đạt được do sử dụng các yếu tố σ khác nhau. Người ta đã chứng minh sự đáp ứng shock nhiệt có liên quan đến việc tổng hợp những protein mới khi có sự gián đoạn nhiệt môi trường nuôi cấy vi khuẩn hay tế bào. Sự đáp ứng này biến thiên rất rộng ở tế bào nhân thật và tế bào nhân nguyên thủy. Những protein mới này, được gọi là protein shock nhiệt, được xem là để bảo vệ tế bào chống lại sự biến đổi nhiệt.

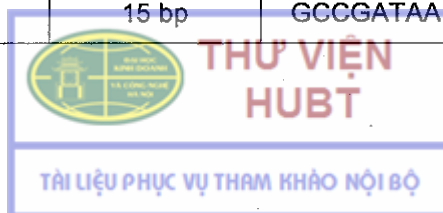
Ở *E. coli* có khoảng 17 protein shock nhiệt mà sự biểu hiện của chúng được kích hoạt bởi sản phẩm của gen *hptR*. Gen này đáp ứng với shock nhiệt và mã hóa cho chuỗi polypeptid có Mr 32.000, đóng vai trò như một yếu tố σ biến đổi. Yếu tố này (được gọi σ^{32}) kết hợp với phần lõi bình thường của ARN polymerase để tạo nên holoenzym $\alpha_2\beta\beta'\sigma^{32}$. Enzym này chỉ bám vào những promoter của gen shock nhiệt, chứ không bám vào promoter của các gen bình thường và có trình tự vùng -35 dài hơn với một số khác biệt nhỏ.

Vi khuẩn *Bacillus subtilis* cũng sử dụng các yếu tố σ khác nhau để biệt hóa từ dạng sinh dưỡng sang dạng bào tử.

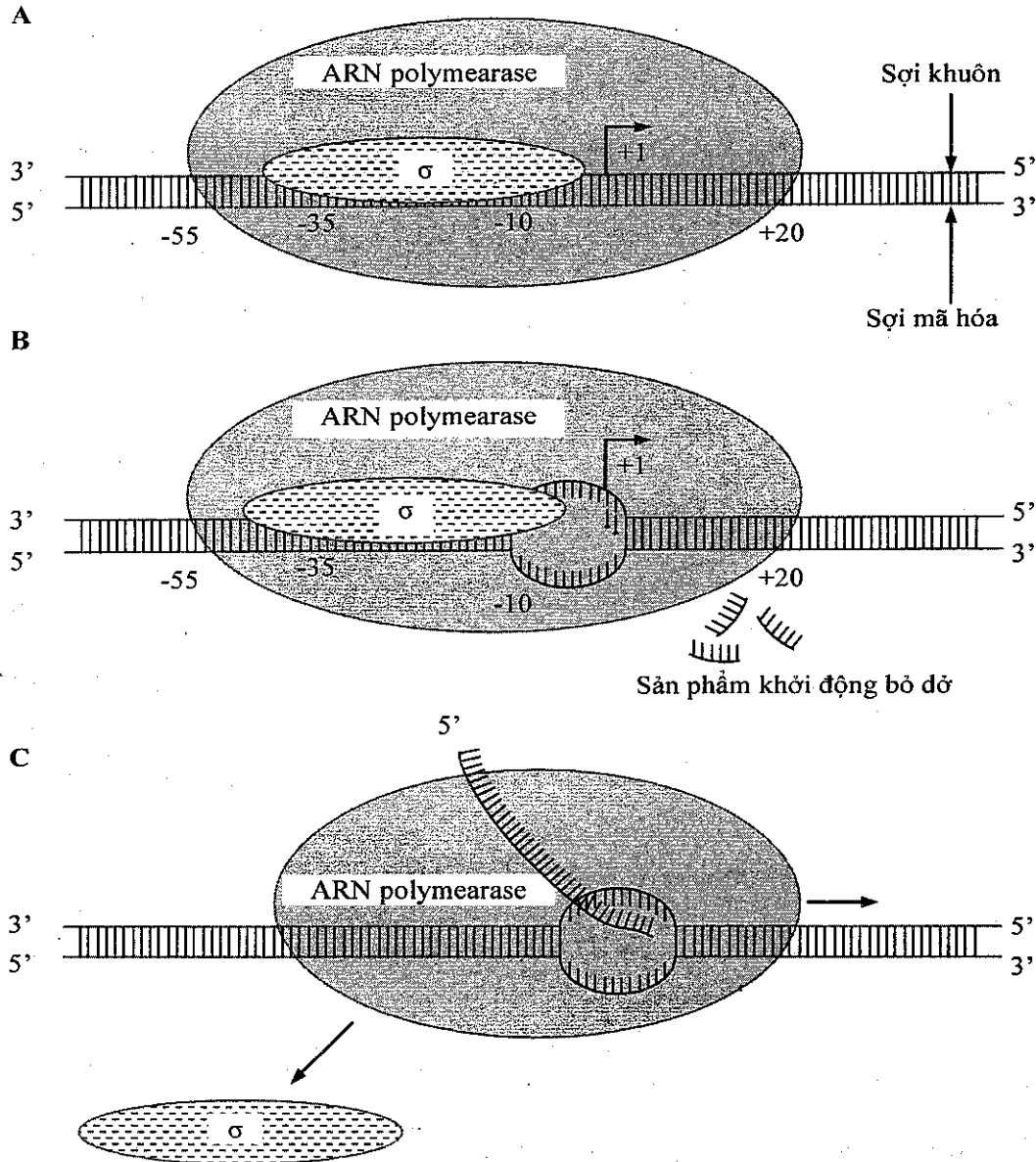
Như vậy, tế bào sử dụng các yếu tố σ khác nhau để phiên mã các họ gen khác nhau nhằm đáp ứng với các điều kiện môi trường hay giai đoạn tăng trưởng khác nhau. Các yếu tố σ có thể được hoạt hóa độc lập theo tín hiệu môi trường hay có liên quan với nhau theo lớp, tức yếu tố σ này cần để phiên mã yếu tố σ tiếp theo.

Bảng 4.2. Một số yếu tố σ của *E. coli*

Yếu tố	Gen	Vùng -35	Khoảng cách	Vùng -10	Chức năng
σ^{70}	<i>rpoD</i>	TTGACA	16-18 bp	TATAAT	Chung, cơ bản
σ^{32}	<i>rpoH</i>	CCCTTGAA	13-15 bp	CCCGATNT	Đáp ứng sốc nhiệt
σ^{54}	<i>rpoN</i>	CTGGNA	6 bp	TTGCA	Đáp ứng thiếu nitơ
σ^{28} (σ^F)	<i>filA</i>	CTAAA	15 bp	GCCGATAA	Tổng hợp tiêm mao



Nhiều ADN thực khuẩn thể mã hóa ARN polymerase riêng, chỉ nhận diện các promoter đặc trưng của chúng. Một số promoter như vậy được sử dụng trong việc tạo dòng plasmid nhằm sản xuất một số lượng lớn mRNA từ một gen quan tâm. Những bản mã sao như vậy có thể được sử dụng như những đoạn mồi hoặc có thể được tổng hợp *in vitro* thành một sản phẩm protein cần thiết.



Hình 4.3. Sự khởi đầu và nối dài chuỗi ARN bởi ARN polymerase của *E. coli*

(A) Holoenzym gắn vào promoter tạo thành phức hợp khởi động đóng. Các trình tự -10 và -35 đóng vai trò chủ yếu cho sự gắn này. (B) Polymerase làm chày tại chỗ chuỗi xoắn kép hình thành nên phức hợp khởi động mở. Triphosphat nucleotide đầu tiên được kết nối và hình thành cặp base tại điểm xuất phát +1. Ban đầu enzym chưa di chuyển và tổng hợp một số đoạn ARN

ngắn, rồi gỡ bỏ, gọi là khởi động từ bỏ. (C) Sự kéo dài chuỗi ARN bắt đầu khi enzym di chuyển và các NTP được polymer hóa vào đầu 3'-OH của đường ribose của chuỗi. Bong bóng phiên mã di chuyển dọc theo sợi khuôn ADN theo hướng 3'→5'. Khi đã kết dính được khoảng 12 nucleotide vào chuỗi thì tiểu đơn vị σ tách ra để lại phần lõi của enzym hoàn tất quá trình nối dài. Tiểu đơn vị sigma lại có thể liên kết với phần lõi tự do trong dịch bào tương để tham gia vào vòng khởi đầu mới.

4.3.3. Khởi đầu và nối dài chuỗi ARN

Việc gắn của holoenzym ARN polymerase vào promoter tạo ra phức hợp khởi động đóng trong đó ADN vẫn ở dạng xoắn kép (Hình 4.3). Phức hợp đóng có tính thuận nghịch. ARN polymerase gắn vào promoter trên một vùng khoảng 75-80 bp, trải từ vị trí +20 đến -55, trong đó các điểm tiếp xúc chủ yếu giữa promoter và enzym nằm ở các vùng -10 và -35. Trong holoenzym, yếu tố σ chịu trách nhiệm gắn trực tiếp với ADN tại các vùng -10 và -35. Yếu tố σ tự do không có khả năng gắn ADN, do vị trí gắn bị che bởi vùng N- tận của chính nó. Khi kết hợp với enzym lõi, yếu tố σ duỗi ra và bộc lộ các vị trí gắn ADN.

Phức hợp đóng nhanh chóng biến thành phức hợp mở do các liên kết hydro của mạch kép ADN bị enzym cắt trên một vùng 12-14 bp xung quanh vùng -10 và điểm xuất phát, tạo thành bong bóng phiên mã. Nucleotid triphosphat đầu tiên (thường là ATP hoặc GTP) gắn vào tiểu đơn vị β của polymerase, hướng đúng vào base bổ sung ở vị trí +1 của sợi khuôn mẫu. Một nucleotid triphosphat thứ 2 tiếp theo được kết nối và hình thành liên kết phosphodiester đầu tiên. Tuy nhiên, ARN polymerase không thực sự bắt đầu phiên mã ngay mà nó thực hiện một số chu kỳ “khởi động bỏ dở” (abortive initiation), nghĩa là kéo dài phân tử ARN mới tối đa đến 9 nucleotid rồi gỡ bỏ và tổng hợp lại từ đầu.

Nếu khởi động thành công, enzym sẽ chuyển cấu hình thành phức kéo dài và di chuyển khỏi promoter dọc theo ADN để tổng hợp sợi ARN mới. Thời điểm ARN polymerase di chuyển khỏi promoter xảy ra lúc sợi ARN mới dài hơn 10 bp. Lúc này yếu tố σ không còn cần thiết và có thể tách ra để tham gia khởi động tại một promoter khác. Như vậy, chuỗi mARN đang tổng hợp mang triphosphat ở đầu 5' và kéo dài theo hướng 5' → 3'. Về điểm này, phiên mã giống với sao chép ADN. Ở vi khuẩn, tốc độ phiên mã khoảng 40 nucleotid/giây.

Ở *E. coli*, hầu hết mARN có thời gian bán hủy được tính bằng phút. Chúng nhanh chóng bị thoái hóa bởi hệ thống enzym endonuclease và exonuclease. Thật ra, phần lớn mARN có trong tế bào chất ở bất kỳ thời điểm nào cũng đều là những chuỗi mới sinh. Chúng gắn vào phức hợp ADN-khuôn-ARN polymerase. Ribosom gắn vào vị trí gắn ribosom ở gần đầu 5' của mARN vừa rời phức hợp ADN-khuôn-ARN polymerase. Do đó, sự phiên mã và dịch mã ở tế bào nhân nguyên thủy được thực hiện gần như

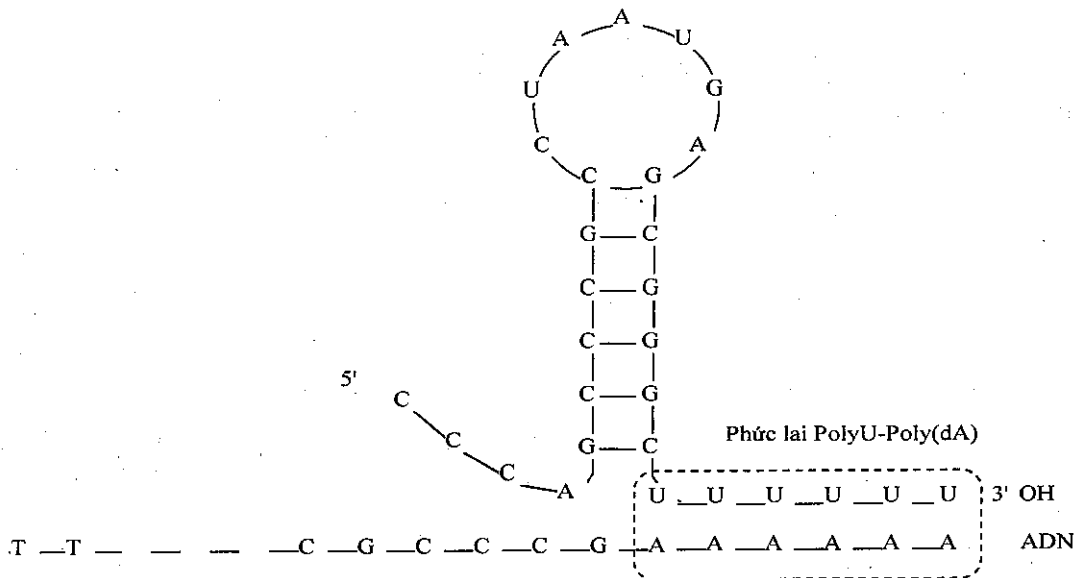
đồng thời. Chính điều này làm xuất hiện một dạng kiểm soát phiên mã gọi là phiên mã suy giảm.

4.3.4. Kết thúc quá trình tổng hợp ARN

Chuỗi ARN được nối dài cho đến khi gặp phải một trình tự chuyên biệt trong chuỗi ADN. Chuỗi này dùng để kết thúc quá trình tổng hợp ARN. *E. coli* có 2 cơ chế kết thúc: một cơ chế phụ thuộc vào protein phụ được gọi là ρ (yếu tố Rho) và một cơ chế khác không phụ thuộc vào yếu tố này.

4.3.4.1. Kết thúc độc lập với yếu tố Rho

Sự kết thúc phiên mã kiểu này chỉ phụ thuộc vào trình tự kết thúc (terminator) sẵn có trên ADN và được ARN polymerase phiên mã thành ARN, do đó được gọi là kết thúc nội tại, gặp ở khoảng một nửa số gen ở *E. coli*. Điểm kết thúc kiểu này đặc trưng bởi sự hiện diện của một trình tự giàu GC trên ADN, có các palindrom để tạo cấu trúc vòng hay kẹp tóc, theo sau bởi 5 hoặc 6 Adenin. ARN được phiên mã từ trình tự này có thể hình thành một nút vòng hay một cấu trúc kẹp tóc do các liên kết hydro nội phân tử giữa các base bổ sung (Hình 4.4). Cấu trúc bậc hai này tách ARN polymerase ra ngay khi nó phiên mã tới trình tự giàu A phía sau. Cả cấu trúc bậc hai lẫn liên kết yếu giữa polyU trên ARN với polyA trên ADN tạo điều kiện cho sự phiên mã kết thúc và ARN tách khỏi ADN.



Hình 4.4. Cấu trúc bậc 2 được giả thiết của bản sao ARN được phiên mã từ điểm kết thúc (terminator) của trp operon ở *E. coli*

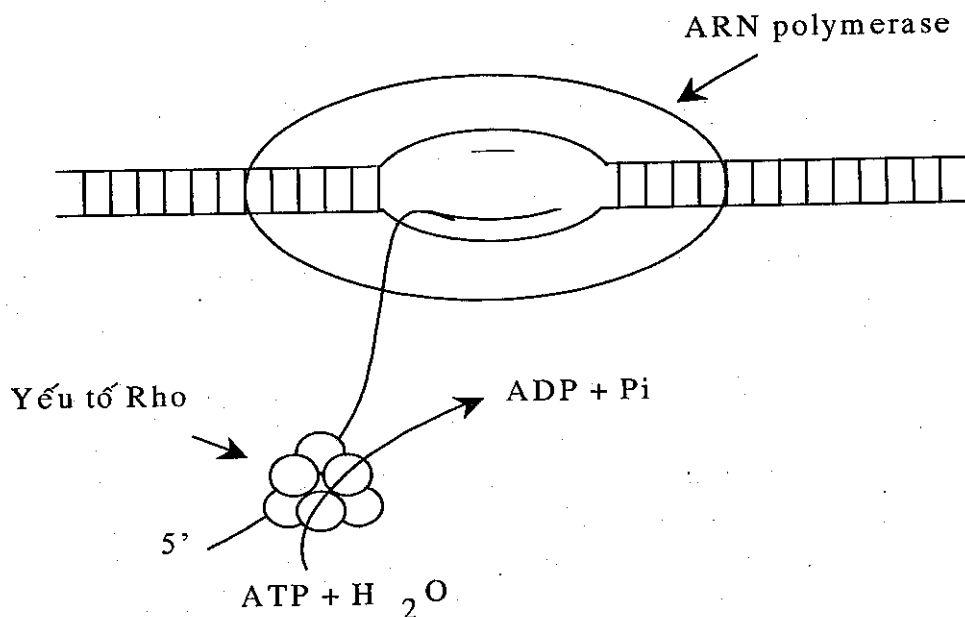
Nút vòng hay kẹp tóc và trình tự polyU tạo nên tín hiệu kết thúc

4.3.4.2. Kết thúc lệ thuộc yếu tố Rho

Sự kết thúc này cũng cần sự có mặt của nút và vòng ngay trước đầu kéo dài của ARN. Nhưng không có trình tự oligo (U). Yếu tố ρ là một protein gồm 6 tiểu đơn vị giống nhau

có Mr 46 000 và có ái lực cao với chuỗi ARN đơn. Yếu tố ρ gắn với ARN tại các vị trí *rut* (viết tắt của rho utilization), có đặc điểm là giàu C và nghèo G, nhưng không gắn vào các mARN đang được dịch mã bởi ribosom. Yếu tố ρ có hoạt tính helicase và sử dụng năng lượng thủy giải ATP để di chuyển trên ARN.

Khi gắn vào ARN, yếu tố ρ thủy giải ATP và năng lượng tự do được giải phóng giúp nó di chuyển dọc theo chuỗi ARN mới sinh, hướng tới bong bóng phiên mã (Hình 4.5). Khi đuổi kịp ARN polymerase, yếu tố ρ tách phức lai ARN-ADN và phóng thích mARN vào tế bào chất.



Hình 4.5. Mô hình giả thiết cho sự kết thúc phiên mã ở *E. coli* bởi yếu tố Rho

4.4. SỰ PHIÊN MÃ Ở TẾ BÀO NHÂN THẬT

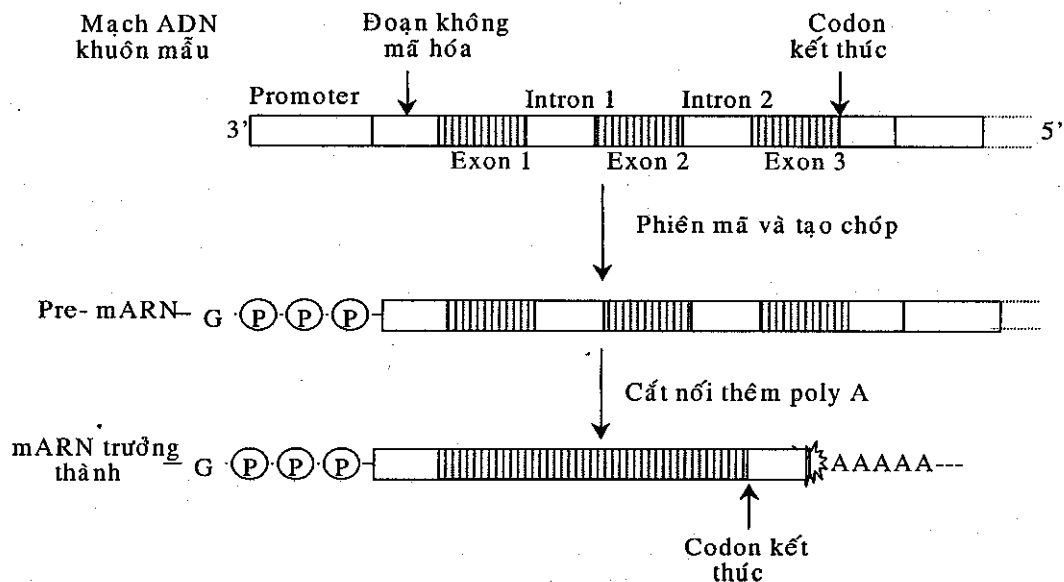
Phiên mã ở tế bào nhân thật có các đặc điểm sau:

- ARN polymerase II chịu trách nhiệm tổng hợp mARN, ARN polymerase I tổng hợp rARN, ARN polymerase III phiên mã cho ra các ARN có kích thước nhỏ, tARN, ARN 5S.

- mARN chứa thông tin của một gen (monocistron mARN).

- Quá trình phiên mã phức tạp hơn. Ở đầu 5' mARN có gắn thêm "chóp" (CAP) là 7- methylguanosine, còn cuối mARN phía 3' có thêm "đuôi" polyadenin dài 100 - 200 nucleotid.

- Bản phiên mã đầu tiên còn gọi là tiền mARN không được sử dụng trực tiếp mà phải qua biến đổi.



Hình 4.6. Phiên mã gen gián đoạn ở tế bào nhân thật

4.4.1. Các gen gián đoạn

Từ năm 1977, người ta phát hiện nhiều gen của tế bào nhân thật có tính gián đoạn. Trên gen, các đoạn mã hoá cho protein được gọi là exon xen kẽ với các đoạn không mã hoá được gọi là intron. Bản phiên mã đầu tiên tức là tiền mRNA chứa cả trình tự phiên mã của exon và intron. Tiếp theo các intron tức các đoạn không mã hoá protein được cắt rời ra, còn các exon mã hoá protein được nối liền lại với nhau để tạo mRNA trưởng thành. Quá trình này được gọi là cắt nối (Hình 4.6).

Khi mạch mRNA đang được nối dài khoảng 20 – 30 nucleotid thì ở đầu 5'P, enzym sẽ nối thêm vào gốc 7 – methylguanylate. Chóp này gắn vào đầu 5'P tạo nên liên kết 5'P – 5'P.

Một đoạn ngắn của đuôi 3' pre – mRNA, tại phần không mã hoá sau bộ kết thúc, bị cắt để nối đuôi poly A.

Phức hợp ribonucleoprotein (snRNP) của nhân tế bào thực hiện quá trình cắt nối (splicing) tạo cấu trúc không gian thuận tiện cho các đầu exon gần nhau và xúc tác phản ứng kết nối.

Sau splicing, mRNA vừa mới trưởng thành không còn intron, qua lỗ màng nhân ra tế bào chất để dịch mã.

Không phải mRNA của tế bào nhân thật nào cũng cần gắn chóp, nối đuôi và cắt nối.

Chỉ có một tỉ lệ nhỏ (khoảng 10%) bộ gen của tế bào nhân thật là phiên mã được. Phần lớn còn lại của bộ gen ở trạng thái sợi nhiễm sắc đóng xoắn mạnh

(dị nhiễm sắc) nên trở với sự phiên mã. Phiên mã chỉ xảy ra ở các nguyên nhiễm sắc xoắn lỏng lẻo.

4.4.2. ARN polymerase của tế bào nhân thật

Nhân của tế bào nhân thật có 3 loại ARN polymerase (I, II, III) (Bảng 4.3).

Bảng 4.3. Các ARN polymerase nhân của tế bào nhân thật

Loại	Vị trí	Bản sao	Hiệu ứng α - amanitin
Polymerase I	Nhân con	Tiền rARN lớn	Không nhạy cảm
Polymerase II	Dịch nhân	Tiền mARN (hn ARN)	Rất nhạy cảm
Polymerase III	Dịch nhân	5S rARN, tARN và các ARN nhỏ khác	Nhạy cảm trung bình

Polymerase I trong nhân, phiên mã các gen ARN ribosom. Bản sao chủ yếu là pre - rARN. Polymerase II tìm thấy trong dịch nhân, rất nhạy cảm với sự ức chế của α - amanitin (khoảng 50% ức chế ở nồng độ 10^{-8} M). Nó phiên mã các gen mã hoá protein và các bản sao sơ cấp không đồng nhất (hnARN), là tiền chất của mARN tế bào chất. Polymerase III cũng ở trong dịch nhân, phiên mã các gen 5S rARN và gen tARN. Nó nhạy cảm trung bình với sự ức chế của α - amanitin (khoảng 50% ức chế ở 10^{-6} M). Phản ứng được xúc tác bởi các polymerase tế bào nhân thật, về mặt hoá sinh cũng giống với các phản ứng được xúc tác bởi các enzym ở *E.coli*.

Cả 3 loại polymerase của tế bào nhân thật đều có phân tử lượng lớn với M vượt quá 500.000. Chúng gồm 2 tiểu đơn vị >100.000, mỗi phần là một polymerase chuyên biệt và có tới 12 tiểu đơn vị nhỏ hơn, trong đó có tới một vài phần không gặp ở cả 3 polymerase. Chức năng chính xác của các tiểu đơn vị khác nhau chưa được biết hết. Khác với polymerase vi khuẩn, polymerase nhân tế bào nhân thật không có khả năng khởi đầu phiên mã ở các vùng tương ứng với các vị trí khởi đầu *in vivo*, khi ử enzym tinh khiết với ADN *in vitro*.

4.4.3. Sự phiên mã do ARN polymerase I

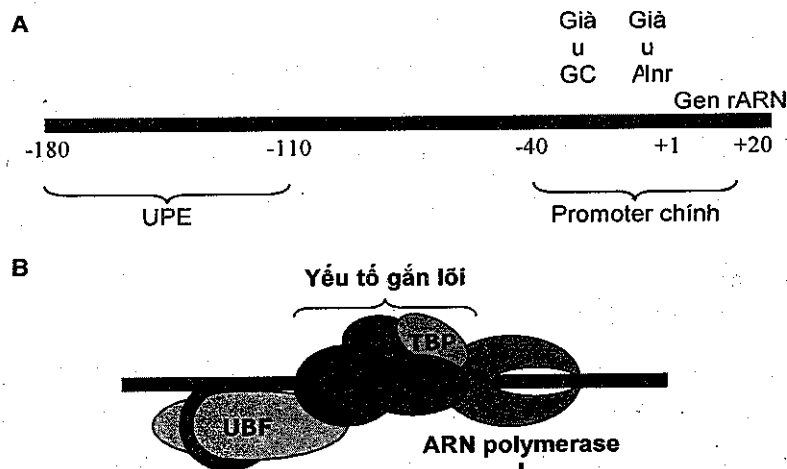
Các gen ARN ribosom thường là các gen được phiên mã nhiều nhất trong tế bào, phản ánh nhu cầu cao về rARN. Sợi nhiễm sắc (chromatin) chứa các gen ribosom, hoàn toàn không tạo nucleosom. Các trình tự gen rARN được bảo lưu ở mức độ cao ở các loài sinh vật, chính vì thế thật không lạ khi tìm thấy có các trình tự chuyên biệt cho từng loài phía trước các điểm xuất phát phiên mã.



Các rARN được phiên mã bởi ARN polymerase I. Bản phiên mã đầu tiên là tiền rARN 45S (pre-rARN), sau đó cắt nối tạo ra 3 loại 28S, 18S và 5,8S. Các gen này không phân tán mà xếp thành cụm, mỗi cụm có thể hơn 200 bản sao. Ở người, các cụm này nằm trên vai ngắn của các nhiễm sắc thể tâm đầu 13, 14, 15, 21 và 22. Các nhóm này xếp quanh yếu tố tổ chức hạch nhân hình thành eo thứ cấp và ở gian kỳ tạo nên hạch nhân.

Ở động vật có vú nói chung, có khoảng 200 bản sao của gen mã hóa cho rARN được bố trí lặp lại liên tiếp, giữa các vùng mã hóa cho rARN là vùng không phiên mã chứa promoter và các yếu tố giúp tăng cường tần suất phiên mã. Promoter của ARN polymerase I bao gồm hai vùng tách biệt (Hình 4.7-A). Promoter chính (core promoter) nằm quanh điểm xuất phát, trải từ -45 đến +20 và đủ hiệu quả để khởi đầu phiên mã. Promoter này giàu GC một cách bất thường, nhưng trong đó có một vùng bảo tồn ngắn giàu AT nằm quanh vị trí xuất phát gọi là Inr. Tuy nhiên, hoạt động của promoter chính được khuếch đại lên nhiều lần nhờ yếu tố nằm trước promoter (Upstream Promoter Element - UPE). UPE cũng giàu GC và nằm trong vùng từ -180 đến -107.

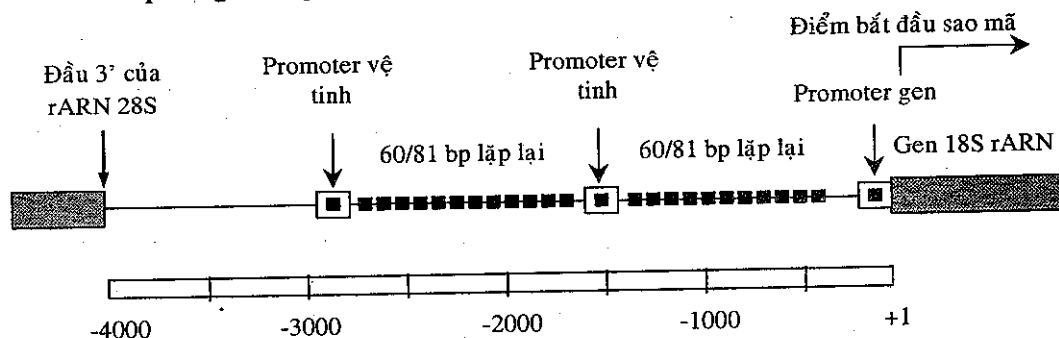
ARN polymerase I cần hai yếu tố phụ trợ. Yếu tố gắn lõi (gọi là SL1 hay TIF-IB hay Rib1 tùy theo loài), cấu tạo từ 4 protein. Một trong các protein này là TBP (TATA-Binding Protein) giống như ở phiên mã với ARN polymerase II và III, tuy nhiên nó không gắn vào vùng giàu GC, nhưng có lẽ tương tác với ARN polymerase I, việc gắn ADN do các protein khác trong yếu tố gắn lõi đảm nhiệm. Yếu tố gắn lõi giúp polymerase I khởi động ở mức cơ bản và giúp định vị ARN polymerase I chính xác ở vị trí xuất phát. Để phiên mã mạnh, ARN polymerase I cần yếu tố UBF (Upstream Binding Factor). Yếu tố này là một polypeptid đơn gắn vào vùng giàu GC của UPE, giúp kéo UPE vào gần promoter chính và gia tăng sự gắn của yếu tố gắn lõi vào promoter chính dẫn đến tăng tần suất khởi động phiên mã (Hình 4.7-B).



Hình 4.7. Cấu trúc promoter và sự khởi đầu phiên mã của ARN polymerase I

A: cấu trúc promoter của ARN polymerase I; B: sự lắp ráp các yếu tố phiên mã và ARN polymerase I

Ở *Xenopus laevis*, có khoảng 500 gen mã hóa cho rARN, cũng được bố trí lặp lại liên tiếp, giữa các vùng mã hóa cho rARN là vùng không phiên mã chứa nhiều bản sao lặp lại của promoter và các yếu tố giúp tăng cường tần suất phiên mã (Hình 4.8). Promoter thực gồm khoảng 150 bp được tìm thấy giữa các vị trí -146 và +6. Tuy nhiên, các bản sao chép hoàn chỉnh đạt độ dài của promoter này được lặp lại ở các vị trí -1.200 và -2.300 bp trong vùng vệ tinh (vùng không phiên mã), và cài rất nhiều bản sao 60 hoặc 81 bp của promoter thực. Các bản sao lặp lại 60/81 bp kích thích rất nhiều sự phiên mã của promoter thực. Các promoter vệ tinh này có khả năng khởi động quá trình tổng hợp các bản mã sao rARN. Các phân tử polymerase “xếp hàng” trong vùng vệ tinh trước khi chuyển động tới promoter thực.



Hình 4.8. Các phân tử promoter trong vùng vệ tinh không phiên mã của các gen rARN *Xenopus laevis*

4.4.4. Sự phiên mã bởi polymerase II

Các gen được phiên mã bởi polymerase II để tạo ra mARN đa dạng hơn nhiều so với các gen rARN đồng nhất.

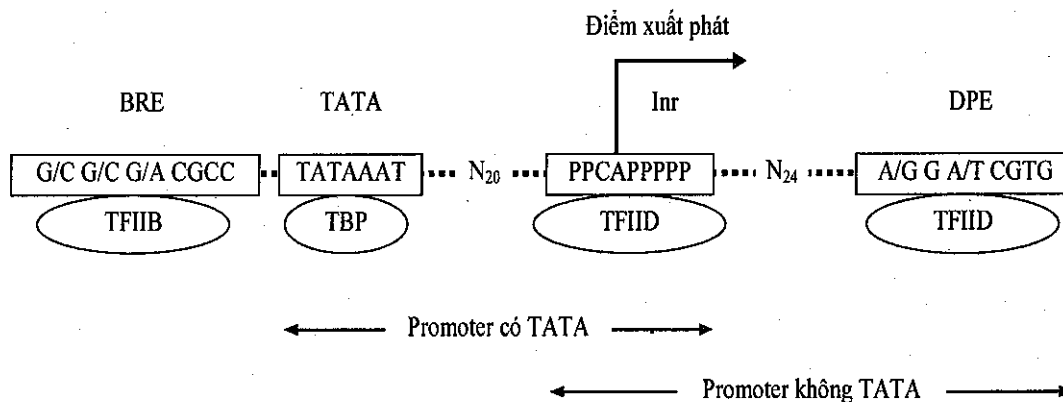
Trừ các gen mã hoá cho protein histon, thì các gen được phiên mã bởi ARN polymerase II đều có trình tự duy nhất hoặc với số ít bản sao. Các gen này hầu như chỉ mã hoá cho một loại protein.

Một số gen có bản sao thứ 2 trong quá trình tiến hoá. Cả 2 bản sao có thể chuyển đổi hỗ trợ nhau, như trường hợp các gen α - globine. Các gen mã hoá protein này biểu hiện thường xuyên ở các mô (gen giữ nhà) và có thể chỉ biểu hiện ở một mô nào đấy của cơ thể đa bào (đa phần là gen điều hoà) hoặc điều hoà bởi sự có hay thiếu cơ chất riêng biệt, như hormon hoặc chất kích thích môi sinh. Các gen liên quan đến sự chuyển hoá galactose trong nấm men là một thí dụ.

Promoter của ARN polymerase II khá phức tạp và ít có điểm chung như ở tế bào nhân nguyên thủy. Nhưng nói chung có thể chia làm hai nhóm (Hình 4.9): có hộp TATA hay không có. Cấu trúc chung của promoter gồm 2 trình tự: trình tự thứ nhất là vị trí xuất phát có trình tự chung Py2CAPy5 (trong đó A thường là base +1) gọi là trình tự Inr, và trình tự thứ hai ở promoter nhóm 1 có hộp TATA (hộp Goldberg-Hogness) ở vị trí -25 với trình tự

chung là TATAAAT, trong khi đó nhóm 2 không có hộp TATA nhưng có yếu tố promoter xuôi dòng (Downstream Promoter Element - DPE) có trình tự chung là A/G G A/T CGTG ở vị trí từ +28 đến + 32. Ngoài ra, một số promoter với hộp TATA có yếu tố nhận diện B (B Recognition Element - BRE) có trình tự chung gồm 7 base là G/C G/C G/A CGCC nằm ở vị trí -35, ngay trước hộp TATA. Các trình tự này là nơi gắn của các yếu tố phiên mã (Hình 4.9).

Ở tế bào nhân thật, ARN polymerase II không trực tiếp gắn vào promoter. Việc nhận diện promoter và khởi động phiên mã do các yếu tố phiên mã đảm nhiệm, sau đó ARN polymerase mới gắn vào vị trí xuất phát thông qua tương tác với các yếu tố phiên mã. Có hai nhóm yếu tố phiên mã: yếu tố phiên mã chung, cần cho hầu hết các gen ở tất cả các loại tế bào khác nhau và yếu tố phiên mã chuyên biệt chỉ có ở một số tế bào được biệt hóa. Nói cách khác, các gen ở tế bào nhân thật chỉ được biểu hiện nếu có sự hiện diện của yếu tố phiên mã thích hợp.



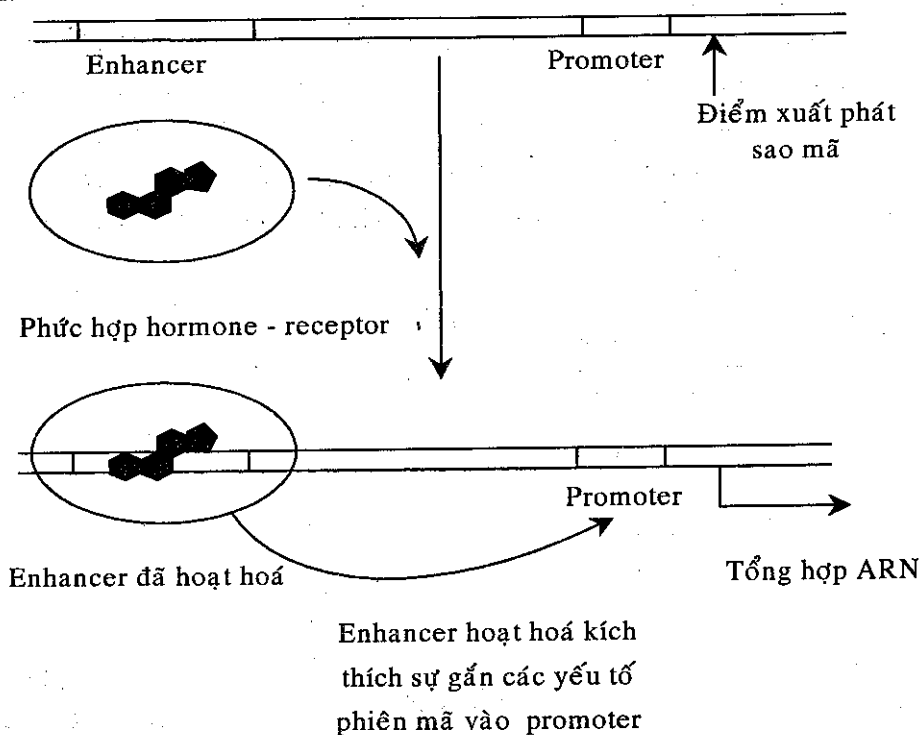
Hình 4.9. Cấu trúc promoter của ARN polymerase II và các yếu tố phiên mã chung có liên quan

Ngoài ra, một số trình tự khác đôi khi gặp ở các gen cần gắn protein hoạt hóa để biểu hiện một cách đầy đủ: hộp CAAT (nằm cách vị trí phiên mã 80 bp), có trình tự chung là GGCAATCT và hộp GC hay thay đổi vị trí, có trình tự chung là GGGCCGGG. Người ta cho rằng các trình tự chung này có chức năng liên kết các yếu tố phiên mã chuyên biệt hơn là gắn trực tiếp với ARN polymerase II. Ví dụ, các tế bào của động vật có vú chứa 1 protein gọi là SP₁ có Mr 100 000, cần cho sự phiên mã của các gen có hộp GC trong promoter. Các trình tự khởi động khác nhau nằm rất gần với điểm bắt đầu phiên mã là cần thiết, nhưng thường vẫn không đủ để biểu hiện các gen ARN polymerase phiên mã. Có các phân tử được đưa thêm vào, gọi là các trình tự tăng cường. Chúng không có khả năng khởi động một mình, nhưng lại có khả năng kích thích phiên mã mạnh và có thể tác động ở các khoảng cách đáng kể (đến vài kilobase). Trình tự tăng cường có hiệu quả như nhau đối với các gen cấu trúc do chúng điều hòa và có thể nằm rất gần đầu 3', 5' hoặc nằm ngay trong một cistron của gen cấu trúc.

Trình tự tăng cường luôn thể hiện tính chuyên biệt, do đó không liên quan tới việc điều hòa biểu hiện gen trong quá trình phát triển chung của cơ thể. Ví dụ, trình tự tăng cường miễn dịch globulin chỉ hoạt động trong β -lymphocyte.

Một trong những ví dụ về hoạt động tăng cường được nghiên cứu trong điều hòa phiên mã gen đáp ứng với hormon steroid. Gắn steroid vào thụ thể protein tan trong nước, để phức hợp này lại gắn vào các trình tự tăng cường làm gia tăng hoạt động các gen đáp ứng steroid (Hình 4.10).

Các bản mã sao của ARN polymerase II đều là những ARN nhân không đồng nhất (hnARN). Chúng có phân tử lượng rất lớn (đến 10 kb), vì chứa các intron. Đa phần các bản mã sao đều chứa trình tự AAUAAA gồm 20 nucleotid từ đầu 3'. Trình tự này rất quan trọng vì: nó được nhận biết bởi một endonuclease đặc hiệu để cắt ARN mới sinh khỏi polymerase II và nó là dấu hiệu được nhận biết bởi enzym poly(A) polymerase nhằm gắn đuôi polyA.



Hình 4.10. Hoạt hóa quá trình phiên mã của các gen đáp ứng hormon steroid (glucocorticoid)

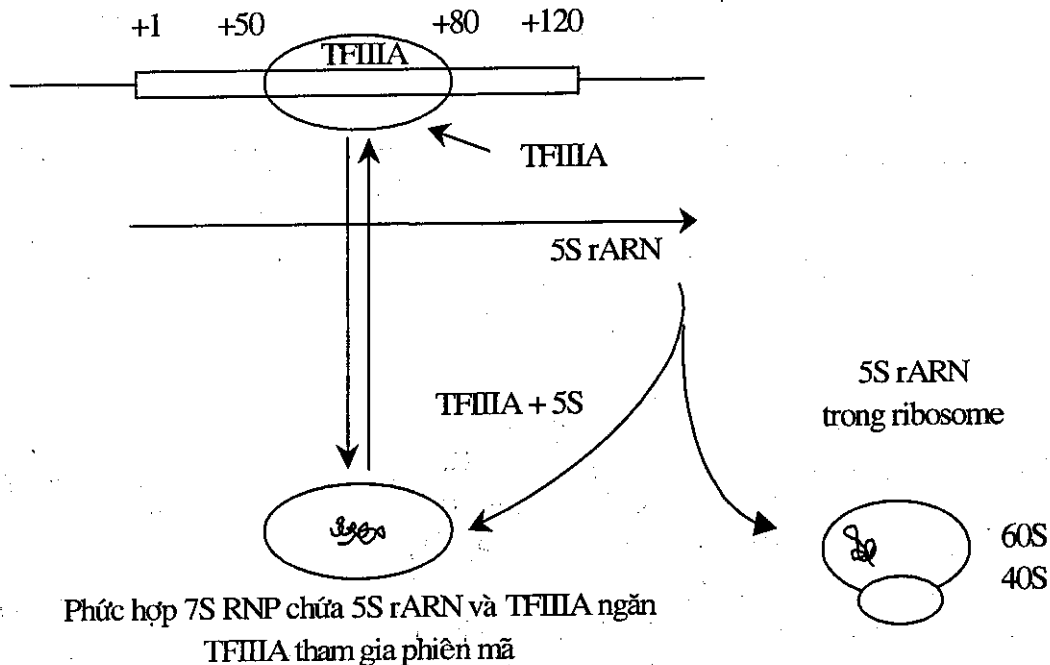
4.4.5. Sự phiên mã bởi ARN polymerase III

ARN polymerase III chịu trách nhiệm phiên mã ra rARN 5S, tARN và một số ARN nhỏ của nhân và của tế bào chất. Các gen do ARN polymerase III phiên mã thường chứa dưới 300 nucleotid. Gen 5S rARN không chứa intron, trong khi đó, nhiều gen tARN chứa các intron nhỏ.

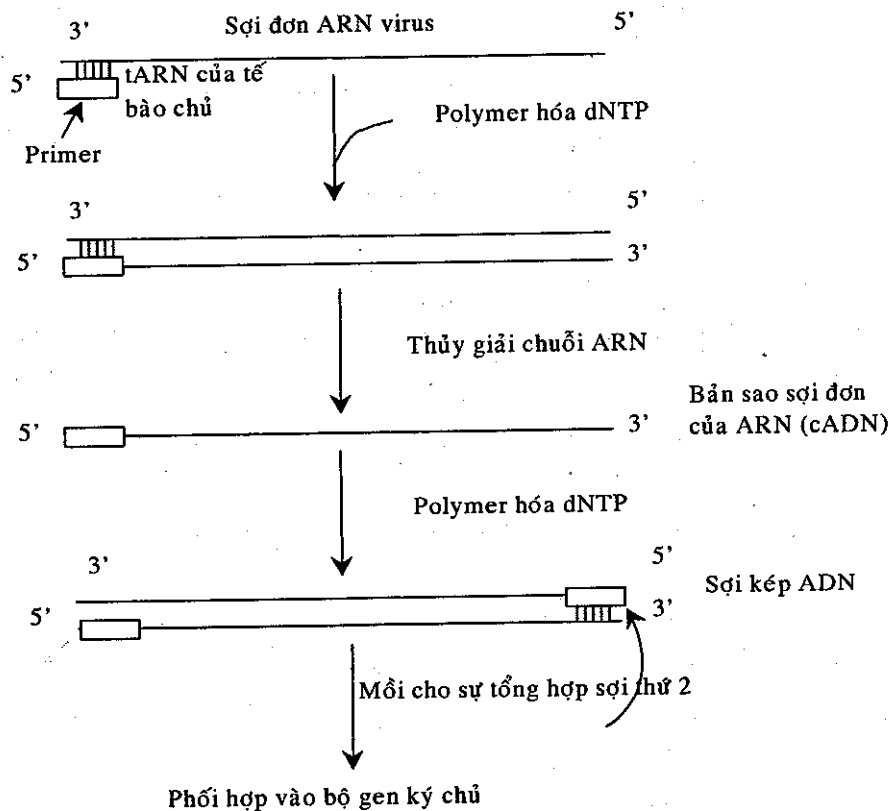
Promoter của ARN polymerase III được chia làm hai nhóm và được nhận diện bởi các nhóm yếu tố phiên mã khác nhau. Promoter của 5S rARN và tARN nằm trong vùng mã hóa (ở vị trí từ +55 đến +80 bp) và khởi động phiên mã tại vị trí xuất phát phía trước nó một khoảng nhất định. Promoter của snARN nằm phía trước điểm xuất phát như thường thấy ở các promoter khác. Trong cả hai trường hợp, promoter mang các trình tự được nhận diện bởi yếu tố phiên mã và ARN polymerase gắn vào các yếu tố phiên mã này.

Các yếu tố phiên mã của ARN polymerase III gồm TFIIIA, TFIIIB và TFIIIC. Trong đó, TFIIIA và TFIIIC là yếu tố lắp ráp, chúng gắn với promoter trước để giúp TFIIIB gắn vào gần vị trí +1. ARN polymerase III sau đó gắn vào TFIIIB và bắt đầu phiên mã. TFIIIB có tiểu đơn vị TBP để tương tác với ARN polymerase III.

TFIIIA (một protein có Mr 40 000) chuyên biệt ở noãn bào tham gia điều hòa tổng hợp các rARN 5S của noãn bào *X. laevis*. TFIIIA thường gắn vào rARN 5S tự do trong tế bào chất noãn bào hình thành nên phức hợp 7S ribonucleoprotein (7S RNP). Nếu việc tổng hợp các thành phần khác của ribosom như các ARN 28S, 18S, 5,8S và protein ribosom diễn ra chậm hơn quá trình tổng hợp rARN 5S, thì gần như tất cả TFIIIA đã gắn trong phức hợp 7S RNP và do đó không còn TFIIIA cho sự phiên mã. Chính vì vậy, sự tổng hợp rARN 5S bị giảm xuống cho đến khi TFIIIA được giải phóng trở lại. Do đó, sự tổng hợp rARN 5S noãn bào được điều hòa để đồng bộ với việc tổng hợp các rARN còn lại cần thiết cho sự lắp ráp thành ribosom (Hình 4.11).



Hình 4.11. Kiểm soát quá trình tổng hợp rARN 5S noãn bào *X. laevis*



Hình 4.12. Sao chép ngược sợi đơn ARN của Retrovirus thành sợi kép ADN

4.5. PHIÊN MÃ NGƯỢC Ở RETROVIRUS

Retrovirus là các virus mà vật liệu di truyền của chúng là ARN, ví dụ virus làm suy giảm miễn dịch (HIV) gây ra hội chứng suy giảm miễn dịch mắc phải ở người (AIDS). Sự tổng hợp ADN từ khuôn mẫu ARN được xúc tác bởi các enzym phiên mã ngược. Enzym này bắt đầu quá trình phiên mã bằng các dNTP như là chất nền (Hình 4.12). Khác với ARN polymerase, enzym phiên mã ngược phải dùng tới môi. Đoạn môi là một tARN của tế bào chủ, gắn kết vào đầu 3' của retrovirus. Chuỗi kép đầu tiên là chuỗi lai ARN - ADN. ARN virus trong chuỗi lai bị huỷ bởi RNaseH. Sau đó, một mạch ADN vừa được tổng hợp để tạo ra chuỗi kép ADN. Chuỗi kép lại tích hợp vào ADN ký chủ.

Enzym phiên mã ngược của HIV là một protein gồm 2 tiểu đơn vị. Mỗi tiểu đơn vị chứa 560 acid amin. Enzym phiên mã ngược có một vai trò quan trọng trong việc tổng hợp ADN bổ sung (cADN) trong quá trình điều khiển gen. Sự tích hợp của virus vào bộ gen tế bào chủ thường được ổn định (các provirus không bao giờ bị cắt rời như prophage). Hiện nay, nhiều người cho rằng các virus có thể tác động đến sự biểu hiện của các gen ung thư.

Bảng 4.4. Mã di truyền gồm 64 codon

Vị trí thứ I	Vị trí thứ II				Vị trí thứ III
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	stop	stop	A
	Leu	Ser	stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

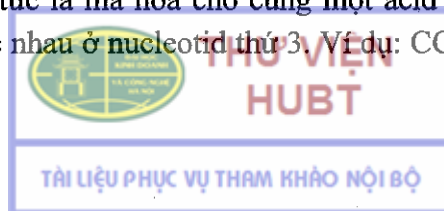
4.6. MÃ DI TRUYỀN

Như đã biết, các acid amin trong phân tử protein được mã hóa bằng nhóm các nucleotid trên phân tử ADN. Có tất cả 4 loại base, nếu các base có nhóm đôi, tức 2 nucleotid mã hóa cho một loại acid amin thì tất cả chỉ có 16 tổ hợp, không đủ cho 20 loại acid amin. Như vậy, đơn vị mã hóa hay còn gọi là codon phải gồm 3 hay nhiều nucleotid hơn.

Năm 1961, F. Crick đã làm thí nghiệm chứng minh rằng nhóm nucleotid mã hóa có 3 hay nói cách khác codon gồm 3 nucleotid kế tiếp nhau. Tất cả sẽ có $4^3 = 64$ tổ hợp. Bảng mã di truyền (Bảng 4.4) cho thấy trong 64 codon có 3 codon UAA, UAG, UGA không mã hóa cho acid amin được gọi là codon vô nghĩa (non-sense), đồng thời là codon kết thúc (stop) tức là “dấu chấm câu”, chấm dứt mạch polypeptid.

Mã di truyền có tính “suy thoái” tức một acid amin có nhiều codon mã hóa, trừ methionin và tryptophan chỉ có một codon.

Các codon đồng nghĩa tức là mã hóa cho cùng một acid amin thường có 2 base đầu tiên giống nhau, nhưng khác nhau ở nucleotid thứ 3. Ví dụ: CCU, CCC, CCA và CCG đều



mã hóa cho prolin. Trên thực tế U và C luôn luôn tương đồng nhau ở vị trí thứ 3, còn A và G tương đồng nhau trong 14 của 16 trường hợp.

Trừ một số ngoại lệ, mã di truyền có tính phổ biến cho tất cả sinh vật.

Một codon chỉ mã cho một loại acid amin, trường hợp ngoại lệ là AUG vừa mã hóa cho Met nội, vừa mã hóa cho acid amin mở đầu N-formyl Methionin trong tế bào tế bào nhân nguyên thủy hoặc methionin trong tế bào nhân thật.

Tóm tắt:

Quá trình phiên mã trong tế bào nhân nguyên thủy và tế bào nhân thật có một vài chi tiết khác nhau đáng kể. Tế bào nhân nguyên thủy chỉ chứa một loại enzym nhân ARN polymerase, mà khi tương tác với một tiểu đơn vị điều hoà (sigma) sẽ tạo ra holoenzym quy định tính đặc hiệu của promoter. Promoter gồm 2 vùng chứa các trình tự nằm ở trung tâm - 10 và - 35 kể từ vị trí bắt đầu phiên mã. Các trình tự này tương tác với holoenzym ARN polymerase làm cho sự phiên mã có hiệu quả và chính xác.

Nhân tế bào nhân thật chứa 3 loại ARN polymerase khác nhau. Polymerase I phiên mã các rARN; polymerase II phiên mã các gen mã hoá protein; polymerase III phiên mã rARN 5S, tARN và các ARN nhỏ khác. Các promoter của những nhóm gen trên có khác biệt đáng kể. Các gen được phiên mã bởi polymerase II có trình tự nhận biết ở ngay trong các promoter của chúng, nhưng cũng có nhiều sự khác biệt. Các promoter thuộc ARN polymerase III khác thường là nằm ngay trong trình tự của gen được phiên mã.

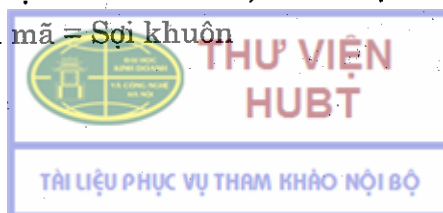
Khác với các promoter của tế bào nhân nguyên thủy, các promoter tế bào nhân thật không là các vị trí gắn trực tiếp ARN polymerase, mà thực hiện chức năng kết dính các yếu tố phiên mã (protein) làm cho nhiễm sắc thể tháo xoắn, cho phép toàn bộ ARN polymerase xâm nhập vào. Ngoài ra, nhiều gen tế bào nhân thật cần được hoạt hoá bởi các yếu tố gắn vào trình tự tăng cường (enhancer element) trước khi phiên mã có hiệu quả.

Mã di truyền gồm bộ 3 các base (codon) trong mARN. Chúng chuyên biệt cho việc kết hợp acid amin vào mạch polypeptid. Mã được tính theo chiều 5' → 3' của mARN, liên tục, phổ quát và suy thoái, trong đó nhiều loại acid amin được quy định bởi 2 hoặc nhiều hơn 1 codon.

CÂU HỎI

1. Khuôn mẫu trong phiên mã:

- a) Sợi phiên mã = Sợi khuôn
- b) Sợi phiên mã = Sợi mã hóa
- c) Sợi không phiên mã = Sợi khuôn
- d) Trình tự ARN giống sợi khuôn
- d) Trình tự ARN giống sợi phiên mã



9. Acid amin nào chỉ có một codon?
- a) Leucin
 - b) Methionin
 - c) Tryptophan
 - d) Alanin
 - e) b và c
10. Tính chất nào **không** phải của mã di truyền?
- a) Có ngoại lệ
 - b) Một chiều, không chồng lên nhau
 - c) Phổ biến ở mọi sinh vật là mã bộ 3
 - d) Đặc hiệu, một codon chỉ mã hoá cho một loại acid amin
 - e) Suy thoái, nhiều bộ ba mã hoá cho một loại acid amin

Bài 5

QUÁ TRÌNH DỊCH MÃ

MỤC TIÊU

- Mô tả được cơ chế dịch mã di truyền.
- Phân biệt được quá trình dịch mã ở tế bào nhân thật và tế bào nhân nguyên thủy.
- Áp dụng Dựa vào cơ chế khác nhau này giải thích được tác động chọn lọc của kháng sinh.
- Mô tả được cách thức đảm bảo sự chính xác của quá trình tổng hợp protein.

5.1. MỞ ĐẦU

Giai đoạn cuối của sự biểu hiện gen xảy ra khi thông tin được mã hóa bằng trình tự nucleotid trên mRNA được dịch mã thành trình tự acid amin tương ứng trong chuỗi polypeptid. Quá trình này được gọi là quá trình dịch mã. Mặc dù, về cơ bản tiến trình dịch mã xảy ra như nhau ở mọi sinh vật, nhưng vẫn có những chi tiết khác nhau giữa tế bào vi khuẩn và tế bào nhân thật. Nhìn chung, quá trình này ở tế bào nhân thật phức tạp hơn. Phần lớn quá trình tổng hợp protein xảy ra trong tế bào chất. Một số bào quan của tế bào nhân thật cũng có bộ máy di truyền riêng (lục lạp, ty thể) và cũng có khả năng tổng hợp được protein mã hóa bởi chúng. Bộ máy dịch mã của các bào quan này giống tế bào nhân nguyên thủy hơn là tế bào nhân thật.

5.2. CÁC YẾU TỐ CẦN THIẾT CHO QUÁ TRÌNH DỊCH MÃ

5.2.1. Ribosom

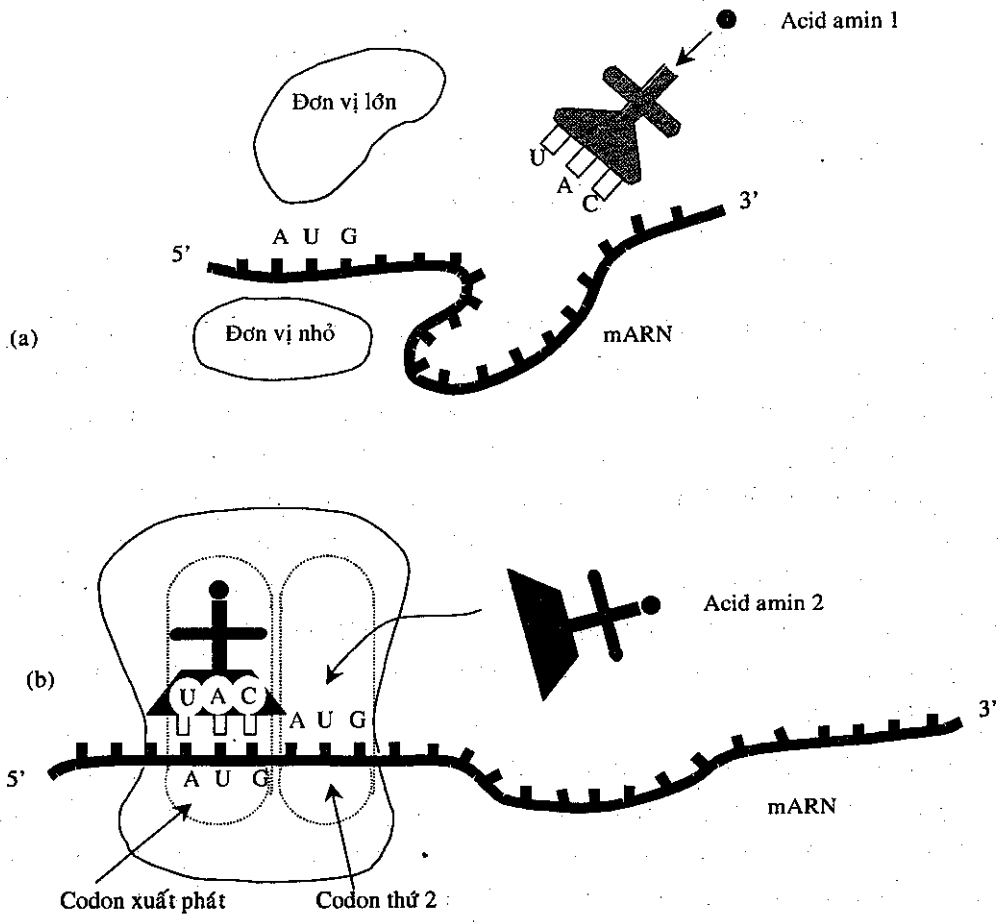
Ribosom là xưởng tổng hợp protein của tế bào. Tuy có sự khác nhau trong cấu tạo và kích thước giữa các giới sinh vật, nhưng đều gồm 2 tiểu đơn vị. Và chỉ có các

ribosom hoàn chỉnh, gồm một tiểu đơn vị lớn và một tiểu đơn vị nhỏ mới có đủ khả năng dịch mã mARN.

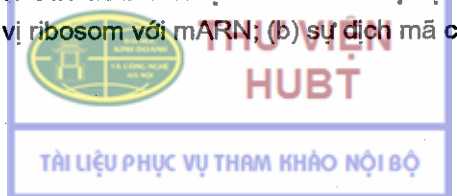
Khi không thực hiện tổng hợp protein, mỗi tiểu đơn vị tồn tại tách rời trong tế bào chất (Hình 5.1). Tiểu đơn vị lớn của tế bào nhân nguyên thủy gồm 2 phân tử rARN và 35 phân tử protein, tiểu đơn vị nhỏ có 1 phân tử rARN và khoảng 20 phân tử protein. rARN có cấu trúc không gian phức tạp do nhiều đoạn bắt cặp với nhau nhờ có trình tự nucleotid bổ sung.

Ở tế bào nhân nguyên thủy, đầu 5' của mARN gắn trước tiên vào tiểu đơn vị nhỏ và lúc đó có khả năng gắn thêm vào tiểu đơn vị lớn. Khi đơn vị lớn gắn xong thì sự dịch mã bắt đầu.

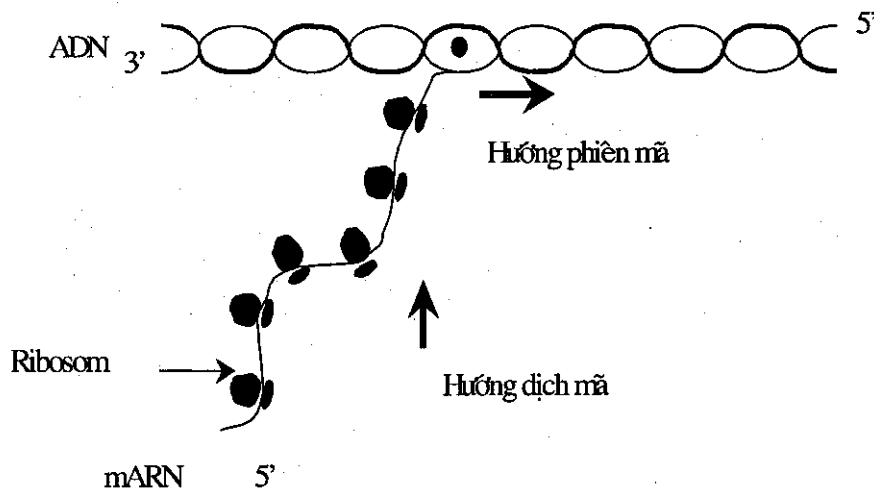
Vì tế bào nhân nguyên thủy không có màng nhân, quá trình phiên mã diễn ra trong tế bào chất, các ribosom có khả năng gắn vào mARN đang được phiên mã và dịch mã tạo ra protein trong khi ARN polymerase vẫn tiếp tục phiên mã từ ADN.



Hình 5.1. Các tiểu đơn vị ribosom và sự dịch mã
 (a) 2 tiểu đơn vị ribosom với mARN; (b) sự dịch mã của ribosom



Ở cả tế bào nhân nguyên thủy lẫn tế bào nhân thật, khi ribosom đầu tiên dịch mã mRNA được một đoạn, thì các ribosom khác có thể gắn vào phía trước để dịch mã. Do đó, các ribosom xếp thành chuỗi trên mRNA tạo nên cấu trúc polyribosom hay còn gọi là polysome (Hình 5.2). Khoảng 15 – 20 ribosom có thể gắn cùng lúc trên mRNA. Nhờ vậy, tốc độ tổng hợp protein tăng nhanh đáng kể.

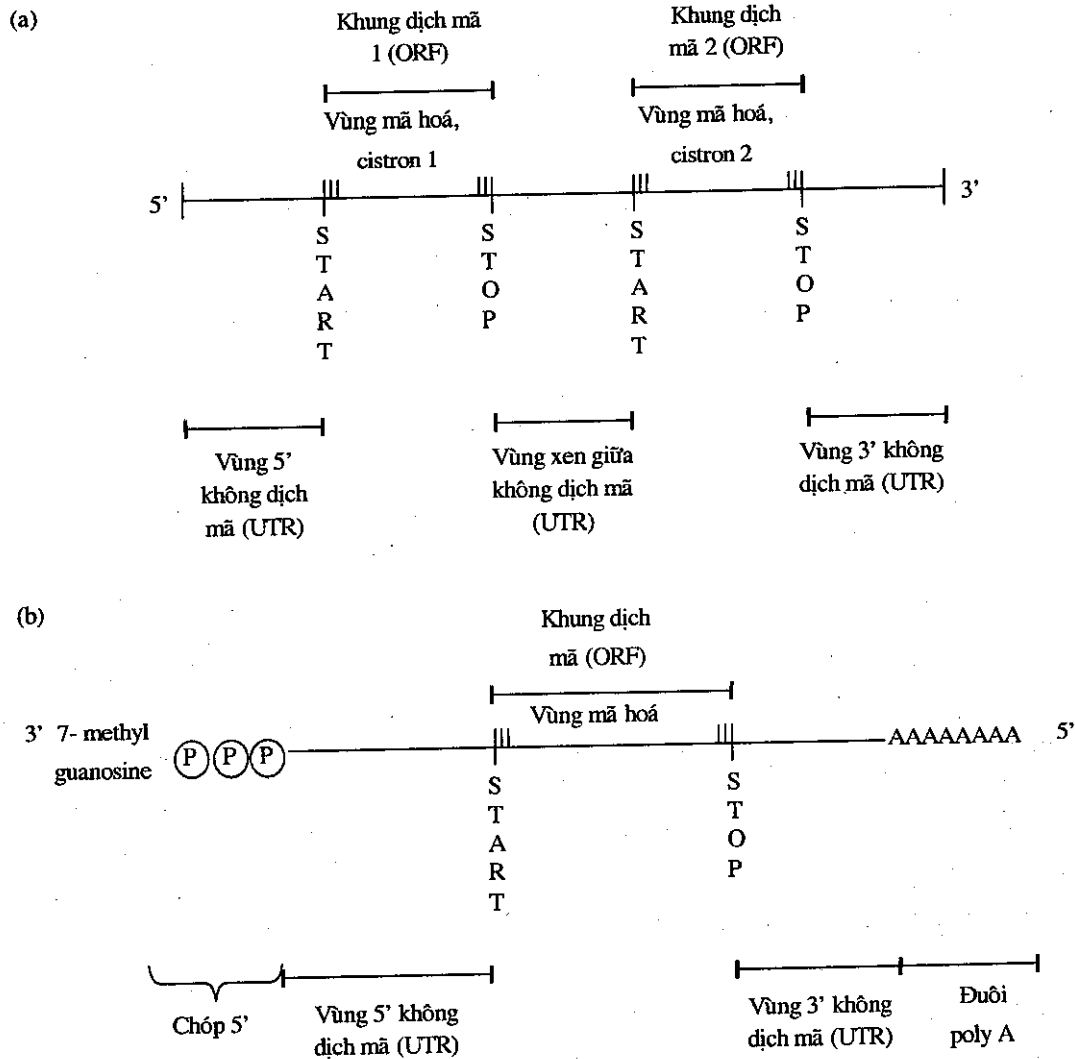


Hình 5.2. Polyribosom

5.2.2. ARN thông tin

Cấu trúc của mRNA điển hình được minh họa ở. Thông tin di truyền là một trình tự liên tục của các base nucleotid. Trình tự trong vùng mã hóa của mRNA chứa các codon quy định trình tự acid amin của protein. Trình tự này được đọc theo hướng 5' đến 3' trên mRNA được mở đầu bằng một codon mở đầu (thường là AUG). Kết thúc là một trong các codon kết thúc (UAA, UGA, hoặc UAG). Do mã di truyền được đọc theo bộ ba nên đối với mỗi trình tự mRNA, có 3 cách đọc khác nhau, tức có 3 khung đọc (Open Reading Frame - ORF) và do đó có thể được dịch thành 3 polypeptid hoàn toàn khác nhau tùy theo việc lựa chọn bộ ba khởi đầu. Điều này cũng giải thích lý do các đột biến mất base hay chèn base có thể dẫn đến việc mã hóa cho các polypeptid khác nhau.

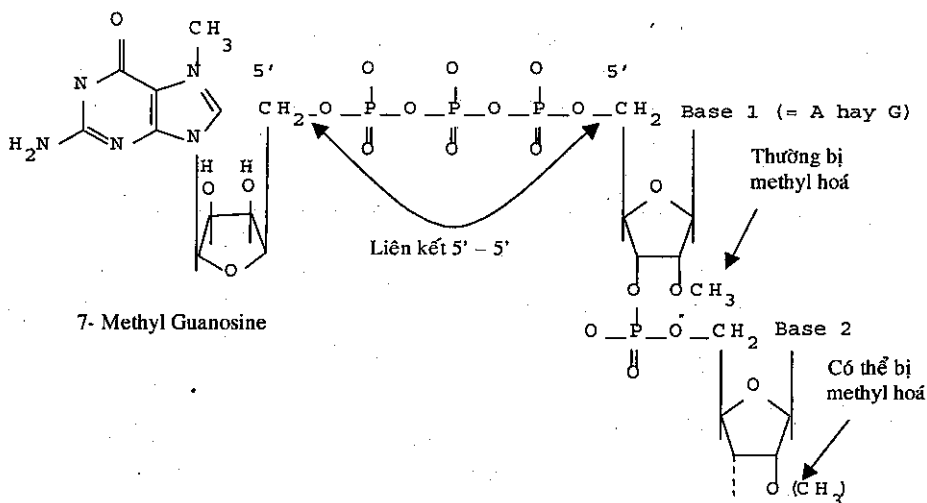
Phần tử mRNA chấm dứt tại một trình tự không mã hóa. Và cũng có một vùng không mã hóa ở đầu 5' được gọi là trình tự dẫn bao gồm một trình tự gắn ribosom ở tế bào nhân nguyên thủy. Các trình tự dẫn đầu 5' rất khác nhau về độ dài trong các loại mRNA khác nhau, nhưng nhìn chung, ở tế bào nhân thật đều dài hơn ở tế bào nhân nguyên thủy. Các trình tự 3' và 5' không mã hóa còn được gọi là các vùng không dịch mã (UTR).



Hình 5.3. Các cấu trúc chi tiết của các phân tử mRNA

(a) mRNA vi khuẩn có chứa 2 cistron; (b) monocistron mRNA tế bào nhân thật

Ngoài các trình tự không mã hóa ở đầu 5' và 3', các mRNA tế bào chất tế bào nhân thật có gắn thêm "chóp" (CAP) 7-methyl Guanosine Triphosphat ở đầu 5' (Hình 5.4). Đặc điểm này liên quan đến sự tương tác giữa mRNA và ribosom. Các trình tự không dịch mã ở đầu 3' có chiều dài rất khác nhau. Hầu hết các phân tử mRNA tế bào nhân thật ở đầu 3' có thêm đuôi 100-200 Adenin, sau khi phiên mã. Đuôi polyA đóng vai trò quan trọng trong việc bảo vệ mRNA khỏi bị thủy giải của exonuclease, làm tăng độ bền của mRNA trong tế bào chất. Trong quá trình dịch mã, đuôi polyA được gắn các protein gắn đuôi và qua đó tương tác với các yếu tố ở đầu 5' để mRNA có cấu trúc vòng, giúp tái sử dụng ribosom nhanh hơn. mRNA vi khuẩn và một vài loại mRNA ở tế bào nhân thật như histon-mRNA không có đuôi polyA.



Hình 5.4. Cấu trúc chóp (CAP) của tế bào nhân thật

5.2.3. ARN vận chuyển và các enzym tổng hợp aminoacyl - tARN

Các ARN vận chuyển hoạt động như một chất kết nối giữa thông tin base-nucleotid của mRNA với trình tự acid amin của phân tử protein tương ứng. Do đó, các tARN phải được nhận biết chuyên biệt đồng thời bởi mRNA và bởi enzym gắn acid amin vào tARN tương ứng (enzym aminoacyl-tARN synthetase). Điều này giúp giải quyết trở ngại không gian trong quá trình dịch mã. Kích thước của codon lớn hơn nhiều so với acid amin, nên nếu một acid amin nhận biết và gắn trực tiếp lên một codon trên mRNA thì nó sẽ cách xa với acid amin tương ứng với codon kế cận để hình thành liên kết peptid.

Hầu hết các loài đều có 20 loại Aminoacyl-tARN synthetase tương ứng với 20 loại acid amin. Đó là Alanyl-tARN synthetase, Arginyl-tARN synthetase, Asparaginyl-tARN synthetase, ... Nhưng một số vi khuẩn (ví dụ *Helicobacter pylori*) không có Glutaminyl-tARN synthetase để aminoacyl hóa tARN^{Gln}. Trong trường hợp này, Glutamat-tARN synthetase được dùng để aminoacyl hóa tARN^{Gln} với Glutamate và sau đó một enzym khác sẽ chuyển glutamat trên glutamat-tARN^{Gln} thành Glutamin. Điều này cho thấy Aminoacyl-tARN synthetase không bao giờ gắn 2 acid amin khác nhau lên cùng một loại tARN.

Mặc dù, có thể có tới một vài loại tARN khác nhau cho mỗi một loại acid amin (để giải mã các codon khác nhau tương ứng với acid amin đó), nhưng chỉ có một aminoacyl-tARN synthetase cho mỗi loại acid amin. Enzym aminoacyl-tARN synthetase phải phù hợp và chọn đúng tARN: các tARN được nhận diện bởi một synthetase thì được gọi là tARN cùng họ (cognate tARN), các loại tARN liên quan đến cùng một acid amin thì được gọi là tARN đồng nhận (isoaccepting tARN). Một số loài có số tARN nhiều hơn số codon (ví dụ *E. coli* có 86 tARN), nghĩa là có nhiều tARN cùng giải mã cho một codon.

Có hơn một codon trong mã di truyền cho từng loại acid amin, ngoại trừ methionin và tryptophan. Những codon đồng nghĩa (synonym) này thường liên quan đến các tARN khác nhau. Các tARN lại có trình tự đối mã (anticodon) khác nhau. Thí dụ, trong trường hợp của serin, mặc dù có tới 6 codon (UCU, UCC, UCG, UCA, AGU, AGC) đều có ý nghĩa như nhau, nhưng không phải tất cả đều được sử dụng thường xuyên như nhau. Thật vậy, đối với một loài chỉ có một vài codon cho một acid amin được sử dụng và có một vài phân tử tARN cho mỗi codon trong tế bào. Hiện tượng này được gọi là sự dị biệt codon và đây là điểm khác nhau quan trọng giữa loài này và loài khác.

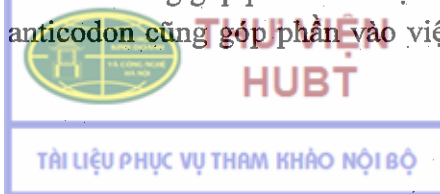
Trong một số trường hợp, tARN mang 1 đối mã nhưng có thể giải mã cho nhiều codon khác nhau, khi đó đối mã chỉ bắt bổ sung với hai vị trí đầu tiên của codon trên mRNA. Điều này cũng lý giải vì sao các codon đồng nghĩa thường chỉ khác nhau ở nucleotid cuối. Do đó, mặc dù có 61 codon có nghĩa nhưng trong tế bào chỉ có khoảng 31 đến 50 loại tARN khác nhau. Ví dụ, ở người có gần 500 gen mã hóa cho tARN nhưng chỉ mang 48 loại đối mã khác nhau.

Trong quá trình dịch mã, ribosom không có khả năng xác định tARN có gắn đúng với acid amin tương ứng hay không, nó chỉ đơn thuần đảm bảo sự ăn khớp giữa codon trên mRNA và anticodon trên tARN, nếu khớp thì acid amin trên tARN dù đúng hay sai đều được nối vào chuỗi polypeptid. Do đó, bước aminoacyl hóa phải chính xác. Bởi vì, nếu một tARN chấp nhận một acid amin sai dẫn đến việc gắn acid amin không đúng vào protein và sẽ không có “bước đọc sửa sai” theo sau bước aminoacyl hóa. Bước đọc sửa sai chỉ được thực hiện ở giai đoạn aminoacyl hóa bởi chính enzym aminoacyl-tARN synthetase. Như vậy, quá trình aminoacyl hóa cùng lúc thực hiện 2 chức năng: Trước tiên, nó cung cấp sự liên hệ duy nhất của mã di truyền (dựa vào acid nucleic) và cấu trúc protein (với các acid amin). Tiếp đến, nó hoạt hóa acid amin trước khi kết nối vào protein. Liên kết ester aminoacyl giữa acid amin và tARN dự trữ một năng lượng tự do tương đối cao, năng lượng này rất cần thiết cho việc hình thành liên kết peptid trong quá trình dịch mã.

Tính chính xác của bước aminoacyl hóa tARN được đảm bảo qua hai vòng: nhận diện chính xác acid amin và nhận diện chính xác tARN cùng họ:

– Aminoacyl-tARN synthetase nhận diện chính xác acid amin nhờ vào sự khác biệt về kích thước, hình dạng, điện tích hay nhóm chức hóa học của nó. Ngoài ra, một số aminoacyl-tARN synthetase có vị trí sửa lỗi trong cấu trúc enzym để loại bỏ các acid amin không được hoạt hóa đúng. Tần suất sai số chung của việc nhận diện acid amin vào khoảng 0,01%.

– Aminoacyl-tARN synthetase nhận diện được các tARN cùng họ nhờ vào việc nhận diện nhánh gắn acid amin trên tARN, đặc biệt là base quyết định nằm trong nhánh này, bên cạnh đó việc nhận diện vòng anticodon cũng góp phần vào việc lựa chọn tARN của enzym. Bản thân anticodon trên vòng anticodon cũng góp phần vào việc nhận diện nhưng không



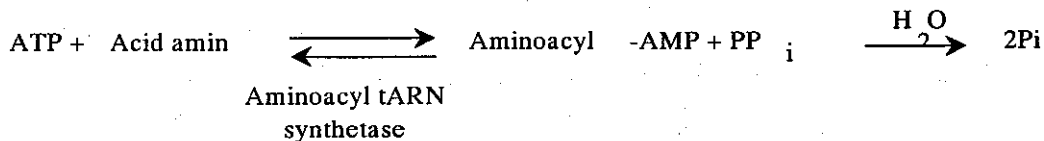
có tính quyết định vì có thể có nhiều anticodon hoàn toàn khác nhau tương ứng cho một acid amin.

5.2.3.1. Gắn acid amin vào tARN bằng phản ứng aminoacyl hóa

Quá trình aminoacyl hóa xảy ra theo 2 bước:

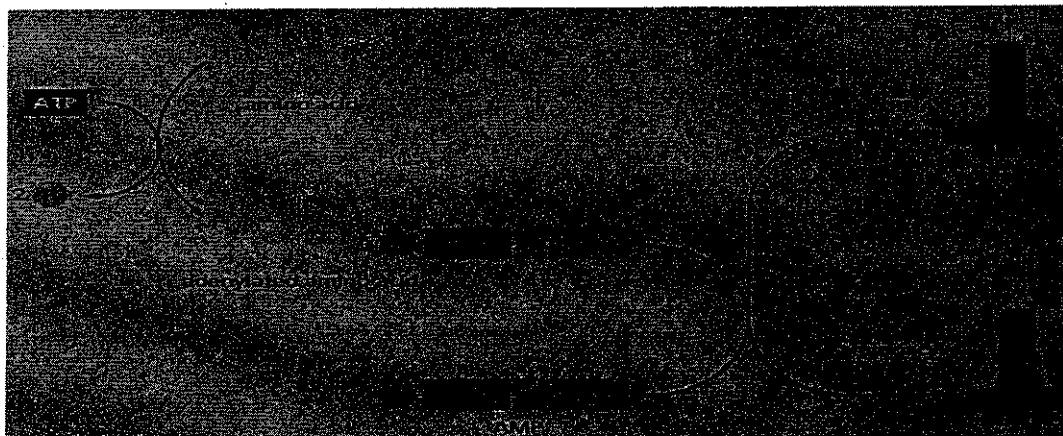
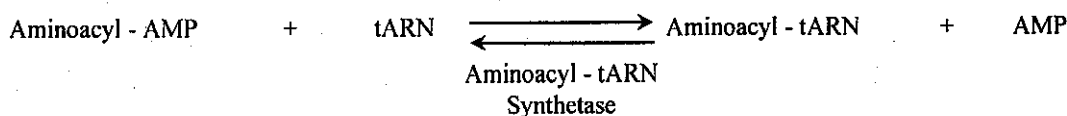
Hình thành aminoacyl-adenylate hoạt hóa:

Acid amin phản ứng với ATP có enzym xúc tác để tạo thành aminoacyl adenylate hoạt hóa. Pyrophosphat tạo ra bị thủy giải thành 2 phân tử phosphat vô cơ. Phản ứng này tạo ra một năng lượng thủy giải tự do tương đối cao, đẩy phản ứng theo chiều tạo aminoacyl-AMP trung gian:



Hình thành aminoacyl-tARN:

Aminoacyl-AMP được chuyển tới phân tử tARN tương ứng bằng synthetase:



Hình 5.5. Aminoacyl hóa tARN Các tARN và codon khởi đầu

5.2.3.2. Các tARN và codon khởi đầu

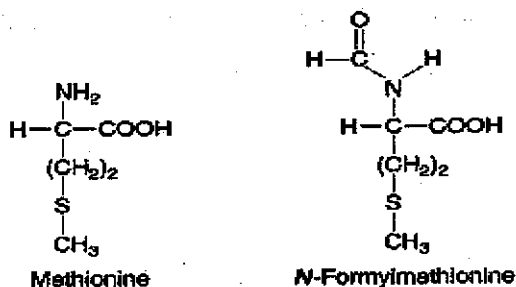
Trình tự mã của phân tử mRNA được khởi đầu với codon AUG, mặc dù ở vi khuẩn còn có codon GUG được sử dụng cho một số mRNA (khoảng 4% trình tự có GUG khởi đầu đã được biết). Rất hiếm khi các codon khác (AUU, UUG) hoạt động như codon khởi đầu.

Dấu hiệu khởi đầu là codon AUG, cũng là codon mã hóa cho methionin (Met). Mặc dù, chỉ có một codon mã hóa cho Met, nhưng trong tế bào chất lại hiện diện 2 loại tARN có khả năng mang Met đến kết hợp với codon đó:

1. tARN^{Met} kết hợp với codon AUG nằm giữa phân tử mARN, có nhiệm vụ gắn Met vào chuỗi polypeptid đang hình thành.

2. tARN^{Met} kết hợp với codon AUG khởi động, mang Met mở đầu vào ribosom trong quá trình khởi đầu dịch mã.

Ở vi khuẩn, tARN^{Met} khởi đầu không vận chuyển methionin mà vận chuyển N-formyl-methionin (fMet), do đó tARN này được gọi là tARN^{fMet} (Hình 5.6).



Hình 5.6. N-formyl-methionin

Sự tổng hợp tất cả protein ở vi khuẩn đều bắt đầu bằng formyl-methionin, mặc dù nhóm formyl, cũng như methionin đầu tiên, thường bị loại bỏ sau khi tổng hợp xong phân tử protein. Ví dụ, ở *E. coli*, chỉ có 45% các protein hoàn chỉnh bắt đầu bằng methionin. Mặc dù codon khởi đầu không phải AUG (GUG chẳng hạn) vẫn được giải mã bởi fMet-tARN^{fMet}. Phân tử tARN liên quan tới codon AUG bên trong mARN làm nhiệm vụ vận chuyển methionin được viết tắt là tARN^{Met}.

Ở tế bào nhân thật, có các tARN chuyên biệt cho các codon AUG khởi đầu và AUG bên trong mARN, những tARN^{Met} mở đầu chỉ vận chuyển methionin khởi đầu được gọi là tARN^{iMet} và tARN^{mMet} cho các codon AUG bên trong (Bảng 5.1).

Bảng 5.1. Các codon khởi đầu và tARN khởi đầu

	Vi khuẩn	Tế bào nhân thật
Codon khởi đầu	AUG (GUG)	AUG
tARN khởi đầu	tARN ^f	tARN ⁱ
Acid amin được vận chuyển	N-formyl-methionin	Methionin

5.2.4. Các yếu tố dịch mã

Nhiều protein và enzym tham gia quá trình dịch mã để đảm bảo sự lắp ráp chính xác bộ máy dịch mã và sự chính xác các bước dịch mã. Các yếu tố này được chia làm yếu tố khởi đầu, nối dài và kết thúc.

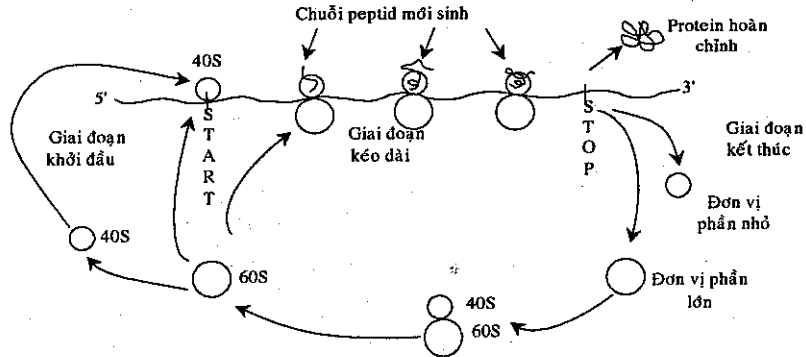
5.3. DIỄN BIẾN DỊCH MÃ Ở RIBOSOM (CHU TRÌNH RIBOSOM)

Quá trình dịch mã mARN được mô tả trong Hình 5.7, bao gồm các giai đoạn: khởi đầu, nối dài và kết thúc.

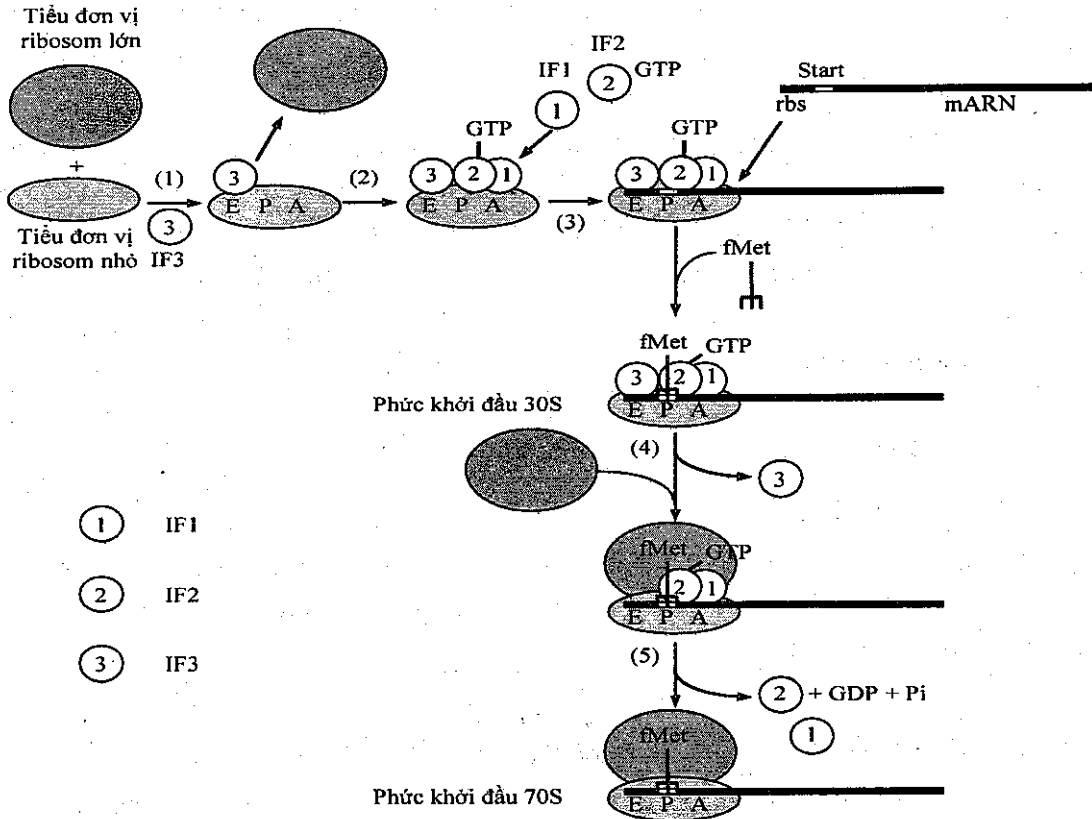


5.3.1. Khởi đầu

Giai đoạn khởi đầu, có sự tham gia của 2 tiểu đơn vị của ribosom, một phân tử mARN (được sắp thẳng hàng với ribosom) và tARN khởi đầu mang methionin hoặc formyl - methionin ở tế bào nhân nguyên thủy (Bảng 5.2). Và các yếu tố khởi đầu như IF₁, IF₂, IF₃ ở vi khuẩn và 10 yếu tố eIF (IF của tế bào nhân thật) như eIF₁, eIF₂, eIF₃, eIF - 4A, eIF - 4B... ở tế bào nhân thật (Bảng 5.2).



Hình 5.7. Chu trình ribosom tế bào nhân thật



Hình 5.8. Sự khởi đầu dịch mã ở vi khuẩn



5.3.1.1. Khởi đầu chuỗi peptid ở tế bào nhân nguyên thủy

Đặc trưng của sự khởi đầu chuỗi peptid ở vi khuẩn là sự tương tác trực tiếp giữa mARN và rARN của tiểu đơn vị nhỏ 30S. Sự tương tác này dẫn đến việc gắn mARN vào ribosom và quan trọng hơn, là nhận diện được codon khởi đầu của trình tự mARN. Các mARN vi khuẩn có chứa một trình tự để gắn ribosom (ribosom binding site-rbs) ở đầu 5' khá ngắn, bao gồm trình tự polypurin ngắn có thể bổ sung với một trình tự giàu pyrimidin ở tận cùng 3' của rARN 16S. Shine và Dalgarno cho rằng việc ghép đôi base giữa các trình tự này có thể là cơ chế làm cho ribosom của vi khuẩn hướng tới đúng codon khởi đầu của mARN. Do đó, người ta gọi trình tự ngắn này là trình tự Shine-Dalgarno.

Sự khởi đầu chuỗi peptid ở vi khuẩn bao gồm sự nhận diện trình tự Shine-Dalgarno, nằm gần codon khởi đầu của trình tự mã hóa mARN. Do đó, sự khởi đầu dịch mã không nhất thiết bắt đầu từ đầu 5' của mARN. Đúng vậy, các mARN tế bào nhân nguyên thủy có chứa tới vài vùng mã hóa, đúng hơn là vi khuẩn thường có polycistron mARN, mỗi cistron có vị trí gắn ribosom và codon khởi đầu riêng. Ví dụ mARN mang mã cho 3 loại protein khác nhau ở *lac operon*.

Khởi đầu dịch mã ở tế bào nhân nguyên thủy bắt đầu với tiểu đơn vị 30S và được xúc tác bởi các yếu tố khởi đầu IF1, IF2 và IF3.

– IF1 ngăn không cho aminoacyl-tARN gắn vào vị trí sẽ trở thành vị trí A của ribosom hoàn chỉnh.

– IF2 là một GTPase (protein gắn và thủy phân GTP). IF2 tương tác với 3 thành phần của bộ máy khởi đầu là tiểu đơn vị 30S, IF1 và fMet-tRNA^{Met}, do đó giúp fMet-tRNA^{Met} gắn vào tiểu đơn vị 30S tại vị trí P và ngăn các aminoacyl-tARN khác gắn vào.

– IF3 gắn vào tiểu đơn vị 30S và ngăn nó tái hợp với tiểu đơn vị 50S hay gắn với aminoacyl-tARN khác. Cần lưu ý là sự khởi đầu dịch mã cần tiểu đơn vị 30S tự do, do đó IF3 sẽ gắn với tiểu đơn vị 30S khi kết thúc quá trình dịch mã để giúp tách ribosom thành các tiểu đơn vị riêng biệt.

Tiến trình khởi đầu gồm các bước (Hình 5.8):

1. IF3 gắn vào tiểu đơn vị nhỏ tại vị trí E.
2. IF1 gắn vào tiểu đơn vị nhỏ tại vị trí A. IF2 (có hoạt tính GTPase) gắn vào IF1 và chồm qua vị trí P để có thể tiếp xúc với tARN^{Met}.
3. Tiểu đơn vị nhỏ của ribosom gắn mARN nhờ bắt cặp base giữa vị trí gắn ribosom (rbs) trên mARN với rARN 16S sao cho codon khởi đầu nằm tại vị trí P. tARN^{Met} gắn vào tiểu đơn vị nhỏ tại vị trí P nhờ sự trợ giúp của IF2. Sự gắn của cả hai ARN (mARN và tARN^{Met}) vào ribosom có thể xảy ra theo thứ tự bất kỳ và độc lập với nhau.

4. Khi codon khởi đầu bắt cặp đúng với anticodon trên tARNf^{Met}, tiểu đơn vị nhỏ thay đổi cấu hình làm cho IF3 không gắn được và bị phóng thích. Khi không có IF3, tiểu đơn vị lớn mới gắn vào tiểu đơn vị nhỏ.

5. Sự gắn tiểu đơn vị lớn kích hoạt GTPase (trên IF2-GTP) thủy giải GTP thành GDP. IF2-GDP có ái lực yếu với ribosom và tARNf^{Met} dẫn đến phóng thích IF2-GDP và IF1 khỏi ribosom và giải phóng vị trí A để tiếp nhận các tARN đã hoạt hóa khác cho giai đoạn nối dài.

Bảng 5.2. Các yếu tố khởi đầu của tế bào nhân nguyên thủy và tế bào nhân thật và chức năng của chúng

Yếu tố ở vi khuẩn	Yếu tố ở tế bào nhân thật	Chức năng
IF-1	eIF-1A	Ngăn không cho aminoacyl-tARN thường gắn vào vị trí A
IF-2	eIF-2	Giúp định vị Met-tARNi (hoặc fMet-tARN) vào vị trí P. eIF-2 gắn trực tiếp với Met-tARNi.
IF-3	eIF-3	Ngăn cản sự kết hợp của các tiểu đơn vị ribosom
	eIF-4F (eIF-4E, eIF-4G, eIF-4A) eIF-4B	Nhóm yếu tố gắn vào chóp 5' của mARN và tháo xoắn cấu trúc bậc 2 tạo điều kiện cho ribosom gắn vào và quét tìm codon khởi đầu
	eIF-5B	Giúp phức eIF2-GTP/tRNA ^{Met} gắn vào ribosom. Giải phóng các yếu tố khởi động khỏi ribosom sau khi gắn tiểu đơn vị lớn

5.3.1.2. Khởi đầu chuỗi peptid ở tế bào nhân thật

Quá trình khởi đầu ở chuỗi peptid ở tế bào nhân thật phức tạp hơn và có nhiều điểm khác biệt với tế bào nhân nguyên thủy. Trong khi, ở *E. coli* chỉ có 3 yếu tố khởi đầu thì ở tế bào nhân thật có tới 10. Các khác biệt chủ yếu trong cách thức gắn kết mARN và xác định codon khởi đầu.

Trình tự dẫn mARN tế bào nhân thật không được xem như trình tự Shine-Dalgarno ở tế bào nhân nguyên thủy. Đầu tận cùng 3' của rARN 18S (tương đương với rARN 16S ở vi khuẩn) không có trình tự giàu pyrimidin. Và mARN ở tế bào nhân thật là monocistron, chúng chỉ chứa một trình tự mã hóa và khởi đầu.

Marilyn Kozak đã đề xuất mô hình “quét” để nhận diện codon khởi đầu ở tế bào nhân thật. Trước tiên tiểu đơn vị nhỏ 40S của ribosom gắn vào đầu tận cùng 5' của mARN, sau đó di chuyển dọc theo mARN tới khi gặp codon khởi đầu AUG. Năng lượng cho sự di chuyển lấy từ sự thủy phân ATP và sự tương tác khởi đầu với mARN có thể liên quan tới

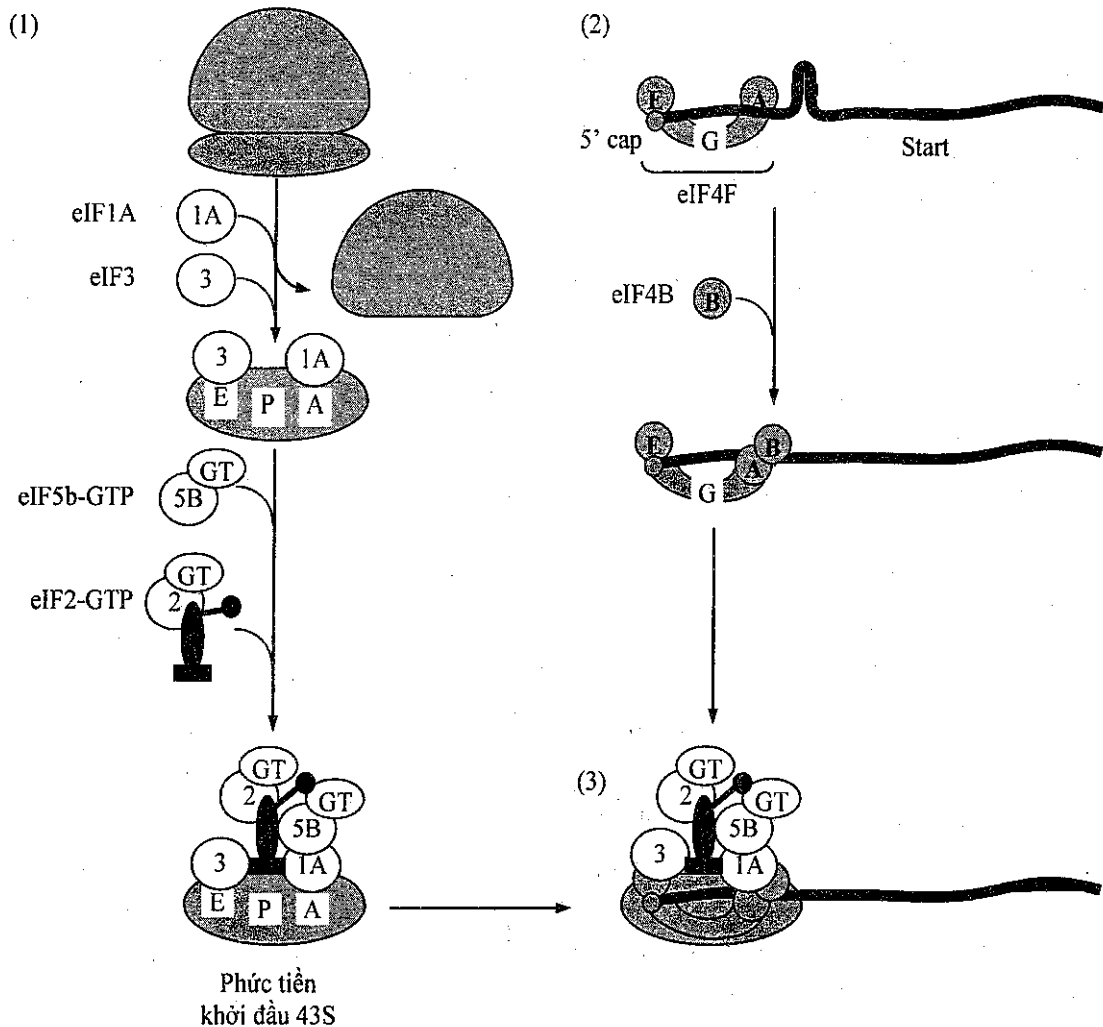
cấu trúc chóp 5' (CAP). Codon AUG có thể chỉ là codon mã hóa bình thường nhưng cũng có thể là codon khởi đầu, nếu nó được định vị trong một “ngữ cảnh” đúng. Ngữ cảnh biết tới nhiều nhất được trình bày ở (Hình 5.9), do Kozak phát hiện, do đó còn được gọi là trình tự Kozak.



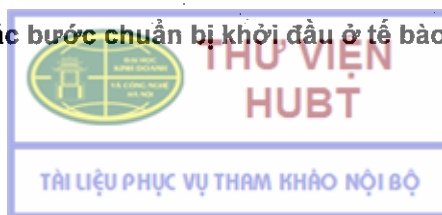
Hình 5.9. Ngữ cảnh nhận diện codon khởi đầu

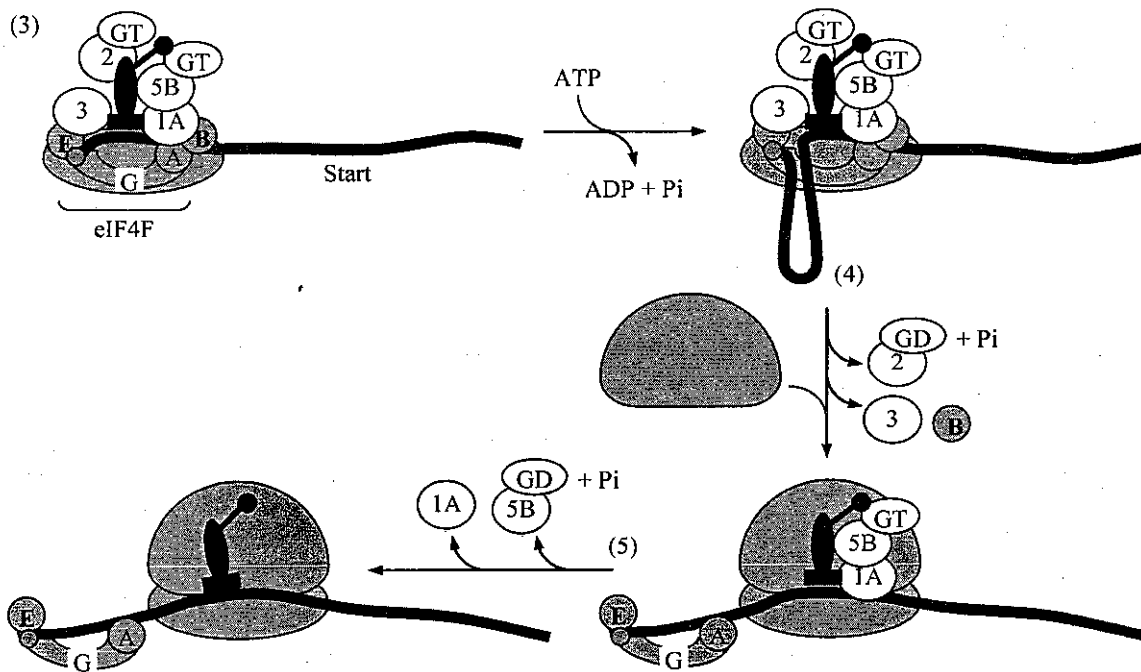
Những nucleotid viết đậm là những nucleotid quan trọng trong trình tự chung.

Tiểu đơn vị 40S sẽ bỏ qua codon AUG bình thường cho đến khi định vị tại codon AUG khởi đầu nằm trong ngữ cảnh tốt. 95% các trường hợp khởi đầu ở tế bào nhân thật xảy ra tại codon khởi đầu AUG, hầu như không dùng tới GUG.



Hình 5.10. Các bước chuẩn bị khởi đầu ở tế bào nhân thật





Hình 5.11. Định vị codon khởi đầu và hoàn tất khởi đầu ở tế bào nhân thật

Nhìn chung quá trình khởi đầu dịch mã ở tế bào nhân thật phức tạp hơn rất nhiều so với tế bào nhân nguyên thủy và có hơn 30 peptid tham gia. Tiến trình khởi đầu dịch mã ở tế bào nhân thật diễn ra như sau (Hình 5.10 và Hình 5.11):

1. Chuẩn bị tiểu đơn vị 40S (Hình 5.10):

- a) eIF1A và eIF3 gắn vào tiểu đơn vị nhỏ tại vị trí A và E.
- b) eIF1A giúp eIF5B-GTP gắn vào tiểu đơn vị nhỏ.

c) eIF5B-GTP giúp phức eIF2-GTP với $tRNAi^{Met}$ gắn vào, eIF5B-GTP và eIF2-GTP định vị $tRNAi^{Met}$ vào vị trí P.

2. Chuẩn bị mRNA (Hình 5.10):

a) eIF4E gắn vào chóp 5' của mRNA, giúp eIF4A và eIF4G gắn vào mRNA, 3 protein này tạo thành eIF4F hoàn chỉnh.

b) eIF4B được gắn tiếp vào và hoạt hóa hoạt tính helicase trên eIF4F để loại bỏ cấu trúc bậc 2 ở đầu 5' của mRNA.

3. eIF4F/B giúp tiểu đơn vị nhỏ đã gắn $tRNAi^{Met}$ tiếp xúc với mRNA từ phía chóp 5' nhờ tương tác giữa eIF4F và eIF3. Tiểu đơn vị nhỏ trượt trên mRNA theo hướng 5'→3' đến khi codon khởi đầu AUG vào đúng vị trí P. Sự trượt của mRNA xảy ra phụ thuộc ATP nhờ hoạt tính helicase trong eIF4F. Việc nhận diện codon khởi đầu liên quan đến ngữ cảnh của trình tự Kozak như đã nói ở trên. Ở đây cần lưu ý sự khác biệt với tế bào nhân nguyên thủy: việc gắn mRNA vào ribosom luôn diễn ra sau khi ribosom đã gắn $tRNAi^{Met}$. (Hình 5.11).

4. Sự bắt cặp chính xác giữa anticodon của tRNA^{Met} và AUG dẫn đến phóng thích eIF2 và eIF3. Mất eIF2 và eIF3 tạo điều kiện cho tiểu đơn vị lớn gắn vào tiểu đơn vị nhỏ. (Hình 5.11).

5. Sự gắn tiểu đơn vị lớn dẫn đến sự phóng thích các yếu tố khởi đầu còn lại do kích thích thủy giải GTP nhờ eIF5B. (Hình 5.11).

5.3.2. Nối dài

Quá trình dịch mã trên mRNA được thực hiện theo hướng 5' → 3' và chuỗi polypeptid được bắt đầu tổng hợp ở đầu tận cùng -N. Quá trình này tương đối giống nhau ở tế bào nhân nguyên thủy và tế bào nhân thật, cả về cơ chế lẫn các yếu tố tham gia.

Sự nối dài là một quá trình có chu kỳ, mỗi chu kỳ giải mã một codon và thêm một acid amin vào chuỗi polypeptid mới sinh. Sự nối dài cần đến các protein, được gọi là các yếu tố nối dài (EF). Ở vi khuẩn, các yếu tố này là EF-Tu, EF-Ts và EF-G còn ở tế bào nhân thật là EF-1α (tương ứng EF-Tu) và EF-2 (tương đương EF-G) (Ở điểm khởi đầu mỗi chu kỳ nối dài, chỉ có vị trí P bị chiếm chỗ bằng một peptidyl-tARN chứa chuỗi polypeptid mới sinh gắn với tARN mang acid amin vừa được thêm vào. Peptid được gắn vào tARN theo cùng một cách như acid amin, bằng nhóm 3'-hydroxyl ở vị trí kết thúc CCA của tARN), (Bảng 5.3).

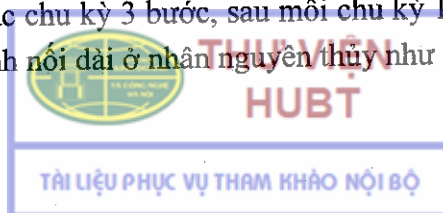
Trên ribosom có 2 vị trí A và P. Vị trí A là vị trí Aminocyl (hoặc Acceptor), và vị trí P là vị trí peptidyl, nối giữ phức hợp peptidyl-tARN, tức là chuỗi polypeptid đang hình thành vẫn còn gắn với tARN trước đó.

Bảng 5.3. Các yếu tố nối dài

Nhân nguyên thủy	Nhân thật	Chức năng
EF -Tu (Mr = 43 000)	EF-1α (Mr = 53 000)	Mang các aminoacyl-tARN tới vị trí A của ribosom, thí dụ, phức hợp (EF-1α-GTP-aminoacyl-tARN).
EF-Ts (Mr = 30 000)	EF-1βγ (Mr = 50/38 000)	Tái hồi EF-Tu hoặc EF-1α (Hình 5.13).
EF-G (Mr = 77 000)	EF-2 (Mr = 100 000)	Tham gia vào bước chuyển vị của chu kỳ nối dài.

Khi bắt đầu nối dài, vị trí P được gióng hàng với codon đã được dịch mã, trong khi đó vị trí A lúc này đang trống và gióng hàng với codon kế tiếp của mRNA. Bất kỳ aminoacyl-tARN (ngoại trừ tARN-khởi đầu) đều vào vị trí A, chúng lưu lại đây nếu anticodon khớp mã với codon tại A. Bằng cách này, aminoacyl-tARN chính xác được lựa chọn và chọn được đúng acid amin cho sự hình thành chuỗi polypeptid mới sinh.

Tiến trình nối dài gồm các chu kỳ 3 bước, sau mỗi chu kỳ 1 acid amin mới được thêm vào mạch polypeptid, tiến trình nối dài ở nhân nguyên thủy như sau (Hình 5.12):



1. Aminoacyl-tRNA có đôi mã đúng được mang vào vị trí A nhờ yếu tố EF-Tu-GTP; khi vào đúng vị trí, hoạt tính GTPase của EF-Tu-GTP được kích hoạt bởi tiểu đơn vị lớn dẫn đến thủy phân GTP thành GDP và phóng thích aminoacyl-tRNA. Sự kích hoạt GTPase chỉ xảy ra nếu tRNA vào đúng vị trí A và có anticodon đúng với mã trên mRNA.

2. Liên kết peptid được hình thành giữa aminoacyl-tRNA ở vị trí A với peptidyl-tRNA ở vị trí P nhờ enzym peptidyl transferase (do rARN 23S đảm nhiệm) và do đó chuyển polypeptid đang tổng hợp từ tRNA ở vị trí P sang aminoacyl-tRNA ở vị trí A.

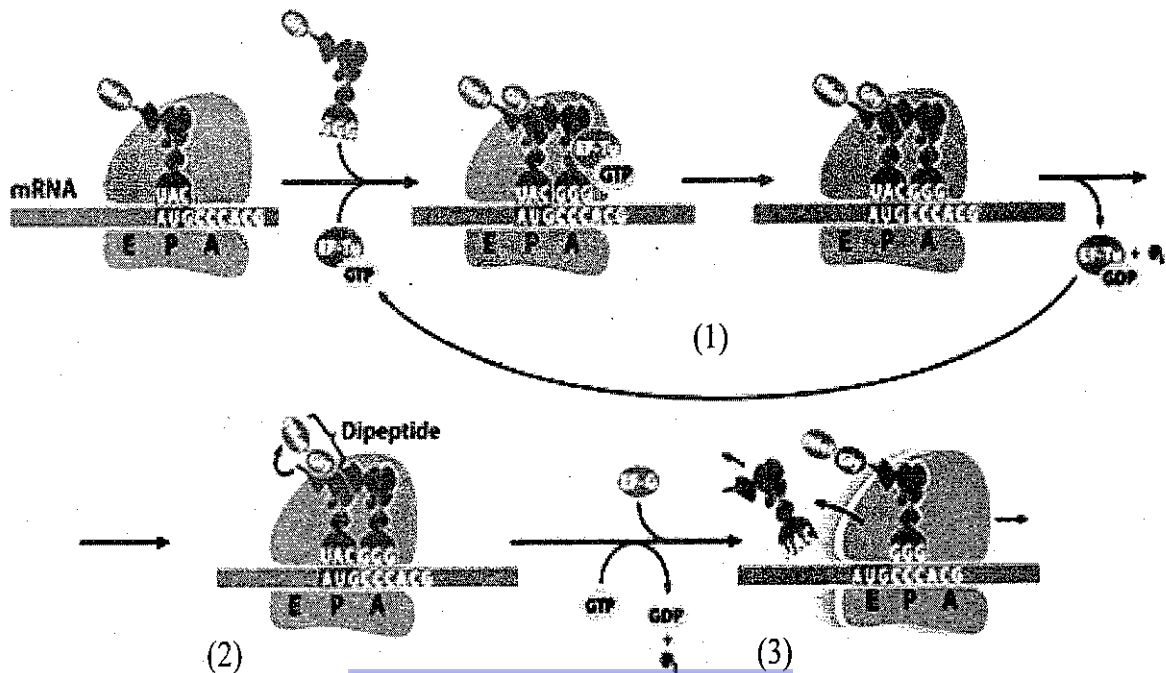
3. Peptidyl-tRNA mới hình thành ở vị trí A được chuyển vị sang vị trí P nhờ hoạt động của yếu tố kéo dài EF-G-GTP sử dụng năng lượng thủy phân GTP thành GDP để giải phóng vị trí A cho 1 chu kỳ mới. Sự chuyển vị có 3 hiện tượng xảy ra đồng thời:

- tARN vận chuyển ở vị trí P được đẩy sang vị trí E và thoát ra ngoài.
- Peptidyl-tARN di chuyển từ vị trí A sang vị trí P.
- Ribosom tiến thêm một codon trên mRNA, do đó, codon tiếp theo được đưa vào giống hệt với vị trí A.

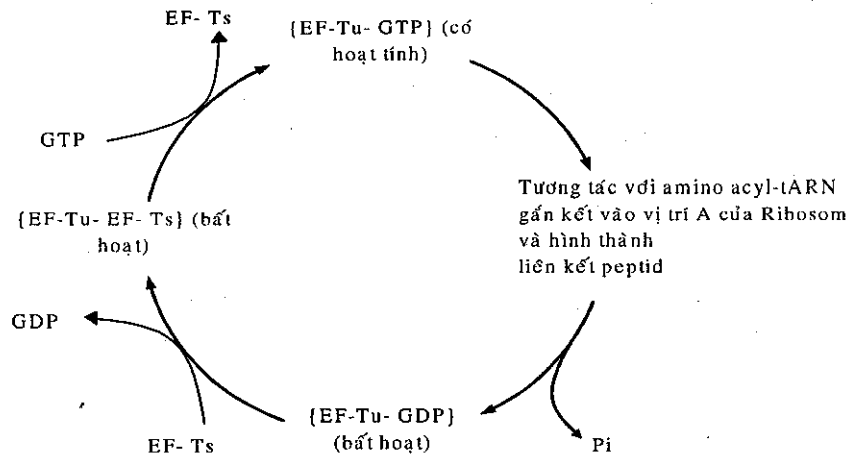
EF-Tu-GDP và EF-G-GDP cần được tái tạo lại GTP để tham gia chu kỳ kéo dài mới, cơ chế tái tạo GTP của các yếu tố nối dài như sau:

a) EF-G có ái lực yếu với GDP, do đó EF-G-GDP dễ dàng mất GDP để nhận GTP mới.

b) EF-Tu-GDP tái tạo GTP nhờ yếu tố EF-Ts: EF-Tu-GDP gắn EF-Ts dẫn đến GDP bị phóng thích, sau đó GTP mới gắn vào phức EF-Tu-EF-Ts dẫn đến phóng thích EF-Ts và tái tạo EF-Tu-GTP (Hình 5.13).



Hình 5.12. Tiến trình của một chu kỳ nối dài

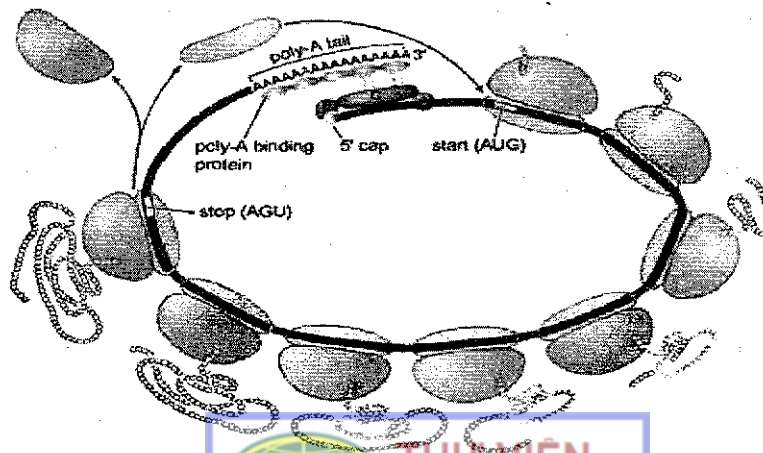


Hình 5.13. Sự tái hồi EF-Tu bởi EF-Ts và đồng phân của protein-G

Ba bước của quá trình nối dài sẽ được lặp lại cho đến khi có một codon kết thúc xuất hiện ở vị trí A. Vai trò của GTP trong quá trình nối dài chưa rõ ràng. Người ta cho rằng sự thủy phân GTP cung cấp năng lượng cho việc tạo liên kết peptid và cho bước chuyển vị, nhưng nó có thể liên quan đến mức độ chính xác của sự chuyển vị.

Tốc độ trung bình của quá trình dịch mã ở tế bào nhân nguyên thủy khoảng 15-20 codon mỗi giây, ở tế bào nhân thật khoảng 2 codon mỗi giây. Khi một ribosom bắt đầu dịch mã, ribosom khác có thể bắt đầu khởi động, do đó cùng lúc trên mRNA có thể có nhiều ribosom liên tiếp dịch mã và tạo thành cấu trúc gọi là polyribosom hay polysom.

Ở tế bào nhân thật, các yếu tố khởi đầu gắn ở chóp 5' có thể tương tác với đuôi polyA ở đầu 3' qua trung gian PABP (polyA Binding Protein) để vòng hóa mRNA. Cấu trúc vòng này giúp cho codon khởi đầu và codon kết thúc ở gần nhau, do đó giúp ribosom sau khi hoàn tất một chuỗi peptid có thể quay lại điểm xuất phát để khởi đầu chu trình dịch mã mới một cách nhanh chóng.



Hình 5.14. Cấu trúc vòng của mRNA khi dịch mã ở nhân thật

5.3.3. Kết thúc

Chu trình dịch mã chấm dứt khi khi một trong các codon kết thúc là UAA, UAG và UGA đi vào vị trí A. Ở bước kết thúc, các mã kết thúc không có anticodon. Thay vào đó là các yếu tố kết thúc (RF) làm kết thúc quá trình dịch mã. Yếu tố kết thúc RF nhận diện stop codon và kích hoạt sự kết thúc dịch mã, phóng thích chuỗi polypeptid.

Có 2 loại yếu tố kết thúc (RF - Release Factor):

– Loại I: nhận diện stop codon và thủy phân chuỗi polypeptid khỏi tARN ở vị trí P. Phản ứng này được xúc tác bởi enzym peptidyl transferase của ribosom và đòi hỏi GTP.

+ Tế bào nhân thật: chỉ có 1 yếu tố eRF1 nhận diện cả 3 codon kết thúc.

+ Tế bào nhân nguyên thủy: 2 yếu tố.

* RF1 nhận diện UAG

* RF2 nhận diện UGA

* UAA được nhận diện bởi cả RF1 và RF2.

– Loại II (RF3 và eRF3): kích hoạt sự tách yếu tố loại I ra khỏi ribosom nhờ năng lượng thủy giải GTP.

Sau khi kết thúc ribosom vẫn gắn mARN và 2 tARN tại vị trí P và E. Yếu tố tái sử dụng ribosom (RRF - Ribosome Recycling Factor) giúp thu hồi bộ máy dịch mã cho chu kỳ mới. Ở Prokaryot:

+ RRF gắn vào vị trí A, giả làm tARN, nó kéo EF-G vào ribosom và đẩy các tARN ra khỏi ribosom theo cách tương tự ở quá trình nối dài.

+ Khi tARN bị loại ra, RFF và EF-G phóng thích khỏi ribosom cùng với mARN.

IF3 có thể tham gia phóng thích mARN và sau đó gắn vào tiểu đơn vị nhỏ để ngăn không có tiểu đơn vị lớn lắp ghép. Kết quả là tiểu đơn vị nhỏ gắn IF3 và tiểu đơn vị lớn tự do để tham gia chu kỳ dịch mã mới.

5.4. NHU CẦU NĂNG LƯỢNG CHO QUÁ TRÌNH SINH TỔNG HỢP PROTEIN

Ở một vài giai đoạn của quá trình sinh tổng hợp protein cần đến năng lượng chuyển hoá và năng lượng này được lấy từ sự thủy phân ATP và GTP (Bảng 5.4).

Bảng 5.4. Nhu cầu năng lượng cho quá trình dịch mã

Giai đoạn	Tế bào nhân nguyên thủy	Tế bào nhân thật
Aminoacyl hoá tARN	ATP → AMP + PPi	ATP → AMP + PPi
Khởi đầu	GTP → GDP + Pi, kết hợp với IF – 2	GTP → GDP + Pi, kết hợp với eIF – 2 ATP → ADP + Pi, kết hợp với việc gắn chóp và tháo xoắn

Giai đoạn	Tế bào nhân nguyên thủy	Tế bào nhân thật
		mARN (chưa rõ có bao nhiêu phân tử ATP được sử dụng)
Nối dài	GTP → GDP + Pi, kết hợp với EF – Tu GTP → GDP + Pi, kết hợp với EF – G	GTP → GDP + Pi, kết hợp với EF – 1 GTP → GDP + Pi, kết hợp với EF – 2
Kết thúc	GTP → GDP + Pi	GTP → GDP + Pi

Năng lượng cần cho các giai đoạn khác nhau, có thể đóng một hoặc một số vai trò sau:

- Cung cấp năng lượng để thực hiện các phản ứng đặc biệt, như hoạt hoá acid amin bằng cách gắn chúng vào tARN.
- Cung cấp năng lượng cho một quá trình nào đó như tháo xoắn cấu trúc bậc 2 của mARN hoặc cho sự chuyển dịch ribosom trong các bước chuyển vị trong quá trình nối dài.
- Duy trì các yếu tố ở trạng thái hoạt tính như EF – Tu.
- Cung cấp năng lượng cho sự biến dạng của ribosom, như sau khi gắn kết aminoacyl – tARN vào vị trí A.

5.5. ĐỘ CHÍNH XÁC CỦA QUÁ TRÌNH DỊCH MÃ

Sai sót trong quá trình dịch mã sẽ tạo ra một protein bất hoạt, hoặc là một protein điều hoà và xúc tác không phù hợp làm thay đổi sinh lý của tế bào bình thường. Các sai sót đó là:

5.5.1. Aminoacyl hoá tARN

Một tARN có thể nhận thêm một acid amin không đúng hoặc enzym aminoacyl-tARN synthetase dùng tARN có anticodon không chuyên biệt cho các acid amin được gắn vào nó. Có một số acid amin có cấu trúc tương tự với acid amin cần được đưa vào, cho nên các enzym aminoacyl-tARN synthetase cũng phải có cách thức chọn đúng. Ví dụ, enzym isoleucyl-tARN synthetase, một mặt phải phân biệt giữa isoleucin và các acid amin nhỏ hơn như Valin, mặt khác phải phân biệt giữa isoleucin với các acid amin kỵ nước, cồng kềnh hơn như leucin, phenylalanin. Cơ chế điều chỉnh xảy ra như sau:

Enzym isoleucyl-tARN synthetase có 2 vị trí tác động, một vị trí xúc tác thành lập aminoacyl-tARN (A) và vị trí thứ 2 (H), ở sát ngay vị trí thứ nhất, thủy phân aminoacyl-tARN để cho ra acid amin tự do và tARN. Các acid amin lớn hơn isoleucin (hoặc có cấu hình không đúng) không thể đi vào vị trí acyl hóa (A) và do đó không được các enzym

synthetase sử dụng. Tuy nhiên, các acid amin nhỏ hơn isoleucin có thể đi vào vị trí A để tạo ra aminoacyl-tARNile và sẽ là nguồn gốc tiềm tàng gây ra các sai sót. Các dẫn xuất aminoacyl cũng có thể đi vào vị trí thủy phân (H), nên bị thủy phân. Vì vậy, chỉ có isoleucyl-tARN là được sản sinh chuyên biệt. 2 bước nhận biết trên làm tăng độ chính xác của sự aminoacyl hóa tARN.

5.5.2. Tương tác codon - anticodon

Phức hợp sai giữa codon và anticodon sẽ dẫn đến việc gắn vào acid amin không đúng. Các nghiên cứu đột biến vi khuẩn cho thấy các protein chuyên biệt trong các tiểu đơn vị của ribosom đóng vai trò quan trọng trong việc gắn đúng các codon và anticodon trong suốt quá trình dịch mã.

5.5.3. Kết thúc sớm, kết thúc muộn

Các sai sót của giai đoạn kết thúc có thể xảy ra do kết thúc sớm (premature termination) hoặc kết thúc muộn do đọc quá. Sự đọc quá do lỗi của bộ máy dịch mã, không nhận ra codon dừng. Tất nhiên, acid amin được đưa vào và dịch mã liên tục qua cả vùng không mã hóa 3'mARN. Sự đọc quá cũng có thể do một đột biến ở codon đối mã nào đó làm cho codon đối mã này giải mã được codon kết thúc.

Người ta ước tính rằng, sai sót tăng lên ở tỉ lệ 1/3000 gốc acid amin được kết hợp. Nghĩa là, nếu một phân tử protein có 300 gốc acid amin thì chỉ 1 phân tử protein trong 10 phân tử có chứa 1 lỗi và thường những lỗi này sẽ được định vị mà không quan trọng đối với cấu trúc hoặc chức năng của protein.

5.6 CÁC YẾU TỐ CỦA QUÁ TRÌNH DỊCH MÃ

Một số lớn các chất được biết ức chế một hoặc nhiều bước trong quá trình sinh tổng hợp protein (Bảng 5.5). Trong số này có những chất ức chế chọn lọc sự tổng hợp protein vi khuẩn mà không ảnh hưởng tới sinh tổng hợp protein tế bào nhân thật. Và chúng có triển vọng sử dụng như những kháng sinh loại trừ sự nhiễm khuẩn mà không ảnh hưởng đến quá trình chuyển hóa protein của tế bào chủ (Bảng 5.6).

Bảng 5.5. Các chất ức chế dịch mã

Chất ức chế	Chọn lọc trên	Tác động ở giai đoạn
Chloramphenicol	Tế bào nhân nguyên thủy	Nối dài: ức chế peptidyl transferase
Cycloheximide	Tế bào nhân thật	Nối dài: chưa rõ cơ chế
Erythromycin	Tế bào nhân nguyên thủy	Nối dài: ức chế sự chuyển peptid hóa
Acid fusidic	Cả 2 giới (tế bào nhân nguyên thủy và tế bào nhân thật)	Nối dài: ức chế sự phóng thích của EF-G (hoặc EF-2): phức hợp GDP

Chất ức chế	Chọn lọc trên	Tác động ở giai đoạn
Kanamycin	Cả 2 giới	Nối dài: gây đọc sai
Neomycin	Cả 2 giới	Khởi đầu và nối dài: mRNA đọc nhầm
Puromycin	Cả 2 giới	Nối dài: gây kết thúc sớm
Sparsomycin	Cả 2 giới	Nối dài: ức chế peptidyl transferase
Spectinomycin	Tế bào nhân nguyên thủy	Nối dài: ức chế chuyển peptid hóa
Streptomycin	Tế bào nhân nguyên thủy	Nối dài: gắn vào 30S và ảnh hưởng tới tương tác codon - anticodon gây ra sự đọc nhầm của mRNA
Tetracyclin	Tế bào nhân nguyên thủy	Nối dài: phong bế việc gắn aminoacyl-tARN vào vị trí A

Ví dụ, chloramphenicol, một kháng sinh có phổ kháng khuẩn rộng. Chất này và các chất ức chế dịch mã được dùng như các kháng sinh, trong khi đó các chất khác thì chỉ được dùng để thăm dò cơ chế của chính sự dịch mã.

Bảng 5.6. Áp dụng trị liệu của các chất ức chế sự tổng hợp protein

Hợp chất	Áp dụng và chỉ dẫn
Chloramphenicol	Phổ kháng khuẩn rộng dùng trong điều trị thương hàn
Erythromycin	Phổ kháng khuẩn giống penicillin và do đó thường dùng điều trị các bệnh nhân nhạy cảm với penicillin
Acid fusidic	Phổ kháng khuẩn hẹp, dùng cho các <i>Staphylococci</i> đề kháng penicillin
Kanamycin	Dùng cho nhiễm trùng Gram (-) nặng
Spectinomycin	Điều trị lao
Nhóm tetracyclin	Kháng sinh có phổ kháng khuẩn rộng
Streptomycin	Kháng sinh có phổ kháng khuẩn rộng
Neomycin	Kháng sinh có phổ kháng khuẩn rộng

Tóm tắt:

Sự dịch mã là một quá trình phức tạp với sự tham gia của nhiều thành phần:

– ARN thông tin: mang thông tin di truyền mã hóa dưới dạng codon, mỗi codon là một tổ hợp 3 nucleotid mã hóa cho một acid amin.

– ARN ribosom: là thành phần của ribosom, nơi hình thành chuỗi polypeptid đang tổng hợp.

– ARN vận chuyển: yếu tố mang các acid amin tương ứng tới các codon trên khuôn ARN đến gắn vào chuỗi polypeptid đang hình thành tại ribosom

– Acid amin và ribosom

Quá trình dịch mã là các chu kỳ ribosom, bao gồm 3 giai đoạn:



– Khởi đầu: trước tiên là hình thành một phức hợp gồm 3 thành phần: (1) tiểu đơn vị nhỏ ribosom, (2) tARN mang methionin (Met-tARN^{Met}), (3) mARN. Một yếu tố khởi đầu sẽ phát hiện codon khởi đầu AUG giúp phức hợp và tiểu đơn vị lớn của ribosom gắn vào, và sự dịch mã bắt đầu.

– Nối dài: các yếu tố nối dài tạo ra các bước chuyển vị. Các acid amin được đưa đến ribosom ở dạng aminoacyl-tARN theo từng codon riêng lẻ.

– Kết thúc: khi các yếu tố kết thúc nhận ra dấu hiệu kết thúc trên mARN, sự dịch mã dừng lại.

Sự khác biệt giữa các thành phần và cơ chế chi tiết của sự dịch mã ở vi khuẩn và tế bào nhân thật là nền tảng cho việc sử dụng rộng rãi các chất kháng sinh trong việc ức chế chọn lọc quá trình sinh tổng hợp protein ở vi khuẩn.

CÂU HỎI

- Năng lượng giải phóng được từ GTP thành GDP + Pi dùng để:
 - Ghép tARN khởi đầu với tiểu đơn vị 50S
 - Ghép tARN khởi đầu với tiểu đơn vị 30S
 - Bước chuyển vị trong nối dài
 - Hoạt hoá acid amin
 - Gắn kết mARN với ribosome
- Trong quá trình dịch mã:
 - Mỗi tARN có một aminoacyl synthetase tương ứng
 - Một aminoacyl synthetase chung cho tất cả acid amin
 - tARN – aminoacyl synthetase kéo dài chuỗi peptid
 - Một aminoacyl synthetase cho mỗi loại acid amin
 - tARN – aminoacyl synthetase xúc tác sự chuyển vị
- Trình tự Shine – Dalgarno là:
 - Vị trí gắn ribosom
 - Yếu tố khởi lắp ráp rARN
 - Yếu tố khởi đầu dịch mã
 - Vị trí kết thúc dịch mã
 - Là trình tự codon khởi đầu
- Acid amin khởi đầu chuỗi peptid ở tế bào nhân nguyên thủy:
 - Formyl – methionin
 - Methyl – Methionin
 - Methionin
 - AUG
 - GUG
- Vai trò của eIF – 2 trong tổng hợp protein ở tế bào nhân thật:
 - Mang aminoacyl – tARN tới vị trí A
 - Gắn Met – tARN vào ribosom

- c) Hoạt hoá acid amin cần để nối dài
 d) Tái hồi EF – Tu
 e) Tái hồi EF – 1 α
6. Chuỗi peptid đang hình thành gắn vào:
 a) mARN
 b) Tiểu đơn vị nhỏ
 c) Tiểu đơn vị lớn
 d) Vị trí P
 e) Vị trí A
7. Sự chuyển vị ribosom có các hiện tượng:
 a) tARN vận chuyển xong được tách khỏi vị trí P
 b) Peptidyl – tARN di chuyển từ vị trí A sang vị trí P
 c) Ribosom tách ra để gắn vào codon kế tiếp
 d) Ribosom chuyển vị từng bước
 e) a, b và d
8. Yếu tố nào **không** tham gia kết thúc dịch mã ở vi khuẩn?
 a) RF – 1
 b) RF – 2
 c) RRF
 d) EF – G
 e) EF – Ts
9. Tác hại của streptomycin trong quá trình dịch mã của vi khuẩn:
 a) Ức chế chuyển peptid hoá
 b) Ức chế peptidyl transferase
 c) Phong bế việc gắn aminoacyl – tARN vào vị trí A
 d) Tương tác codon – anticodon gây đọc nhầm của mARN
 e) Gây kết thúc sớm quá trình dịch mã
10. Ảnh hưởng của tetracyclin trong dịch mã ở vi khuẩn:
 a) Ức chế chuyển peptid hoá
 b) Ức chế peptidyl transferase
 c) Phong bế việc gắn aminoacyl – tARN vào vị trí A
 d) Tương tác codon – anticodon gây đọc nhầm của mARN
 e) Gây kết thúc sớm quá

Bài 6

ĐIỀU HOÀ HOẠT ĐỘNG GEN

MỤC TIÊU

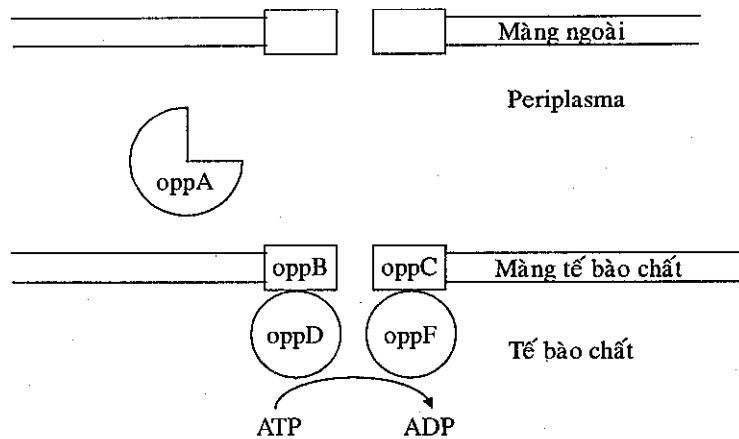
- Tóm lược được các phương thức điều hoà hoạt động gen ở tế bào nhân nguyên thủy.
- Mô tả được cơ chế điều hoà của tế bào vi khuẩn để đáp ứng với các thay đổi trong môi trường sống.
- Giải thích được cách thức tương tác giữa các protein và các trình tự ADN đặc biệt và các phân tử khác trong quá trình điều hoà hoạt động gen.

6.1. MỞ ĐẦU

Sự điều hoà hoạt động gen vi khuẩn chính là sự điều hoà trao đổi chất, là điều khiển tốc độ của các quá trình sinh hoá bằng cách thay đổi thuận nghịch hàm lượng các chất có bản chất protein tham gia vào các quá trình này hoặc là điều hoà hoạt tính của chúng.

Trong đa số các quá trình sinh hoá, các chất có dẫn xuất protein đều là những chất xúc tác các quá trình sinh học, gọi là enzym. Song có một vài quá trình, ví dụ vận chuyển qua màng sinh học được thực hiện bởi protein, trong đó protein không đóng vai trò xúc tác cho sự biến đổi hoá học mà "nhận biết" sự vận chuyển (translocation) cơ chất.

Ở vi khuẩn có ba loại enzym permease vận chuyển qua màng: permease dipeptid, permease tripeptid và oligopeptid permease. Oligopeptid permease có oppA, B, C, D, F lần lượt chuyển vận chất lẫn nhau và qua màng, trong đó oppA có tính đặc hiệu.



Hình 6.1. Các permease ở vi khuẩn

Có hai mức độ điều hoà cơ bản: mức độ điều hoà sinh tổng hợp protein và mức độ điều hoà hoạt tính của chúng trong quá trình thực hiện chức năng.

Trong phần này, chúng ta chỉ xét đến sự điều hoà sinh tổng hợp protein.

Mức độ điều hoà quá trình sinh tổng hợp protein gồm nhiều giai đoạn: các thời kỳ chuẩn bị sao chép bộ gen và phiên mã nghĩa là sinh tổng hợp các ARN, thời kỳ kết thúc – dịch mã tạo ra những phân tử protein, mất xích di truyền cuối cùng của tế bào.

6.2. ĐIỀU HOÀ QUÁ TRÌNH SAO CHÉP

Quá trình sao chép, phiên mã và dịch mã đều gồm ba giai đoạn: khởi đầu, kéo dài và kết thúc.

Những đặc điểm cơ bản của quá trình sao chép ADN là cơ chế bán bảo tồn, sự gián đoạn (đoạn Okazaki) và sự kéo dài của mỗi ARN trong mỗi đoạn sao chép, luôn luôn xảy ra theo hướng 5' đến 3'. Kết thúc giai đoạn sao chép, các đoạn mỗi (ARN_o) bị phân huỷ và các đoạn Okazaki nối lại với nhau nhờ enzym ligase. Dĩ nhiên, sợi sớm sẽ được tổng hợp liên tục theo hướng 5' đến 3'.

Ở tế bào nhân nguyên thuỷ có hệ thống đa enzym: ADN – polymerase I, II, III, ADN – ligase. Ngoài ra, để lơi xoắn và tháo xoắn có helicase và topoisomerase I, phục hồi dạng siêu xoắn có topoisomerase II (ADN – gyrase) phụ thuộc ATP.

Ngoài sao chép bộ gen, còn có sự sửa chữa và tái tổ hợp, đồng thời có sự sao chép những yếu tố di truyền ngoài nhân – plasmid.

Người ta xác định được khoảng 15 locus di truyền (*Dna*) kiểm soát sự sao chép ADN. Mỗi một gen như vậy mã hoá cho một polypeptid cần thiết cho quá trình sao chép. Ví dụ, *DnaA* cần cho sự tổng hợp primer (ARN_o) và đóng vai trò điều hoà dương trong quá trình này. Song việc tạo ra sản phẩm này lại tổn hại đến quá

trình tự điều chỉnh. Một số protein có tính đa chức năng. Ví dụ, protein do *DnaC* tổng hợp lại cần thiết cho việc gắn primer vào chuỗi ADN và sau đó protein này trở thành thành phần của phức hệ sao chép và tham gia vào quá trình kéo dài.

Tất nhiên sự điều hoà chặt chẽ hơn cả xảy ra ở giai đoạn khởi động sao chép. Bên cạnh các protein của các cistron *DnaA*, *DnaC* vừa kể trên, còn có sự tham gia của các protein khác sinh ra từ các locus *DnaB*, *DnaI* và *DnaP*. Sự sao chép ADN chịu sự kiểm soát dương và kiểm soát âm của các protein – điều hoà.

Nói một cách tổng quát, kiểm soát dương là sự tích lũy chất hoạt hoá cho đến ngưỡng đủ để khởi động một chu trình sao chép mới, cần thiết cho sự cân bằng với việc nhân đôi của sinh khối tế bào.

Ngược lại sự kiểm soát âm, khi chất kiềm hãm cần được tổng hợp một cách giới hạn tiếp ngay khi bắt đầu chu trình sao chép trước (chất kiềm hãm như vậy có thể là sản phẩm của gen nằm gần vị trí bắt đầu sao chép, mà sự phiên mã của nó chỉ được thực hiện trong giai đoạn sao chép của vùng ADN này).

Trong quá trình tăng trưởng của tế bào, chất kiềm hãm bị yếu dần và trong thời điểm tế bào nhân đôi thì mức độ kiềm hãm của nó hạ thấp dưới mức đẳng hạn.

6.3. ĐIỀU HOÀ QUÁ TRÌNH PHIÊN MÃ

Trong mỗi tế bào, không phải gen nào cũng hoạt động như gen nào. Trong khi sản phẩm của gen này cần được tổng hợp liên tục, thì sản phẩm của gen khác lại chỉ được tổng hợp trong những giai đoạn cần thiết tùy thuộc môi trường. Thậm chí, khi gen đã “mở” thì hàm lượng protein tương ứng cũng bị kiểm soát. Do đó hoạt động của hầu hết các gen đều được điều hoà trong một hay một vài con đường làm cho hiệu ứng sử dụng của tế bào về mặt năng lượng đạt được cao nhất. Cơ chế điều hoà này đối với sự biểu hiện gen có thể xảy ra ở các mức độ khác nhau. Sự điều hoà gen có thể xảy ra ngay tại thời điểm phiên mã hay trong quá trình dịch mã. Sau quá trình dịch mã, một số protein có thể bị biến đổi cấu hình để thành protein chức năng.

Các gen phiên mã ARN được gọi là các gen cấu trúc. Các protein được dịch mã từ mARN có thể là enzym hay không phải enzym. Trong số các protein không phải enzym có các *protein điều hoà*. Chúng sẽ gắn với những trình tự ADN chuyên biệt được gọi là operator, để kiểm soát quá trình phiên mã của các gen đặc hiệu. Những gen tổng hợp nên các phân tử protein điều hoà được gọi là các gen điều hoà. Mỗi gen cấu trúc (hay một nhóm gen cấu trúc) đều đứng sau một trình tự (gọi là trình tự khởi động) được nhận biết bởi ARN polymerase. Khi polymerase gắn vào vùng khởi động thì sự phiên mã xảy ra: Sợi có nghĩa của ADN sẽ được phiên mã thành mARN. Operon là đơn vị phiên mã bao gồm tối thiểu là vùng khởi động và các gen mã hóa mARN cho một hay một vài chuỗi polypeptid. Một operon có thể chứa một hay nhiều vị trí điều hoà hơn so với số vùng khởi động.

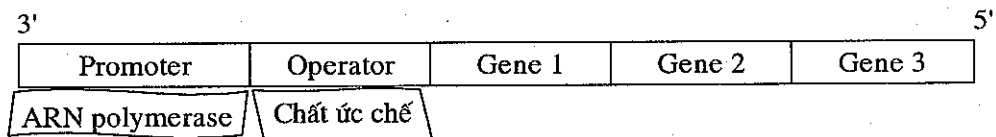


Hoạt động phiên mã của gen có thể không được điều hoà nếu sản phẩm của chúng không liên quan tới điều kiện môi trường. Có thể nói những sản phẩm không phụ thuộc vào môi trường như vậy được tổng hợp liên tục. Tuy nhiên, số lượng của những sản phẩm "không được điều hoà" này vẫn có thể thay đổi phụ thuộc vào ái lực tương đối với promoter của ARN polymerase. Các promoter có ái lực cao sẽ tạo nhiều sản phẩm của gen hơn là có ái lực thấp.

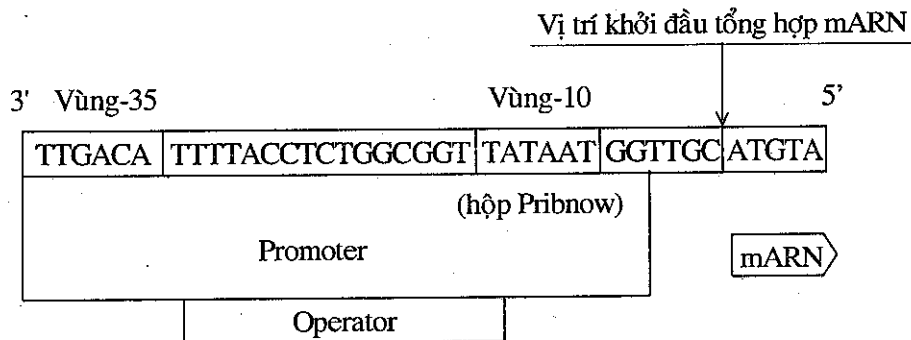
Đối với những protein đòi hỏi một điều kiện môi trường nhất định, thì sự hoạt động của các gen tạo ra chúng thường được điều hoà bởi một hay một số protein điều hoà.

Operator là một trình tự ADN của operon có một protein điều hoà được gắn vào đó và protein này được gọi là **protein ức chế**.

Operator có thể nằm cạnh promoter:

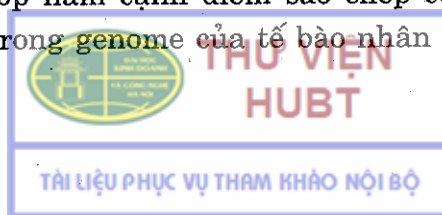


Operator có thể phủ lên một vùng của promoter: ví dụ operator của λ - phage.



Sự gắn của protein ức chế vào operator sẽ ngăn cản sự phiên mã của một số gen cấu trúc trong cùng operon. Gen nào có dạng điều hoà này được gọi là gen bị điều hoà âm. Các operon của vi khuẩn thường sản sinh ra các mARN, gọi là đa cistron mang mã di truyền cho một số chuỗi polypeptid; thế nhưng, hầu như toàn bộ mARN của tế bào nhân thật (trừ một số cơ quan) đều là đơn cistron.

Các protein cần thiết cho sự biểu hiện của một operon được gọi là protein hoạt hoá (activator). Chúng có thể gắn vào vị trí khởi đầu nằm trong promoter của operon hay vào trình tự tăng cường (enhancer) nằm cách xa operon nhằm tăng cường khả năng phiên mã của các gen nằm trên cùng ADN. Trình tự tăng cường đầu tiên tìm ra chứa 72 bp nằm cạnh điểm sao chép của virus khỉ 40 (SV 40). Enhancer được tìm thấy trong genome của tế bào nhân thật và trong retrovirus.



Hoạt động của enhancer làm tăng số lượng ARN polymerase khi ADN phiên mã. Enhancer có thể nằm cách gen cấu trúc mà nó tăng cường.

Hiệu quả tăng cường được thực hiện qua một loại protein gắn vào một trình tự ADN đặc hiệu. Khi protein gắn đặc hiệu này gắn vào vị trí tăng cường, các nucleotid xen giữa nơi nút ra đưa vị trí tăng cường lên tiếp xúc với promoter của operon cần tăng cường. Cấu trúc mở nút này giúp cho ARN polymerase gắn được vào gen phiên mã.

Khi protein hoạt hoá gắn vào vị trí khởi động hay vị trí tăng cường sẽ kích thích sự phiên mã của gen cấu trúc trong operon. Người ta gọi cơ chế này sự **kiểm soát dương**. Sự kích thích do các gen điều hoà đáp ứng có thể xảy ra từ những phân tử nhỏ (đường, acid amin...) đến những chất lớn hơn như các phức hợp steroid của hormon (tế bào nhân thật) và các protein thụ thể của nó.

Chất "mở" gen, làm cho gen phiên mã được gọi là **chất cảm ứng**. Chất "đóng" sự phiên mã được gọi là **chất ức chế**. Các **gen có thể cảm ứng** thường tham gia trong các phản ứng dị hoá, ví dụ như khi phân giải polysaccharid thành đường đơn.

Các **gen có thể ức chế** thường tham gia trong các phản ứng **đồng hoá**.

Như vậy có 4 tổ hợp tham gia kiểm soát sự phiên mã:

- Kiểm soát cảm ứng âm.
- Kiểm soát ức chế âm.
- Kiểm soát cảm ứng dương.
- Kiểm soát ức chế dương.

Ta lần lượt xem xét từng trường hợp.

6.3.1. Kiểm soát cảm ứng âm

Kiểu khởi nguyên của sự kiểm soát âm được thực hiện bởi **operon có thể cảm ứng**, đó là "hệ thống lactose" của *E. coli*.

β -galactosidase là một enzym có hai chức năng. Chức năng chủ yếu của nó là phân giải lactose thành glucose và galactose. Chức năng thứ hai của nó là biến đổi liên kết 1 - 4 glycosid của glucose và galactose trong đường lactose thành liên kết 1 - 6 trong đường allolactose. Bình thường, khi lactose thiếu hụt trong môi trường thì enzym này không nhiều lắm. Ngay sau khi thêm lactose vào môi trường trong điều kiện không có glucose, enzym này lập tức được sản sinh. Protein vận chuyển có tên là galactoside permease cần thiết cho sự vận chuyển có hiệu quả lactose qua màng tế bào. Protein này cũng đột ngột tăng lên sau khi cho lactose vào môi trường. "Hệ lactose" hoang dại gồm gen điều hoà (I^+) và operon chứa trình tự promoter (p^+), operator locus (O^+) và 3 gen cấu trúc của β -galactosidase (Z^+), permease (y^+) và transacetylase (a^+). Đột biến của từng locus tương ứng đã được tìm thấy trong tự nhiên.

Một số ký hiệu:

Các allen của promoter:

- p^+ : promoter kiểu hoang dại, hoạt động bình thường với ARN polymerase
- p^- : promoter đột biến không được nhận biết bởi ARN polymerase
- ps : tăng cường sự nhận biết của ARN polymerase, nâng cao mức độ phiên mã của operon.
- $picr$: ảnh hưởng tới vị trí gắn của CRP – cAMP làm giảm mức biểu hiện của operon *lac* dưới 10% kiểu hoang dại (icr = insensitive to catabolite repression).

Các allen của operator:

- O^+ : allen điều khiển này cho phép các cistron cạnh nó (tức là vị trí cis) biểu hiện, nghĩa là được phiên mã. Allen này miễn cảm với **chất ức chế**, chất này ngăn chặn sự phiên mã phần còn lại của operon.
- O_c : allen điều khiển ổn định; nó không miễn cảm với chất ức chế và để cho các cistron cạnh kề phiên mã liên tục.

Các allen β – galactosidase:

- Z^+ : mã hoá một β – galactosidase hoạt động khi operon "mở".
- Z^- : đột biến có nghĩa, sinh ra một sản phẩm biến tính không hoạt tính enzym gọi là Cz protein.
- Z^{ns} : đột biến vô nghĩa, không có operon *lac* nào biểu hiện.

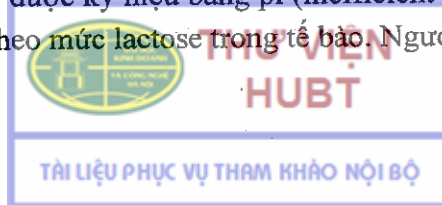
Các allen permease:

- y^+ : mã hoá cho một β – galactosidase permease hoạt động.
- y^- : không mã hoá permease, có thể là đột biến vô nghĩa.

Các allen regulator:

- i^+ : tạo ra chất ức chế hai mặt, có thể kìm hãm khả năng tổng hợp của bất kỳ operon O^+ nào khi không có lactose; khi có lactose thì chất ức chế bị bất hoạt.
- i^- : gen điều hoà bị hỏng, không có khả năng sinh chất ức chế hoạt tính do đột biến vô nghĩa hay lệch nghĩa.
- i_s : tạo chất siêu ức chế không nhạy cảm với lactose và không hoạt hoá với bất kỳ operon O^+ nào, nghĩa là không bị bất hoạt bởi lactose.

Đôi khi có sự chông chéo giữa promoter và các vị trí operator của "hệ *lac*"; trong một vài operon người ta thấy các locus operator này phủ hoàn toàn lên promoter. **Operon điều hoà** thường xuyên sinh ra các **protein ức chế** ở mức độ thấp bởi vì nó có một promoter kém hiệu suất. Promoter này được ký hiệu bằng p_i (inefficient promoter). Việc tổng hợp ra chất ức chế không thay đổi theo mức lactose trong tế bào. Ngược lại, promoter bình thường

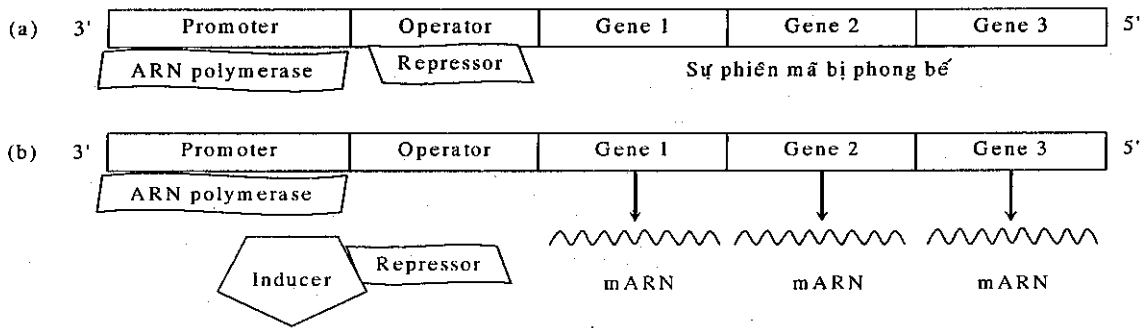


(P⁺) của operon lactose (*lac* - p) lại gắn với ARN polymerase rất hiệu quả. Khi thiếu lactose (điều kiện không cảm ứng) thì protein ức chế có hoạt tính (sinh ra bởi i⁺) sẽ gắn vào operator O⁺ (*lac* O). ARN polymerase lúc bấy giờ có thể hoặc là không gắn được vào promoter hoặc là không đọc được operator, bởi vì protein ức chế đã bao vây cả vùng promoter. Lúc bấy giờ, sự phiên mã của 3 gen cấu trúc của *lac* operon bị ngăn cản.

Khi lactose xuất hiện (điều kiện cảm ứng có) song lại không được chuyển qua màng tế bào vào tế bào chất một cách hiệu quả bởi vì lúc bấy giờ chỉ mới có vài phân tử permease xuất hiện. Bên trong tế bào, một vài phân tử lactose bị biến đổi thành allolactose nhờ β -galactosidase. Allolactose là chất cảm ứng của *lac* operon. Nó gắn vào protein ức chế, làm thay đổi cấu hình của protein ức chế và làm thay đổi vị trí gắn của protein ức chế vào operator. Sự thay đổi cấu hình trong phân tử protein do một phân tử khác gắn vào, được gọi là **sự biến hình dị lập thể**. Phức hợp allolactose - protein ức chế không thể gắn lâu vào operator được và nó bị rời khỏi ADN ngay (Hình 6.3). Lúc này, ARN polymerase có thể đọc operator để phiên mã các gen cấu trúc trên operon. Khi lượng permease tăng lên, lactose được chuyển qua màng nhiều lên và đường lactose lại bị phân hủy bởi β -galactosidase. Khi lactose trong môi trường trở nên cạn kiệt, protein ức chế vừa được tổng hợp sẽ được giải phóng khỏi allolactose, do đó chúng (chất ức chế) có thể gắn chặt vào operator, làm ngừng phiên mã của các gen cấu trúc trong operon *lac*. Chất ức chế nói chung là các protein có cấu trúc oligomer, trong đó mỗi một tiểu đơn vị của nó đều có hai tâm đặc thù. Trong hai tâm nói trên, tâm nào có ái lực cao với trình tự của toàn bộ promoter thì được coi là tâm gắn chất ức chế, còn tâm còn lại sẽ là tâm gắn chất cảm ứng. Sự gắn chất cảm ứng vào chất ức chế ở tâm biến cấu thứ hai sẽ làm thay đổi cấu hình của chất ức chế, chính vì vậy mà làm giảm ái lực của tâm gắn chất ức chế đối với operator và giúp operator tự giải thoát khỏi chất ức chế.

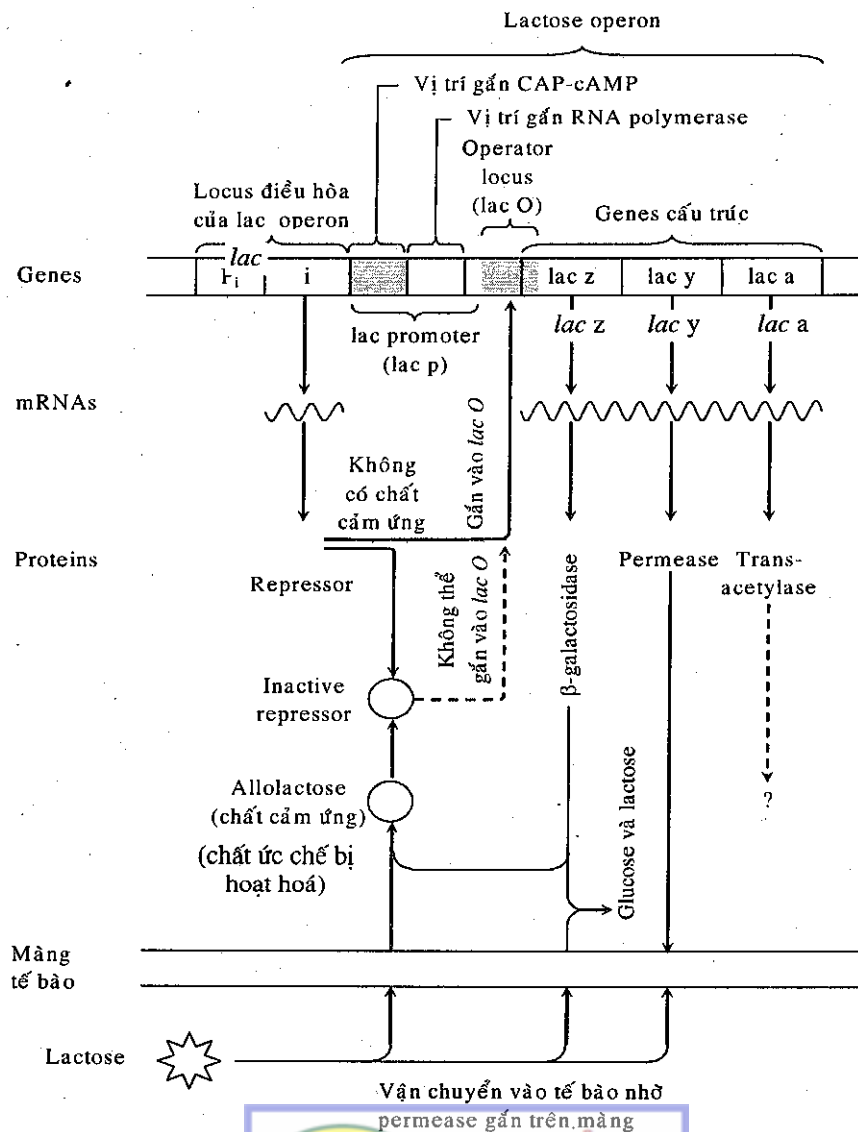
Allolactose luôn có xu hướng tách khỏi protein ức chế, ngay sau khi mức lactose trong tế bào xuống thấp. Thậm chí khi "hệ *lac*" bị ức chế, đôi khi protein ức chế tách khỏi operator một cách nhanh chóng. ARN polymerase có thể lẻn vào operator mở operon và tổng hợp nên phân tử mARN đa cistron. Điều này giải thích được mức độ rất thấp của permease và β -galactosidase luôn luôn thường trực trong tế bào vi khuẩn.

Các phân tử mARN vi khuẩn có thời gian bán huỷ rất ngắn (chỉ một vài phút), do đó sự tổng hợp protein ngừng lại rất nhanh sau khi phiên mã bị ức chế. Mặt khác, protein rất bền vững, nhưng lại dễ phân tán mỗi khi tế bào phân chia. Operon chứa gen điều hoà (i) trong hệ lactose bao gồm promoter - i và gen cấu trúc của protein ức chế. Bình thường promoter này có hiệu lực rất kém và chỉ có một ít phân tử protein ức chế - *lac* tồn tại trong tế bào. Tuy nhiên, ở các operon nằm sát với promoter của gen điều hoà và **sự tự điều hoà** có thể xảy ra.

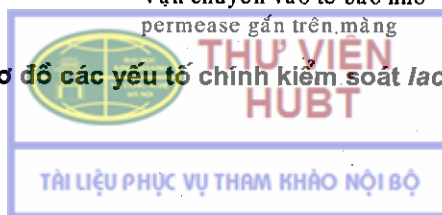


Hình 6.2. Quá trình cảm ứng enzym

(a) protein ức chế gắn vào operator và phong bế sự hoạt động của ARN polymerase; (b) chất cảm ứng gắn vào protein ức chế và bất hoạt nó. ARN polymerase tiến hành phiên mã các gen 1,2,3 trong operon



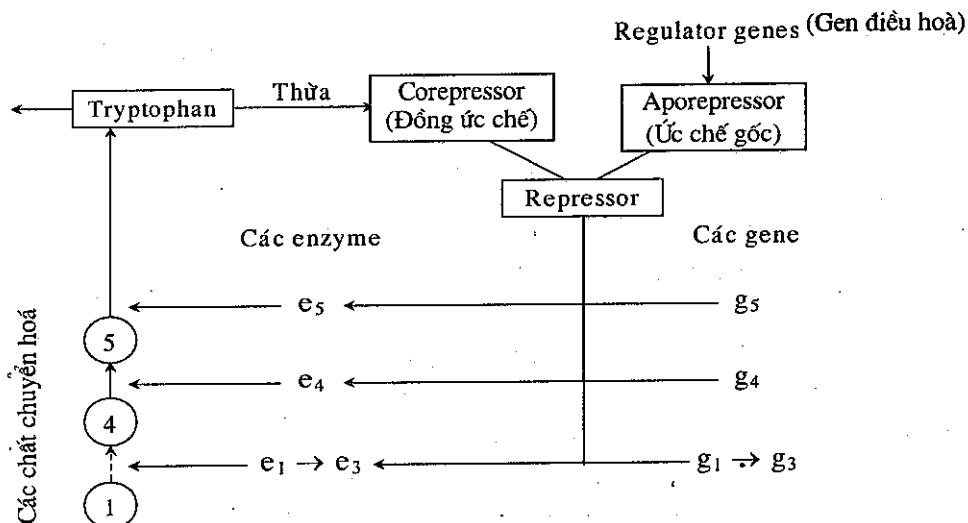
Hình 6.3. Sơ đồ các yếu tố chính kiểm soát lac operon



6.3.2. Kiểm soát ức chế âm

Ngoài cơ chế cảm ứng âm (hoặc giải ức chế) vừa nêu trên, còn có sự thích ứng enzym âm hay còn gọi là ức chế âm quá trình tổng hợp enzym. Trong trường hợp này, sự tổng hợp enzym thay vì được gây cảm ứng bởi cơ chất, thì nó bị ức chế bởi chính bản thân sản phẩm mà nó xúc tác. Hiện tượng này được quan sát thấy trong trường hợp tổng hợp acid amin tryptophan và histidin. Tryptophan được tổng hợp theo 5 bước, mỗi bước được xúc tác bởi một enzym đặc hiệu (Hình 6.4). Các gen đáp ứng cho 5 enzym nằm trên một operon, thông thường theo một trật tự giống như trật tự của sản phẩm enzym của chúng trong quá trình sinh tổng hợp.

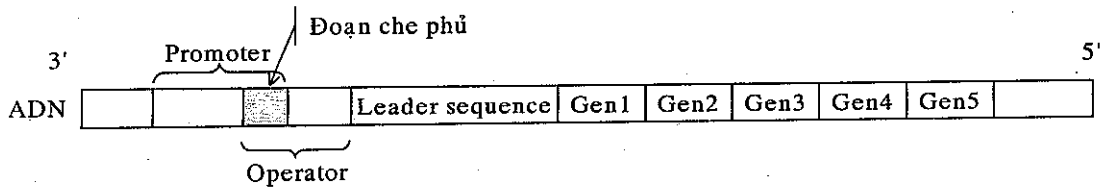
Gen điều hòa cho hệ thống này luôn luôn sản sinh ra một loại protein không chức năng, được gọi là chất **ức chế gốc** (apo-repressor). Khi tryptophan sản xuất dư thừa, thì sự quá tải của tryptophan sẽ tác động như một chất **đồng ức chế** (co-repressor). Chất đồng ức chế gắn vào chất ức chế gốc hình thành nên một phức hợp ức chế chức năng. Phức hợp ức chế chức năng này gắn vào operator *trp*, và ức chế sự phiên mã của 5 gen cấu trúc trên operon. Các vùng promoter và operator của operon này nhiều khi phủ lên nhau và gắn được phức hợp co-aporepressor, do đó ARN polymerase khó chen vào được. Khi nồng độ tryptophan giảm xuống, tryptophan tách khỏi chất ức chế gốc và chất ức chế gốc rời khỏi operator, thì ARN polymerase lại tổng hợp được mRNA đa cistron cho 5 enzym tổng hợp tryptophan.



Hình 6.4. Kiểm soát ức chế âm trong hệ tryptophan của *E.coli*

6.3.3. Cơ chế điều hoà suy giảm

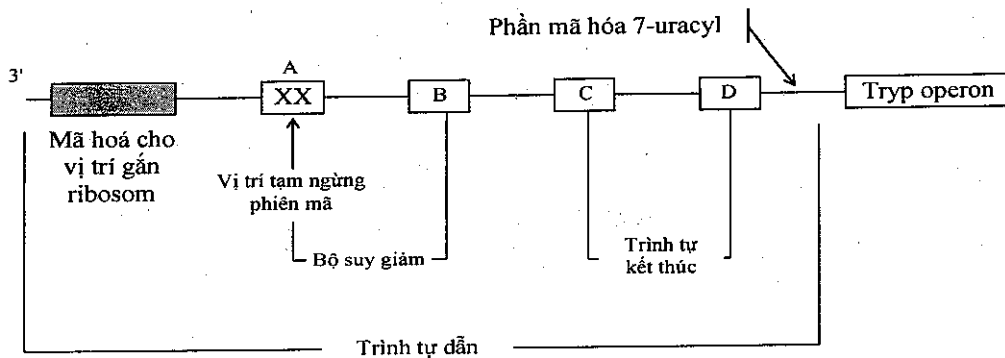
Cơ chế điều hoà thứ hai này được phát hiện ở hệ tryptophan. Tại đầu 3' của operon phiên mã mRNA đa cistron tương ứng g_1, g_2, g_3, g_4, g_5 có một trình tự gồm 162 base; phân bố giữa operator và gen cấu trúc thứ nhất, được gọi là **trình tự dẫn** (Hình 6.5, Hình 6.6).



Hình 6.5. Sự phân bố gen của tryptophan operon

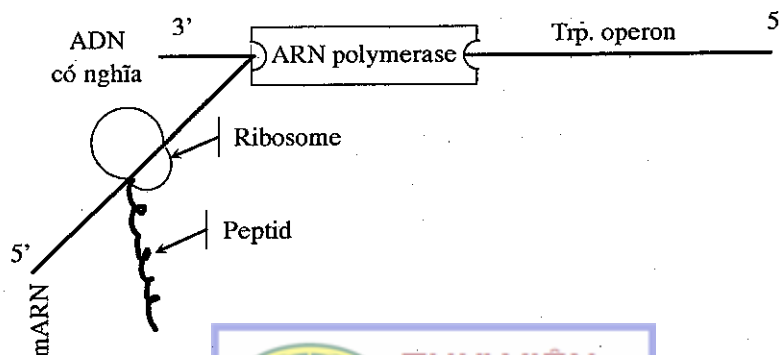
Trình tự dẫn chứa một phần đóng vai trò bộ suy giảm, chính bộ suy giảm này mã hóa phân tử peptid dẫn chứa 14 acid amin.

Trình tự dẫn chứa 4 vùng, trong đó vùng A chứa hai codon mã hóa cho 2 codon mã cho tryptophan (UGGUGG) được ký hiệu XX. Phần D trước khi nối với Tryptophan operon có một trình tự mã hóa cho 7 uracyl trên mARN. Phần C và D có thể tạo nút kẹp tóc với nhau và cùng với 7 Uracyl tạo thành cấu trúc của tín hiệu chấm dứt phiên mã (terminator). Trình tự của vùng A có thể tạo nút kẹp tóc với vùng B, và trình tự vùng B có thể tạo nút kẹp tóc với vùng C.



Hình 6.6. Sắp xếp các vùng nucleotid trong trình tự dẫn

ARN polymerase chuyển động dọc sợi có nghĩa ADN theo chiều $3' \rightarrow 5'$ và phân tử mARN tương ứng sẽ được tổng hợp theo chiều $5' \rightarrow 3'$. Gần đầu 5' của mARN là vị trí gắn ribosom (Hình 6.7).



Hình 6.7. Sự chuyển dịch của ribosom trên mARN

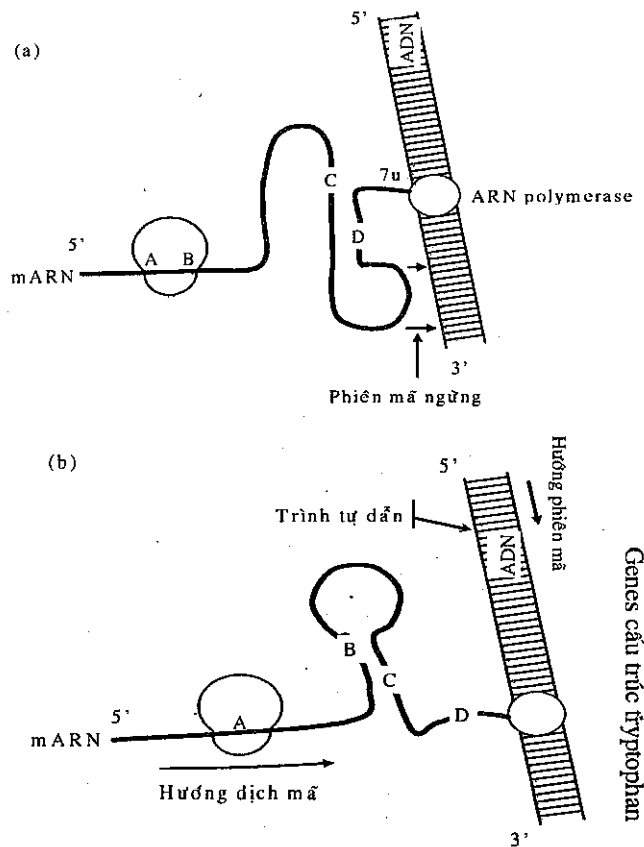


Ribosom gần như chuyển dịch liên tiếp theo sau ARN polymerase trong quá trình phiên mã và giải mã. Ribosom gắn và trượt trên mARN. Các vùng A và B có thể bắt cặp với nhau vì có chứa những trình tự bổ sung và vùng C, D cũng vậy. Nhưng nếu B cũng bắt cặp được vị trí với C thì sẽ có thể hình thành nên 3 nút kẹp tóc A/B, B/C và C/D.

Trong điều kiện thừa tryptophan thì ribosom dịch mã ra sợi peptid dẫn, do đó vùng B và C của mARN không tạo nút được. Nhờ đó, vùng C và D bắt cặp được với nhau tạo thành nút kẹp tóc kết thúc phiên mã trước khi ARN polymerase đọc đến vùng gen mã hóa cho các enzym tổng hợp tryptophan.

Nếu môi trường thiếu tryptophan thì ribosom sẽ dừng lại ở các codon tryptophan của vùng A. Vì ribosom phủ lên A, nên B bắt cặp với C, do vậy C không thể bắt cặp với D nên không hình thành được kẹp tóc ngừng. Nhờ đó, ARN polymerase tiếp tục phiên mã các gen cấu trúc của tryptophan operon tạo nên mARN của các enzym tổng hợp tryptophan.

Sự phiên mã suy giảm cũng gặp ở một số operon liên quan đến việc chuyển hóa các acid amin khác như histidin và leucin.



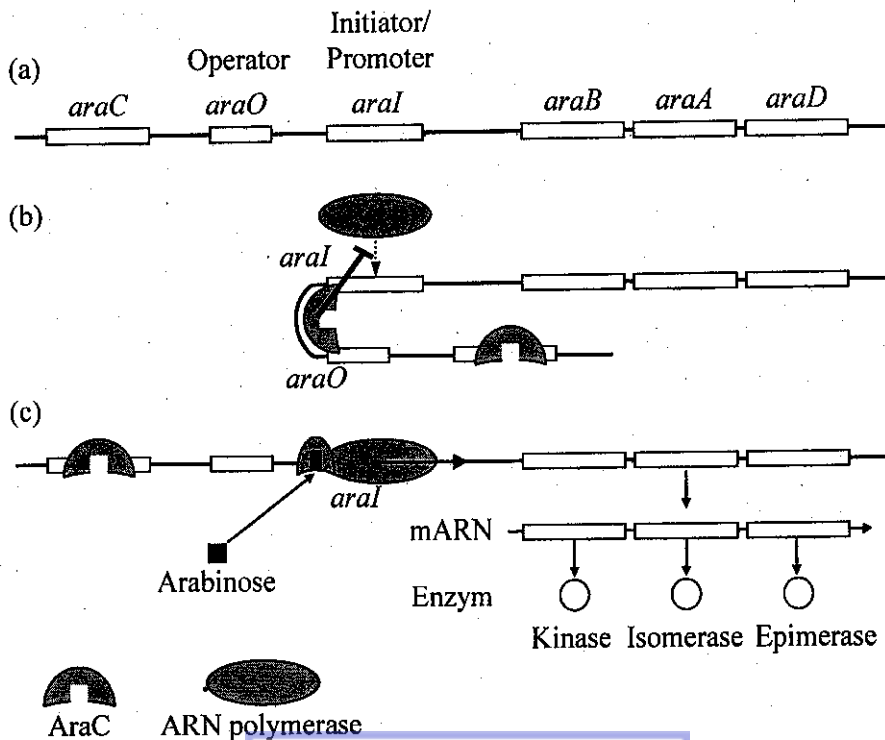
Hình 6.8. Mô hình phiên mã suy giảm của operon tryptophan *E. coli*

(a) Thừa tryptophan: ngừng; (b) Thiếu tryptophan: không ngừng

6.3.4. Kiểm soát cảm ứng dương

Đa số các hệ thống tổng hợp enzym cảm ứng được biết cho đến nay điều hoạt động nhờ sự điều hòa âm tính, giống như đã nêu trong thí dụ dị hóa lactose. Khác với sự điều hòa tổng hợp enzym âm tính trong quá trình trao đổi lactose, các enzym của quá trình trao đổi arabinose ở *E. coli* nhờ sự kiểm soát của repressor có khả năng thực hiện điều hòa dương tính và âm tính. Arabinose là đường mà sự chuyển hóa của nó đòi hỏi 3 enzym được mã hóa bởi các gen *araB*, *araA*, *araD*. Khi môi trường không có arabinose thì 3 enzym trên chỉ có một lượng rất nhỏ. Khi bổ sung arabinose vào môi trường nuôi cấy hàm lượng 3 enzym này cùng tăng lên. Quá trình cảm ứng được điều hòa bởi một gen *araC* nằm gần 3 gen trên nhưng cách chúng bởi phức hợp điều hòa (gồm operator (*araO*) và initiator (*araI*), *araI* chứa promoter) của nhóm BAD.

Sản phẩm protein của *araC* (AraC) là chất ức chế của nhóm BAD khi môi trường không có arabinose. Nó luôn luôn gắn vào *araO* làm ADN có dạng loop nên ARN polymerase không gắn vào được để phiên mã. Tuy nhiên, sau khi gắn với arabinose thì repressor này bị biến đổi cấu hình thành **phức hợp hoạt hoá** rời khỏi operator (*araO*) và gắn vào initiator (*araI*), tạo điều kiện cho ARN polymerase gắn được vào promoter của BAD, nhờ đó cảm ứng được sự phiên mã của operon. Bản thân AraC tự điều hoà chính nó bằng cách gắn vào *araC*.



Hình 6.9. Kiểm soát cảm ứng dương ở operon arabinose



Cần lưu ý rằng, operon của arabinose được coi là một trong số các operon nhạy cảm đối với glucose và do đó sự tổng hợp mRNA của nó cũng phụ thuộc vào liên kết của CAP là một protein hoạt hoá gen dị hoá với promoter BAD. Nói cách khác để cho operon của arabinose hoạt động mạnh cần có hai tín hiệu điều hoà dương tính, tín hiệu thứ nhất thông báo sự có mặt của arabinose, tín hiệu thứ hai sự vắng mặt của glucose.

Adenosine monophosphat vòng (cAMP), protein hoạt hoá gen dị hoá (CAP) và protein nhận AMP vòng (cyclic AMP receptor protein = CRP) tham gia vào sự điều hoà của hệ thống arabinose. AMP vòng có khả năng hoạt hoá cho những promoter có ái lực yếu nhưng bản thân nó không gắn được vào promoter. Muốn làm điều này nó phải liên kết với CAP hoặc với CRP. Khi AMP gắn được vào promoter nó sẽ hoạt hoá ARN polymerase và ARN polymerase sẽ gắn được vào promoter.

6.3.5. Đa kiểm soát

Một locus di truyền có thể được kiểm soát bởi nhiều cơ chế khác nhau. Khi có mặt glucose thì vi khuẩn *E. coli* không cần thiết phải phân giải một loại đường nào khác và các gen mã hoá cho các enzym phân giải các đường ấy cũng đóng lại. Ví dụ, nếu không có glucose nhưng có lactose trong môi trường thì *lac* operon được cảm ứng. Nhưng nếu có glucose thì sự cảm ứng của *lac* operon sẽ không xảy ra. Hiện tượng này được đặt tên là hiệu ứng glucose (glucose effect), ngày nay người ta gọi hiện tượng này là ức chế dị hoá (catabolite repression). Phức hệ của 2 chất hoạt hoá ức chế dị hoá là cAMP và CAP.

Với *lac* operon (Hình 6.2) ở đây có vị trí gắn phức hệ cAMP – CAP. ARN polymerase chỉ gắn có hiệu quả vào promoter nếu phức hệ cAMP – CAP cũng được gắn vào vị trí này. Khi mức glucose tăng lên trong tế bào, thì lượng cAMP giảm xuống và sẽ có ít phức hệ cAMP – CAP hơn có thể dùng để hoạt hoá *lac* operon. Bản thân cAMP không gắn được vào promoter để ARN polymerase dễ gắn vào ADN, nó phải liên kết với CAP làm thay đổi cấu hình của CAP. Lúc đó phức hệ cAMP – CAP gắn được vào promoter. Sự điều hoà này là sự điều hoà cảm ứng dương tính.

CAP được sinh ra ở mức độ thấp do một locus di truyền đặc hiệu. Enzym adenylate cyclase (adenylcyclase) biến adenosine triphosphat (ATP) thành adenosine monophosphat vòng (cAMP).

6.4. KIỂM SOÁT SAU DỊCH MÃ

Sự biểu hiện của gen có thể được điều hoà sau khi protein hình thành.

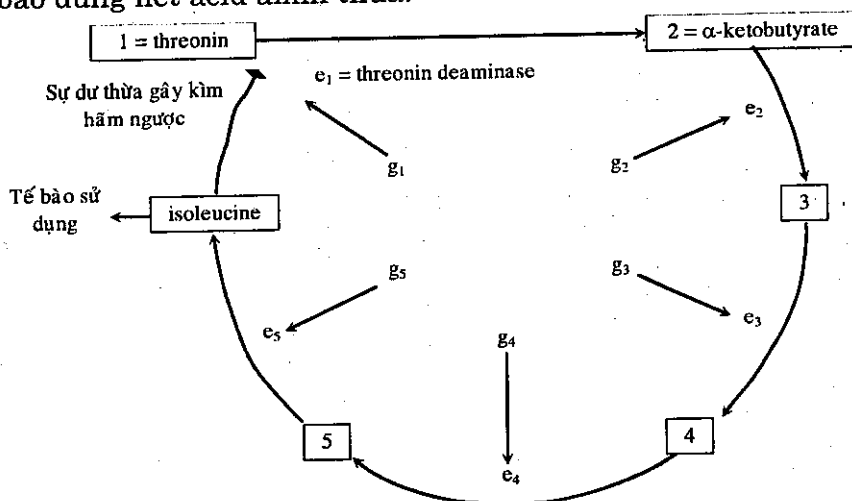
Sự kìm hãm ngược còn gọi là sự ức chế bởi sản phẩm cuối, cũng là cơ chế điều hoà có sự tham gia ức chế của enzym. Sản phẩm cuối cùng của quá trình sinh tổng

hợp (thường là phân tử nhỏ như acid amin chẳng hạn) có thể liên kết chặt chẽ (nếu ở nồng độ cao) với enzym đầu tiên được tổng hợp ra. Sự liên kết này không xảy ra tại vị trí xúc tác (trung tâm hoạt động) của enzym mà ở vị trí gắn chất kim hãm (vị trí điều hoà) làm biến tính cấu trúc bậc 3 hay bậc 4 của enzym, do đó bất hoạt vị trí xúc tác.

Sự biến hình dị lập thể của enzym sẽ phong bế khả năng xúc tác của enzym và ngăn cản sự sản xuất quá độ của sản phẩm cuối cùng trong quá trình sinh tổng hợp các chất chuyển hoá trung gian.

Ví dụ, sản phẩm cuối cùng isoleucin trong *E. coli* khi xuất hiện ở nồng độ cao, liên kết với enzym đầu tiên (e_1) của quá trình sinh tổng hợp và do vậy kim hãm toàn bộ con đường này cho đến khi isoleucin trở lại mức bình thường thông qua mức tiêu thụ của tế bào.

Chất khởi đầu của chuỗi gồm 5 phản ứng này là threonin. Phản ứng đầu tiên threonin biến thành acid α -ketobutyric do enzym threonin - desaminase, qua một loạt quá trình chuyển hoá tạo ra sản phẩm cuối là isoleucin. Vào một thời điểm nào đó, isoleucine không được tế bào sử dụng đến sẽ liên kết trở lại với threonin-desaminase và kim hãm tác dụng enzym này. Các enzym tuy vẫn được hình thành nhưng không hoạt động được. Việc tổng hợp isoleucin bị đình chỉ cho đến khi tế bào dùng hết acid amin thừa.



Hình 6.10. Sơ đồ kim hãm ngược trong *E. coli*
(e = enzym; g = gen; 1,2,3,4,5 = các chất trung gian)

Tóm tắt:

Vi khuẩn sống trong một môi trường cạnh tranh gay gắt, chúng cần phải thay đổi nhanh chóng trong cơ chế chuyển hoá để sống sót. Nhiều vi khuẩn tạo ra các phương thức chuyển hoá để thích ứng với sự thay đổi này, các enzym trong quá trình chuyển hoá chỉ được tổng hợp khi cần thiết. Sự kiểm soát như vậy được thực hiện ở mức độ gen.

Các enzym cần cho các phương thức dị hoá chỉ được tạo ra khi có sẵn một chất dinh dưỡng đặc biệt như lactose. Các protein ức chế kìm hãm sự biểu hiện gen và tác động ức chế bị kìm hãm bởi các chất cảm ứng, các chất cảm ứng bản thân nó là các phân tử dinh dưỡng hay các phân tử có nguồn gốc từ chất dinh dưỡng. Các kiểm soát này bị thay đổi nếu có sự lựa chọn giữa glucose hay lactose thì glucose được ưu tiên hơn.

Ở phương thức đồng hoá, các enzym cần cho sự sản xuất một acid amin đặc biệt như tryptophan chỉ được tổng hợp khi tryptophan không có sẵn ở vi khuẩn. Trong trường hợp này, các chất ức chế là sản phẩm cuối cùng của quá trình.

Các phương thức khác của sự kiểm soát biểu hiện gen là bằng các protein hoạt hoá, đây là sự kiểm soát dương. Có sự tham gia của cAMP vào việc điều chỉnh sự biểu hiện gen ở vi khuẩn.

CÂU HỎI

1. Protein ức chế khác với protein hoạt hoá ở chỗ:
 - a) Gắn vào operator ngăn cản phiên mã gen cấu trúc
 - b) Thuộc dạng điều hoà âm
 - c) Gắn vào vị trí khởi đầu trên promoter
 - d) Gắn vào vị trí tăng cường
 - e) a, b
2. Hai chức năng chủ yếu của enzym β – galactosidase thể hiện ở:
 - a) Phân giải lactose thành glucose và galactose
 - b) Phân giải lactose thành glucose và fructose
 - c) Biến đổi liên kết 1 – 4 glycosid của glucose và galactose thành 1 – 6 trong allolactose
 - d) Biến đổi liên kết 1 – 6 glycosid trong allolactose thành liên kết 1 – 4 trong lactose
 - e) a, c
3. "Hệ lactose" hoang dại của *E. coli* có:
 - a) Gen điều hoà (I_s)
 - b) Operon chứa promoter (P_{lac}) và operator (O_c)
 - c) Operon gồm vùng promoter và enhancer
 - d) 3 gen cấu trúc: β – galactosidase (Z^+), permease (y^+), transacetylase (a^+)
 - e) a, b, d

4. "Hệ lactose" thường xuyên sản sinh protein ức chế ở mức thấp vì:
- E. coli* chuộng lactose hơn glucose
 - Promoter của operon này kém hiệu suất
 - Promoter của operon này gắn với ARN – polymerase
 - Chất ức chế được tổng hợp trong tế bào có thay đổi
 - a, b đúng
5. Chất ức chế gốc là:
- Protein không chức năng sinh ra do gen điều hoà của hệ tryptophan
 - Tryptophan
 - 5 enzym tổng hợp tryptophan
 - Trình tự dẫn
 - 5 chất chuyển hoá do 5 enzym tham gia tổng hợp tryptophan
6. Operon của arabinose được coi là operon nhạy cảm đối với glucose vì
- AMP tăng khi hàm lượng glucose tăng
 - AMP vòng có khả năng hoạt hoá promoter yếu
 - AMP muốn gắn vào promoter phải nhờ CAP
 - AMP gắn được vào promoter sẽ hoạt hoá ARN polymerase
 - b, c và d
7. Sự kìm hãm ngược **không** chấp nhận trường hợp:
- Sự ức chế bởi sản phẩm cuối cùng
 - Cơ chế điều hoà có sự tham gia ức chế của enzym
 - Sự liên kết giữa sản phẩm cuối cùng với enzym ở vị trí điều hoà của enzym làm bất hoạt vị trí xúc tác
 - Sự liên kết giữa sản phẩm cuối cùng với enzym ở vị trí xúc tác của enzym
 - Sự biến hình dị lập thể của enzym sẽ phong bế khả năng xúc tác của enzym
8. Operon gồm:
- Vùng khởi động (promoter)
 - Các gen cấu trúc
 - Vị trí điều hoà
 - Chóp GMP
 - a, b và c
9. Operator là:
- Đoạn mARN gắn được protein điều hoà



- b) Đoạn ADN chuyên biệt gắn được protein điều hoà
 - c) Đoạn ADN nằm trước promoter
 - d) Đoạn ADN nằm sau promoter
 - e) Gen tổng hợp protein
10. Kiểm soát dương khác với kiểm soát âm vì cần phải:
- a) Loại bỏ tích cực phân tử ức chế
 - b) Hoạt hoá quá trình khởi đầu của ARN – polymerase
 - c) Đưa vào co – repressor
 - d) Loại bỏ co – repressor
 - e) a và d

Bài 7

BỘ GEN TẾ BÀO NHÂN THẬT

MỤC TIÊU

- So sánh được bộ gen của tế bào nhân thật và tế bào nhân nguyên thủy.
- Mô tả được một số loại trình tự ở bộ gen tế bào nhân thật.
- Nêu được sự biểu hiện của bộ gen tế bào nhân thật.
- Nêu được các cách kiểm soát sự biểu hiện của bộ gen.

7.1. MỞ ĐẦU

Các bộ gen tế bào nhân thật phức tạp hơn tế bào nhân nguyên thủy. Số gen mã hoá cho protein cũng nhiều hơn gấp trăm lần. Sự dư ADN được thể hiện ở nhiều loại trình tự khác nhau. ADN của tế bào nhân thật ở dạng thẳng, còn ở tế bào nhân nguyên thủy thì ở dạng vòng.

Điểm khác nữa là sự phiên mã và dịch mã ở tế bào nhân thật được phân ra làm 2 giai đoạn tạm thời về mặt không gian: sự phiên mã xảy ra trong nhân, còn sự dịch mã xảy ra trong tế bào chất. Ngược lại, với tế bào nhân nguyên thủy: sự dịch mã của phân tử mARN có thể bắt đầu trước khi tổng hợp mARN được hoàn thành.

7.2. TỔ CHỨC BỘ GEN Ở TẾ BÀO NHÂN THẬT

7.2.1. Kích thước của bộ gen tế bào nhân thật – giá trị C

Giá trị C (C – value) được dùng để diễn tả kích thước bộ gen của một loài, nó biểu thị số cặp nucleotid (thường biểu diễn bằng đơn vị Mb hay Mbp – megabase pair) của bộ gen đơn bội của một sinh vật tế bào nhân thật. Các tế bào ở tế bào nhân thật chứa nhiều ADN hơn là các tế bào nhân nguyên thủy (Bảng 7.1).



Bảng 7.1. Bộ gen chứa acid nucleic

Bộ gen	Số lượng gen	Số cặp base
Thực vật	<50 000	<10 ¹¹
Người	25 000	3.10 ⁹
Giun	14 000	10 ⁸
Ruồi	12 000	1,6.10 ⁸
Nấm	6 000	1,3.10 ⁷
Vi khuẩn	2 000 – 4 000	<10 ⁷
Mycoplasma	500	<10 ⁶
ssADN virus		
Parovirus	5	5 000
Phage fx174	11	5 387
dsARN virus		
Coronavirus	7	20 000
Influenza	12	13 000
TMV	4	6 400
Phage MS2	1	1 300
Viroid		
PSTV ARN	0	359

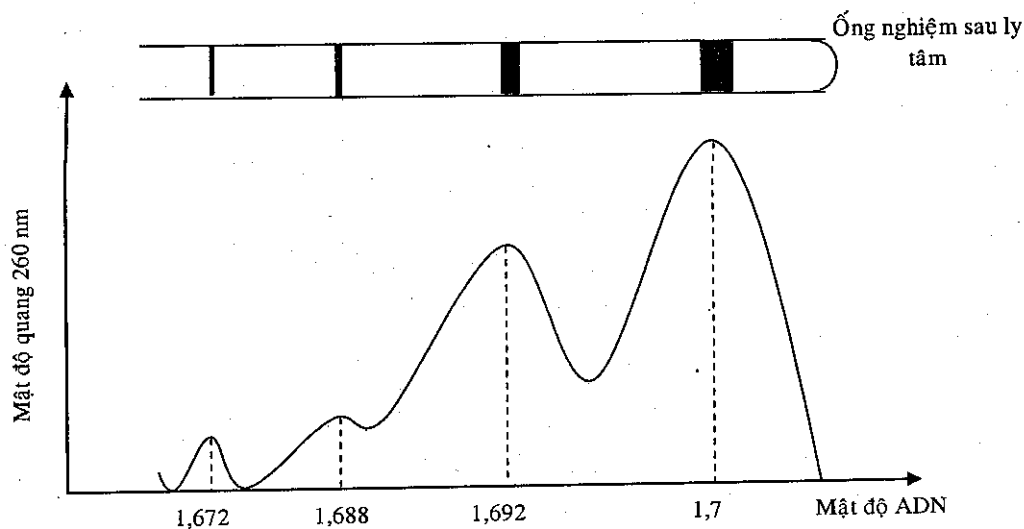
Nghịch lý giá trị C là một khái niệm phản ánh sự không tương xứng giữa kích thước bộ gen (giá trị C) với số gen của một loài. Ví dụ, người có bộ gen lớn hơn giun đến 30 lần nhưng chỉ gấp đôi về số gen. Như vậy phần lớn trình tự tồn tại trong bộ gen không liên quan đến gen.

7.2.2. Sơ đồ khái quát về các loại trình tự ADN

7.2.2.1. Sự không đồng nhất của ADN tế bào nhân thật

Khi tách ADN của tế bào nhân thật bằng siêu li tâm trên thang nồng độ Cesium chlorid, có 3 vệt được ghi nhận (Hình 7.1).





Hình 7.1. Sự xuất hiện của ADN vệ tinh sau khi siêu ly tâm

Sự không đồng nhất của ADN tế bào nhân thật thể hiện rõ hơn khi thực hiện phản ứng tái hợp (reassociation). ADN được cắt nhỏ và cho biến tính rồi sau đó cho hồi tính. Khi hồi tính, các đoạn có trình tự bổ sung dễ tái hợp với nhau, nhờ vậy nhận biết các trình tự lặp lại.

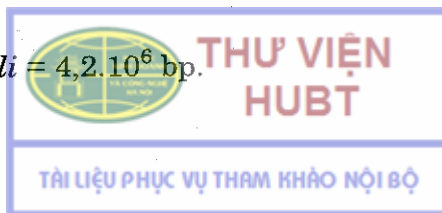
Động học tái hợp thành cặp của ADN tế bào nhân thật khác với tế bào nhân nguyên thủy (Hình 7.2). Đường cong tái hồi của tế bào nhân nguyên thủy có dạng hình sigma điển hình, chứng tỏ sự đồng nhất của các trình tự trong tái hợp. Ở tế bào nhân thật đường cong phức tạp hơn, kéo dài một khoảng rộng các giá trị C_{0t} (đơn vị của nó là số mol nucleotid/lit/giây) và có chứa 3 thành phần lặp lại ở mức độ khác nhau:

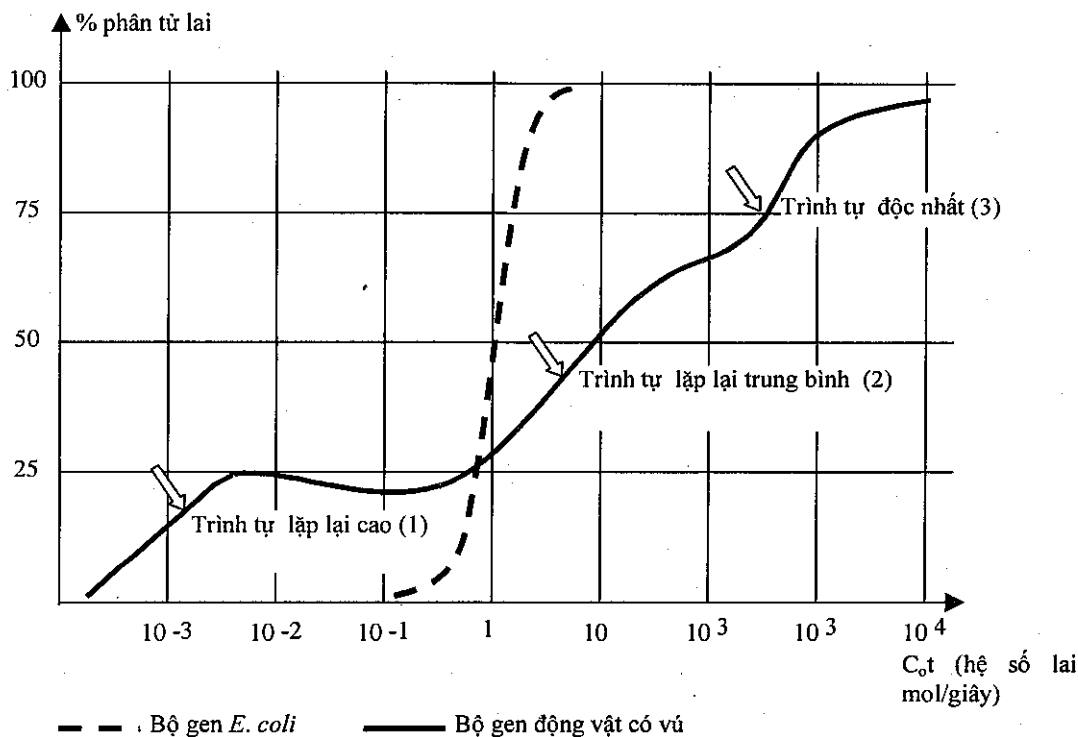
- ADN lặp lại nhiều (tái hợp rất nhanh), chiếm khoảng 10 – 15% bộ gen.
- ADN lặp lại trung bình (tái hợp nhanh vừa), chiếm khoảng 25 – 40% bộ gen.
- ADN duy nhất (tái hợp rất chậm), chiếm khoảng 60% bộ gen.

Phức hợp của một vài bộ gen đã biết kích thước có thể được tính toán bằng cách so sánh đường cong C_{0t} của nó với đường cong C_{0t} của một ADN ở một phức hợp đã được biết trước (thường là ADN *E. coli* với kích thước bộ gen có tổng chiều dài của các trình tự là $4,2 \cdot 10^6$ bp). Cách so sánh như vậy sử dụng $C_{0t} \frac{1}{2}$ của một dân số ADN, đó là giá trị C_{0t} ở thời điểm 50% ADN đã tái tổ hợp.

$C_{0t} \frac{1}{2}$ của ADN từ bộ gen thử nghiệm = phức hợp hiện diện của bộ gen thử nghiệm.

$$C_{0t} \frac{1}{2} \text{ của ADN } E. coli = 4,2 \cdot 10^6 \text{ bp.}$$





Hình 7.2. Đường cong tái hợp của ADN động vật có vú

1. ADN có khả năng bắt cặp tức thời; 2. ADN bắt cặp chậm hơn $10^{-3} < C_0t < 10$;
3. ADN bắt cặp rất chậm $10 < C_0t < 10^3$

7.2.2.2. Các trình tự ADN lặp lại ở mức độ cao

Các trình tự này cũng được biết như là ADN vệ tinh (satellite DNA), bởi vì khi ADN của bộ gen được cắt ra và ly tâm trong gradient tỉ trọng Cesium chloride, các chuỗi lặp lại cao thường hình thành các dải băng vệ tinh cách xa đỉnh chính (khá rộng) so với các chuỗi lặp lại trung bình hay duy nhất. Điều này cho thấy rằng ADN lặp lại cao thường có một thành phần base khác với thành phần của bộ gen (và vì vậy tỷ trọng các phần nổi khác nhau).

Các trình tự này chiếm 10 - 15% bộ gen của động vật có vú và không mã hóa cho protein. Chúng liên quan đến các chất dị nhiễm sắc cơ cấu (constitutive heterochromatin).

Phân tích trình tự nucleotid cho thấy chúng gồm 3 loại:

- Họ ADN vệ tinh: gồm các đoạn nhỏ của trình tự lặp lại liên tiếp có trình tự đơn giản (thường 1 - 4 bp) và nằm rải rác khắp bộ gen. Trong trường hợp lặp lại 1 nucleotid, A và T thường gặp, chiếm khoảng 10 Mb hay 0,3% bộ gen tế bào nhân thật. Trái lại, G và C rất hiếm gặp. Trong trường hợp lặp lại 2 nucleotid, đoạn lặp lại CA (GT trên sợi bổ sung) thường gặp, chiếm khoảng 6% bộ gen và thường đa hình. Đoạn lặp lại CT/GA cũng phổ biến, thường xảy ra trên mỗi 50 kb và chiếm khoảng 2% bộ gen, nhưng CG/GC rất hiếm. Điều này do C bổ sung với G ở đầu 3' bị methyl hóa, kết quả tạo TpG (hay CpA của sợi bổ

sung). Đoạn lặp lại 3 và 4 nucleotid khá hiếm, thường đa hình và được nghiên cứu dùng làm chất đánh dấu đa hình. Điểm nổi bật của ADN vi vệ tinh chưa được biết. Các đoạn lặp lại xen kẽ purin-pyrimidin, như các đoạn lặp lại liên tục của cặp 2 nucleotid CA/GT, có khả năng phù hợp cấu tạo của ADN xen kẽ, Z-ADN, *in vitro*, nhưng ít có bằng chứng chúng cũng làm như vậy trong tế bào.

– Loại thứ hai: tương ứng với các trình tự lặp lại liên tiếp nhưng với các đoạn dài hơn (100 - 200 bp).

– Loại thứ ba: có nhiều trình tự lặp lại mức độ cao, phân tán phía ngoài chất dị nhiễm sắc như trình tự CEN và trình tự TEL.

+ Trình tự CEN: các trình tự CEN lặp lại cao là của tâm động (centromere). Ở nấm men *Saccharomyces cerevisiae* gồm:

* Trình tự đầu có 9 cặp base 5'TCACATGAT.

* Trình tự giữa có 80 – 90 cặp base, rất giàu A và T (>90%).

* Trình tự 11 cặp base ở đầu 3'TGATTTCCGAA.

Các trình tự này ở ADN người phức tạp hơn.

+ Các trình tự TEL:

* Các trình tự TEL thuộc các nhóm telomere (đầu mút của nhiễm sắc thể) với nhiều vai trò khác nhau như bảo vệ đầu mút nhiễm sắc thể khỏi bị cắt, giữ chiều dài của nhiễm sắc thể khi sao chép, gắn với màng nhân và kim hãm sự biểu hiện của các gen ở đầu mút.

* Các trình tự TEL có tính bảo tồn trong tiến hóa. Chúng có số lần lặp lại cao, giàu C và A: CC(C)ACACA(CA) ở nấm men, CCCTAA ở người. Đầu mút của nhiễm sắc thể giàu G gấp lại hình kẹp tóc có cấu trúc 4 mạch. Cấu trúc này bảo vệ đầu mút nhiễm sắc thể khỏi bị cắt bởi nuclease, đồng thời khi sao chép giữ đầu mút khỏi bị mất các trình tự mã hóa.

Các chuỗi vệ tinh lặp lại cao thường thấy trong tất cả các nhiễm sắc thể của bộ gen. Ví dụ, 25% bộ gen của *Drosophila virilis* có trình tự ACAACT, hiện diện không ít hơn trong 10^7 bản sao.

7.2.2.3. Các trình tự lặp lại ở mức độ trung bình

Loại trình tự này chiếm 25 – 40% bộ gen người. Chúng cũng gồm các trình tự lặp lại nhưng dài hơn (100 – 1000 bp) và đa dạng hơn nhiều so với loại lặp lại cao (vệ tinh). Các trình tự này phân tán trong bộ gen và nếu cắt bộ gen thành đoạn 20 – 40 kb thì có 90 đến 100% số đoạn có trình tự lặp lại trung bình. Trong số các trình tự lặp lại trung bình, có các họ trình tự đặc hiệu là SINE và trình tự LINE.

Một tế bào nhân thật có thể có hàng trăm nhóm ADN lặp lại trung bình, với mỗi nhóm có khoảng từ $50 - 10^5$ đoạn lặp lại.



Có một số lượng lớn các nhóm khác ADN lặp lại trung bình trong bộ gen người, như chuỗi 6400 cặp base được lặp lại từ 3000 – 4600 lần được tính cho khoảng 1% trong tổng số ADN người.

Các trình tự SINE (còn gọi là Alu ở người):

Một trong những nhóm lớn nhất của ADN lặp lại trung bình ở người là phức hợp của các chuỗi được gọi là họ Alu. Sở dĩ có tên này vì nó có chứa vị trí nhận diện cho enzym cắt giới hạn AluI. Chuỗi Alu xuất hiện khoảng 3×10^5 lần trong bộ gen đơn bội của người, vì vậy chiếm khoảng 3% ADN người.

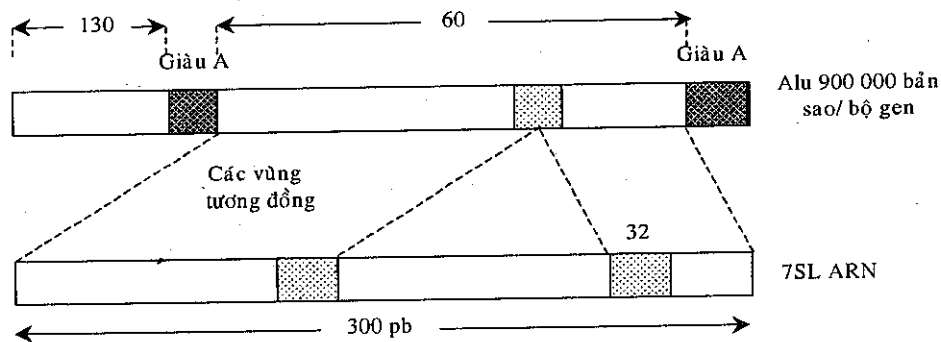
Alu chứa một lượng lớn GC và mặc dù nằm rải rác khắp chất nhiễm sắc hoạt động (euchromatin) của bộ gen, nhưng chúng thường ở trên nhiễm sắc thể R và thấy trên dải xanh khi dùng thuốc nhuộm Giemsa chuẩn và chứa các vùng có khả năng phiên mã của bộ gen. Chiều dài toàn bộ của đoạn lặp lại Alu khoảng 280 bp và thường tiếp theo bởi các trình tự thuận (6 – 18 bp) (có cùng hướng), giàu A tại một sợi và T tại sợi bổ sung. Tuy nhiên, có sự không đối xứng tại các đoạn lặp lại không liên tục: một đơn vị lặp lại chứa trình tự 32 bp bên trong nhưng đoạn khác không có. Các monomer chứa một trong hai đoạn lặp lại liên tục và kiểu gen khác nhau của dimer và monomer.

Sự chuyển vị nghịch tạo bản sao của các phân tử ở vị trí mới, trong khi phân tử cho ban đầu vẫn giữ nguyên cấu trúc không biến đổi. Do vậy, chuyển vị nghịch tạo nên một ít đoạn đứt và sự tái cấu trúc của bộ gen tế bào chủ. Những biến đổi này sẽ làm ngừng hay hoạt hoá các gen, mà một số biến đổi này có thể gây ra ung thư.

Hai đơn vị lặp lại của trình tự Alu giống nhau về trình tự 7SL ARN, thành phần cấu tạo của dấu hiệu nhận diện, vận chuyển protein qua màng lưới nội sinh chất. Do điều này người ta tin rằng trình tự Alu tăng lên là nhờ sự chuyển vị 7SL ARN, do vậy tạo thành gen giả cụt 7SL ARN. Quá trình chuyển vị trình tự Alu xảy ra nhờ enzym phiên mã ngược mã hoá bởi LINE – 1 (Kpn) và có thể là nguyên nhân gây các vấn đề lâm sàng. Có thể số lượng lớn các bản sao gen giả cụt liên quan đến sự hiện diện trình tự promoter trong trình tự 7SL ARN (gen 7SL ARN, giống gen tARN được mã hoá bởi ARN polymerase II từ một promoter nội). Ngược lại, gen giả cụt từ bản sao ARN polymerase II thiếu trình tự promoter và chúng chỉ có thể biểu hiện nếu chúng ở gần trình tự promoter có chức năng.

Hiện nay chức năng của trình tự Alu chưa được biết. Mặc dù tần số thường gặp là 1 bản sao trên 4 kb, nhóm các đoạn lặp lại Alu xảy ra tại các vùng bất kỳ. Do chúng có mặt khắp nơi, trình tự Alu được xem như thúc đẩy quá trình tái tổ hợp không tương đồng, nguyên nhân gây bệnh trong một số trường hợp nhưng cũng có thể là các tiến hoá do thúc đẩy nhân đôi ADN. Mặc dù thường thiếu một

trình tự mã hoá, trình tự Alu thường tìm thấy tại các vị trí không mã hoá nội sinh, tại intron và các trình tự không dịch mã. Do vậy, chúng thường hiện diện trong bản phiên mã ARN đầu tiên từ gen mã hoá polypeptid, thường là mRNA. *In vitro*, trình tự Alu cũng có thể được sao chép từ promoter nội nhờ ARN polymerase III và bản sao của một vài trình tự Alu do sự tích lũy các ARN tế bào chất có thể liên kết với protein của dấu hiệu nhận diện.

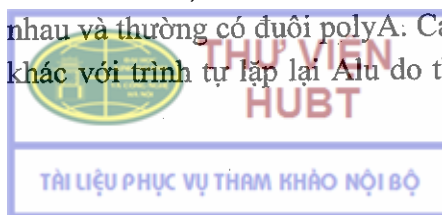


Hình 7.3. Các trình tự SINE

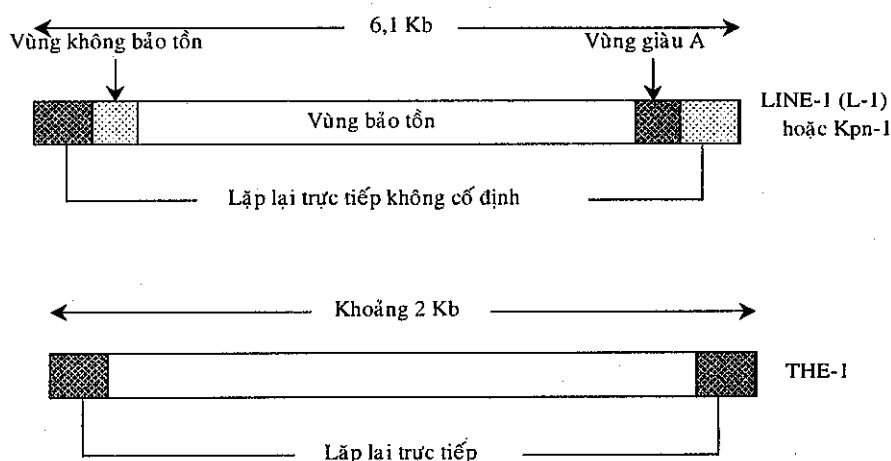
Các trình tự LINE:

Chúng gồm các họ LINE 1 hay *KpnI* và THE 1, xảy ra ở mỗi 50 kb trên bộ gen người. Các trình tự LINE được tạo ra do cắt bộ gen với enzym giới hạn *KpnI*. Chúng có chiều dài khoảng 6 và 7 Kb với gần 5 000 bản sao nguyên vẹn và 100.000 bản sao từng phần rải khắp bộ gen người. Chúng là những trình tự lặp lại không mã hóa dài nhất và thường ở vùng giàu AT, giống như các trình tự Alu (Hình 7.4). Các bản sao mã trình tự LINE gắn với protein tạo thành phức hợp ribonucleoprotein. Ở một dòng tế bào người bị ung thư (teratocarcinome), người ta quan sát thấy các ribonucleoprotein này. Sự xen đoạn LINE vào các vị trí khác nhau có thể gây hậu quả nhất định, như trong trường hợp bệnh máu không đông A (hemophilia A).

Trong họ LINE 1, nhiều đoạn có khả năng chuyển vị. Chiều dài toàn bộ là 6,1 kb và có 2 ORF, tuy nhiên chúng không hiện diện trong hầu hết trình tự riêng rẽ. ORF1 ở gần cuối (đầu tận cùng 5') và mã hóa protein chưa rõ chức năng p40 (trọng lượng phân tử khoảng 40 kDa). ORF2 có những vùng giống với trình tự nucleotid mã hóa các enzym phiên mã ngược khác nhau và các protein virus khác. Cấu trúc hoàn chỉnh của LINE 1 chứa một promoter trong vùng ADN không dịch mã trước ORF1 (gọi là 5'-UTR) trong khi tại đầu 3' có trình tự (A)_n/(T)_n, thường có đuôi poly A. Trong trường hợp các cấu trúc khác có thể chuyển vị và các cấu trúc LINE-1 bị kẹp bởi các trình tự đôi ngắn. Cấu trúc LINE-1 hoàn chỉnh khá hiếm (chỉ khoảng 3.500 bản sao) và hầu hết các đoạn lặp lại bị cụt ở đầu 5' do vậy chúng có chiều dài khác nhau và thường có đuôi polyA. Các cấu trúc LINE-1 thường ở vùng chất nhiễm sắc nhưng khác với trình tự lặp lại Alu do thích ở nhánh G tối (Giemsa



ương) của nhiễm sắc thể kì giữa. Giống Alu, chúng thiếu các trình tự mã hóa nhưng có thể tìm thấy các trình tự không mã hóa nội sinh. Do vậy, chúng hiện diện trong bản sao ARN đầu tiên của các gen lớn nhưng không có trong mARN.



Hình 7.4. Trình tự LINE

Các gen của rARN, tARN, ARN 5S và ARN 7SL:

Một loạt các gen giữ vai trò quan trọng trong sinh tổng hợp protein như các gen ribosome, tARN, ARN 5S và ARN 7SL có sự lặp lại hàng nghìn lần.

– Gen nhóm I: các gen của rARN:

Các gen rARN được phiên mã bởi ARN polymerase I. Bản phiên mã đầu tiên là pre-rARN, sau khi cắt nối tạo ra 3 loại rARN là rARN 28S, rARN 18S và rARN 5,8S. Các gen này không phân tán mà xếp thành cụm (cluster), mỗi cụm có thể hơn 200 bản sao. Ở người, các cụm đó được tìm thấy trên vai ngắn của các nhiễm sắc thể tâm đầu (acrocentric) 13, 14, 15, 21 và 22, chiếm khoảng 0,4% bộ gen. Các nhóm này xếp quanh yếu tố tổ chức hạch nhân hình thành kiểu cấu trúc đặc biệt trên nhiễm sắc thể và ở gian kỳ tạo nên các hạch nhân (nucleolus).

Các gen 28S, 5,8S và 18S rARN bất thường do các gen nhân chiếm số lượng lớn, được phiên mã riêng rẽ, đầu tiên chúng được biểu hiện như là các bản sao đa gen, theo cách của gen ty thể. Đoạn sao mã 13 kb biểu hiện tiền 45S rARN sau đó trải qua nhiều biến đổi khác nhau để tạo thành 28S, 5,8S và 18S rARN trưởng thành. Giống như các sản phẩm của gen ty thể, các sản phẩm rARN riêng rẽ của nhóm rADN là các thành phần có liên quan với một chức năng riêng biệt. Do vậy, việc sử dụng bất thường các bản sao ARN không khác so với các sản phẩm dịch mã đầu tiên từ từ được tách ra thành 2 hay nhiều chuỗi polypeptid có chức năng thông thường. Ví dụ, phân tử insulin có 2 chuỗi polypeptid được tạo thành từ sản phẩm dịch mã duy nhất. Cách di truyền này có liên quan đến các sản phẩm



không trùng lặp từ tiền phân tử duy nhất hiếm khi xảy ra: các thành phần riêng rẽ của đa số protein đa tiểu đơn vị ở người được mã hóa bởi các gen khác nhau, thường trên các nhiễm sắc thể khác nhau.

– **Gen nhóm III: các gen tARN, ARN 5S và ARN 7SL:**

Các gen này được phiên mã bởi ARN polymerase III nên gọi là nhóm III. Nhóm này gồm các gen mã hoá cho một số ARN nhỏ tìm thấy trong nhân và tế bào chất.

Ở người có hơn 200 gen mã hoá cho ARN 5S. Các gen này tập trung thành cụm ở một số điểm nhất định của chất nhiễm sắc thể. Ở một số loài như *Xenopus* số bản sao của ARN 5S có thể hơn 20 000.

7.2.2.4. Các trình tự độc nhất

Các trình tự độc nhất chiếm phần lớn bộ gen (chiếm khoảng 60%), gồm 2 loại:

- Các gen mã hoá cho protein.
- Các gen giả.

Các gen mã hoá cho protein thuộc nhóm II vì nó được phiên mã do ARN polymerase II. Trừ các gen mã hoá cho protein histone, các gen nhóm II có trình tự độc nhất hoặc với số ít bản sao. Các gen nhóm II này hầu như chỉ mã hoá cho một loại protein.

Đa số chúng có trình tự độc nhất hay gần như độc nhất. Một số gen có bản sao thứ hai trong quá trình tiến hoá. Cả hai bản sao này có thể chuyển đổi hỗ trợ nhau, như trường hợp các gen α của globine. Đôi khi, hai bản sao có sự phân hoá nhỏ trong tiến hoá ra hai protein nhưng ít khác nhau.

Một gen có thể tạo ra được nhiều bản sao rất sớm trong quá trình tiến hoá. Mỗi bản sao có sự phân hoá độc lập nhau do hiện tượng nhân đôi (duplication). Sự phân hoá này dẫn đến hàng loạt gen mã hoá cho các protein tương đồng. Sự biểu hiện của các gen này phụ thuộc vào kiểu hay trạng thái của tế bào. Các gen như vậy hiện diện ở dạng lặp lại thành họ. Ví dụ, nhiều họ gen đã được biết như:

- Họ các gen Globine.
- Họ các gen Actine.
- Họ các gen Myosine.

Các gen Globine tập hợp thành cụm: phức hợp của cụm α Globine nằm trên vai ngắn của nhiễm sắc thể 16 và phức hợp β Globine nằm trên vai ngắn của nhiễm sắc thể 11. Trình tự sắp xếp của gen tương ứng với trình tự biểu hiện trong quá trình phát triển cá thể.



Hemoglobine là protein có cấu trúc bậc bốn. Ở động vật có vú trưởng thành, cấu trúc bậc bốn gồm 2 mạch polypeptide α và 2 mạch β .

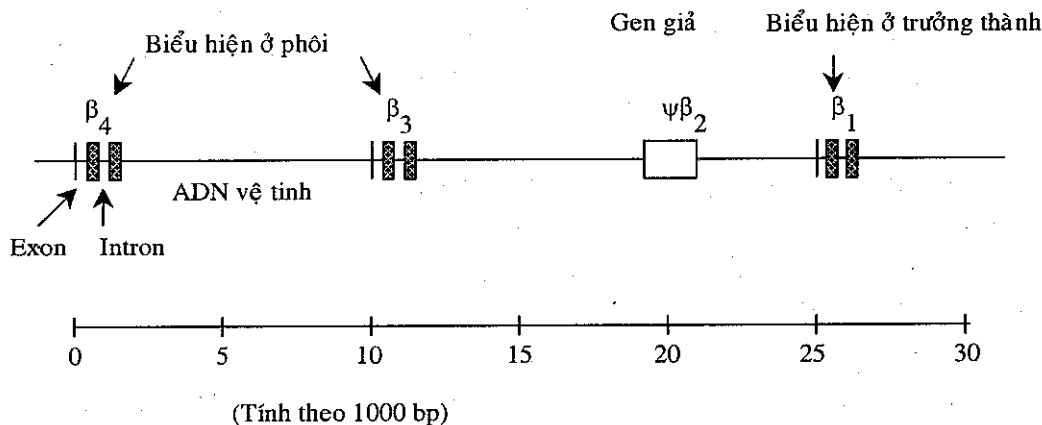
Trong quá trình phát triển cá thể có sự thay đổi các dạng cấu phần như (Bảng 7.2).

Bảng 7.2. Mạch Globine ở người trong quá trình phát triển cá thể

Peptide globine	Phôi	Thai	Người lớn
α - tương tự	ξ (xi)	α	α
β - tương tự	ϵ (epsilon)	$G\gamma, A\gamma$	δ, β

Gen giả được xác định bởi việc chứa các trình tự có liên quan mật thiết với các gen cấu trúc, nhưng không dịch mã thành protein chức năng được.

Sự sắp xếp của cụm các gen β - tương tự globine của thỏ được trình bày ở Hình 7.5. Có 4 chuỗi β - tương tự globine trong nhóm này. Gen β_1 được biểu hiện ở các thể trưởng thành, trong khi đó các gen β_3 và β_4 thì có hoạt tính chủ yếu trong quá trình phát triển của phôi.



Hình 7.5. Sự sắp xếp các gen β tương tự

Gen giả $\Psi\beta_2$ của thỏ cũng chứa intron và exon liên quan mật thiết với gen cấu trúc β_1 . Nhưng việc mất một cặp base ở codon thứ 20 của nó đã gây nên hiện tượng đột biến lệch khung làm chấm dứt ngay sự dịch mã sau đó. Gen giả cũng có thể được tạo ra khi 2 intron có xen giữa là một exon và không có intron nào được cắt ra khỏi exon khi gen phiên mã.

Phân tích trình tự β_2 cho thấy rằng mặc dù gen β_2 có quan hệ mật thiết với các trình tự khác trong cụm, nhưng nó lại tích tụ một số đột biến bao gồm lệch khung

và chuỗi kết thúc, ngăn chặn β_2 mã hoá protein ở bất kỳ phương diện nào. Do đó, β_2 của thỏ là một ví dụ của gen giả. Gen giả được biểu thị theo chữ Hy Lạp là Ψ , vì vậy gen này được ký hiệu là $\Psi\beta_2$. Loại gen giả này được xem là kết quả tiến hoá của một gen đã có khi biểu hiện nhưng bây giờ lại là gen đột biến lặn. Có một vài ví dụ khác của các gen giả loại này, như các gen tương tự β - tubulin ở người, các gen rARN 5S ở *Xenopus*.

7.3. CÁC MỨC ĐỘ ĐIỀU HOÀ BIỂU HIỆN GEN

Ở tế bào nhân nguyên thủy, cũng như ở tế bào nhân thật, các cơ chế điều hoà sự biểu hiện của gen có thể tác động ở một hay nhiều mức độ khác nhau. Sự điều hoà có thể ở mức độ ngay bản thân gen, bằng sự kiểm soát thời gian và tốc độ phiên mã. Các cơ chế khác có thể hoạt động lúc dịch mã hoặc sau dịch mã (Hình 7.6).

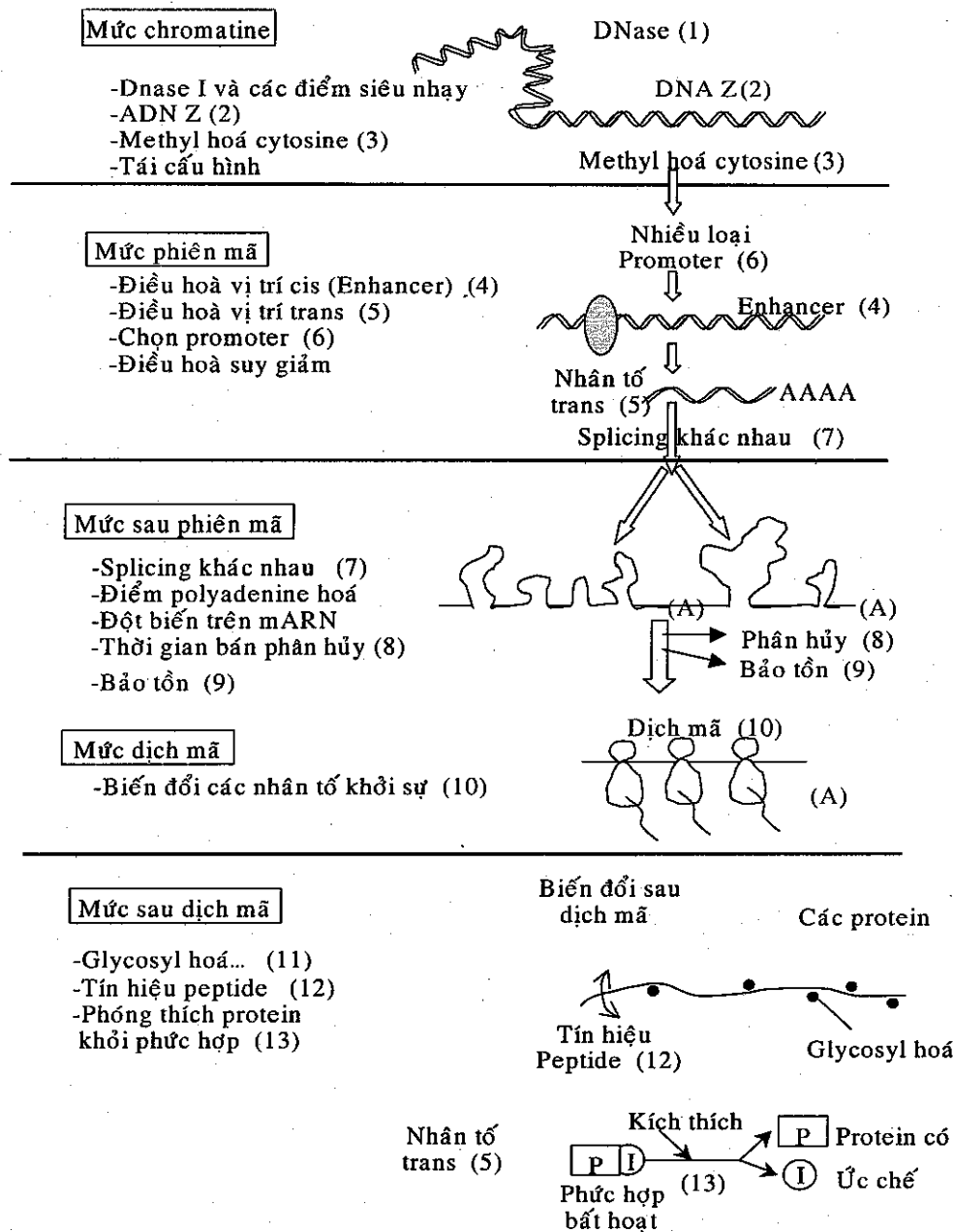
7.3.1. Mức độ chất nhiễm sắc

Ngay trên sợi nhiễm sắc có thể thực hiện các kiểu:

DNaseI cắt ADN bộ gen ở một số vùng làm tháo xoắn để cho các gen biểu hiện (1) của Hình 7.6.

Hai vùng được lưu ý: nhạy cảm và tăng nhạy cảm. Các vùng nhạy cảm có liên quan đến các gen có hoạt tính cao và những gen trước đây đã biểu hiện (gen hoạt động ở phôi).

- ADN Z là dạng cấu trúc siêu xoắn có thể liên quan đến đóng mở gen (2).
- Sự Metyl hoá các base (3) ở tế bào nhân nguyên thủy xảy ra ở A và C, còn ở tế bào nhân thật ở C vị trí 5': Sự metyl hoá làm gen ngừng hoạt động. Ví dụ, nhiễm sắc thể X bất hoạt động ở người thuộc loại siêu metyl hoá.
- Sự thay đổi cấu hình có thể ảnh hưởng đến sự biểu hiện gen.



Hình 7.6. Sơ đồ về các mức điều hoà khác nhau

7.3.2. Mức độ phiên mã

Đây là sự điều hoà ảnh hưởng trực tiếp đến việc mở hoặc đóng gen. Kiểu điều hoà này thường gặp trong điều hoà trao đổi chất, cũng như trong biệt hoá tế bào (xem Hình 7.6).

- Sự tác động của trình tự *cis* (cùng phía) nằm trên cùng mạch ADN như enhancer (đoạn tăng cường) (4).

- Điều hoà bởi các yếu tố *trans* (trái phía) do các yếu tố không nằm cùng trên một mạch ADN (5).
- Chọn lựa promoter thích hợp (6).
- Điều hoà suy giảm (Attenuation).

7.3.3. Mức độ sau phiên mã

Điều hoà khi mARN cắt bỏ các intron và gắn các exon lại với nhau để tạo mARN trưởng thành. Như vậy, các hệ thống ảnh hưởng đến sự trưởng thành của mARN có thể kiểm tra gián tiếp biểu hiện của gen tương ứng. Các mARN của tế bào nhân thật còn có những đoạn không mã hoá liên quan tới thời gian tồn tại và ra khỏi nhân đi vào tế bào chất (xem Hình 7.6).

- Splicing (cắt nối) khác nhau (7).
- Điểm polyadenin hoá khác nhau.
- Đột biến trên phân tử mARN.
- Bán chu kỳ phân huỷ của mARN (8).
- Sự bảo tồn các ARN trong tế bào (9).

7.3.4. Mức độ dịch mã

Sự biến đổi của các yếu tố khởi đầu IF (10).

7.3.5. Mức độ sau dịch mã

Sau khi mạch polypeptid được tổng hợp, các protein có thể phải trải qua các biến đổi thứ cấp trước khi có hoạt tính. Ví dụ, trypsin là enzym phân giải protein trong dạ dày chỉ có được hoạt tính sau khi tiền chất của nó (proenzym) không có hoạt tính, bị cắt mất một đoạn polypeptid.

Các protein có thể chịu những biến đổi lập thể (allosteric) như sự kết hợp các enzym với một số sản phẩm đặc biệt có thể làm thay đổi cấu trúc không gian làm cho chúng có hoạt tính (xem 7.6).

- Glycosyl hoá, phosphoryl hoá... (11): gắn thêm các nhóm chất như đường, phosphat để protein có hoạt tính.

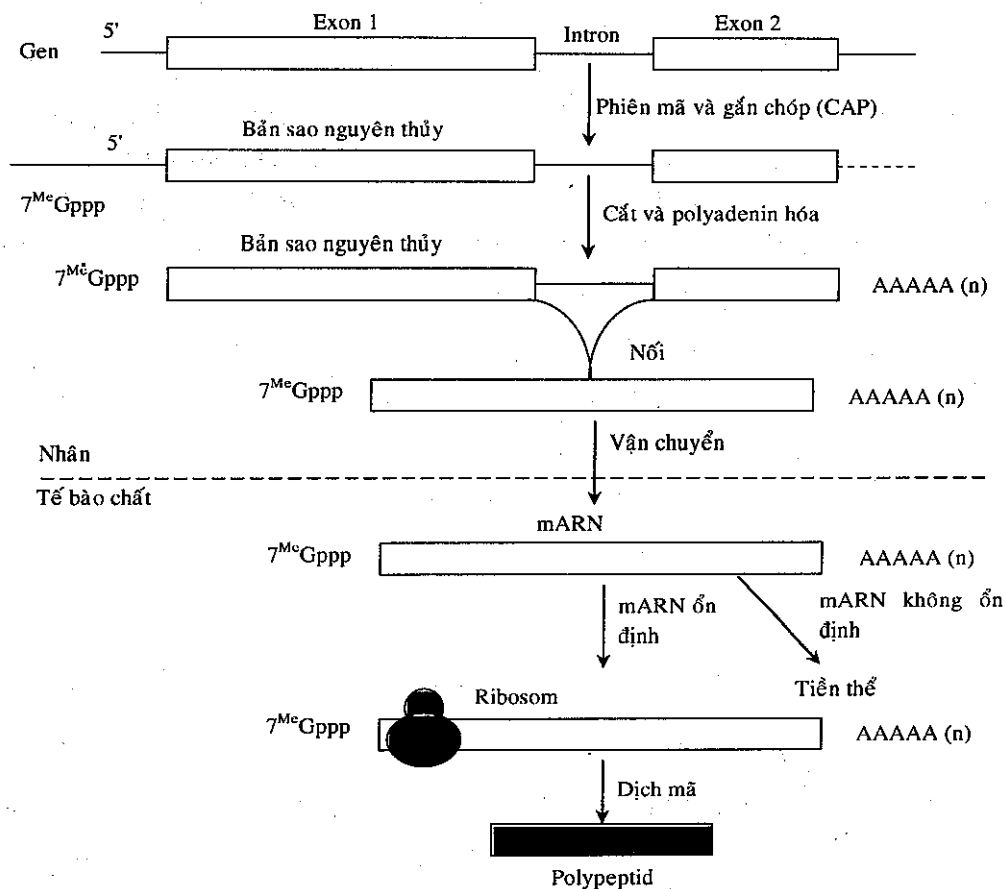
- Tín hiệu peptide (12) là đoạn gồm khoảng 20 acid amin nằm gần phía đầu N của polypeptid, có vai trò gắn polypeptid và ribosome đang được tổng hợp mạch này với lưới nội chất. Trong bộ máy Golgi, polypeptid được phóng thích ra ngoài.

- Sự phóng thích ra protein có hoạt tính từ một phức hợp như từ proinsulin thành insulin.



7.4. ĐIỀU HOÀ HOẠT TÍNH GEN CỦA TẾ BÀO NHÂN THẬT

Trong các tế bào nhân thật có một số các điểm, mà ở đó sự biểu hiện gen có thể được điều hoà (Hình 7.7).



Hình 7.7. Các điểm của gen tế bào nhân thật có thể điều hoà biểu hiện

Một số đặc điểm của điều hoà hoạt động gen ở tế bào nhân thật cần lưu ý

– Ở các operon của tế bào nhân nguyên thủy, các gen điều hoà và các promoter thường nằm gần nhau, nhưng ở tế bào nhân thật thì các gen điều hoà ít khi nằm gần các promoter do chúng kiểm soát.

– Các trình tự tăng cường (enhancer) là những trình tự cùng nằm trên một phân tử với các promoter, có thể có hàng trăm cặp base ở phía trước hoặc ở phía sau promoter mà chúng kích thích.

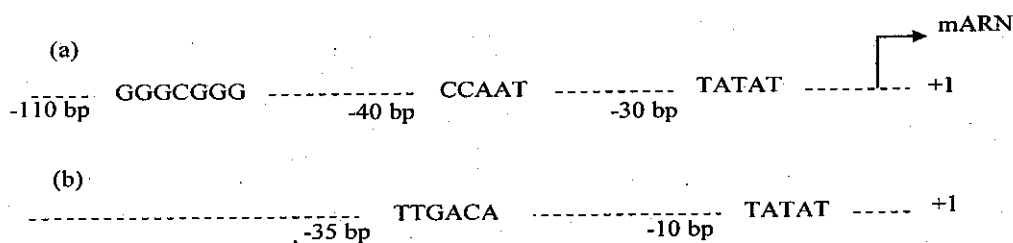
– Trình tự điều hoà 5' ở phía trước các promoter ở tế bào nhân thật thường rất dài, có khi hàng chục kb.

– Có nhiều kiểu điều hoà ở dạng các yếu tố có tác động *trans* là các protein.

- Sự phiên mã có thể được kích thích bởi các tín hiệu khác nhau. Sự điều hoà hoạt động các gen ở tế bào nhân nguyên thuỷ phần lớn đáp lại tín hiệu ngoại sinh. Ngược lại, phần lớn sự điều hoà ở tế bào nhân thật là đáp lại các tín hiệu nội sinh.

7.4.1. Các promoter

Tương tự như ở tế bào nhân nguyên thuỷ, các promoter của tế bào nhân thật cũng nằm phía trước điểm xuất phát của mRNA và cũng có những trình tự chung trong tiến hoá. Hộp TATA định hướng cho ARN polymerase bắt đầu phiên mã, ở phía ngược chiều với chiều phiên mã khoảng 30 bp ở ADN động vật có vú và 60 đến 120 bp ở ADN nấm men. Hộp TATA hoạt động có hiệu quả cùng với hai trình tự tương ứng phía trước khoảng 40 bp là CCAAT và 110 bp là trình tự giàu GC (Hình 7.8).



Hình 7.8. Trình tự nucleotide của promoter

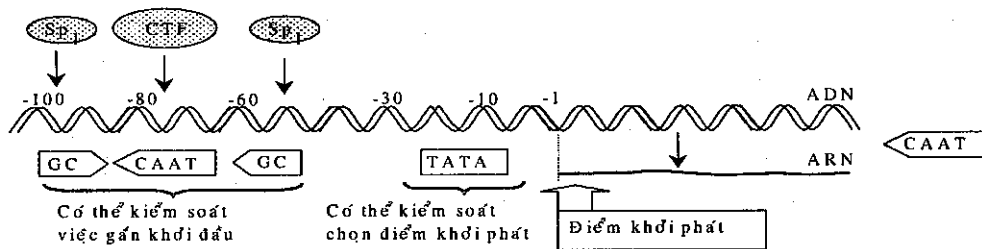
(a) Tế bào nhân thật; (b) Tế bào nhân nguyên thuỷ

Sự thay đổi hộp TATA làm giảm tốc độ phiên mã. Sự giảm tốc độ phiên mã được đo bằng sự thay đổi của từng base trong promoter. Các thay đổi base ngoài hộp TATA và các trình tự phía trước không gây tác động đối với sự phiên mã. Khác với promoter của tế bào nhân nguyên thuỷ, các promoter của tế bào nhân thật khó đảm bảo sự nhận biết các tín hiệu một cách đầy đủ để ARN polymerase khởi đầu phiên mã *in vivo*. Do đó, xoá bỏ hộp TATA không hoàn toàn làm huỷ bỏ sự biểu hiện gen mà làm thay đổi vị trí bắt đầu phiên mã.

Hộp TATA và các trình tự phía trước phải được nhận biết bởi các protein điều hoà. Chính các protein này gắn với các điểm nhất định trên hộp TATA và hoạt hoá sự phiên mã.

Hộp TATA có ở hầu hết các gen tế bào nhân thật. Tuy nhiên, hầu hết các gen mã hóa cho các protein cơ bản nồng độ thấp (low-abundance housekeeping protein) lại không có hộp TATA. Hai trình tự chung khác được tìm thấy ở nhiều gen đó là: hộp CCAAT (phát âm là "cat") thường thấy ở tế bào nhân thật bậc cao như β -globine thỏ trưởng thành, người ta tìm thấy khoảng 70 bp ngược chiều với vị trí bắt đầu phiên mã và có trình tự chung là CCAAT. Sự khiếm khuyết hộp CCAAT làm giảm phần lớn tỷ lệ phiên mã, nhưng không làm thay đổi vị trí khởi đầu phiên mã.

Trình tự thứ hai được biết đến là hộp GC hay hộp Sp1, cũng hiện diện ngược chiều ở một số các gen tế bào nhân thật, thường là ở trong các bản sao nhiều lần. Ví dụ, ở trình tự promoter của gen Thymidine kinase có 2 hộp GC ngược chiều với hộp TATA (Hình 7.9).



Hình 7.9. Promoter của Thymidine kinase có chứa các trình tự tăng cường khác nhau

Hộp GC này cũng có kiểu hình trội, được tìm thấy ở promoter của các gen bảo vệ không có hộp TATA. Cũng như hộp CCAAT, hộp GC cần thiết cho sao mã.

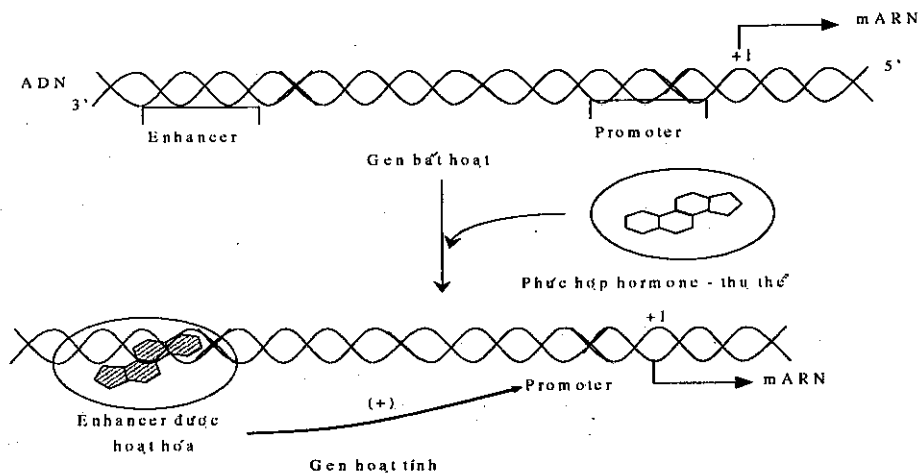
Đáng lưu ý là, cả 2 hộp CCAAT và GC hoạt động theo cả 2 hướng. Hộp GC tương tác với một protein được gọi là Sp1, còn hộp CCAAT tương tác với protein ký hiệu là CTF (CTF = CCAAT binding Transcription Factor). Nó là yếu tố điều hoà *trans* gắn trên hộp CCAAT và có thể là với các protein khác.

7.4.2. Các yếu tố tăng cường (Enhancer)

Enhancer là các trình tự có tác động *cis* (cùng phía), chúng tăng tốc độ phiên mã đáng kể từ ngay các promoter nằm ngay trên cùng một phân tử ADN. Tính độc đáo của enhancer là ở chỗ chúng có khả năng thực hiện tác động cách xa đến vài nghìn cặp base. Ngoài ra, chúng có thể hoạt động ở bất kỳ hướng nào, dù ở phía trước hay phía sau promoter.

Trình tự tăng cường đầu tiên được mô tả ở virus SV40, được định vị khoảng 200 bp ngược chiều với vị trí bắt đầu phiên mã của gen sớm. Nếu loại bỏ trình tự này sẽ làm giảm đáng kể khả năng phiên mã từ trình tự promoter của gen sớm. Tuy nhiên, sự tái chèn của trình tự tăng cường vào bất kỳ nơi nào của bộ gen SV40 sẽ giữ lại mức độ biểu hiện bình thường. Hơn nữa, nếu trình tự tăng cường SV40 được nối với gen β -globine thì sẽ đạt ngay sự biểu hiện ở cách xa 200 bp về phía trước hay phía sau promoter β -globine và ngay cả khi trình tự khuếch đại SV40 nằm cách vị trí khởi đầu phiên mã tới vài nghìn cặp base. Gần đây, người ta tìm thấy các trình tự tăng cường có ở các gen chuỗi nặng globine miễn dịch ở động vật có vú và gen albumine chuột. Khoảng cách giữa một trình tự tăng cường và một trình tự khởi động lớn nhất gặp ở gen albumine chuột là 10 kb.

Xét một số mặt nào đó, các enhancer tương tự các promoter. Chúng được tổ chức thành một dãy các trình tự có tác động *cis* để nhận biết các yếu tố tác động (Hình 7.10).



Hình 7.10. Sự tương tác của hormon steroid với enhancer

Hormon glucocorticoide gắn vào protein thụ thể hoà tan trong phức hợp này sau đó lại gắn vào enhancer kích thích phiên mã.

Các trình tự tại *cis* có đặc điểm chung là chúng thường có cấu trúc gồm hai phần đối xứng nhau. Ví dụ, trình tự đáp ứng với hormon tuyến giáp (TRE – Thyroid hormone Response Element) dưới đây:



Sở dĩ như vậy là vì các trình tự này thường tiếp nhận các protein điều hoà (yếu tố *trans*) dưới dạng dimer (cấu tạo từ hai tiểu đơn vị).

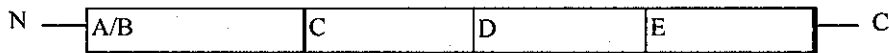
7.4.3. Các yếu tố protein tham gia vào quá trình điều hoà biểu hiện gen, các yếu tố *trans*

Đặc điểm chung của các yếu tố *trans* là gồm ít nhất 2 vùng cấu trúc – chức năng chính:

- Vùng gắn yếu tố *trans* vào ADN.
- Vùng tác động lên sự phiên mã.

Các vùng cấu trúc – chức năng này độc lập với nhau.

Tiến hành ghép hai trình tự ADN có nguồn gốc khác nhau: vùng tác động phiên mã lấy từ một yếu tố *trans* ở động vật có vú và vùng gắn vào ADN lấy từ yếu tố *trans* ở nấm men. Kết quả tạo ra một yếu tố "lai" có khả năng hoạt hoá các gen đích tương ứng ở nấm men. Gen đích là gen mang trình tự tiếp nhận yếu tố *trans*. Ngoài hai vùng kể trên, nhiều yếu tố *trans* còn mang một số vùng khác như: vùng gắn hormon, các ion... (Hình 7.11).



Hình 7.11. Cấu trúc một yếu tố trans

Hình 7.11 cho thấy cấu trúc của một thụ thể tuyến giáp (THR – Thyroid Hormon Receptor) bao gồm nhiều vùng cấu trúc và chức năng, trong đó hai vùng được biết rõ nhất là:

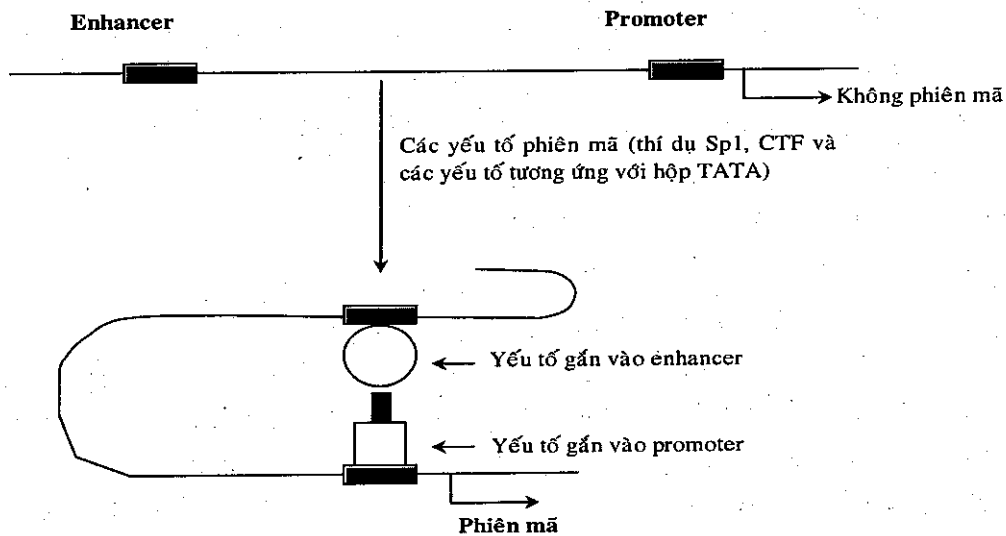
- C: vùng gắn lên ADN.
- E: vùng tiếp nhận hormon tuyến giáp.

Hormon tuyến giáp khi gắn lên vùng E của thụ thể sẽ làm biến đổi cấu hình thụ thể. Lúc đó, vùng C của thụ thể có thể gắn lên vùng ADN điều hoà của các gen đích và kích thích sự phiên mã của các gen này.

Các kết quả nghiên cứu cho thấy các yếu tố *trans* có thể phân vào bốn nhóm cấu trúc:

- Ngón tay kẽm (zinc – finger).
- Xoắn – vòng – xoắn (helix – turn – helix).
- Xoắn – cuộn – xoắn (helix – loop – helix).
- Dây kéo leucine (leucine – zipper).

Như vậy, các gen tế bào nhân thật được hoạt hoá bởi hai trình tự ADN có tác động *cis* là promoter và enhancer. Chúng được nhận biết bởi các yếu tố protein có tác động *trans*. Các yếu tố này cho phép ARN polymerase khởi sự phiên mã và đạt tốc độ phiên mã tối đa (Hình 7.12).



Hình 7.12. Cơ chế "tiếp xúc trực tiếp" của trình tự enhancer

7.4.4. Hormon

Hormon là một trong các chất điều hoà nội tại của hoạt tính gen. Đó là những chất được tạo ra do một loạt tế bào có tác động đến các tế bào khác. Các hormon chỉ tác động đến các tế bào có thụ thể tương ứng. Cấu trúc thụ thể bao gồm nhiều vùng trong đó có 2 vùng được biết rõ nhất:

- Vùng gắn vào ADN với hai cấu trúc "ngón tay kẽm" (zinc finger).
- Vùng gắn vào hormon.

Sự tương tác hormon với thụ thể gây ra tín hiệu tác động đến các vùng đặc hiệu của ADN, làm hoạt hoá gen hoặc nhóm gen tương ứng.

- Các hormon có thể kích thích phiên mã bởi một trong các cơ chế sau:
 - Tách ADN khỏi histone và tạo điều kiện cho ARN polymerase bắt đầu phiên mã.
 - Tác động như một chất cảm ứng, gây bất hoạt phân tử ức chế.
 - Gắn trực tiếp với đoạn ADN đặc hiệu, tạo thuận lợi cho ARN polymerase gắn vào để phiên mã.
 - Hoạt hoá protein hiệu ứng, tạo thành phức hợp hoạt hoá, kích thích ARN polymerase gắn vào để phiên mã.
 - Gắn với protein tạo phức hợp hoạt hoá, làm cho ARN polymerase gắn vào phiên mã.

7.5. SỰ KIỂM SOÁT CÁC CHẤT THƯỜNG GẶP TRONG NHÂN

Một số nhóm phân tử hiện diện nhiều trong tế bào nhân thật như histone, các thành phần của bộ máy dịch mã, các thành phần của màng, tế bào,.... Để duy trì số lượng lớn chúng trong tế bào phải có các chương trình:

- Phiên mã liên tục và lặp lại trong chu trình tế bào.
- Lặp lại các gen.
- Khuếch đại ngoài nhiễm sắc thể của các trình tự gen đặc hiệu.

7.5.1. Sự dồi dào ARN

Không phải tất cả các gen đều biểu hiện ở cùng một mức độ. Điều này dẫn đến sự khác nhau đáng kể về mức độ tập trung của các mARN và các protein riêng lẻ. mARN được phân chia một cách gượng ép thành 3 nhóm: dư ít, trung bình và nhiều. Một tế bào ổn định có dưới 100 mARN dư, một vài tế bào chỉ có 1 hoặc 2 mARN dư. Trong một tế bào, trung bình có khoảng 33% tổng số mARN là mARN dư nhiều, mỗi mARN dư này có thể chiếm tới 1% trở lên trong tổng số mARN của tế bào.

Kiểu hình của một tế bào lệ thuộc rất nhiều vào quá trình tổng hợp của một hoặc vài protein dư, vì vậy cần phải có các phân tử mRNA dư. Ví dụ, globin trong tế bào lưới nội chất, sợi cơ và myosin trong các tế bào cơ. Vì vậy, các mRNA dư nhiều thường là các bản sao của các gen chuyên biệt của tế bào.

Một tế bào chứa vài trăm mRNA dư trung bình, chiếm khoảng 33% tổng số mRNA tế bào, nhưng mỗi loại mRNA dư chỉ chiếm 0,1 – 0,5% tổng số mRNA. Phần mRNA còn lại của tế bào gồm các bản mã sao dư ít. Có khoảng 10000 bản mã sao như vậy trong một tế bào điển hình, mỗi bản sao đại diện cho ít hơn 0,01% tổng số mRNA.

Bảng 7.3. Các nhóm mRNA dư ở vôi trứng gà và gan chuột nhắt

Loại tế bào	Nhóm dư	% của tổng mRNA trong nhóm	Số lượng các loại mRNA khác
Vôi trứng gà	Nhiều	50	1
	Trung bình	15	7
	ít	35	12.500
Gan chuột nhắt	Nhiều	22	9
	Trung bình	41	700
	ít	37	11.500

Các nhóm gen dư ít hay dư trung bình đều nhạy cảm không kém so với dư nhiều.

Các loại ARN khác cũng có hiện tượng dư. Ở *Xenopus* có vùng tổ chức hạch nhân chứa 450 bản sao của ADN mã hoá cho rARN 18S và 28S. Ngược lại, trong mỗi nhân có từ 20000 bản mã sao của các gen mã hoá cho rARN 5S và các gen này không nằm trong vùng tổ chức hạch nhân.

Các gen mã hoá cho histone cũng có nhiều bản sao được lặp lại hàng trăm lần.

Sự dồi dào của các gen nêu trên đảm bảo đủ số lượng cần thiết cho dịch mã khi cần phải tổng hợp hàng triệu phân tử protein.

7.5.2. Sự khuếch đại gen

Một ví dụ về khuếch đại gen trong nhân là các chỗ phình (puff) của nhiễm sắc thể khổng lồ hay đa sợi ở tuyến nước bọt *Drosophilla*. Ở các chỗ phình này, ADN nguyên nhiễm sắc được khuếch đại khoảng nghìn lần.

Trong nhân của tế bào trứng *Xenopus* có tới hàng trăm nhân con ngoài nhiễm sắc thể với kích thước khác nhau. Mỗi nhân con chứa các vòng rARN có kích thước khác nhau, vai trò chưa rõ. Các ADN vòng này sản sinh nhiều rARN để lắp ráp thành nhiều ribosome.

Tóm lại, sự biểu hiện của các gen ở tế bào nhân thật rất phức tạp. Tuy vậy, hàng nghìn promoter được điều hoà để tạo ra các mARN ở mức độ hợp lý.

Tóm tắt:

Kích thước và cấu trúc bộ gen tế bào nhân thật phức tạp hơn nhiều so với bộ gen của tế bào nhân nguyên thủy. ADN của bộ gen tế bào nhân thật có mức độ lặp lại cao, trung bình và đơn độc.

Sự điều hoà biểu hiện gen được thực hiện ở tế bào nhân thật ở nhiều mức độ khác nhau:

– Ở mức độ nhiễm sắc thể, phiên mã, sau phiên mã và dịch mã. Bộ máy di truyền không những có các gen cấu trúc mà còn có những trình tự khác tham gia biểu hiện gen.

– Ở tế bào nhân thật có nhiều cơ chế điều hoà hoạt động gen phức tạp như: các yếu tố *cis* (enhancer), các yếu tố *trans*, các hormon. Ngoài ra, còn có sự dư dôi của một số bản mã sao và sự khuếch đại một số gen cần thiết cho hoạt động của tế bào.

CÂU HỎI

1. Giá trị C là:
 - a) Tổng số nucleotid có trong tế bào
 - b) Kích thước nhiễm sắc thể
 - c) Kích thước gen
 - d) Kích thước bộ gen lưỡng bội
 - e) Kích thước bộ gen đơn bội
2. Nghịch lý giá trị C
 - a) Là giá trị nghịch đảo của giá trị C
 - b) Sự khác biệt giữa giá trị C của sinh vật bậc cao với vi khuẩn
 - c) Sự khác biệt giữa giá trị C của người với sinh vật bậc cao khác
 - d) Sự không tương xứng giữa kích thước bộ gen với số gen
 - e) Sự không tương xứng giữa bộ gen đơn bội và lưỡng bội
3. Các trình tự lặp lại ở ADN tế bào nhân thật gồm:
 - a) SINE
 - b) LINE
 - c) CEN
 - d) TEL
 - e) Tất cả



4. Trình tự TEL
- a) Thuộc nhóm telomer
 - b) Bảo vệ đầu mút nhiễm sắc thể
 - c) Kiểm hãm sự tiến hoá
 - d) Không gắn với màng nhân
 - e) a và b
5. Trình tự SINE
- a) Còn gọi là Alu
 - b) Không chứa nhóm Alu
 - c) Ngắn hơn LINE
 - d) Dài hơn LINE
 - e) a và d
6. Gen giả là:
- a) Trình tự lặp lại thấp
 - b) Trình tự lặp lại cao
 - c) Trình tự độc nhất
 - d) Không mã hoá cho protein
 - e) c và d
7. Các hormon có thể kích thích phiên mã bằng cách:
- a) Tách ADN khỏi histone
 - b) Tác động như một chất cảm ứng
 - c) Hoạt hoá protein hiệu ứng
 - d) Gắn với protein tạo phức hợp hoạt hoá
 - e) Tất cả
8. Các mức điều hoà biểu hiện gen ở tế bào nhân thật
- a) Chất nhiễm sắc
 - b) Phiên mã
 - c) Hậu phiên mã
 - d) Dịch mã và hậu dịch mã
 - e) Tất cả
9. Enhancer
- a) Có tác động *cis*
 - b) Hoạt động theo kiểu *trans*
 - c) Có khả năng thực hiện tác động cách xa đến vài nghìn cặp base
 - d) Hoạt động theo hướng 5' → 3'
 - e) a và c
10. Điều hoà hoạt động gen ở tế bào nhân thật khác tế bào nhân nguyên thủy
- a) Trình tự điều hoà 5' thường rất dài
 - b) Trình tự điều hoà 3' thường rất dài
 - c) Chủ yếu đáp ứng lại tín hiệu ngoại sinh
 - d) Chủ yếu đáp ứng lại tín hiệu nội sinh
 - e) a và d.

Bài 8

ĐỘT BIẾN GEN

MỤC TIÊU

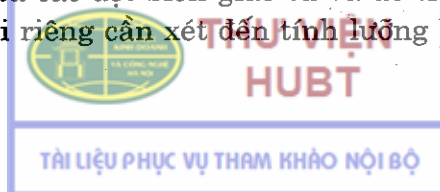
- Trình bày được nguyên nhân và hậu quả của đột biến.
- Trình bày được cách tế bào sửa chữa ADN hư hỏng để tránh đột biến.
- Nhận diện được một số bệnh di truyền ở người do đột biến gây ra.
- Trình bày được tác động của các yếu tố gây đột biến.
- Mô tả được việc sử dụng các yếu tố gây đột biến trong phòng thí nghiệm và các phương pháp chọn lọc đột biến ở vi sinh vật.

8.1. MỞ ĐẦU

Đột biến theo nghĩa rộng chỉ các biến đổi di truyền xảy ra đột ngột. Một đột biến gen là một sự thay đổi trong trình tự nucleotid trên gen, dẫn đến kết quả là tạo ra một protein đột biến với trình tự acid amin thay đổi hay tác động đến quá trình điều khiển sự tổng hợp của các sản phẩm gen. Nếu protein đột biến có chức năng khác với dạng tự nhiên thì sẽ tạo ra một thay đổi tương ứng ở những đặc tính có thể quan sát được và tạo ra thể đột biến. Sự khác biệt có thể do hoạt tính hay tính ổn định của enzym. Một số đột biến có thể tồn tại trong quần thể bên cạnh các dạng tự nhiên làm thành một thể đa hình. Trong thể đa hình, không thể phân biệt giữa dạng đột biến và dạng tự nhiên.

Đột biến xảy ra ở các tế bào không sinh sản được gọi là đột biến sinh dưỡng và tạo ra các thay đổi có tính cục bộ, không di truyền được như ở các u sắc tố ngoài da người, tạo ra nốt ruồi. Những đột biến sinh dưỡng cũng có thể là nguyên nhân gây ra ung thư và có thể là nguyên nhân của sự lão hoá.

Đột biến xảy ra trong giao tử được gọi là đột biến dòng mầm và được truyền cho đời sau. Khi nghiên cứu các đột biến giao tử và do tỉ lệ đột biến ở sinh vật đa bào nói chung và người nói riêng cần xét đến tính lưỡng bội. Hầu hết các đột biến



gen là đột biến lặn và sẽ không được phát hiện nếu hợp tử không có hai bản của allele đột biến. Do đó việc phát hiện và đo tần số đột biến phải dựa vào quan sát các đột biến trội đang xảy ra trên nhiễm sắc thể thường, các đột biến lặn và trội trên nhiễm sắc thể X. Ví dụ, chứng loạn sản sụn xảy ra rải rác là bởi các đột biến mới trong gen của thụ thể yếu tố tăng trưởng nguyên bào sợi. Trong một năm có bảy trẻ sơ sinh bị loạn sản sụn trên tổng số 242 257 trẻ sơ sinh, tỉ lệ là $7/242\ 257 \times 1/2$ (2 allele/hợp tử) = $1,4 \times 10^{-5}$.

Ở người, các gen có tỉ lệ đột biến cao (khoảng 10^{-4}) là bệnh u xơ thần kinh (NF) và bệnh loạn dưỡng cơ Duchenne (DMD), có tỉ lệ đột biến thấp (khoảng 10^{-6}) là gen bệnh Huntington. Sự cách biệt hàng trăm lần cho thấy tỉ lệ đột biến gen có thể khác nhau thật sự.

Có hai giải thích cho sự khác nhau về tỉ lệ đột biến của các gen:

- Kích thước đích: gen NF và DMD có vùng mã hoá protein rất lớn, có nhiều base có thể bị thay đổi hoặc làm mất chức năng của gen.

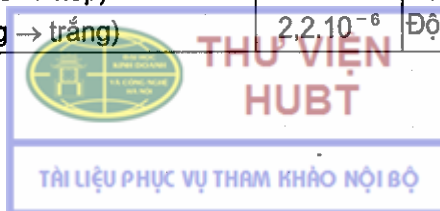
- Điểm nóng: một số gen nằm trong vùng nhiễm sắc thể nhạy cảm hơn với sự phá huỷ/ thay đổi di truyền hoặc gen có chứa các trình tự dễ bị thay đổi bởi các đột biến tự phát hơn. Ví dụ, gen chứng loạn sản sụn có chứa điểm nóng (trình tự CpG).

Trong tự nhiên, dù trong điều kiện nào, tất cả các gen đều có đột biến được gọi là đột biến tự nhiên hay đột biến tự phát. Các đột biến tự nhiên thường xuất hiện rất ít. Các gen khác nhau của cùng một sinh vật có thể có tần số đột biến khác nhau nhưng tần số đột biến tự nhiên đối với mỗi gen là ổn định.

Tần số đột biến được đánh giá căn cứ trên một lần sao chép, một lần phân bào hay trên một giao tử và trên một tế bào trong một thế hệ (Bảng 8.1).

Bảng 8.1. Tần số đột biến tự nhiên của một số gen

Sinh vật	Đột biến	Tần số	Căn cứ đánh giá
Thực khuẩn thể T_2	Kim hãm tạo $r \rightarrow r^+$	10^{-8}	Đột biến gen/ lần sao chép
<i>E. coli</i>	Lên men Lactose $Lac^- \rightarrow Lac^+$	2.10^{-7}	Tế bào đột biến/lần phân bào
	$Leu^- \rightarrow Leu^+$	7.10^{-7}	Đột biến / tế bào
	$Arg^+ \rightarrow Arg^-$	4.10^{-9}	Đột biến / tế bào
	$Trp^+ \rightarrow Trp^-$	6.10^{-8}	Đột biến / tế bào
	$Ara^+ \rightarrow Ara^-$	2.10^{-6}	Đột biến / tế bào
	$StrS \rightarrow StrR$ (nhảy \rightarrow kháng Str)	4.10^{-10}	Đột biến / tế bào
<i>Neospora crassa</i>	Adenin: $ade^- \rightarrow ade^+$	4.10^{-8}	Đột biến / bào tử vô tính
Bắp	$R \rightarrow r$ (đỏ \rightarrow trắng)	$4,92.10^{-8}$	Đột biến / giao tử
	$C \rightarrow c$ (màu \rightarrow không màu)		Đột biến / giao tử
	$Wx \rightarrow wx$ (tẻ \rightarrow nếp)		Đột biến / giao tử
	$Y \rightarrow y$ (vàng \rightarrow trắng)	$2,2.10^{-6}$	Đột biến / giao tử



Sinh vật	Đột biến	Tần số	Căn cứ đánh giá
<i>Drosophila</i>	$W \rightarrow w$	10^{-5}	Đột biến / giao tử
	$e^+ \rightarrow e$ (màu đen hổ phách)	2.10^{-5}	Đột biến / giao tử
	$ey^+ \rightarrow ey$ (có mắt \rightarrow không mắt)	6.10^{-5}	Đột biến / giao tử

Tần số đột biến có thể hiểu một cách đảo ngược. Ví dụ, ruồi dấm thân xám bình thường (e^+) đột biến thành thân có màu đen hổ phách e (ebony) với tần số 2.10^{-5} , nghĩa là trong 10^5 giao tử có hai đột biến từ $e^+ \rightarrow e$. Hoặc ví dụ *E. coli* đột biến nhạy cảm với streptomycin thành kháng streptomycin với tần số 4.10^{-10} đột biến trong một tế bào ở một thế hệ. Để dễ hiểu ta có thể tính ngược lại, tức là trong 10 tỉ tế bào của một thế hệ có 4 đột biến StrR xuất hiện ngẫu nhiên.

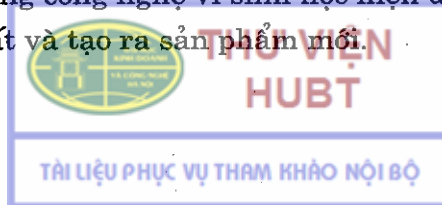
Tuy tần số đột biến của từng gen là rất thấp, nhưng tổng các đột biến của nhiều gen là một số đáng kể, có ý nghĩa quan trọng trong tiến hoá. Các kết quả nghiên cứu *in vivo* và *in vitro* trên tế bào người cho thấy tỉ lệ đột biến ở người khoảng 10^{-6} /gen/thế hệ. Các vi sinh vật tế bào nhân nguyên thủy và tế bào nhân thật cũng có tỉ lệ đột biến tương tự. Có thể ước tính các thay đổi xảy ra trong mỗi thế hệ dựa vào tỉ lệ đột biến người như sau: 10^{-6} đột biến/gen $\times 5.10^4$ gen/bộ gen đơn bội bằng 5.10^{-2} đột biến trên giao tử (5/100 hoặc 1/20). Mỗi hợp tử có (1/20). 2 giao tử trên hợp tử bằng 1/10 cơ hội mang đột biến mới đầu đó trong bộ gen. Con số này có vẻ rất cao nhưng cần nhớ rằng hầu hết các đột biến là lặn và như vậy sẽ không được biểu hiện trong điều kiện dị hợp tử.

Sự thay đổi các base có thể xảy ra trong vùng không mã hoá của phân tử ADN và các acid amin bị thay đổi cũng có thể không phải lúc nào cũng làm thay đổi đặc tính protein. Do vậy, thay đổi trình tự nucleotid có thể không biểu hiện thành sự khác biệt tính trạng và do đó không phải là đối tượng của chọn lọc. Các đột biến lặn tích tụ trong bộ gen.

Hầu hết các đột biến là bất lợi và do đó các cơ chế đặc hiệu thường được dùng để giảm thiểu các biến đổi trong chuỗi ADN hoặc phục hồi trình tự gốc bằng cách sửa chữa các đột biến.

Trong phòng thí nghiệm, đột biến cung cấp công cụ có giá trị cho việc nghiên cứu sự sắp xếp các gen, chức năng của gen và cách kiểm soát hoạt động của chúng. Thực hiện điều này ở vi sinh vật dễ hơn là ở sinh vật bậc cao.

Kỹ thuật cảm ứng đột biến đã được sử dụng trong cây trồng và vật nuôi để tạo ra các chủng mới và trong công nghệ vi sinh học hiện đại để tạo ra các chủng vi sinh vật làm tăng năng suất và tạo ra sản phẩm mới.



Kỹ thuật tái tổ hợp ADN cho phép chèn một đột biến đặc hiệu vào một gen quy định trước (đột biến điểm định hướng) có thể tạo ra các đột biến tốt và hữu ích.

8.2. CÁC LOẠI ĐỘT BIẾN

8.2.1. Đột biến điểm

Đột biến điểm là kết quả của thay đổi một base trong chuỗi ADN gây bởi tác nhân gây đột biến hoặc sai sót trong sao chép ADN.

8.2.1.1. Thay thế cặp nucleotid

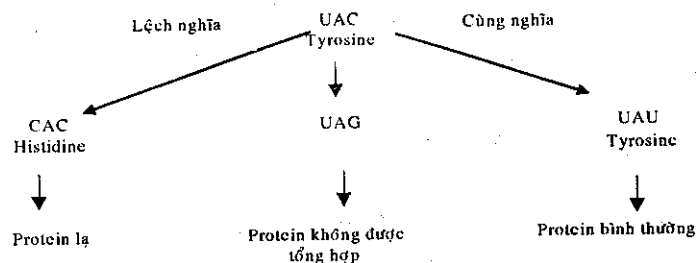
Phổ biến nhất là chuyển vị, pyrimidin được thay thế bởi pyrimidin khác hay purin bởi purin khác ($A \leftrightarrow G, C \leftrightarrow T$). Sự chuyển vị, bắt cặp sai, có thể gây bởi acid nitơ hoặc các đồng đẳng base như 5 - bromo - 2 - deoxyuridin (BrdU). Cách đột biến điểm ít phổ biến hơn là đảo chuyển, pyrimidin được thay thế bởi purin hay purin bởi pyrimidin ($C/T \leftrightarrow A/G$). Đột biến điểm có khuynh hướng đặc trưng là đột biến lùi (hồi biến), nghĩa là chuyển lại dạng nguyên thủy. Điều này xảy ra vì lần đột biến tiếp theo tại cùng một điểm có một phần ba cơ hội quay lại chuỗi ban đầu. Có ba loại đột biến điểm tùy theo codon mang đột biến mã hoá:

- Đột biến lặn hay còn gọi là đột biến cùng nghĩa: mã hoá cho cùng acid amin, không có ảnh hưởng đến protein cuối. Ví dụ, codon GCA hoặc GCG trong mRNA đều nghĩa là arginin (loại này thường đúng với đột biến chuyển vị ở vị trí thứ ba - base "linh hoạt" của codon).

- Đột biến lệch nghĩa: mã hoá cho acid amin khác, có thể làm protein không có chức năng. Ví dụ, trong chuỗi protein β - globin gây bệnh thiếu máu hồng cầu hình liềm, GAG mã hoá cho glutamat; nếu thay bằng GUG sẽ mã hoá cho valin. Đột biến lệch nghĩa có thể gây hậu quả rất nghiêm trọng như thiếu máu tế bào lưỡi liềm, hoặc nhẹ như hemoglobin C (acid amin vị trí 6 của β - globin bị thay thế bởi acid amin khác) hoặc không có kiểu hình.

- Đột biến vô nghĩa: mã hoá cho codon kết thúc, có thể làm cụt protein. Tác động của các đột biến vô nghĩa tùy thuộc protein bị cụt bao nhiêu và mức độ cần cho hoạt động của protein.

Đột biến thay thế base có thể xảy ra trong promoter hoặc các vùng điều hoà 5' của gen hoặc trong intron và có thể ảnh hưởng đến sự phiên mã, dịch mã hoặc mức xoắn bện của ADN. Nhiều β - thalassemia là hậu quả của các đột biến không cấu trúc này ảnh hưởng mức biểu hiện các gen globin. Mức độ ảnh hưởng tùy thuộc nucleotid bị thay thế và vị trí của nó trong ADN.



Hình 8.1. Các loại đột biến điểm

8.2.1.2. Đột biến lệch khung

Các đột biến này do chèn hoặc mất một hoặc nhiều nucleotid trong vùng mã hoá của gen. Điều này làm thay đổi khung đọc, vì các codon là các nhóm ba nucleotid nên có thể có ba khung đọc nhưng chỉ một cái được dùng. Đột biến loại này thay đổi tất cả acid amin phía sau đột biến và rất dễ tạo sản phẩm không chức năng vì khác hẳn với protein bình thường. Ngoài ra, các khung đọc sai thường chứa các codon stop, sẽ làm quá trình tổng hợp protein dừng sớm.

Việc thêm hoặc bớt một nucleotid vào khung đọc sẽ làm thay đổi acid amin kể từ đó trở đi. Nếu đột biến thêm hoặc bớt cùng xảy ra trên một phân tử ADN thì do trao đổi chéo giữa các base tương đồng mà đột biến âm có thể sửa chữa lại khung đọc sai do đột biến dương gây ra. Nhờ đó chỉ có acid amin nào nằm giữa hai điểm đột biến bị sai mà thôi. Như vậy, đột biến âm đã tác động như một đột biến kìm hãm ngoài gen đối với đột biến dương hoặc ngược lại.

8.2.2. Đột biến đa điểm

Các đột biến đa điểm là những thay đổi tác động đến hơn một base, thay đổi từ hai đến hàng ngàn base, nhưng thường nhất là chỉ vài base. Hậu quả thường là xoá bỏ một phần của trình tự mã hoá. Đột biến đa điểm không có tính hồi biến, nghĩa là đột biến được ổn định.

Sự xoá mất đoạn lớn sẽ làm thay đổi hoàn toàn protein tạo ra, thậm chí xoá bỏ đoạn ngắn cũng gây hậu quả nghiêm trọng nếu nó làm thay đổi khung đọc, nghĩa là nếu số lượng base bị xoá không chia hết cho ba.

Sự chèn đa điểm có thể xảy ra do gen nhảy hoặc các lỗi khi sao chép của các yếu tố lặp lại (ví dụ các lặp lại AT). Hầu hết khi gen bị chèn sẽ bị lệch khung hoặc thay đổi kết nối trong mRNA, cả hai trường hợp này đều làm thay đổi sản phẩm của gen. Gen nhảy dài tới vài ngàn base (kb) và vì vậy, nó làm ngưng quá trình phiên mã và dịch mã của bất kỳ gen nào chèn xen vào. Gen nhảy có hai loại, retrotransposon và transposon. Các transposon có thể di chuyển trực tiếp từ vị trí này sang vị trí khác trong bộ gen, retrotransposon trước hết được phiên mã thành ARN và sau đó trở lại

thành ADN nhờ enzym phiên mã ngược và chèn lại vào bộ gen. Transposon rất có ích cho các nhà nghiên cứu làm công cụ để biến đổi ADN trong cơ thể sống.

8.3. NGUYÊN NHÂN ĐỘT BIẾN

Có hai loại đột biến là đột biến tự nhiên (xảy ra một cách tự phát) và đột biến cảm ứng, gây ra bởi các tác nhân đột biến.

8.3.1. Đột biến tự nhiên

Các đột biến xảy ra trong tự nhiên một cách ngẫu nhiên với tần số nhất định và không xác định được nguồn gốc. Bất kỳ một tiến trình nào dẫn đến sự gắn sai base trong quá trình sao chép ADN sẽ tạo ra một đột biến điểm. Tần suất sai sót tự nhiên vào khoảng 10^{-10} base. Tỷ lệ này thấp đạt được nhờ khả năng sửa lỗi của các ADN polymerase chủ yếu. Các enzym này kéo dài ADN theo hướng 5' → 3' nhưng có hoạt tính exonuclease 3' → 5'. Nếu lỗi nằm giữa base được chèn cuối cùng với sợi khuôn được phát hiện, enzym loại bỏ base sai và thay thế base đúng vào. Chuỗi mã của một protein trung bình 300 acid amin là khoảng 1000 base chiều dài. Tỷ lệ đột biến tự phát là một trong 10^7 bản mã sao của gen sẽ có một đột biến điểm, nghĩa là 10^{-7} tế bào sẽ có đột biến về gen đó.

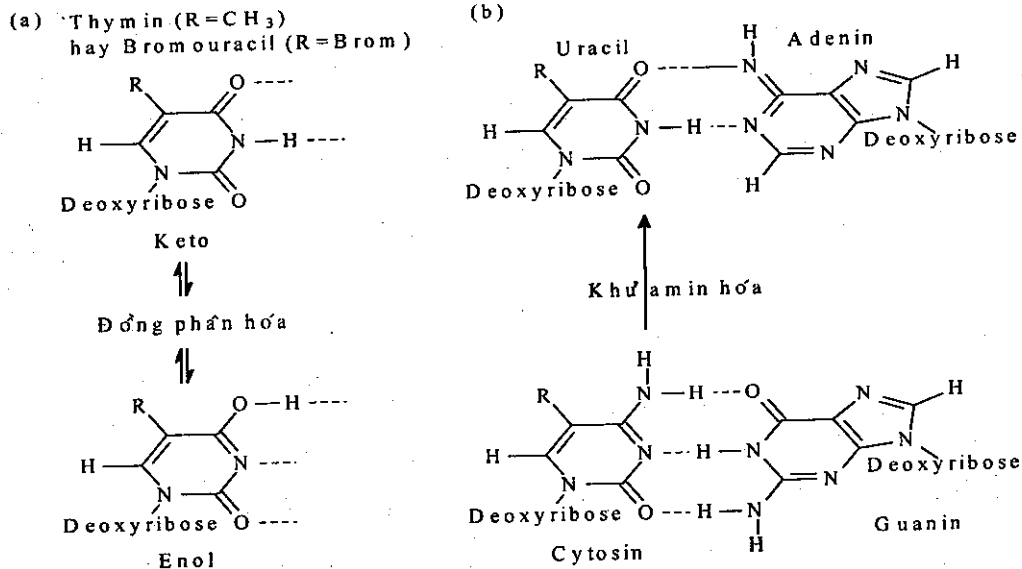
Đột biến tự nhiên ở mức phân tử gồm có hồ biến, khử amin, chuyển vị, đảo chuyển, đột biến lệch khung, oxy hoá phá huỷ.

– Hồ biến: 4 loại base trong ADN có thể tồn tại ở dạng hồ biến, tức là có khả năng tồn tại hoán chuyển giữa hai dạng (Keto ↔ Enol và Amino ↔ Imino). Base pyrimidin bình thường có dạng keto (C=O), nhưng đôi khi chúng cũng có dạng enol (C-OH). Ở dạng enol, thymin có thể bắt cặp với guanin và enol cytosin có khả năng bắt cặp với adenin (Hình 8.2-a) tương tự base purin bình thường tồn tại dưới dạng amino (NH₂) nhưng cũng có dạng hồ biến imino (=NH) hiếm gặp, có thể dẫn đến sự bắt cặp nhầm. Nếu trong khi ADN sao chép, G ở dạng enol, polymerase sẽ thêm T vào thay vì bình thường là C vì quy luật bắt cặp bị thay đổi (không phải do lỗi của polymerase). Kết quả là sự chuyển vị G-C thành A-T; sự hồ biến chỉ gây các đột biến chuyển vị.

– Khử amin: Các phản ứng khử amin thông thường là phản ứng thủy phân cytosin thành uracil, quá trình này phóng thích ammonia và khử amin 5-methylcytosin thành thymin và ammonia (Hình 8.2-b). Trong ADN, sự khử amin cytosin tự phát được chỉnh lại bằng cách loại bỏ uracil (vì uracil không có trong ADN) và thay thế bằng cytosin, còn sự khử amin 5-methylcytosin không được chỉnh vì cơ chế sửa chữa không nhận diện Thymin là lỗi. Trong ADN bộ gen cytosin tại các trình tự CG (CpG) thường bị methyl hóa, điều này nghĩa là ở đâu CpG xảy ra trong gen, nó là “điểm nóng” cho đột biến. Mới đây người ta tìm thấy trong gen liên quan đến loạn sản sụn có những điểm nóng như vậy.

- Đột biến lệch khung (chèn hoặc mất trên một sợi), thường là lỗi của polymerase khi sao chép các đoạn lặp lại của một nucleotid. Mỗi polymerase có độ chính xác khác nhau. Yếu tố chính ảnh hưởng đến độ chính xác của polymerase là hoạt tính exonuclease đọc sửa 3'-5', loại bỏ các base bất cặp sai chèn vào bởi polymerase.

- Oxy hóa phá hủy do các gốc oxy. Các gốc oxy tăng trong tế bào do chuyển hóa oxy hóa (và cũng được tạo bởi các tác nhân vật lý như tia phóng xạ). Sản phẩm oxy hóa quan trọng là 8-hydroxy guanin bất cặp sai với A, tạo đảo chuyển G-C thành T-A.



Hình 8.2. Hiệu quả đột biến của sự thay đổi tính chất cặp base bổ sung
(a) đồng phân của thymin và 5 - bromouracil; (b) sự khử amin hoá cytosin thành uracil.

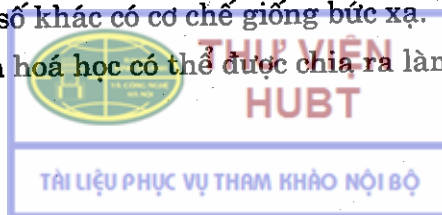
8.3.2. Đột biến cảm ứng

Các tác nhân làm tăng tần số đột biến cao hơn mức tự nhiên được gọi là các tác nhân gây đột biến. Các tác nhân này có thể làm thay đổi cấu trúc hoặc trình tự ADN. Các tác nhân vật lý gây đột biến như phóng xạ, tia X, tia tử ngoại... Nhiều hoá chất là các tác nhân gây đột biến như các đồng đẳng của các base nitơ, acid nitơ (HNO₂), các chất alkyl hoá mạnh.

8.3.2.1. Tác nhân hoá học gây đột biến

Người ta bắt đầu biết đến tác động gây đột biến của chất hoá học khi sử dụng khí ngạt nitơ trong Chiến tranh thế giới lần thứ nhất và hai, chất này có thể gây đột biến trong tế bào. Sau đó, nhiều chất hoá học gây đột biến khác đã được xác định và được kiểm soát nghiêm ngặt. Có thể phân biệt các tác nhân đột biến hoá học dựa vào kiểu tác dụng của chúng, một số chất gây đột biến theo cơ chế tương tự đột biến tự nhiên, một số khác có cơ chế giống bức xạ.

Tác nhân gây đột biến hoá học có thể được chia ra làm 4 nhóm:



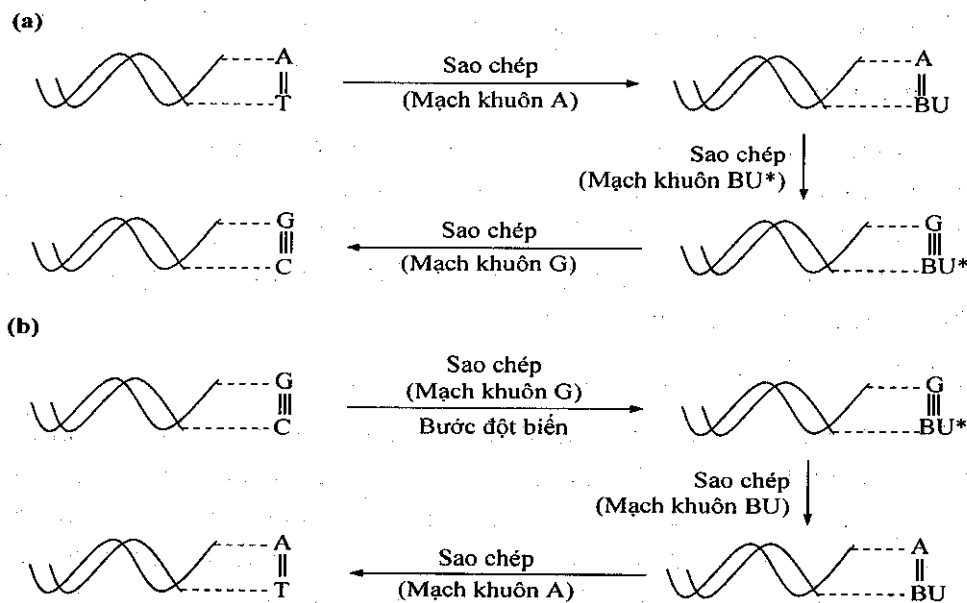
- Nhóm base đồng đẳng.
- Các chất hoá học thay đổi cấu trúc và đặc tính bắt cặp của các base, gồm: các chất khử amin và các chất alkyl hoá.
- Nhóm chèn vào ADN (tác nhân làm lệch khung).
- Các tác nhân thay đổi cấu trúc ADN.

Base đồng đẳng:

Các chất hoá học này có cấu trúc hoá học giống các purin và pyrimidin và có thể được gắn vào ADN tại vị trí của các base bình thường khi sao chép ADN.

- **Bromouracil (BU)** là hợp chất nhân tạo dùng rộng rãi trong nghiên cứu. Giống với thymin (có nguyên tử Br thay cho nhóm methyl) và sẽ gắn vào ADN và bắt cặp với Adenin giống như Thymin. Có nhiều khả năng hỗ biến thành dạng enol (BU*) và có khả năng bắt cặp với Guanin. Như vậy, BU có thể xâm nhập và bắt cặp bổ sung với Adenin thay cho Thymin tạo A-BU và ở vòng sao chép tiếp theo BU lại hỗ biến thành BU* và bắt cặp với Guanin tạo BU*-G, do đó A-T được thay thế thành G-C (Hình 8.3-a). Trong khi đó BU* có thể xâm nhập sai, khi gắn thay chỗ cho Cytosine tạo G-BU*, sau đó BU* lại hỗ biến thành BU và tiếp theo bắt cặp với Adenin tạo A-BU và ở vòng sao chép tiếp theo Adenin bắt cặp với Thymin thành A-T (Hình 8.3-b).

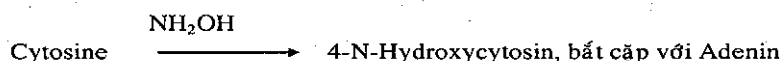
- **Aminopurin** là đồng đẳng của Adenin, có thể bắt cặp với Thymin hoặc (rất ít) với Cytosin, gây chuyển vị A - T thành G - C hoặc G - C thành A - T. Các đồng đẳng base gây chuyển vị như gây hỗ biến tự phát.



Hình 8.3. Tác động gây đột biến của bromouracil (BU)

Các chất hoá học thay đổi cấu trúc và đặc tính bắt cặp của các base:

– **Các chất khử amin.** Ví dụ điển hình: **acid nitơ** tạo bởi sự phân giải các nitrit (chất bảo quản) trong thực phẩm biến cytosin thành uracil, methylcytosin thành thymin, và adenin thành hypoxanthin khử amin, hypoxanthin trong ADN bắt cặp với C gây chuyển vị. Methylcytosin hiện diện trong các gen của sinh vật bậc cao có thể ảnh hưởng tới hoạt tính phiên mã. Các base ngoại như xanthin, hypoxanthin và uracil có thể khởi động cơ chế sửa chữa. Thymin, dẫn xuất từ methylcytosin nội sinh và nhiều sự chuyển vị có thể xảy ra theo cách này. Các gốc 5-methylcytosin, sau đó trở thành "điểm nóng" của sự đột biến. Hydroxylamin (NH_2OH) chỉ khử amin cytosin, do đó có tác động rất giới hạn:



– **Các chất alkyl hoá:** Ví dụ điển hình: Nitrosoguanidin (NTG), Methyl Methane Sulphonat (MMS), Ethyl Methane Sulphonat (EMS) là các tác nhân hoá học gây đột biến phản ứng với các base và thêm các nhóm Methyl hoặc Ethyl nên còn gọi là các tác nhân alkyl hoá (Bảng 8.2). Tùy yếu tố bị tác động, chúng có thể gây đột biến ít nhất bằng 3 cách:

+ Thêm nhóm Methyl ($-\text{CH}_3$) hay ethyl ($-\text{C}_2\text{H}_5$) vào Guanin tạo base đồng đẳng của Adenin dẫn đến sự bắt cặp bổ sung sai (Hình 8.4).

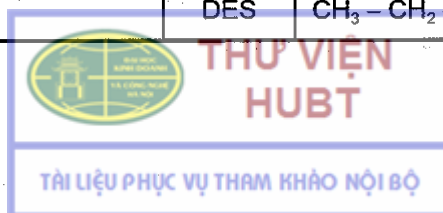
+ Mất purin do Guanin đã bị alkyl hoá tạo lỗ hổng trên ADN, khi sao chép có thể làm đứt mạch (Hình 8.4).

Liên kết chéo giữa các mạch của một hoặc các phân tử ADN khác nhau làm mất nucleotid.

Guanin là base có thể dễ bị tấn công nhất, vòng pyrimidin, đặc biệt là các nhân pyrin dễ dàng bị alkyl hoá. Các vị trí này là các vị trí liên quan trực tiếp đến việc ghép cặp như 6 – keto oxygen của guanin hay 4 – keto oxygen của Thymin và các vị trí cách xa vị trí ghép cặp như N – 7 của Guanin.

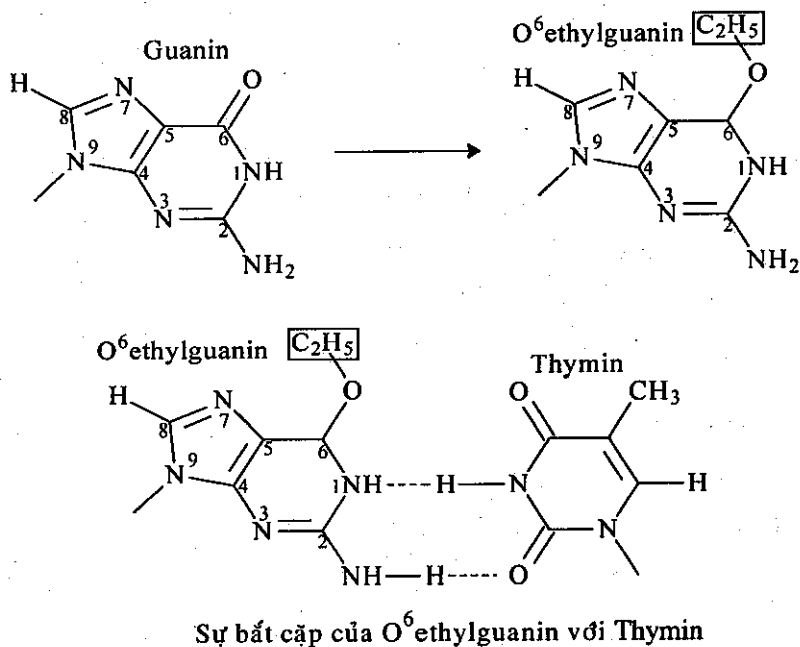
Bảng 8.2. Một số tác nhân alkyl hoá được sử dụng để gây đột biến cảm ứng

Tên hoá học	Viết tắt	Công thức hoá học
Methyl methan sulphonat	MMS	$\text{CH}_3 - \text{SO}_2 - \text{O} - \text{CH}_3$
Ethyl methan sulphonat	EMS	$\text{CH}_3 - \text{SO}_2 - \text{O} - \text{CH}_2 - \text{CH}_3$
Dimethyl sulphonat	DMS	$\text{CH}_3 - \text{O} - \text{SO}_2 - \text{O} - \text{CH}_3$
Diethyl sulphat	DES	$\text{CH}_3 - \text{CH}_2 - \text{O} - \text{SO}_2 - \text{O} - \text{CH}_2 - \text{CH}_3$



N - methyl - N' - nitro - N - nitrosoguanidin	MNNG	$CH_3 - N.NO - C:NH - NH - NO_2$
Nitrogen mustard (khí ngạt nitơ)		$CH_3 - N - (CH_2 - CH_2 - Cl)_2$
Sulphur mustard (khí ngạt sulphur)		$S - (CH_2 - CH_2 - Cl)_2$
Ethylen oxid	EO	$O - (CH_2)_2$
Ethylen imin	EI	
Nitrosoguanidin	NG	

N - ethyl - N - nitrosourea (ENU, công thức hoá học $C_3H_6N_3O_2$) là tác nhân alkyl hoá gây đột biến mạnh trên chuột. Tỷ lệ đột biến của nó là 1 đột biến mới tại một gen riêng biệt trong 700 giao tử. ENU tác động bằng cách chuyển nhóm ethyl của nó vào base nitơ (chủ yếu là Thymin) trong acid nucleic. Đích chính của nó là các tế bào mầm nguyên bào tinh, từ đó mà tinh dịch trưởng thành.



Hình 8.4. Hiệu quả dự đoán khi xử lý Guanin với tác nhân alkyl hoá (EMS hay MMS)

Các tác nhân chèn vào ADN (làm lệch khung):

Nhóm này gồm các chất Proflavin, cam Acridin, Ethidium bromide dùng trong phòng thí nghiệm làm chất nhuộm. Tất cả đều là các phân tử đa vòng phẳng tương tác với các base của ADN và chèn vào giữa chúng. Sự chèn này làm "dãn" sợi đôi ADN và ADN polymerase bị "đánh lừa" chèn base vào đối diện nhiều hơn bình thường, làm lệch khung ADN tạo thành.

Các tác nhân thay đổi cấu trúc ADN:

Các tác nhân này có thể là:

- Các phân tử lớn (kênh cang) gắn vào base trong ADN và làm chúng trở thành không mã hoá. Ví dụ, N - acetoxy - 2 - acetylaminofluorene (NAAAF).
- Các tác nhân gây liên kết chéo trong và giữa các sợi. Ví dụ, các psoralen có trong một số rau và dùng trong điều trị một số bệnh về da.
- Các chất hoá học gây đứt sợi ADN. Ví dụ các peroxid.
- Các hydrocarbon đa vòng. Ví dụ các benzopyrene có trong xăng.

8.3.2.2. Tác nhân bức xạ gây đột biến

Bức xạ là tác nhân đột biến được biết đầu tiên (năm 1920). Bức xạ được phát minh vào những năm 1890: Roentgen phát hiện tia X vào 1895, Becquerel phát hiện hoạt tính phóng xạ vào 1896 và Marie với Pierre Curie phát hiện các nguyên tố phóng xạ vào 1898.

Tác dụng gây đột biến của bức xạ có 2 đặc điểm:

- Không có ngưỡng tác dụng tức là không có liều lượng vô hại.
- Số lượng đột biến tỉ lệ thuận với liều lượng phóng xạ, không phụ thuộc vào cường độ và thời gian chiếu xạ.

Các bức xạ điện từ:

Bước sóng của các bức xạ này thay đổi tùy theo loại bức xạ và tỉ lệ nghịch với năng lượng của chúng.

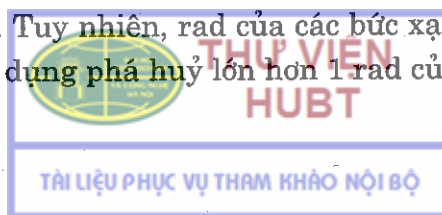
Các sóng dài nhất (radio AM) có năng lượng ít nhất, trong khi các sóng ngắn hơn là FM, TV, viba, hồng ngoại, khả kiến, tia tử ngoại (UV), tia X và tia gamma lần lượt có năng lượng gia tăng. Tia tử ngoại (còn gọi tia cực tím) có năng lượng phóng xạ cao nên có ý nghĩa sinh học.

Bức xạ ion hoá:

Tia X, tia gama, hạt phóng xạ α và β có thể làm ion hoá các phân tử trên đường đi của chúng, do đó có tác động mạnh đến các phân tử sinh học.

Bức xạ UV không ion hoá nhưng có thể phản ứng với ADN và các phân tử sinh học khác, và cũng là tác nhân đột biến quan trọng.

Đơn vị dùng cho tất cả các loại bức xạ ion hoá hiện nay là rem (roentgen equivalent man = 0,01J/kg người bị chiếu xạ), do đó 1 rem của bất cứ bức xạ ion hoá nào cũng có tác dụng sinh học như nhau. Đơn vị được dùng trước đây là rad (radiation absorbed dose). Tuy nhiên, rad của các bức xạ khác nhau sẽ khác nhau: 1 rad của các hạt α có tác dụng phá huỷ lớn hơn 1 rad của các tia gamma, nghĩa là



các hạt α có RBE (relative biological effectiveness) lớn hơn 1 rad của các tia gamma. Mỗi tương quan giữa các đơn vị này là: $\text{rad} \times \text{RBE} = \text{rem}$.

Ngoài ra, cần xét đến loại năng lượng và liều bức xạ toàn phần, tỉ lệ liều: cùng số rem, nếu chiếu tập trung (tỉ lệ liều cao) gây bỏng và phá huỷ da, trong khi chiếu nhẹ trong thời gian dài (tỉ lệ liều thấp) sẽ chỉ làm tăng nguy cơ đột biến và ung thư.

Nguồn bức xạ:

Các nguồn bức xạ tự nhiên sản xuất bức xạ nền. Nguồn bức xạ tự nhiên gồm có các tia vũ trụ từ Mặt Trời và không gian, từ các nguyên tố hoạt động phóng xạ trong đất, trên mặt đất và trong không khí (radon). Mức độ bức xạ nền thay đổi tùy vùng địa lý.

Ngoài ra, con người tạo ra nguồn bức xạ nhân tạo, góp phần tăng mức bức xạ. Trong đó có các thử nghiệm y tế (ví dụ chẩn đoán bằng tia X), thử hạt nhân và các nhà máy năng lượng hạt nhân, các sản phẩm khác (tivi, máy dò khói, kiểm tra bằng tia X ở sân bay).

Tác dụng sinh học của bức xạ:

Bức xạ ion hoá sản sinh các gốc tự do của nước (gốc hydroxyl), gây nhiều hư hỏng đối với tế bào và sinh vật. Các gốc tự do có các điện tử không bắt cặp và rất hoạt động về mặt hoá học, sẽ tương tác với ADN, protein, lipid trong màng tế bào,... Như thế tia X có thể gây hư hỏng ADN và protein, có thể làm hỏng cơ quan, ngăn phân chia tế bào hoặc làm chết tế bào. Các loại tế bào phân chia nhanh (tế bào tạo máu, tế bào tuỷ xương, niêm mạc tiêu hoá) bị ảnh hưởng bởi bức xạ ion hoá nhiều nhất và độ nghiêm trọng của tác động tùy thuộc liều đã nhận.

Ảnh hưởng di truyền của bức xạ:

Bức xạ ion hoá gây nhiều ảnh hưởng trên ADN thông qua các gốc tự do và tác động trực tiếp:

- Đứt một hoặc cả hai sợi (có thể dẫn đến sắp xếp lại, xoá mất, mất nhiễm sắc thể, tế bào chết nếu không sửa chữa).
- Phá huỷ các base (đột biến).
- Liên kết chéo trong ADN hoặc ADN với protein.

Tia cực tím (UV):

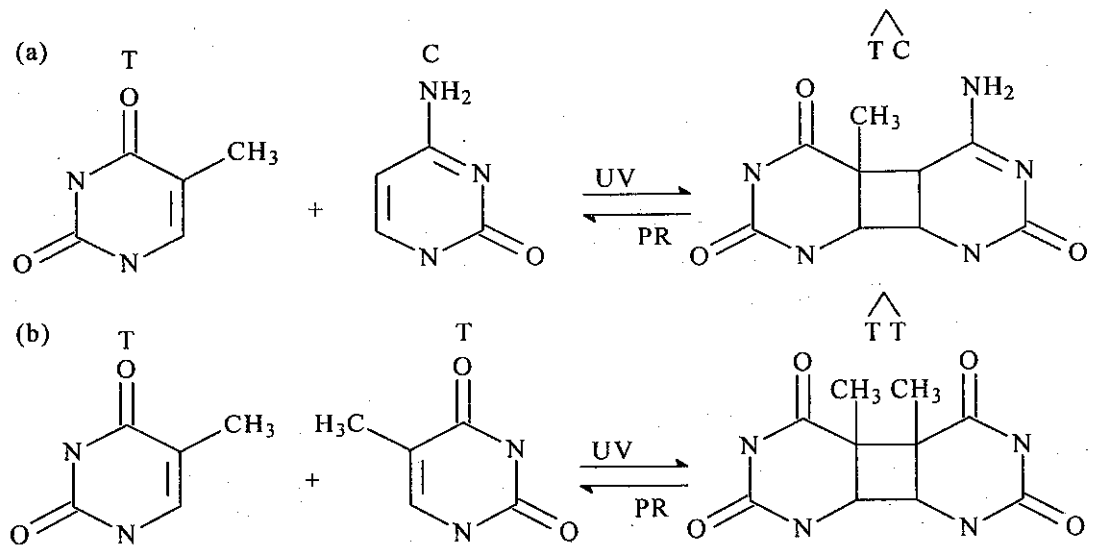
Bức xạ UV ít năng lượng và do đó không ion hoá, nhưng năng lượng của nó được ưu tiên hấp phụ bởi các base của ADN và bởi các acid amin thơm của các protein, vì thế nó cũng có ảnh hưởng sinh học và ảnh hưởng di truyền quan trọng.

UV được phân loại theo bước sóng: UV-C (180 - 290 nm) có tính sát trùng mạnh nhất và gây chết, không có trong ánh sáng mặt trời vì nó được hấp phụ bởi tầng ozon; UV-B (290-320 nm) là phân đoạn gây chết/đột biến chính của ánh sáng mặt trời; UV-A (320 nm - khả kiến) là “cận UV”, cũng có tác dụng xóa mờ (vì nó tạo các gốc oxy) nhưng nó sản xuất rất ít dimer pyrimidin. Các tổn thương gây chết chính là các dimer pyrimidin trong ADN (tạo bởi UV-B và UV-C), các dimer này là liên kết đồng hóa trị giữa các pyrimidin kề nhau trong một sợi. Các dimer này làm ngăn chặn sự phiên mã và sao chép ADN và gây chết nếu không sửa chữa. Chúng có thể kích thích đột biến cũng như sắp xếp lại nhiễm sắc thể.

Trong tế bào, các chất hữu cơ có mạch vòng chủ yếu như purin và pyrimidin hấp thụ trực tiếp UV. Tia tử ngoại có khả năng xuyên thấu thấp nên chỉ tác động lên các sinh vật đơn bào và giao tử. ADN hấp thụ tia tử ngoại mạnh nhất ở bước sóng 257 nm (≈ 260 nm), đây chính là bước sóng làm tăng tần số đột biến ở hạt phấn cây bắp. Dưới tác động của tia tử ngoại, cytosin gắn thêm phân tử nước vào liên kết C=C của mạch vòng (Hình 8.5) và thymin bị đứt liên kết C=C mạch vòng, nối 2 phân tử thành thymin dimer.

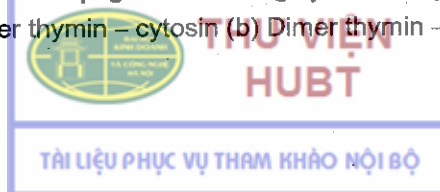
Stone và các cộng sự đã nhận thấy tần số đột biến tăng lên ở *Staphylococcus aureus* khi môi trường nuôi chúng được chiếu tia UV trong thời gian ngắn trước khi cấy vào. Đây là tác động gián tiếp của tia tử ngoại.

Hiện tượng quang phục hồi là một đặc trưng trong tác động của tia UV. Sau khi chiếu tia tử ngoại lên tế bào, nếu để ngoài ánh sáng, thì các sai hỏng phân lớn được phục hồi. ánh sáng có tác động hoạt hoá enzym sửa sai, cắt đứt các thymin dimer.



Hình 8.5. Tác động của tia tử ngoại lên các pyrimidin

(a) Dimer thymin – cytosin (b) Dimer thymin – thymin



8.4. CÁC CƠ CHẾ CHỐNG LẠI ĐỘT BIẾN

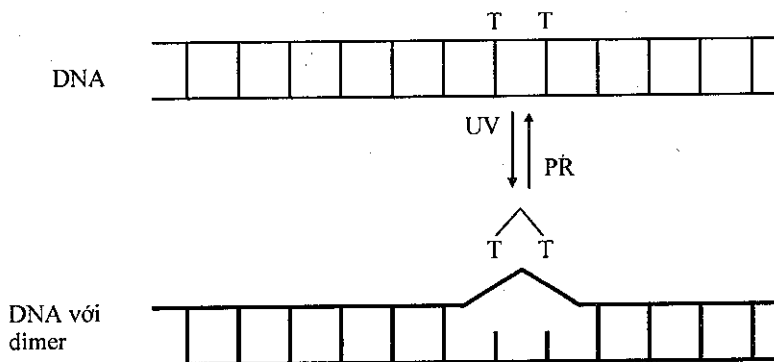
Vì sự sai hỏng ADN xảy ra tự phát và là hậu quả của các tác nhân môi trường ở mọi nơi nên hầu hết các sinh vật có một vài khả năng sửa chữa ADN của chúng và ADN là đại phân tử duy nhất được sửa chữa bởi tế bào. Có thể chia các cơ chế "sửa chữa" thành ba loại:

- Đảo nghịch sai hỏng: cơ chế này đơn giản nhất, enzym hoạt động phục hồi cấu trúc bình thường không có gãy khung.
- Loại bỏ sai hỏng: cắt bỏ và thay thế base hoặc phân nucleotid sai hỏng hoặc không thích hợp.
- Dung nạp sai hỏng: không sửa chữa thật sự, chép sai hỏng và tiếp tục sống.

8.4.1. Đảo nghịch sai hỏng

8.4.1.1. Quang phục hồi

Đây là hệ thống sửa chữa đơn giản nhất và có lẽ xưa nhất. Có sự tham gia của một enzym có khả năng cắt các dimer pyrimidin (làm đứt các liên kết đồng hoá trị) khi có ánh sáng (Hình 8.6).



Hình 8.6. Quang phục hồi

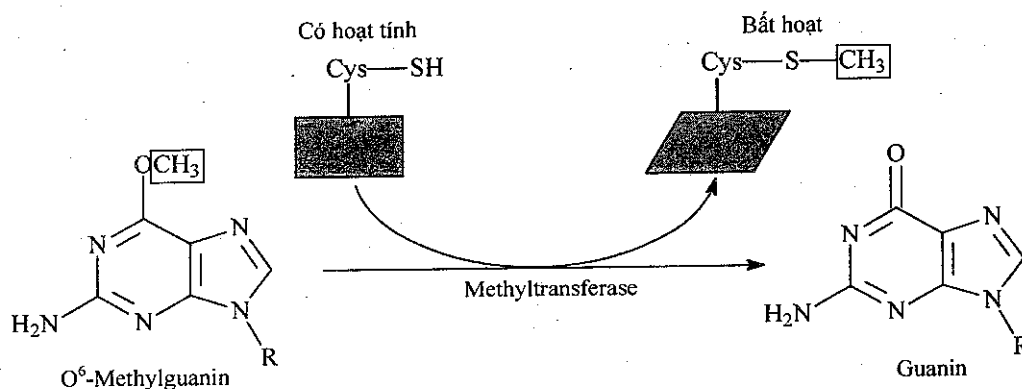
Enzym photolyase xúc tác phản ứng này, nó có nhiều trong vi khuẩn, tế bào nhân thật bậc thấp, côn trùng và thực vật. Không có trong tế bào động vật. Gen (*phr*) có trong động vật có vú nhưng mã hoá cho protein có chức năng hỗ trợ loại sửa chữa khác.

8.4.1.2. Sửa sai bằng cách làm mất nhóm alkyl (dealkylation)

Sửa chữa trực tiếp các tổn thương bằng O^6 -methylguanin-ADN methyltransferase. Enzym này chuyển nhóm alkyl từ nguyên tử oxi trên ADN sang gốc Cystein trên enzym và phản ứng không thuận nghịch và enzym bị mất hoạt tính (Hình 8.7). Các lỗi được sửa chữa gồm O^6 -methylguanin, O^4 -methylthymin, và các methylphosphotriester gây ra bởi các tác nhân alkyl hoá như

N – methyl – N' – nitro – N – nitrosoguanidine (MNNG). Enzym này cũng sửa chữa các guanin bị alkyl hoá khác như O⁶ – ethylguanin và O⁶ – butyl – guanin, mặc dù hiệu suất thấp hơn nhiều. Các MTase được tìm thấy ở tất cả sinh vật được nghiên cứu.

Ở *E. coli* gen mã hoá cho enzym này (39 kDa) là *ada*. Vị trí Cys69 trên enzym nhận gốc methyl từ phosphodiester và thay đổi cấu hình cho phép nó hoạt hoá promoter của các gen *ada*, *alkA* (methyladenin DNA glycosylase), *alkB*, và *aidB*, nhờ đó tăng khả năng đề kháng với các tác nhân alkyl hoá. Trong khi vị trí Cys321 nhận gốc methyl từ O⁶ – MeGuanin và O⁴ – MeThymin, để khôi phục base bình thường.



Hình 8.7. Cơ chế sửa sai bằng methyltransferase

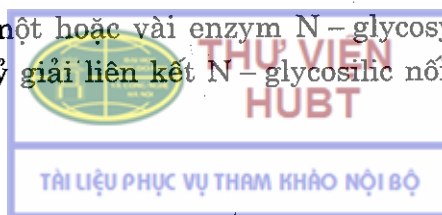
8.4.1.3. Nối các chỗ đứt của sợi đơn

Tia X và một số hoá chất như peroxid có thể gây ra những chỗ đứt trên khung ADN. ADN ligase nhanh chóng sửa các chỗ đứt đơn giản trên một sợi. Các vi sinh vật đột biến mất ligase có mức tái tổ hợp cao hơn vì đầu ADN sinh tái tổ hợp được hoạt hoá cao. Ở người chỉ mới biết mã 46BR có đột biến trên cả hai gen ADN ligase I, làm tăng trưởng chậm, suy giảm miễn dịch và nhạy cảm với Mặt Trời, chết trẻ do u lympho. Các tế bào nguyên bào sợi ở người có 46B dễ bị diệt bởi các tác nhân làm hỏng ADN, trong đó có bức xạ ion. Ngoài ra, hội chứng Bloom cũng liên quan với khiếm khuyết ADN ligase (mặc dù protein của hội chứng Bloom là ADN helicase); tế bào nuôi cấy của bệnh nhân này có mức biến dạng nhiễm sắc thể và đột biến tự phát cao.

8.4.2. Loại bỏ hư hỏng

8.4.2.1. Sửa sai bằng cách cắt base (glycosylase)

Loại bỏ chỗ sai nhờ một hoặc vài enzym N – glycosylase. N – glycolase nhận biết base biến đổi và thuỷ giải liên kết N – glycosylic nối base với đường pentose.



Điều này tạo ra một vị trí AP (không có purin hoặc pirimidin), nghĩa là có mạch đường – phosphat nhưng bị thiếu base nitơ. Vị trí AP sau đó được cắt bỏ bởi AP endonuclease và ADN polymerase chép đầy lại dựa vào khuôn là mạch bổ sung đối diện và được ligase nối liền lại.

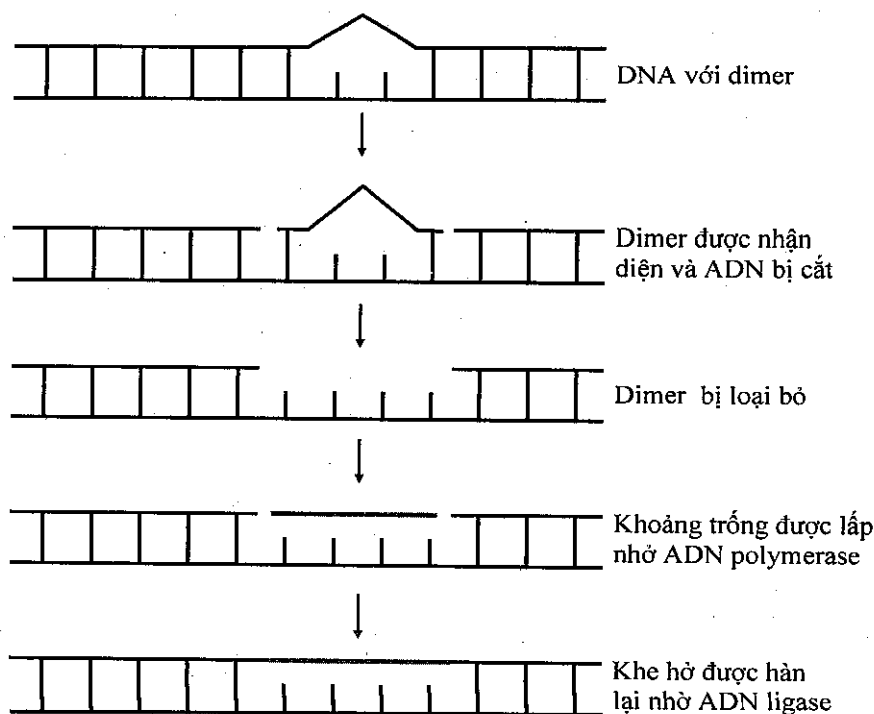
8.4.2.2. Sửa chữa lệch đôi

Quá trình này xảy ra sau sao chép ADN như là "đánh vần kiểm tra" cuối cùng sự chính xác của sao chép. ở *E. coli*, cơ chế này làm tăng độ chính xác của sao chép lên thêm 100 – 1000 lần. Sửa chữa này được tiến hành bởi một nhóm các protein có thể quét ADN và tìm các base bắt cặp không chính xác (hoặc các base không bắt cặp) làm sai lệch hai chiều trong xoắn kép. Nucleotid không chính xác phần dưới sẽ được loại bỏ và ADN polymerase sẽ hoạt động lần thứ hai để có trình tự đúng.

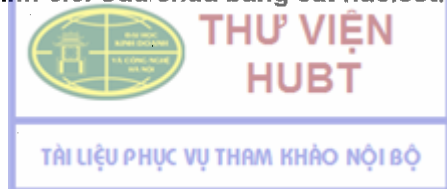
Các protein sửa chữa lệch đôi ở người mới được xác định và rất giống ở tế bào nhân nguyên thủy như *E. coli* và ở tế bào nhân thật đơn giản như nấm men.

8.4.2.3. Sửa chữa bằng cắt nucleotid – NER

Hệ thống này làm việc trên "khối" ADN sai hỏng, làm ngưng sao chép ADN và phiên mã. Quá trình gồm cắt sợi ADN chứa sai hỏng bởi endonuclease ở phía sai hỏng, tiếp theo loại bỏ đoạn ngắn chứa vùng sai hỏng bởi exonuclease, khoảng trống tạo thành được lấp vào bởi ADN polymerase (Hình 8.8).



Hình 8.8. Sửa chữa bằng cắt nucleotid



Nhiều sinh vật bị đột biến khiếm khuyết NER, dễ bị diệt và đột biến bởi UV và các hoá chất hoạt động giống UV. Người có bệnh di truyền xeroderma pigmentosum nhạy cảm với ánh sáng mặt trời, họ có nguy cơ cao bị ung thư da ở những vùng cơ thể phơi dưới Mặt Trời và có các khiếm khuyết trong các gen tương đồng với hệ thống NER ở tế bào nhân thật đơn giản. Các đột biến NER ở sinh vật bậc thấp nhạy với UV và tăng mức đột biến và tái tổ hợp cảm ứng bởi UV (vì chúng không thể dùng phương pháp NER chính xác để loại bỏ các dimer pyridin và phải dùng các hệ thống gây đột biến hoặc gây tái tổ hợp).

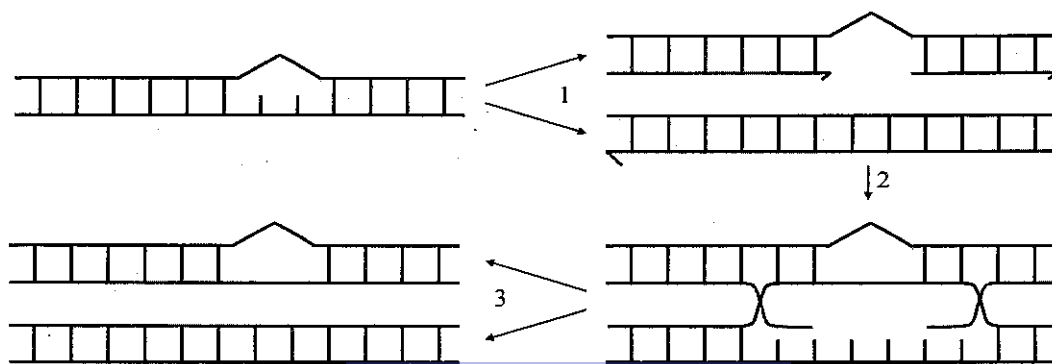
8.4.3. Dung nạp ADN sai hỏng

Không phải tất cả các ADN sai hỏng được loại bỏ ngay mà một số có thể tồn tại một thời gian. Nếu tại chạc ba sao chép ADN có ADN sai hỏng như dimer pyridin, quá trình sao chép sẽ ngừng lại.

Tuy nhiên ở tế bào nhân thật, sao chép khởi động ở nhiều vị trí và có thể hồi phục lại sợi dưới của dimer, chừa lại "khoảng trống" của sợi đơn ADN không sao chép. Khe này có thể nguy hiểm nếu không dimer nữa khi tế bào phân chia. Khoảng trống có thể được sửa bằng tái tổ hợp với tương đồng hoặc với thanh nhiễm sắc em, tạo hai phân tử con nguyên vẹn, một cái vẫn còn chứa dimer.

8.4.3.1. Sửa chữa bằng tái tổ hợp

Đây là cơ chế sửa chữa thúc đẩy tái tổ hợp để lấp khuyết sợi con không dimer và là cách chép đối phó tổn thương lưu trữ không mã hoá trong ADN (Hình 8.9). Loại tái tổ hợp sửa chữa này nói chung chính xác (mặc dù nó có thể gây tính đồng hợp của các allele lặn có hại) và đòi hỏi sự tương đồng của sợi nhiễm sắc em. Sản phẩm của các gen *BRCA1* và *BRCA2* nhạy cảm với ung thư vú ở người có thể tham gia sửa chữa bằng tái tổ hợp cùng với tương đồng của các gen *RAD51* và *RAD52* của nấm men.

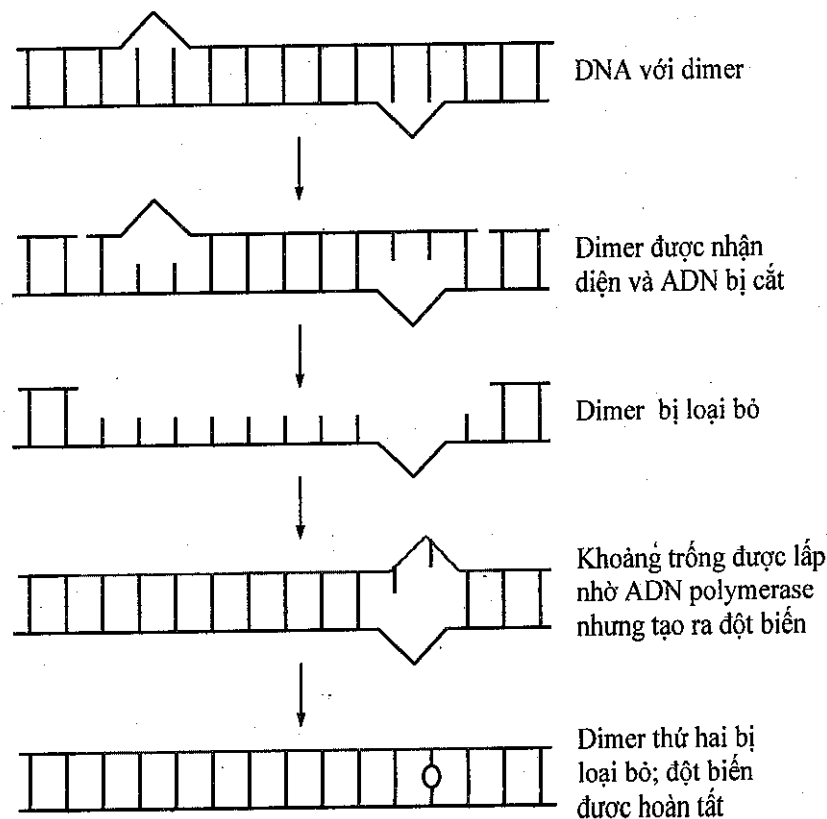


Hình 8.9. Các biến cố trong sửa chữa bằng tái tổ hợp

Loại sửa chữa bằng tái tổ hợp thứ hai là phản ứng liên kết đầu không tương đồng để sửa các đầu ADN bị đứt như gãy do bức xạ ion hoá và các chất hoá học có tác động gây đột biến tương tự. Hệ thống sửa chữa này cũng được sử dụng bởi các tế bào B và T của hệ thống miễn dịch để tái sắp xếp di truyền cần cho hoạt động của chúng. Protein Ku70, Ku80 và các protein kinase phụ thuộc ADN cần cho liên kết đầu không tương đồng. Các dòng tế bào của loài gặm nhấm có các đột biến trong các gen này rất dễ bị diệt bởi bức xạ ion hoá và khiếm khuyết tái sắp xếp hệ thống miễn dịch.

8.4.3.2. Sửa chữa bằng đột biến

Một cách khác để ADN polymerase bị ngăn tại dimer là thay đổi tính đặc hiệu của nó, nhờ đó có thể chèn bất cứ nucleotid đối diện dimer nào và tiếp tục sao chép (cách "đột biến hoặc là chết") (Hình 8.10). Cách này xảy ra ở vi khuẩn, người ta nghĩ nó có thể xảy ra ở tế bào nhân thật tuy chưa biết rõ cơ chế. Đây là lý do tại sao đôi khi sửa chữa có thể gây đột biến.



Hình 8.10. Sửa chữa bằng đột biến



8.4.4. Sửa sai nhờ hệ thống SOS (cấp cứu)

Hệ thống sửa chữa này được kích hoạt nếu các cơ chế trên bị quá tải bởi các đột biến lan rộng. Trong trường hợp ADN bị tổn thương ngừng sao chép, thì phản ứng SOS hồi phục sao chép và chuyển sai hỏng thành sửa sai trong khi nhân đôi. Ở *E.coli* đã quan sát sự phá huỷ ADN làm mở ra khoảng 20 gen của hệ thống SOS, được kiểm soát âm bởi chất kìm hãm *lexA*. Chất này gắn vào hộp SOS chống lấp các promoter của các gen SOS.

Khi có nhiều sai hỏng cấp cứu, *lexA* bị kích thích, thay đổi cấu hình, tự cắt và mất hoạt tính kìm hãm. Lúc đó, các gen SOS được mở ra. Nếu sửa sai không kịp, tế bào phải chấp nhận hoặc bị đột biến hoặc bị chết.

8.5. CÁC TÍNH TRẠNG ĐỘT BIẾN VÀ PROTEIN ĐỘT BIẾN

Một sinh vật bị đột biến đơn giản là một sinh vật biểu hiện những tính trạng bất thường di truyền được. Hầu hết những tính trạng này có thể thấy được như những đột biến sắc tố gây nên bệnh bạch tạng, thay đổi màu sắc bên ngoài ở động vật và màu sắc hoa bất thường.

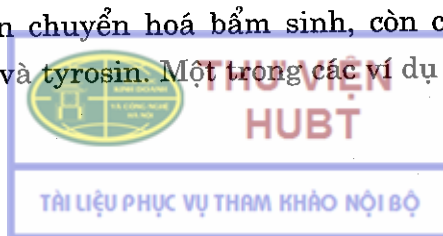
Trong y khoa, đã tìm thấy các loại đột biến khác nhau ở người. ứng dụng các định luật di truyền Mendel, người ta đã chứng minh rằng một vài bệnh ở người được di truyền như là các bệnh do một gen riêng lẻ quy định.

Năm 1908, hơn 30 năm trước khi có giả thiết "một gen – một enzym" của Beadle và Tatum, nhà vật lý người Anh Sir Archibald Garrod đã trình bày một loạt quan niệm về dị tật ở người. Trong đó, có một dị tật được ông gọi là "rối loạn bẩm sinh về trao đổi chất". Garrod cho rằng, một số dị tật gây nên do cơ thể mất khả năng thực hiện các quá trình hoá học mang bản chất di truyền.

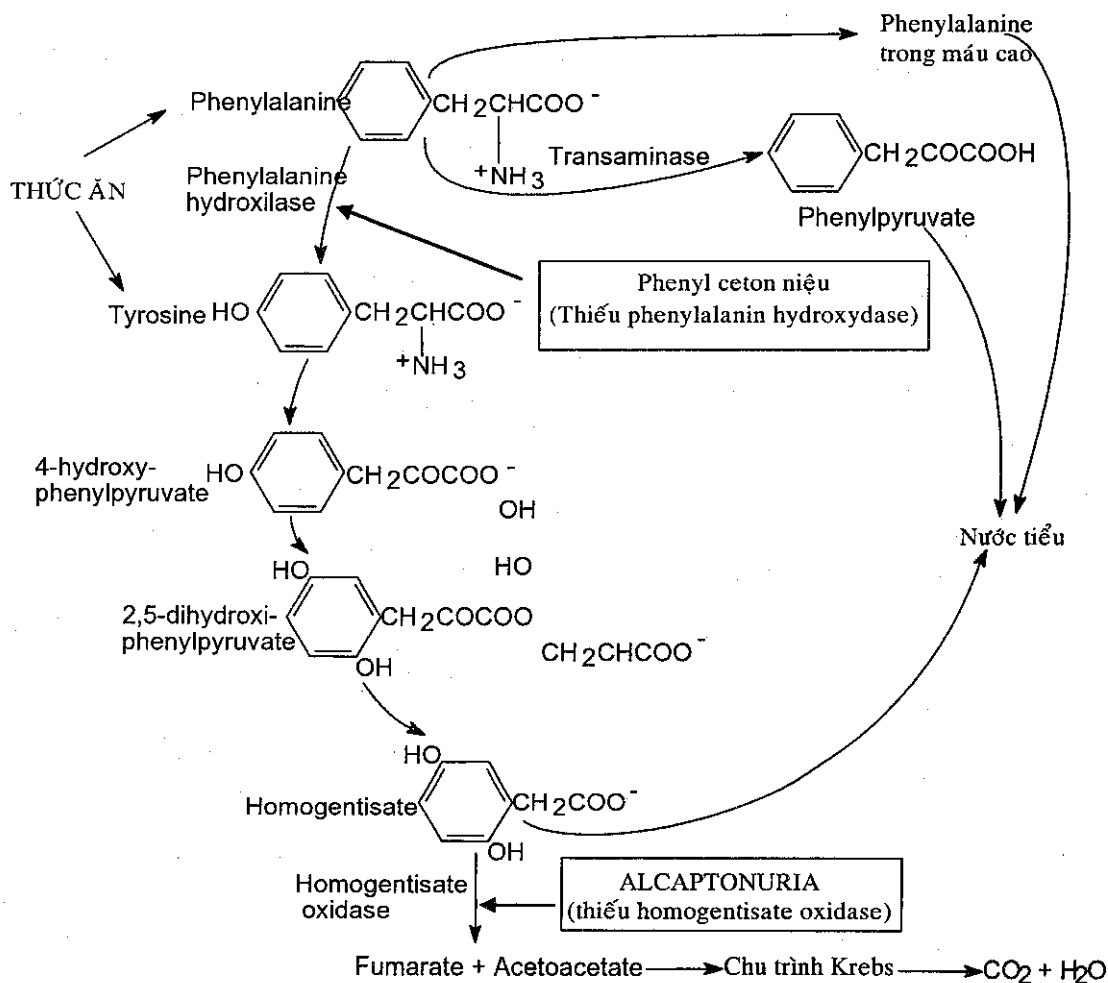
Các rối loạn chuyển hoá bẩm sinh còn được gọi là các bệnh phân tử và bệnh có nguồn gốc là những sai sót ở mức độ phân tử, những biểu hiện trực tiếp của bệnh là những phân tử protein.

Bệnh đột biến đầu tiên do Garrod nghiên cứu là bệnh Alkapton niệu. Nước tiểu của bệnh nhân trở nên đen khi để ngoài không khí. Bệnh xảy ra do thiếu enzym homogentisate oxydase trong gan nên không phân huỷ được acid homogentisic (tên cũ là alcapton) thành một sản phẩm không màu là acid acetoacetic. Lúc bấy giờ acid homogentisic tích tụ lại trong nước tiểu và có màu đen khi gặp không khí.

Liên quan tới rối loạn chuyển hoá bẩm sinh, còn có bệnh thiếu hụt enzym chuyển hoá phenylalanin và tyrosin. Một trong các ví dụ quen thuộc về dị tật bẩm



sinh được di truyền do yếu tố gen lặn Mendel ẩn trên nhiễm sắc thể thường với tần số cao (khoảng 1/10.000) là Phenyl ceton niệu. Bệnh này là do thiếu enzym phenylalanine hydroxylase trong gan để có thể oxy hoá phenylalanine thành tyrosin (Hình 8.11). Hậu quả trực tiếp là tích tụ phenylalanine trong máu và bài tiết thái quá từng phần qua nước tiểu dưới dạng các chất chuyển hoá bất thường như acid phenyl pyruvic.



Hình 8.11. Chuỗi phản ứng sinh hoá của bệnh Alkapton niệu và Phenylketon niệu, hai sự rối loạn di truyền của sự thoái quá acid amin thơm

8.5.1. Đột biến sinh dưỡng ở vi sinh vật

Năm 1945, các đột biến liên quan đến sinh dưỡng được xác định đầu tiên ở nấm *Neurospora*, sau đó là vi khuẩn. Nhiều vi sinh vật bình thường sẽ phát triển được trên môi trường tối thiểu có chứa đường, ion NH_4^+ và một vài ion cần thiết

khác. Những vi sinh vật này là những vi sinh vật hoang dại và được gọi là thể nguyên dưỡng. Sau khi được xử lý với những tác nhân gây đột biến cảm ứng, chúng thu được không còn phát triển trên môi trường giản đơn nữa trừ khi thêm vào một chất khác, thường là acid amin, vitamin nucleotid hoặc một phần tử vi lượng nào đó. Những cá thể mang đột biến này được gọi là thể khuyết dưỡng.

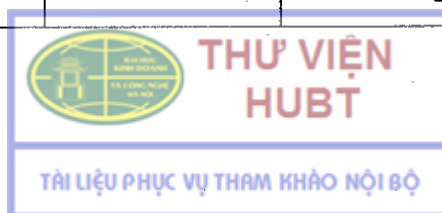
Như vậy, thể nguyên dưỡng có thể tổng hợp tất cả các thành phần tế bào phức tạp từ các nguyên liệu thô sơ, đơn giản, nhưng vi sinh vật khuyết dưỡng thì đã mất đi khả năng này. Vì hệ tổng hợp các phân tử phức hợp nói chung có liên quan đến enzym và sự đột biến ở sinh vật khuyết dưỡng có thể do làm mất hoạt tính enzym.

8.5.2. Đột biến và bệnh ở người

Vài năm sau phát minh của Garrod, hàng ngàn bệnh được tìm ra có thể di truyền được do đột biến gen. Nhiều bệnh có liên quan đến các enzym bị bất hoạt hay protein bị thay đổi chức năng (Bảng 8.3).

Bảng 8.3. Một số bệnh di truyền ở người có liên quan đến protein chuyên biệt hay khiếm khuyết enzym

<i>Loại rối loạn</i>	<i>Tên</i>	<i>Triệu chứng điển hình</i>	<i>Hoạt tính sinh hoá bị ảnh hưởng hay khiếm khuyết</i>	<i>Tần số tính trên 1000 người</i>
Máu	Hồng cầu lưỡi liềm (sickle cell anaemia)	Hồng cầu biến dạng, hình lưỡi liềm	Hemoglobin có xu hướng ngưng kết và kết tủa	10 – 20 (Afro – Caribbean)
	Bệnh thiếu máu Địa Trung Hải (Thalassaemia)	Thiếu máu nặng, lách lớn	Thiếu một trong các chuỗi polypeptid của hemoglobin người trưởng thành	10 – 20 (dân địa Trung Hải)
	Thiếu máu huyết giải hồng cầu bia (Non – spherocytic hemolytic anaemia)	Thoái hoá tế bào hồng cầu, vàng da, lách lớn	Có nhiều dạng tất cả đều mất 1 enzym trong nhánh chuyển hoá HMP	0,5 (Pyruvat kinase)
	Bệnh ưa chảy máu A	Máu không đông được	Yếu tố VIII trong huyết tương	0,1 (đàn ông)
Cơ	Teo cơ Duchenne	Thoái hoá cơ	Dystrophin, một protein liên quan đến màng tế bào cơ	0,1 (đàn ông)



Loại rối loạn	Tên	Triệu chứng điển hình	Hoạt tính sinh hoá bị ảnh hưởng hay thiếu hụt	Tần số tính trên 1000 người
Chuyển hoá carbohydrate	Bệnh Von Gierke	Đường huyết thấp, sưng gan	Glucose – 6 – phosphatase trong gan	0,01
	Không dung nạp lactose	Đầy hơi, tiêu chảy	β – Galactosidase trong ruột	100
Chuyển hoá acid amin	Phenylketon niệu	Mức phenylalanine, mức phenylacetic acid trong huyết tương và nước tiểu cao; trí tuệ chậm phát triển	Phenylalanine hydroxylase trong gan	0,1
Chuyển hoá purin	Hội chứng Lesh – Nyhan	Tổn thương nơron thần kinh; thường có hành vi tự huỷ hoại	Hypoxanthine guanin phosphoribosyl transferase	0,1 (đàn ông)
Chuyển hoá Hem	Protoporphyrin	Da nhạy cảm với ánh sáng, gan bị tổn thương, protoporphyrin trong huyết tương	Enzym Ferro – chelatase của quá trình sinh tổng hợp Hem	Hiếm
Vận chuyển lipoprotein	Cao cholesterol máu gia đình	Xơ vữa động mạch (Atherosclerosis)	Các thực thể lipoprotein tỉ trọng thấp (LDL) trong màng tế bào LDL và cholesterol không được hấp thu	1
Thoái hoá Mucopolysaccharid	Bệnh Sanfilippo B	Chậm trí tuệ, tổn thương gan và khớp xương	Acetylglucosaminidase trong lysosom	0,01
Thoái hoá Gluosephingolipid	Bệnh Tay – Sachs	Bị thoái hoá thần kinh, tích lũy gangliosid trong tế bào thần kinh	Hexoseaminidase A trong lysosom	0,3 (Ashkenazi Jews)

8.6. ĐỘT BIẾN GEN VÀ UNG THƯ

Rất nhiều nguyên nhân gây ung thư khác nhau, nhưng ở cơ chế phân tử đều liên quan đến các biến đổi di truyền trên ADN làm tổn thương tiến trình tăng sinh bình thường.

8.6.1. Các nhóm gen liên quan tới ung thư

Liên quan tới sự tăng sinh của tế bào, có 2 con đường dẫn tới ung thư:

Thứ nhất, các biến đổi làm tăng hoạt gen kích thích, kiểu đột biến này có hiệu quả trội. Alen biến đổi được gọi là oncogen (gen ung thư), còn alen bình thường là proto – oncogen (gen tiền ung thư).

Thứ hai, các biến đổi làm bất hoạt gen kìm hãm, kiểu đột biến này thường có hiệu quả lặn. Gen kìm hãm được gọi là gen ức chế khối u.

8.6.2. Gen ung thư và gen tiền ung thư

8.6.2.1. Các gen ung thư do retrovirus tải nạp

Các retrovirus bằng con đường sao chép ngược đã tích hợp được ADN của mình vào nhiễm sắc thể tế bào chủ. Chúng có thể ngẫu nhiên mang các gen ung thư làm biến đổi tế bào chủ bằng cách xen đoạn ADN của mình cạnh gen tiền ung thư của tế bào chủ. Hàng loạt các gen ung thư đã được xác định từ retrovirus gây biến đổi tế bào (Bảng 8.4).

Bảng 8.4. Một số oncogen được xác định nhờ retrovirus gây biến thể

Họ gen ung thư	Chức năng của gen tiền ung thư	Nguồn gốc virus	U do virus gây ra
<i>Abl</i>	Protein kinase (tyrosin)	Chuột, mèo	Pre – B – calleukemia
<i>erb – B</i>	Protein kinase (tyrosin): thụ thể của yếu tố tăng trưởng thượng bì (EGF)	Gà	Erythroleukemia Fibrosarcoma
<i>Sis</i>	Yếu tố tăng trưởng từ tiểu cầu (PDGF)	Khỉ	Sarcoma, Carcinoma, Myelocytoma
<i>Src</i>	Protein kinase	Gà	Sarcoma

abl = Bạch huyết Abelson; *erb* = Erythroblastose (tăng nguyên hồng cầu);
Sis = Simian sarcoma virus; *Src* = Sarcoma.

Một số gen ung thư khác được phát hiện không do virus. Hiện nay, đã phát hiện được hơn 60 gen tiền ung thư.

Các gen tiền ung thư hầu như liên quan đến các hệ thống tín hiệu của tế bào như các protein, các thụ thể xuyên màng, các gen điều hoà...

8.6.2.2. Các cách biến đổi gen tiền ung thư thành gen ung thư

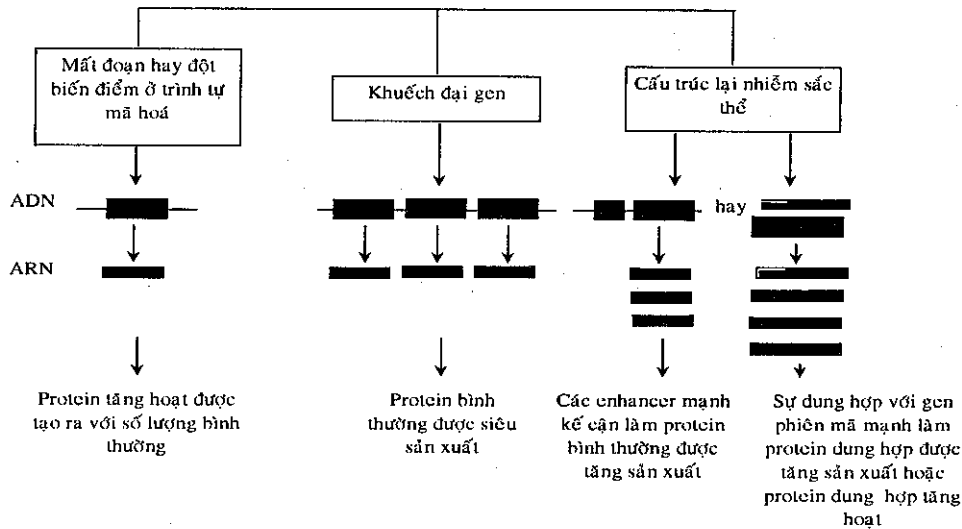
Ba cách chủ yếu trong các cách đó là:

- Mất đoạn hoặc đột biến điểm trong trình tự mã hoá.
- Khuếch đại gen.



– Cấu trúc nhiễm sắc thể.

Các gen có thể bị biến đổi do đột biến điểm, mất đoạn, chuyển đoạn nhiễm sắc thể hay do xen đoạn bộ gen của retrovirus (Hình 8.12).



Hình 8.12. Ba cách biến đổi gen tiền ung thư thành gen ung thư

8.6.3. Các đột biến ở gen ức chế khối u

Quá trình ức chế là kết quả của đột biến thứ cấp đảo ngược tác động của đột biến sơ cấp. Quá trình này xảy ra ở mức độ phiên mã và chỉ có hiệu quả một phần, lưu trữ đủ protein có hoạt tính để cho tế bào sống mà thường là một loại giả hoang dại bất thường.

Các đột biến ở gen ức chế khối u thường là lặn, nên tế bào chỉ mất kiểm soát khi cả 2 allele đều bị đột biến. Có 2 đột biến được phát hiện:

- Gen Rb tạo protein retinoblasma ức chế sự phát triển một dạng ung thư ở mắt (retinoblastoma). Rb còn có tác dụng ức chế đối với một số dạng ung thư khác.
- Gen p53 tạo protein p53 ức chế nhiều dạng ung thư. Các nghiên cứu gần đây cho thấy protein p53 tham gia hệ thống cấp cứu, sửa chữa nhiều tổn thương của tế bào đang phân chia.

Các virus ADN như Papilomavirus và SV40 có thể gây ung thư bằng cách cô lập các sản phẩm của gen ức chế khối u như protein etinoblastoma hay protein p53.

8.7. CÁC HỆ THỐNG CHỌN LỌC ĐỘT BIẾN Ở VI SINH VẬT

Các hệ thống chọn lọc đột biến giúp phát hiện có hiệu quả các đột biến hiếm hoi trong nguồn đột biến rộng lớn. Có nhiều hệ thống chọn lọc đột biến và áp dụng tùy thuộc vào các đột biến (Bảng 8.5).

Bảng 8.5. Cơ sở phát hiện ba loại đột biến dựa vào kiểu hình

Kiểu hình Kiểu gen	Đột biến có điều kiện		Đột biến khuyết dưỡng		Đột biến đề kháng	
	Nhiệt độ thường	Nhiệt độ cao	Không bổ sung	Có bổ sung	Không có tác nhân	Có tác nhân
Hoang dại	Bình thường	Bình thường	Mọc	Mọc	Mọc	Không
Đột biến	Bình thường	Đột biến	Không	Mọc	Mọc	Mọc

Hệ thống chọn lọc đột biến được đánh giá bằng năng lực phân giải, là khả năng phát hiện các đột biến rất hiếm so với không đột biến. Năng lực phân giải càng lớn thì phát hiện được các đột biến càng hiếm. Ví dụ, phát hiện các đột biến đề kháng có độ phân giải cao vì khi cấy số lượng rất lớn tế bào lên môi trường chọn lọc có chứa các tác nhân thì phần lớn tế bào chết, chỉ có rất ít tế bào có đột biến đề kháng mọc được thành khuẩn lạc.

8.7.1. Phương pháp đề kháng

Dùng tác nhân là thuốc hoặc phage để chọn lọc vi khuẩn. Các đột biến dễ dàng được phát hiện ở các khuẩn lạc mọc trên môi trường thạch có thuốc hay phage.

8.7.2. Phương pháp làm giàu chậm

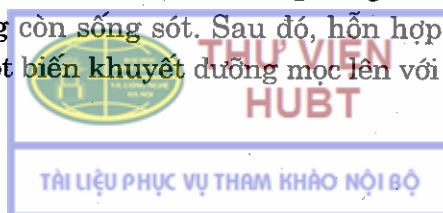
Dùng để phát hiện được các đột biến khuyết dưỡng. Dung dịch vi khuẩn pha loãng được cấy trên bề mặt môi trường thạch tối thiểu để mọc thành các khuẩn lạc rời. Phủ một lớp mỏng môi trường tương tự lên mặt môi trường thạch vừa rời, ủ để các khuẩn lạc bình thường mọc lên. Sau đó lại phủ thêm một lớp môi trường dinh dưỡng có bổ sung và ủ tiếp cho chất bổ sung khuếch tán. Các đột biến khuyết dưỡng sẽ mọc sau khi có chất bổ sung, nên khuẩn lạc nhỏ hơn do mọc chậm.

8.7.3. Phương pháp làm giàu hạn chế

Đây là dạng đơn giản hơn của phương pháp làm giàu chậm. Các vi khuẩn được cấy trên môi trường tối thiểu có một ít chất bổ sung. Trong điều kiện này, các đột biến khuyết dưỡng mọc đến khi hết chất bổ sung thì dừng, do đó tạo khuẩn lạc nhỏ. Các vi khuẩn bình thường tiếp tục tạo khuẩn lạc to.

8.7.4. Phương pháp làm giàu nhờ penicillin

Penicillin có tác động diệt khuẩn dạng phân chia. Vi khuẩn được cho vào môi trường tối thiểu có penicillin. Các vi khuẩn đang tăng trưởng bị diệt, chỉ các tế bào đột biến không tăng trưởng còn sống sót. Sau đó, hỗn hợp được cấy lên môi trường không có penicillin thì các đột biến khuyết dưỡng mọc lên với tỉ lệ tương đối cao hơn.

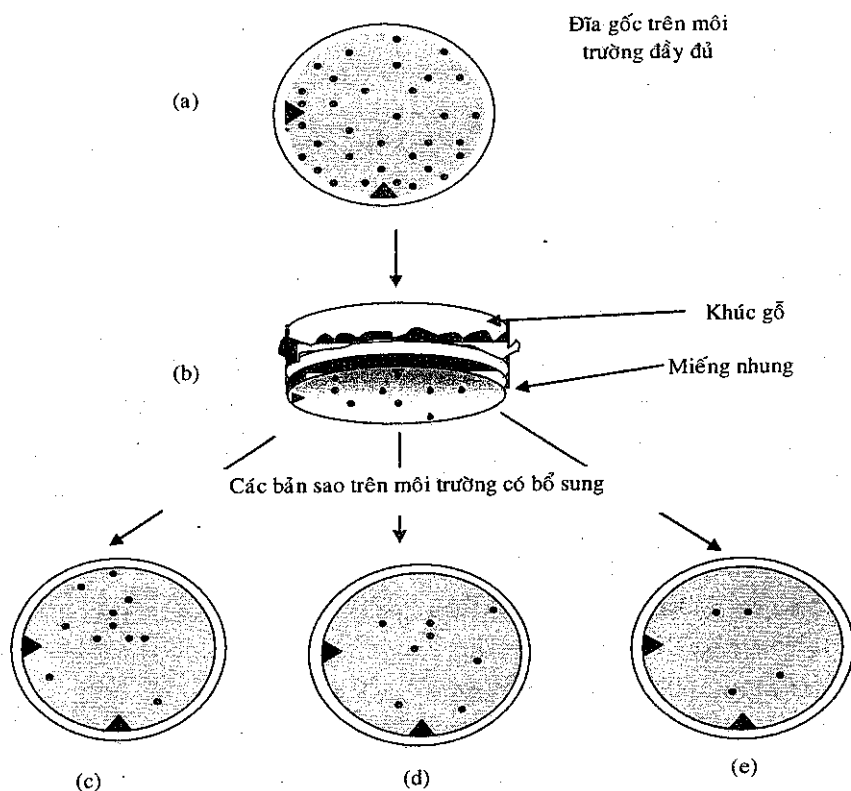


8.7.5. Phương pháp lọc

Dùng để chọn lựa các đột biến khuyết dưỡng ở nấm sợi. Dung dịch bào tử được nuôi trong môi trường dinh dưỡng thiếu chất bổ sung. Các đột biến thiếu chất bổ sung không mọc được, các dạng bình thường mọc ra thành nhiều sợi. Khi lọc qua màng lọc sợi thủy tinh, các dạng bình thường nhiều sợi bị giữ lại, các dạng đột biến đi qua màng lọc. Dung dịch có nhiều dạng đột biến được cấy trên môi trường có chất bổ sung và kiểm tra tìm các dạng đột biến.

8.7.6. Phương pháp in

Vi khuẩn được cấy để mọc rồi từng khuẩn lạc trên môi trường dinh dưỡng. Dùng miếng nhung (có nhiều lông mịn bọc trên khúc gỗ) in đúng vị trí khuẩn lạc trên môi trường tối thiểu. Các khuẩn lạc đột biến không mọc lên được. Căn cứ vị trí khuẩn lạc không mọc ở bản sao, tách các đột biến khuyết dưỡng (Hình 8.13).



Hình 8.13. Phương pháp in

- (a) Đĩa gốc trên môi trường đầy đủ; (b) Khúc gỗ bọc miếng nhung;
(c) Môi trường có bổ sung chất A và B; (d) Môi trường có bổ sung chất A, B và C;
(e) Môi trường có bổ sung chất A và C

Tóm tắt:

Đột biến là một quá trình có thể được nghiên cứu ở mức độ sinh hoá tế bào và ở mức độ cơ thể sinh vật. Đột biến là nguyên nhân đầu tiên của các biến dị trong tự nhiên và trong thí nghiệm, biến dị cũng có thể xảy ra bằng cách sử dụng các hoá chất và tia phóng xạ gây đột biến.

Ở mức độ phân tử, một đột biến là một sự thay đổi trong trình tự base của ADN do một sai sót trong sự sao chép hay sửa chữa. Đột biến tự nhiên và tác động của các tác nhân gây đột biến có thể được quy về sự bất cặp sai của các base, trong khi đó các lỗi do các hệ thống sửa chữa lỗi tự do thông thường lại tương ứng với sự biến dạng trong chuỗi xoắn kép. Sự hư hại lớn ADN do tia cực tím có thể hoạt hoá một hệ thống sửa chữa lỗi phát sinh.

Các biến dị đột biến ở mức độ ADN có thể là sự thay thế một base riêng lẻ, hay chèn hay xoá một base (thay đổi khung) hoặc của vài base (đa điểm). Sự thay thế một base riêng lẻ có thể dẫn đến sự thay đổi của một acid amin trong phân tử protein tạo thành và tùy vào bản chất và vị trí của các biến đổi này mà protein có thể mất hoạt tính hoàn toàn, hay mất hoạt tính một phần, hay có hoạt tính nhưng không ổn định. Nếu sự biến đổi một base lại tạo ra codon kết thúc thì protein sẽ không được tổng hợp nữa và thường là bất hoạt. Các đột biến ức chế cũng có thể bảo toàn được hoạt tính trong một số trường hợp nhờ sự bù trừ cho đột biến gốc trong quá trình dịch mã.

Ở mức độ cơ thể, các kết quả đột biến thể hiện ở một tính trạng đột biến nào đó. Hầu hết các đột biến đều có hại và các đột biến này hầu như được biết ở các bệnh di truyền ở người. Hy vọng giải thích sinh hoá sẽ làm sáng tỏ các nguyên nhân gây ung thư.

Các hệ thống chọn lọc đột biến ở vi sinh vật đã hoàn thiện các phương pháp chọn lọc đột biến trong phòng thí nghiệm.

CÂU HỎI

1. Kiểu hình nào **không** là hậu quả của đột biến ở vi khuẩn?
 - a) Biến đổi từ nguyên dưỡng sang khuyết dưỡng và ngược lại.
 - b) Đột biến thành thường biến và ngược lại.
 - c) Mất hoặc tạo thành phần cấu trúc của tế bào.
 - d) Nhạy cảm hay đề kháng thuốc.
 - e) Nhạy cảm hay đề kháng phage.



- e) Bị cảm ứng bởi hoá chất
9. Cơ chế chống lại đột biến
- a) NER – cắt bỏ nucleotid b) Mtase – khử alkyl
c) Glycosylase – cắt bỏ base sai d) Dung nạp AND sai
e) Tất cả
10. Gen kích thích trở nên tăng hoạt tính gây ra bệnh:
- a) Ung thư b) Tiền ung thư
c) Bình thường d) Không ung thư
e) a và b

Bài 9

MỘT SỐ PHƯƠNG PHÁP VÀ KỸ THUẬT CƠ BẢN TRONG SINH HỌC PHÂN TỬ

MỤC TIÊU

- Mô tả được đặc điểm một số công cụ cơ bản của sinh học phân tử và cách sử dụng chúng.
- Trình bày được một số phương pháp cơ bản của sinh học phân tử.
- Vận dụng được các phương pháp và công cụ này trong thực hành.

9.1. CÔNG CỤ CƠ BẢN

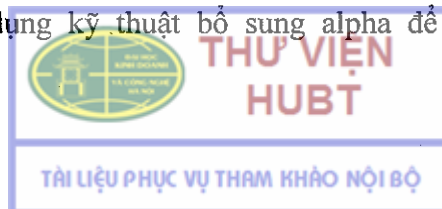
9.1.1. Chủng vi sinh vật

Các nghiên cứu sinh học phân tử thường được tiến hành trên *E. coli* vì nó có bộ máy di truyền đã được nghiên cứu khá đầy đủ và có tốc độ tăng trưởng nhanh trong khi khả năng gây bệnh lại thấp. Dưới đây là một số chủng xuất phát từ chủng *E. coli* K12 thường sử dụng:

- *HB101*: là chủng thường dùng cho việc nhân bản các dòng và dòng con của plasmid. Chủng này phát triển tốt và một đột biến ở gen *recA*, làm cho nó mất khả năng thực hiện sự tái tổ hợp giữa các trình tự tương đồng. Đột biến này làm giảm khả năng tái sắp xếp các trình tự đã đồng hóa trong quá trình nhân bản.

- *Y1090*: là chủng bị mất β -galactosidase được thiết kế đặc biệt cho việc biểu hiện các protein dung hợp với β -galactosidase. Nó chứa plasmid pMC9, chịu trách nhiệm tổng hợp lượng lớn *lac* repressor và ngăn sự biểu hiện của fusion protein, trừ khi *lac* repressor bị bất hoạt bởi isopropyl-1-thio- β -D-galactoside (IPTG). Chủng này cũng mang một đột biến *lon* (làm bất hoạt một protease của *E. coli*), cho phép biểu hiện mạnh các protein dung hợp ngoại lai. Phát triển trên môi trường LB có bổ sung ampicillin để duy trì sự ổn định của plasmid pMC9.

- *JM103*: dùng để nhân bản các vector của phage M13, cũng có thể dùng với plasmid pUC, là các plasmid sử dụng kỹ thuật bổ sung alpha để chọn lọc dòng tái tổ hợp.



Chủng này bị mất gen β -galactosidase trên nhiễm sắc thể, nhưng có mang gen cho đoạn omega của β -galactosidase trên plasmid F, cho phép nó bị nhiễm bởi các phage dạng sợi như M13. Nếu mất plasmid F sẽ làm mất khả năng này, do đó phải duy trì chủng này trên môi trường tối thiểu.

– *JM109*: tương tự JM103, ngoại trừ có đột biến trên gen *recA*, làm cho nó mất khả năng tái tổ hợp tương đồng. Chủng này được ưa dùng hơn JM103 nếu tái tổ hợp là các đoạn chèn thay đổi. Tuy nhiên, hiệu suất thu ADN của M13 của nó thấp hơn JM103.

– *DH5- α* : chủng này cũng có đặc tính bổ sung alpha như JM103 hay JM109. Chủng này mang gen *recA1* đột biến. Ngoài ra, nó còn mang đột biến *deoR* làm dễ dàng cho việc thu nhận các mảnh ADN lớn.

– *DH5- α F'*: chủng này rất tốt cho việc nhân bản vector M13 mp. Nó cũng cho phép tạo ra ADN sợi đơn từ plasmid có origin f1 (như vector M13) khi bị nhiễm đồng thời một phage trợ giúp (helper phage). Khác với JM103, F' episome, cần để nhân bản phage sợi đơn, ổn định, không cần sử dụng môi trường tối thiểu để duy trì nó.

– *DMI*: chủng này bị mất cả *dam* và *dcm* methylase, enzym metyl hóa điểm nhận diện của RE có trình tự GATC, CCAGG hay CCTGG. Do vậy, ADN được nhân bản trong chủng này có thể được cắt ra bởi nhiều RE mà ở các chủng các không thể.

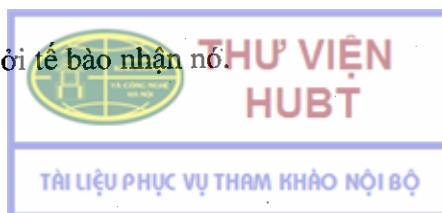
– *SURE*: chủng này được chọn để nhân bản các dòng plasmid có xu hướng tái sắp xếp và là chủng đa dụng cho việc nhân bản plasmid. Nó mang một loạt các đột biến nhằm ngăn chặn một số cách sửa chữa ADN có liên quan đến sự tái sắp xếp tạo ra các cấu trúc không xác định như Z-ADN. SURE cũng chứa episome F' để nhân bản các phage dạng sợi như M13, cũng như các plasmid kháng ampicillin hay chloramphenicol. Nó cũng kháng được tetracyclin và kanamycin, nên không dùng để chọn lọc các thể biến nạp. Giống như *DH5- α* nó cũng sử dụng kỹ thuật bổ sung alpha.

– *XLI-Blue MRF'*: rất tốt để nhân bản các thư viện cADN hay genome cũng như các phage như M13. Nó mang các đột biến làm bất hoạt hầu hết các RE của *E. coli*, do đó ngăn việc cắt các ADN có mang C hoặc A được metyl hóa, hay gập trong các cADN hay ADN chần hạch. Nó có mang đột biến *hsdR*, ngăn sự cắt ADN đã tạo dòng bởi *EcoK*. Nó cũng mang *recA*, F' và kỹ thuật bổ sung alpha. Chủng này mang gen kháng tetracyclin nên không dùng để chọn lọc dòng sau biến nạp.

9.1.2. Vector tạo dòng

Vector tạo dòng là vật liệu di truyền trung gian để chuyển gen giữa các tế bào do chúng có khả năng được sao chép và phân phối vào các tế bào mới trong quá trình phân bào, nhờ đó gen đang nghiên cứu cũng được sao chép và hiện diện ở tất cả các tế bào xuất phát từ tế bào nhận gen ban đầu. Vector tạo dòng là các ADN có thể truyền được và đáp ứng các yêu cầu sau:

- Có thể được sao chép bởi tế bào nhận nó



– Có trình tự nhận diện của enzym giới hạn để cắt và đưa gen ngoại lai vào.

Mang một hay nhiều yếu tố chọn lọc nào đó để phân biệt hoặc chọn lọc tế bào nhận được vector với tế bào không nhận được.

Có bốn loại vector cơ bản dùng trong nghiên cứu sinh học phân tử là plasmid, các vector thực khuẩn tiêu giải có nguồn gốc từ phage lambda, các cosmid và vector từ các phage dạng sợi như M13, tuy nhiên loại sau cùng hiện nay ít được sử dụng.

9.1.2.1. Plasmid

Plasmid là một phân tử ADN sợi đôi, dạng vòng kín, nhỏ có thể nhân bản độc lập với nhiễm sắc thể trong tế bào. Tất cả các plasmid được dùng trong nghiên cứu sinh học phân tử phải thỏa mãn ba yêu cầu:

– Có yếu tố đánh dấu để chọn lọc dòng tái tổ hợp, thường là một hay nhiều gen kháng kháng sinh.

– Có điểm khởi đầu sao chép (Ori) cho phép nhân bản plasmid trong tế bào chất không phụ thuộc nhiễm sắc thể. Các plasmid được sử dụng hiện nay thường có điểm Ori có thể kiểm soát được, ví dụ cho phép khuếch đại plasmid khi có mặt chất ức chế tổng hợp protein (như chloramphenicol). Điều này giúp tăng số bản sao của plasmid trong tế bào chủ và cho hiệu suất cao khi chiết tách plasmid với độ tinh khiết khá tốt. Trình tự của Ori cũng quyết định số bản sao của plasmid trong một tế bào.

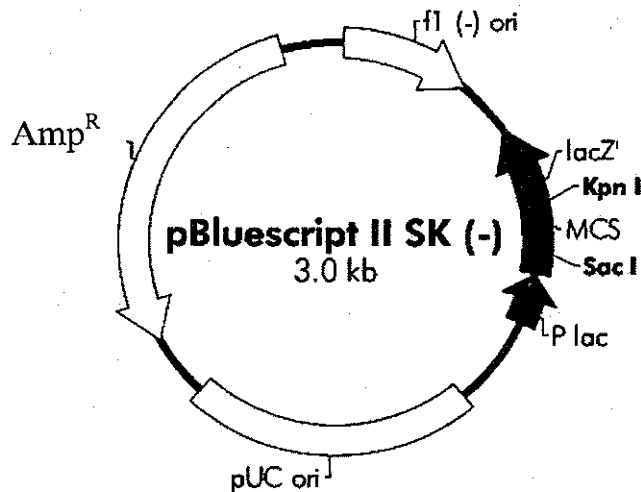
– Có vùng tạo dòng (Multiple Cloning Site - MCS): là một trình tự ADN ngắn, mang nhiều trình tự nhận diện duy nhất của các RE khác nhau nằm liên tiếp (polylinker), cho phép cắt để mở vòng plasmid và nối đoạn gen mong muốn vào.

Ngoài ra, một số plasmid có thể mang yếu tố đánh dấu chèn để nhận biết plasmid sau khi đưa vào tế bào đã được nối với gen ngoại lai trong vùng tạo dòng hay chưa.

Plasmid được đưa vào tế bào chủ bằng cách biến nạp (transformation) hay thẩm điện (electroporation).

pUC:

Là một dãy các plasmid có mang một số đặc tính của vector M13 và pBR, làm thuận lợi hơn cho thao tác. Kích thước nhỏ (2,7 kb), có gen kháng ampicillin, điểm Ori của pBR322 và một phần gen *lacZ* của *E. coli*. MCS (tương tự M13) nằm trong đoạn *lacZ* nên có thể dùng kỹ thuật bổ sung alpha để làm yếu tố đánh dấu chèn. Cho số bản sao cao, và có thể được khuếch đại bằng chloramphenicol.



Hình 9.1. Sơ đồ một plasmid điển hình

pBR322:

Plasmid dài 4363 bp (base pair - cặp base), dùng để tạo dòng nhanh và đơn giản các đoạn ADN. Có gen kháng tetracycline và ampicillin. Hiệu suất tạo dòng thấp hơn các vector sau này.

pGEM và pBluescript II:

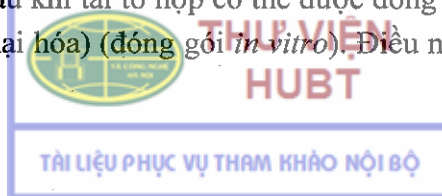
Mang nhiều yếu tố thuận tiện cho thao tác. Kích thước khá nhỏ (2,9 kb), cho số bản sao cao trong tế bào *E. coli* thích hợp (SURE). MCS nằm giữa 2 promoter T7 và T3 cho phép phiên mã từ trình tự được chèn vào.

9.1.2.2. Các vector từ phage lambda

Đây là một phage tiêu giải, nghĩa là sau khi xâm nhập tế bào chủ, phage sẽ nhân bản ADN của mình và tạo các virion, ly giải tế bào chủ rồi phóng thích phage. Nếu tế bào chủ được nuôi cấy trong môi trường bán lỏng, pha nhiễm vào và nhân lên sẽ làm tế bào bị ly giải tạo ra vùng ly giải (plaque) xung quanh vị trí của phage ban đầu, khi đó muốn thu hoạch phage người ta sẽ hút phần môi trường tại plaque.

Khi được dùng làm vector, người ta cắt bỏ một phần bộ gen của phage không liên quan đến quá trình tiêu giải và thay thế bằng đoạn ADN cần tạo dòng. Do đặc tính của phage, nên sau tái tổ hợp kích thước có thể vector phải lớn hơn 78% (38 kb) và nhỏ hơn 102% (51 kb) so với dạng hoang dại (50 kb) thì mới đóng gói dễ dàng thành phage được. Vì vậy, khi sử dụng làm vector tái tổ hợp phải chú ý đến kích thước đoạn gen muốn chèn, nếu đoạn gen mong muốn không đủ kích thước thì người ta sẽ nối thêm ADN độn khác để đủ kích thước có thể đóng gói được.

Có hai điểm thuận lợi khi dùng loại vector này để xây dựng các thư viện tái tổ hợp. Thứ nhất, bộ gen của phage sau khi tái tổ hợp có thể được đóng gói bằng cách trộn với các protein vỏ (đã được thương mại hóa) (đóng gói *in vitro*). Điều này rất có ý nghĩa vì khi đó



chỉ cần một lượng ít vector và ADN cần chèn. Đóng gói *in vitro* là cách đơn giản nhất để đưa ADN tái tổ hợp vào tế bào *E. coli* chủ. Thứ hai, các thư viện phage có thể dễ dàng phát hiện với mẫu dò, khuếch đại và bảo quản trong một thời gian dài.

λgt10:

Vector sợi đôi thẳng, dài 43,3 kb. Dùng để tạo dòng những mẫu ADN ngắn (< 7 kb) được cắt bằng *EcoRI*. Việc chèn này sẽ làm bất hoạt *cI* (chất ức chế) gen (*i434*), tạo ra dòng phage *cI* âm. Phage không tái tổ hợp sẽ có *cI*⁺ và cho plaque đục trên tế bào chủ thích hợp như *E. coli* c600. Phage tái tổ hợp có *cI*⁻ tạo plaque trong dễ dàng nhận ra. Hơn nữa, trong *E. coli* có đột biến *hfl*, chỉ có phage *cI*⁻ là tạo được plaque.

λgt11:

Vector sợi đôi thẳng, dài 43,7 kb được thiết kế chỉ để tạo dòng các mảnh ADN nhỏ được cắt bởi *EcoRI* (< 7 kb). Vị trí cắt duy nhất của *EcoRI* nằm gần đầu tận của *lacZ*. ADN ngoại lai khi tạo dòng bằng vector này có thể được biểu hiện thành protein dung hợp (fusion protein) với β -galactosidase nếu được chèn đúng khung (frame) với trình tự mã hóa *lacZ*, và có thể phát hiện bằng kháng thể đánh dấu hay mẫu dò ADN. Ngoài ra, việc chèn ADN không đúng khung sẽ tạo ra phage *lac*⁻ nên tạo plaque không màu trên ký chủ có *gal*⁻ thích hợp như Y1090, trong khi phage không tái tổ hợp sẽ có *lac*⁺ và cho plaque xanh (bổ sung alpha).

Vector này cũng có các đặc điểm di truyền giúp cho việc biểu hiện cADN được tạo dòng. Nó có chất ức chế tiêu giải nhạy với nhiệt (*cI857*), ở 32°C chất ức chế hoạt động và phage có dạng tiềm ẩn bền, ở 42°C, nó bị bất hoạt và phage sẽ chuyển thành tiêu giải.

λZAP II:

Được thiết kế để tạo các thư viện cADN. Nó có thể nhận đoạn chèn đến 10 kb với MCS có 6 vị trí chèn. Nó cũng có bổ sung alpha và khả năng biểu hiện giống *λgt11*.

9.1.2.3. Vector cosmid

Cosmid là một dạng lai giữa phage và plasmid và có được các ưu thế của cả hai dạng vector. Nó có thể dùng để tạo dòng các đoạn gen có kích thước lớn đến 50 kb. Cosmid vector mang yếu tố chọn lọc, điểm Ori của plasmid, vị trí tạo dòng và các trình tự của phage lambda mã hóa cho vị trí *cos*. Để tái tổ hợp, người ta cắt vector bằng RE tại vị trí tạo dòng và trộn với đoạn gen muốn chèn (có chiều dài từ 35 - 45 kb), vector và đoạn chèn sẽ được nối lại bằng ligase, tạo thành một sợi ADN thẳng với đoạn chèn bị kẹp giữa 2 phần của vector. Sợi ADN này khi ủ với các protein vỏ sẽ tạo thành phage và được gây nhiễm vào *E. coli*. Sau khi vào tế bào chủ, ADN tái tổ hợp sẽ tạo thành vòng nhờ vị trí *cos* và hoạt động như một plasmid.

9.1.2.4. Vector từ phage dạng sợi

Vector này có ADN sợi đơn xuất phát từ phage M13. Chúng có khả năng độc nhất vô nhị là có thể tạo dòng ADN cả sợi đơn lẫn sợi kép. Trên thực tế, vector dạng sợi đôi nhân bản được dạng vòng có kích thước 7 kb của phage M13 được tinh chế và ADN ngoại lai được nối bằng ligase vào MCS. ADN tái tổ hợp được đưa vào tế bào chủ bằng cách biến nạp qua trung gian CaCl_2 , khi đó dạng sợi đôi được nhân bản cùng với sợi đơn dạng vòng từ một trong 2 mạch. ADN sợi đơn bây giờ sẽ được đóng gói thành virion và được xuất ra môi trường mà không có sự tiêu giải của tế bào *E. coli*, do vậy chúng sẽ tích lũy với số lượng rất lớn trong môi trường. Các virion này có thể được tinh chế từ môi trường và tách lấy ADN sợi đơn để làm các thí nghiệm như định trình tự hoặc làm khuôn mẫu để gây đột biến đặc hiệu điểm. Phần ADN sợi đôi cũng có thể được tách và tinh chế như các plasmid.

M13: Đây là một phage dạng sợi của *E. coli* có chứa bộ gen là ADN vòng sợi đơn 7,2 kb. Dòng vector M13 mp được cải tạo từ phage M13 hoang dại, cho phép tạo dòng một cách dễ dàng nhiều loại ADN. Vị trí tạo dòng được đưa vào ở đầu amin tận của *lacZ*. Nếu phage không tái tổ hợp nhiễm vào các tế bào chủ có cơ chế bổ sung alpha (gal^-) như JM103, JM109 hay DH5- α F sẽ cho phép vi khuẩn chuyển hóa X-gal thành chất có màu xanh. Ngược lại nếu phage tái tổ hợp nhiễm sẽ làm mất khả năng này và plaque vi khuẩn sẽ không có màu xanh.

9.1.3. Các enzym làm biến đổi acid nucleic

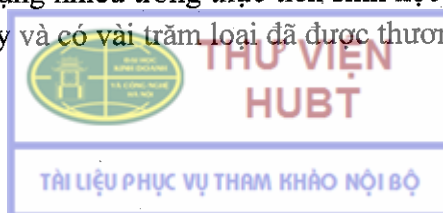
9.1.3.1. Enzym cắt hạn chế / methylase

Chức năng:

Enzym cắt hạn chế (restriction enzym - RE) là các endonuclease có khả năng cắt ADN tại hay gần một trình tự nhận diện đặc hiệu. Song song với chúng tế bào có các enzym methylase xúc tác việc methyl hóa các trình tự nhận diện đặc hiệu và ngăn cản tác động của enzym cắt hạn chế. Các loài khác nhau có cặp RE/methylase khác nhau. Hậu quả là nếu các ADN ngoại lai (ví dụ ADN phage) chưa bị methyl hóa sẽ bị cắt và thoái hóa, trong khi ADN của tế bào chủ được bảo vệ. Điều này giúp cho bộ gen của vi khuẩn được ổn định.

Phân loại RE:

Có ba loại RE: I, II và III. Loại I và III mang một phức hệ gồm cả 2 hoạt tính cắt giới hạn và methyl hóa, loại này cần ATP để có năng lượng và nó cắt ADN tại vị trí khá xa điểm nhận diện. Hai loại này ít được dùng trong các thực hành sinh học phân tử. Loại II chỉ có hoạt tính endonuclease, không cần ATP để hoạt động và cắt ADN tại hoặc rất gần vị trí nhận diện. Loại II được sử dụng nhiều trong thực tiễn sinh học phân tử và hiện người ta đã biết khoảng 1000 RE loại này và có vài trăm loại đã được thương mại hóa.

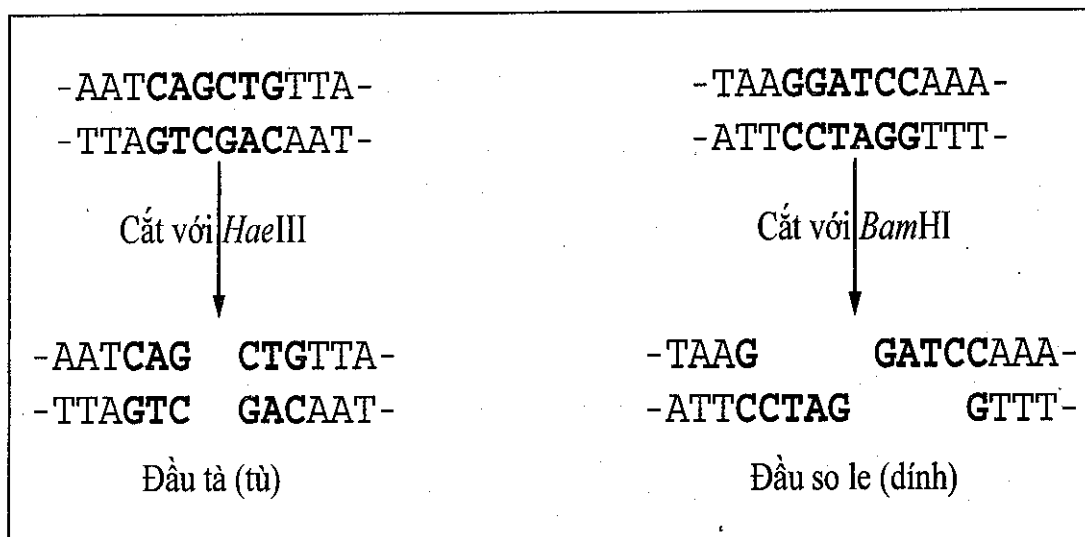


Để gọi tên RE người ta dùng ba ký tự Latin (in nghiêng) trên cơ sở tên chi và loài (đôi khi sử dụng thêm ký tự thứ tư để chỉ chủng (strain) hoặc plasmid) và một số La mã để chỉ thứ tự phát hiện ra enzym. Ví dụ: *Bam*HI: là RE của *Bacillus amyloliquefaciens* chủng H, được phát hiện đầu tiên.

Vị trí nhận diện / cắt:

Vị trí nhận diện của đa số RE loại II thường chứa 4 - 6 nucleotid và thường có giống nhau khi đọc theo cả hai hướng (trình tự palindrome). Có 4 kiểu sắp xếp của trình tự nhận diện (xem Bảng 9.1): palindrome (ví dụ *Bam*HI, *Eco*RI); palindrome bị ngắt (*Hin*fI, *Xmn*I); trình tự bị thoái biến, thường có phần còn lại của palindrome (*Eco*RII, *Hinc*II); không phải palindrome (*Mbo*II).

RE chỉ cắt ADN sợi đôi và tạo ra 2 đầu tận 5' phosphate và 3' hydroxyl. Đầu cắt có thể tà (tù) hoặc so le (còn gọi là đầu dính) (Hình 9.2). Nếu trình tự nhận diện có phần nào tính đối xứng thì điểm cắt thường nằm giữa nó và ở vị trí giống nhau cho mỗi sợi. Ngược lại, thì vị trí cắt cách khoảng so với nó.



Hình 9.2. Đầu cắt tà và đầu cắt so le

Bảng Bảng 9.1. Một số RE thường dùng

Tên RE	Vị trí nhận diện / cắt ²	Vi khuẩn
<i>Aa</i> II	GACGT*C	<i>Acetobacter aceti</i>
<i>Al</i> I	AG*CT	<i>Athrobacter luteus</i>
<i>Ball</i> (<i>Msc</i> I) ¹	TGG*CCA	<i>Brevibacterium albidum</i>
<i>Bam</i> HI	G*GATCC	<i>Bacillus amyloliquefaciens</i>

<i>Bgl</i> II	A*GATCT	<i>Bacillus globigii</i>
<i>Bcl</i> I	T*GATCA	<i>Bacillus caldolyticus</i>
<i>Bsp</i> 106I (<i>Cla</i> I) ¹	AT*CGAT	<i>Bacillus sphaericus</i> / <i>Caryophanon latum</i>
<i>Eco</i> RI	G*AATTC	<i>E. coli</i>
<i>Eco</i> RII (<i>Bst</i> NI) ¹	*CC(A/T)GG	<i>E. coli</i> / <i>Bacillus</i> <i>stearothermophilus</i>
<i>Hae</i> III	GG*CC	<i>Haemophilus aegyptius</i>
<i>Hinc</i> II	GTY*RAC	<i>Haemophilus influenzae</i> Rc
<i>Hind</i> III	A*AGCTT	<i>Haemophilus influenzae</i> Rd
<i>Hinf</i> I	G*ANTC	<i>Haemophilus influenzae</i> Rf
<i>Hpa</i> I	GTT*AAC	<i>Haemophilus parainfluenzae</i>
<i>Kpn</i> I	GGTAC*C	<i>Klebsiella pneumoniae</i>
<i>Mbo</i> II	GAAGANNNNNNN* CTTCTNNNNNNN*	<i>Moxarella bovis</i>
<i>Mbo</i> I (<i>Sau</i> 3A) ¹	*GATC	<i>Moxarella bovis</i> / <i>Staphylococcus</i> <i>aureus</i> 3A
<i>Msp</i> I (<i>Hpa</i> II) ¹	C*CGG	<i>Moxarella species</i>
<i>Not</i> I	GC*GGCCGC	<i>Nocardia otitidis-caviarum</i>
<i>Pst</i> I	CTGCA*C	<i>Providencia stuartii</i>
<i>Pvu</i> I	CGAT*CG	<i>Proteus vulgaris</i>
<i>Pvu</i> II	CAG*CTG	<i>Proteus vulgaris</i>
<i>Sac</i> I (<i>Sst</i> I) ¹	GAGCT*C	<i>Streptomyces achromogenes</i>
<i>Sac</i> II (<i>Ssf</i> II) ¹	CCGC*GG	<i>Streptomyces achromogenes</i>
<i>Sal</i> I	G*TCGAC	<i>Streptomyces albus</i>
<i>Sma</i> I	CCC*GGG	<i>Serratia marcescens</i>
<i>Sph</i> I	GCATG*C	
<i>Stu</i> I	AGG*CCT	<i>Streptomyces tubercidicus</i>
<i>Taq</i> I	T*CGA	<i>Thermus aquaticus</i>
<i>Xba</i> I	T*CTAGA	<i>Xanthomonas campestris</i> pv.

		<i>badrii</i>
<i>Xhol</i>	C*TCGAG	<i>Xanthomonas holcicola</i>
<i>XmaI</i>	C*CCGGG	<i>Xanthomonas malvacearum</i>

¹ có cùng vị trí nhận diện nhưng khác nguồn gốc. ² trình tự nhận diện theo chiều 5' - 3', * là vị trí cắt. A: adenin, C: cytosin, G: guanin, N: bất kỳ base nitơ nào, R: purin, T: thymin, Y: pyrimidin

9.1.3.2. Polymerase

Polymerase là các enzym chịu trách nhiệm tổng hợp acid nucleic từ các đơn vị nucleotid.

Đặc tính chung:

– **Hoạt tính tổng hợp 5'- 3'**: Tất cả ADN polymerase đều dùng deoxynucleotid-5'-triphosphate (dNTP) để gắn thêm 1 deoxynucleotid-5'-monophosphate vào đầu 3' hydroxyl của ADN mới (cần lưu ý rằng một số trường hợp ARN có thể đóng vai trò mới) và phóng thích pyrophosphate NTP. Tùy theo enzym, phản ứng cần có khuôn mẫu ADN, ARN, hay không.

– **Hoạt tính exonuclease 3'- 5'**: Đặc tính này có ở một số ADN polymerase như là chức năng “sửa lỗi”. Nghĩa là loại bỏ nucleotid gắn sai. Khả năng này đối với mạch đơn mạnh hơn. Trên mạch kép nếu có đủ dNTP, hoạt tính tổng hợp mạnh hơn. Hoạt tính này cũng mạnh lên khi tăng nhiệt độ.

– **Hoạt tính exonuclease 5'- 3'**: Một số ADN polymerase có khả năng này nghĩa là loại bỏ các nucleotid hay oligonucleotid từ đầu 5', do đó có thể phân hủy đoạn mới hay khuôn mẫu.

– **Hoạt tính ribonuclease H**: Chỉ có ở một số polymerase, nó phân hủy đặc hiệu ARN khi có sự hiện diện của phức lai ARN/ADN.

– **Dịch chuyển sợi**: Một số polymerase có khả năng chuyển sợi cũ từ vai trò khuôn mẫu thành sợi mới đang được tổng hợp. Nếu không có hoạt tính exonuclease 5'- 3', sự dịch chuyển sợi sẽ tạo ra sự tổng hợp chồng chéo trên cùng một khuôn mẫu. ADN sau khi tách ra sẽ bị phân hủy bởi exonuclease 3'- 5'.

– **Khả năng xử lý**: Là giá trị chỉ khả năng của 1 phân tử polymerase có thể tiếp tục tổng hợp trên sợi khuôn trước khi tách ra và có thể bị thay thế bởi một polymerase khác.

– **Tần suất lỗi**: Khi sử dụng *in vitro* tần suất này thường cao hơn *in vivo* và phụ thuộc vào điều kiện tiến hành. Bên cạnh đó nó cũng tùy vào khả năng “sửa lỗi” và xử lý của polymerase.

ADN polymerase phụ thuộc ADN:

E. coli ADN polymerase I:

Được phân lập từ *E. coli*, nhưng ngày nay đã được sản xuất từ chủng vi khuẩn có khả năng siêu tổng hợp enzym này.

Mục đích sử dụng:

- Chuyển khóa (nick translation): ứng dụng trong đánh dấu acid nucleic.
- Tổng hợp sợi thứ 2 của cADN.

Đặc tính:

- Hoạt tính tổng hợp 5'-3': dùng đoạn mồi ADN hay ARN. Hoạt động tốt trên khuôn mẫu ADN. Khả năng xử lý thấp và cần tối thiểu 1 μ M dNTP.
- Exonuclease 3'-5': yếu hơn polymerase của T₄ và T₇.
- Exonuclease 5'-3'.

Đoạn Klenow (đoạn lớn của polymerase I của E. coli):

Thu được nhờ tác động của enzym thủy phân protein trên polymerase I của *E. coli* hoặc nhờ siêu tổng hợp ở dòng tái tổ hợp gen quy định polymerase đã được biến đổi. Đoạn Klenow không có hoạt tính exonuclease 5'-3'. Gần đây với công nghệ di truyền đã tạo được dòng sản xuất đoạn Klenow không có hoạt tính exonuclease (exo⁻ Klenow enzym; Stratagene).

Sử dụng:

- Làm tà đầu 3' bị hụt (đầu dính).
- Loại bỏ đầu 3' nhô ra.
- Đánh dấu đầu 3' của ADN sợi đôi, đặc biệt đối với đầu 3' bị hụt.
- Tổng hợp sợi thứ 2 của cADN, có ưu điểm hơn polymerase I vì không có hoạt tính exonuclease 5'-3'.
- Đánh dấu bằng kỹ thuật mồi ngẫu nhiên.
- Giải trình tự ADN bằng phương pháp dideoxy.
- Tổng hợp ADN từ ADN sợi đơn với mồi đặc hiệu.

ADN polymerase T₄.

Được ly trích từ *E. coli* bị nhiễm phage T₄. Có đặc tính gần như đoạn Klenow nhưng có hoạt tính exonuclease 3'-5' mạnh hơn, không làm dịch chuyển sợi và không tác động trên ADN bị cắt khóa.

- Sử dụng: Làm tà đầu 3' bị hụt.



- Loại bỏ đầu 3' nhô ra.
- Đánh dấu đầu 3' của ADN sợi đôi.
- Tổng hợp ADN từ ADN sợi đơn với mỗi đặc hiệu.

ADN polymerase T₇

Gồm 2 tiểu đơn vị: lớn: có hoạt tính enzym do phage T₇ mã hóa; nhỏ: là một thioredoxin protein do tế bào chủ *E. coli* mã hóa. Khả năng xử lý thấp, nhưng nếu kết hợp với thioredoxin thì tăng lên rất nhiều. Khi thay đổi phân thioredoxin thì hoạt tính exonuclease 3'-5' bị mất.

Sử dụng:

- Giải trình tự ADN bằng phương pháp dideoxy.
- Tổng hợp các ADN cực dài.
- Làm tà đầu 3' bị hụt.
- Loại bỏ đầu 3' nhô ra và đánh dấu đầu 3' của ADN.
- Tổng hợp sợi thứ 2 của cADN, có ưu điểm hơn polymerase I vì không có hoạt tính exonuclease 5'-3'.
- Đánh dấu bằng kỹ thuật môi ngẫu nhiên.

Taq ADN polymerase:

Ly trích từ *Thermus aquaticus*, hiện nay đã có dạng tái tổ hợp. Hoạt động tối ưu ở 75 - 80°C, hoạt tính giảm 10 lần ở 37°C. Hoạt tính enzym vẫn còn hơn 50% sau 2 giờ ở 95°C.

Sử dụng:

- Khuếch đại ADN bằng phản ứng chuỗi polymerase.
- Giải trình tự ADN bằng phương pháp dideoxy. Có khả năng thực hiện ở nhiệt độ cao nên tránh được các cấu trúc thứ cấp của khuôn mẫu và có thể tiến hành với lượng mẫu rất nhỏ.

Polymerase phụ thuộc ARN (hoặc ADN):

Còn gọi là enzym sao chép ngược (reverse transcriptase). Có 2 loại: AMV (từ Avian Myeloblastosis Retrovirus) và MMLV (từ Mouse Moloney retrovirus)

Sử dụng:

- Tổng hợp sợi thứ nhất của cADN.
- Định trình tự ADN bằng phương pháp dideoxy. Trong vài trường hợp nó tốt hơn các polymerase khác.

ADN Polymerase không phụ thuộc khuôn mẫu:

Còn gọi là terminal (deoxynucleotidyl) transferase. Ly trích từ thymus của bê. Có khả năng thêm nucleotid vào nhóm 3'hydroxyl của mỗi ADN mà không cần khuôn mẫu.

ARN polymerase phụ thuộc AND:

Là các enzym phiên mã, một số enzym thông dụng như ARN polymerase của phage SP6, T₃, T₇. Có khả năng phiên mã ARN từ khuôn mẫu ADN. Ưu điểm của nó là: chỉ phiên mã từ vị trí khởi động (promoter) đặc hiệu của phage; tổng hợp nhanh các phân tử ARN dài hàng nghìn nucleotid; có thể cho ra 5 - 30 bản ARN từ một khuôn mẫu ADN.

9.1.3.3. Ligase

Là enzym xúc tác sự tổng hợp liên kết cộng hóa trị phosphodiester giữa đầu 3'-hydroxyl và 5'-phosphate của ADN, cho phép nối 2 phân tử ADN hoặc đóng vòng ADN. ADN ligase của T₄ và *E. coli* cần ADN sợi đôi, trong khi ARN ligase của T₄ có thể dùng ARN hoặc ADN sợi đơn.

ADN ligase của *E. coli* cần NAD để có năng lượng và có thể dùng sửa chữa chỗ bị cắt khóa hay nối các đầu dính. ADN ligase của T₄ cần ATP để hoạt động và có thể dùng để sửa chỗ cắt khóa, nối đầu dính và cả đầu tù, do đó nó được sử dụng rộng rãi nhất hiện nay.

9.2. MỘT PHƯƠNG PHÁP VÀ KỸ THUẬT CƠ BẢN

9.2.1. Chiết tách và tinh chế vật liệu di truyền

Một kỹ thuật cơ bản trong sinh học phân tử và công nghệ gen là chiết tách vật liệu di truyền từ tế bào sinh vật. Có ít nhất 3 kiểu khác nhau. Đầu tiên là toàn bộ ADN tế bào, thường được sử dụng khi cần nguồn vật liệu di truyền nhằm thu các gen để nghiên cứu. Kiểu thứ hai là ADN plasmid, nó bao gồm những bước cơ bản như chiết ADN toàn bộ, nhưng có những điểm khác nhau chủ yếu là ở một số giai đoạn ADN plasmid phải được tách ra từ hỗn hợp ADN nhiễm sắc thể. Cuối cùng, ADN của phage (thực khuẩn), nó được chiết tách từ thực khuẩn chứ không từ tế bào chủ nên vấn đề tách ADN thực khuẩn và ADN nhiễm sắc thể không quan trọng, tuy nhiên, cần có một số kỹ thuật đặc biệt để loại bỏ capsid của thực khuẩn.

9.2.1.1. Nguyên tắc chung

- Tế bào bị phá vỡ để phóng thích vật liệu di truyền cùng với tế bào chất.
- Tế bào chất được xử lý để kết tủa và để lại ADN tan.
- Dung dịch ADN được cô đặc và thu hồi ở dạng tinh khiết.

Phá vỡ tế bào:

Việc phá vỡ tế bào vi sinh vật nhằm phóng thích vật liệu di truyền có thể thực hiện bằng phương pháp vật lý hay hóa học.



Đối với việc chiết tách ADN từ vi khuẩn, người ta thường dùng phương pháp hóa học. Phần ngoài tế bào vi khuẩn gồm thành tế bào và màng tế bào chất, khi được xử lý với lysozyme (enzym thủy phân thành tế bào) hoặc EDTA (một phức vòng có khả năng đánh bắt các ion như Mg^{2+} , làm mất ổn định cấu trúc thành tế bào), sẽ bị yếu đi và tế bào có thể bị phá vỡ dễ dàng hoặc với các tác nhân như chất tẩy (detergent) hay dung dịch ưu trương.

Chiết tách bằng Phenol/Chloroform:

Một trong những phương pháp phổ biến và hữu dụng để phân lập và cô đặc ADN và ARN từ dung dịch nước là chiết bằng phenol/chloroform sau đó tủa bằng etanol. Trong quá trình chiết bằng dung môi hữu cơ, các protein bị biến tính và phân tách cùng với pha hữu cơ hay trong phần phân giới giữa hai pha hữu cơ và pha nước. Phenol được sử dụng trong phương pháp phải được đệm thích hợp để ngăn các sản phẩm oxi hóa của nó làm hư acid nucleic. Cần lưu ý các ống làm bằng polystyren không chịu được phenol/chloroform.

Trong trường hợp lượng protein của tế bào quá lớn người ta có thể xử lý trước với 1 protease như Pronase hay Proteinase K.

Các ARN, đặc biệt là ARN thông tin, cũng bị loại bỏ khi chiết bằng phenol, tuy nhiên để loại hoàn toàn ARN, người ta có thể dùng ribonuclease.

Phần tủa của xác tế bào và protein sẽ được tách ra bằng cách ly tâm, khi đó pha nước bên trên sẽ chứa ADN và được rút ra.

Thu hồi acid nucleic:

Thông thường dung dịch ADN thu được cần cô đặc và loại tạp bằng cách kết tủa ADN rồi hòa tan lại trong lượng nước tối thiểu.

Quy trình kết tủa kinh điển với etanol:

Với sự hiện diện của muối (cation phải có hóa trị 1) và nhiệt độ thấp, cồn tuyệt đối sẽ làm tủa acid nucleic. Trong quá trình tủa bằng etanol, muối và các chất tan khác như phenol và chloroform thừa và nằm lại trong dung dịch trong khi acid nucleic tạo thành một kết tủa trắng dễ dàng thu gom bằng cách ly tâm.

Các mảnh acid nucleic và oligonucleotid dài hơn 15 nucleotid có thể được kết tủa hiệu quả bằng cách này. Để kết tủa được hiệu quả, nồng độ acid nucleic phải ít nhất là 10 $\mu\text{g/ml}$. Nồng độ thấp hơn vẫn có thể được kết tủa nhưng độ hồi phục không xác định được.

Để tủa các nồng độ acid nucleic thấp hơn, ủ ở -20°C (hay trong nước đá khô) từ 4 giờ đến qua đêm, và sau đó ly tâm trong 30 phút để thu tủa. Một cách khác để tủa lượng acid nucleic hàng nanogram một cách hiệu quả là thêm tARN nấm men vào dung dịch để thu được nồng độ acid nucleic cao hơn 10 $\mu\text{g/ml}$ trước khi bắt đầu quy trình chiết xuất và tủa. Sự hiện diện của tARN không phải là vấn đề lớn, ngoại trừ khi nó tương tác với các thao tác enzym sau đó trên mẫu.

Muối thường được sử dụng trong phương pháp này là natri acetat 0,3 M (nồng độ cuối), muối này tan tốt hơn trong etanol so với NaCl 0,3 M, do đó không tủa cùng với acid nucleic. Đối với mẫu chứa sodium dodecyl sulfat (SDS), nồng độ muối nên dùng là NaCl 0,2 M vì SDS sẽ tan trong etanol trong điều kiện như vậy. Amonium acetat 2 M có thể dùng thay natri acetat 0,3 M, (nhằm mục đích loại các triphosphat) nhưng nó không nên dùng nếu acid nucleic sau đó được phosphoryl hóa bằng T₄ kinase hay ghép đuôi ở đầu 3' với terminal transferase, vì amonium sẽ ức chế các enzym này. Có thể dùng LiCl 8M thay cho natri acetat 0,3 M. Vì LiCl rất tan trong etanol nên tủa thu được sẽ tương đối ít muối hơn. Nhưng không nên dùng LiCl nếu ARN thu được dùng để làm khuôn mẫu cho phiên mã ngược.

Quy trình kết tủa khác:

Nếu muốn tủa thể tích dung dịch ADN lớn hơn, isopropanol đồng thể tích với pha nước được dùng thay cho etanol. Tuy nhiên, một số muối khó tan trong isopropanol và nó cũng khó bay hơi nên đôi khi cần thực hiện tiếp một quy trình tủa kinh điển với etanol ngay sau khi tủa bằng isopropanol để loại bỏ vết isopropanol và muối.

Các mARN và rARN có thể được tinh chế từ ADN sợi đôi và các ARN nhỏ (tARN và 5S ARN) bằng hai lần kết tủa chọn lọc với LiCl 8 M mà không có etanol. Dùng đồng thể tích LiCl 8 M với dung dịch acid nucleic chứa ARN, trộn kỹ và ủ ở - 20°C trong ít nhất 2 giờ. Tủa các ARN lớn bằng cách ly tâm 10000 vòng/phút trong 10 phút ở 4°C. Bỏ phần nước nổi (có chứa ARN nhỏ và ADN sợi đôi), phân tán cặn ARN ở nồng độ khoảng 1 µg/ml, rồi lặp lại quy trình tủa một lần nữa. Cuối cùng loại bỏ LiCl bằng cách thực hiện quy trình kinh điển.

Thu hồi bằng hấp phụ:

Do tích điện âm, ADN bị hấp phụ bởi silica trong sự hiện diện của các tác nhân biến tính (ví dụ muối guanidin), sau khi rửa, ADN được giải hấp phụ bằng dung dịch có nồng độ ion thấp. Silica có thể được cho trực tiếp vào dung dịch ADN hoặc được làm thành cột để dung dịch ADN chảy qua. ADN thu hồi bằng hấp phụ trên silica sạch hơn so với phương pháp kết tủa, tuy nhiên cách đó không hiệu quả đối với các acid nucleic có kích thước nhỏ.

9.2.1.2. Chiết tách ADN từ các dạng tế bào khác

Đối với tế bào thực vật hay nấm men, về nguyên tắc đều giống như đối với vi khuẩn, tuy nhiên có một số điểm khác biệt về cách phá vỡ tế bào. Thường sử dụng các phương pháp vật lý như nghiền tế bào đã được đông đá trong cối. Tế bào động vật thường không có thành tế bào nên chỉ cần xử lý với chất tẩy là đủ.



Điểm khác biệt thứ 2 là thành phần tế bào chứa rất nhiều chất khác ngoài protein như là các đường, lipid, ... và không bị tủa bởi phenol/chloroform. Để khắc phục, dịch chiết tế bào được xử lý với cetyltrimethylammonium bromid (CTAB), chất này tạo phức không tan với acid nucleic trong khi đường và protein tan. Phức hợp CTAB-ADN được ly tâm để tách rồi xử lý với NaCl để phóng thích ADN. Bước tiếp theo được thực hiện theo quy trình tủa bằng etanol.

Một cách khác để tách ADN từ hỗn hợp chứa nhiều đường dựa vào đặc điểm tích điện âm mạnh của ADN. Dịch chiết tế bào được cho hấp phụ với các resin hoặc chạy qua cột có điện tích dương, ADN bị giữ lại trong khi các thành phần khác đi qua. ADN gắn trên resin được thổi ra bằng dung dịch muối có nồng độ cao.

9.2.1.3. Plasmid

Khi cần ADN plasmid để nghiên cứu, điều cần thiết là phải tách nó ra khỏi phần ADN nhiễm sắc thể. Người ta đã phát triển một số phương pháp cho phép loại bỏ hữu hiệu ADN nhiễm sắc thể trong quá trình tinh chế plasmid. Những phương pháp này dựa trên một vài khác biệt về vật lý giữa ADN plasmid và ADN nhiễm sắc thể, mà chủ yếu là về kích thước. Plasmid lớn nhất cũng chỉ bằng 8% kích thước của nhiễm sắc thể *E. coli*. Hơn nữa ngoài kích thước chúng còn khác nhau về cấu dạng (conformation). Cấu dạng nghĩa là cấu hình không gian tổng quát của phân tử, mà hai cấu dạng đơn giản nhất là thẳng và vòng. Cả plasmid và nhiễm sắc thể đều có dạng vòng, tuy nhiên, trong quá trình phá vỡ tế bào nhiễm sắc thể sẽ bị gãy thành dạng thẳng, còn plasmid vẫn ở dạng vòng.

Tách dựa trên kích thước:

Khi tiến hành phá vỡ tế bào trong điều kiện được kiểm soát chặt chẽ và êm dịu, ADN nhiễm sắc thể không bị đứt đoạn nên có kích thước rất lớn so với plasmid và ADN nhiễm sắc thể thường gắn với màng tế bào nhờ đó có thể loại bỏ bằng cách ly tâm. Tế bào được phá vỡ nhẹ nhàng bằng cách ly giải một phần thành tế bào (cell wall) với EDTA và lysozym có sự hiện diện của đường (để tạo dung dịch đẳng trương), khi đó tế bào sẽ chuyển thành dạng sphaeroplast, sau đó sphaeroplast được phá vỡ bằng một tác nhân xà phòng không ion hóa như Triton X-100.

Tách dựa trên cấu dạng - phương pháp thủy giải kiềm:

Tế bào bị phá vỡ với tác nhân tẩy ion hóa như SDS, khi đó các ADN lớn như nhiễm sắc thể bị đứt đoạn. pH dung dịch được nâng lên 12 - 12,5 bằng NaOH làm cho các ADN thẳng bị tách thành sợi đơn, trong khi ADN plasmid ở dạng supercoil (siêu xoắn) nên không bị tác động. Khi thêm acid để làm hạ pH, ADN sợi đơn lại tái hợp, nhưng vì sợi quá dài nên kết hợp không chính xác tạo nên bó rối dày đặc, không tan và có thể dễ dàng loại bỏ bằng ly tâm. Phương pháp này còn thuận lợi hơn vì trong điều kiện ly giải như vậy, hầu hết protein và ARN cũng trở thành không tan và bị loại bỏ.

9.2.1.4. ADN phage

Đối với phage lambda, môi trường nuôi cấy để thu phage được ly tâm để loại tế bào chủ. Dịch thu được chứa phage đem xử lý với PEG (polyethylen glycol) và muối (NaCl), trong điều kiện này phage bị tủa và tách ra bằng ly tâm.

Thông thường tinh chế phage theo cách này vẫn còn lẫn protein, ADN vi khuẩn và xác tế bào. Để có phage tinh khiết hơn người ta có thể dùng phương pháp ly tâm với gradient tỷ trọng CsCl.

Đối với phage dạng sợi như M13, ADN sợi đôi (dạng phage đang nhiễm) được tách gần giống phage lambda. ADN sợi đơn thì đơn giản hơn vì nó được tế bào chủ tiết vào môi trường với số lượng khá lớn, ta có thể thu môi trường đem ly tâm loại tế bào và tủa phage bằng PEG, sau đó loại capsid bằng phenol, rồi tủa bằng etanol.

9.2.1.5. Tinh chế acid nucleic

ADN được chiết tách như trên là một hỗn hợp không đồng nhất của nhiều loại acid nucleic, do đó cần tinh chế thêm nếu muốn thu được acid nucleic tinh khiết hơn. Việc tinh chế có thể thực hiện bằng các phương pháp:

– Ly tâm phân đoạn: dung dịch acid nucleic được siêu ly tâm trong thang tỷ trọng đường saccarose hay CsCl₂ khi đó các phân tử khác nhau sẽ tách theo tỷ trọng (tỷ lệ với kích thước phân tử). Cách này được áp dụng khi lượng acid nucleic cần tinh chế lớn.

– Điện di gel và sắc ký lọc gel: tách acid nucleic theo kích thước phân tử ở qui mô nhỏ (điện di) hay lớn (lọc gel). Điện di với gel polyacrylamid (PAGE) có thể tinh chế các acid nucleic với khác biệt 1 nucleotid.

– HPLC: được áp dụng chủ yếu đối với các oligonucleotid do độ phân giải lên đến 1 nucleotid.

– Sắc ký hấp phụ: ARNm thường tận cùng bằng polyA, do đó được hấp phụ bằng resin có gắn oligo dT hay dU.

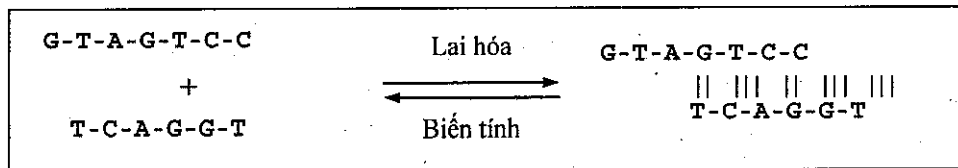
9.2.2. Kỹ thuật lai acid nucleic

Hai phân tử acid nucleic sợi đơn trong điều kiện thích hợp có thể bắt cặp bổ sung với nhau bằng các liên kết hydro để tạo thành phân tử sợi đôi, quá trình này gọi là lai hóa acid nucleic (hybridization) hay đơn giản là “lai phân tử”. Phân tử sợi đôi thu được gọi là phân tử lai (hybrid) hay duplex. Như vậy, lai phân tử là sự tương tác không hóa trị, đặc hiệu trình tự giữa hai hay nhiều trình tự acid nucleic bổ sung để tạo thành một phân tử lai.

Đây là một phản ứng thuận nghịch, quá trình ngược lại khi duplex tách ra thành các sợi đơn được gọi là sự biến tính (denaturation). Do tương tác giữa hai sợi acid nucleic

trong phân tử lai là do liên kết hydro, chúng không bền với nhiệt và các yếu tố ảnh hưởng đến liên kết hydro như lực ion, pH,... nên sự biến tính có thể được thực hiện bằng nhiệt độ hay hóa học.

Phản ứng lai được xem là đặc hiệu khi sự bắt cặp bổ sung hoàn toàn của một phân tử với phân tử còn lại diễn ra trên một vùng liên tục giữa hai mạch, ngược lại khi có các điểm không bắt cặp tồn tại giữa phân tử lai hoặc một phân tử chỉ bắt cặp một phần với phân tử kia thì sự lai hóa được xem là không đặc hiệu hay không hoàn toàn. Phân tử lai đặc hiệu bền hơn và khó biến tính hơn phân tử lai không đặc hiệu.



Hình 9.3. Lai hóa và biến tính

9.2.2.1. Các yếu tố ảnh hưởng

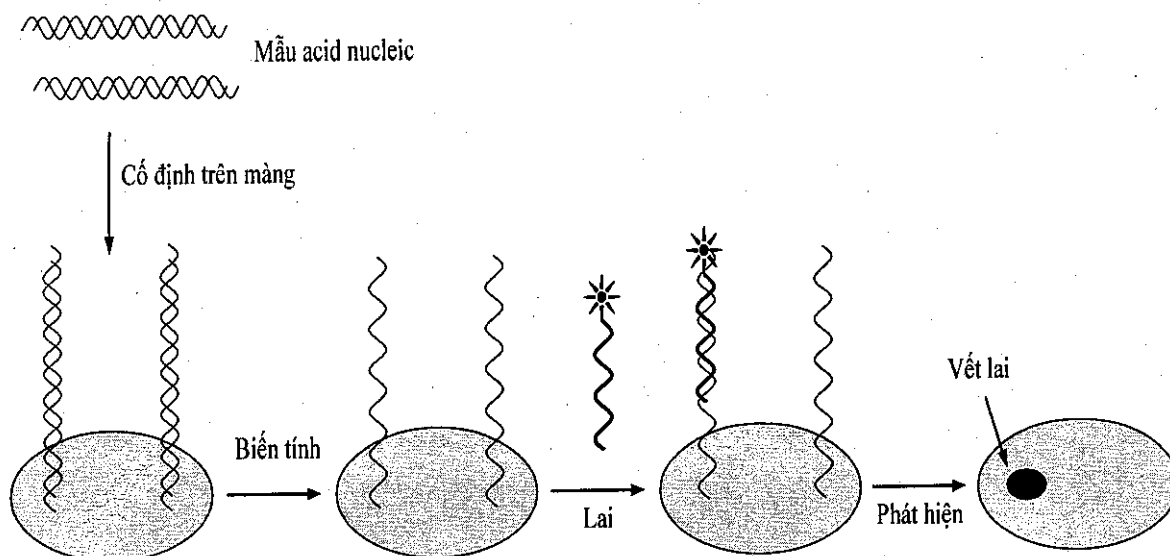
Do bản chất của phản ứng lai là sự hình thành các liên kết hydro nên cân bằng phản ứng, độ đặc hiệu và độ bền của duplex phụ thuộc vào các yếu tố sau:

- Nhiệt độ: là yếu tố quyết định độ bền của liên kết hydro.
- Độ dài các trình tự lai: độ dài duplex càng lớn thì càng bền và thời gian phản ứng càng lâu.
- Tỷ lệ GC trong duplex: tỷ lệ GC càng cao thì duplex càng bền.
- Nồng độ ion: ảnh hưởng đến độ bền của liên kết hydro.
- Nồng độ các phân tử và thời gian phản ứng.

Nhiệt độ chảy (melting temperature, T_m): là nhiệt độ ở đó có 50% số duplex được hình thành. Nhiệt độ này phụ thuộc vào chiều dài và tỷ lệ GC của duplex, ngoài ra nồng độ ion cũng có ảnh hưởng lớn, do đó cách tính T_m khá phức tạp và thường phải xác nhận bằng thực nghiệm. Phản ứng lai sẽ diễn ra ở nhiệt độ nhỏ hơn T_m của duplex, ngược lại ở nhiệt độ lớn hơn T_m , sự biến tính của duplex sẽ chiếm ưu thế. Công thức tính T_m đơn giản nhất như sau: $T_m (^{\circ}\text{C}) = 4(\text{GC}) + 2(\text{AT})$.

9.2.2.2. Các kỹ thuật lai và ứng dụng

Trong ứng dụng, phản ứng lai diễn ra giữa một phân tử đích (phân tử cần được phát hiện) và một đoạn dò (probe). Đoạn dò là phân tử oligonucleotid có trình tự bổ sung với trình tự của phân tử đích, được đánh dấu bằng phóng xạ, huỳnh quang, enzym hay các kỹ thuật đánh dấu khác nhờ đó ta phát hiện được sự tồn tại của phân tử lai hay phản ứng lai có diễn ra hay không, nói cách khác phân tử đích có tồn tại hay không.



Hình 9.4. Kỹ thuật lai để phát hiện acid nucleic

Trong sinh học phân tử, kỹ thuật lai được ứng dụng để phát hiện sự tồn tại của acid nucleic đích, nghiên cứu sự biểu hiện gen. Trong y học nó được ứng dụng trong chẩn đoán phát hiện bệnh dựa trên acid nucleic và nghiên cứu bệnh học. Để đảm bảo tính chính xác của kết quả phát hiện, người ta phải tiến hành phản ứng lai trong các điều kiện sao cho sự lai diễn ra một cách đặc hiệu, nghĩa là có sự bắt cặp đặc hiệu hoàn toàn giữa đoạn dò và phân tử đích. Khi đó duplex sẽ bền và được phát hiện. Ngược lại, nếu đoạn dò chỉ bắt cặp không hoàn toàn với phân tử đích, duplex sẽ không bền và bị biến tính, do đó không tạo tín hiệu lai.

Về mặt kỹ thuật phản ứng lai có thể diễn ra trên màng rắn, trong dung dịch hay lai tại chỗ (*in situ* hybridization).

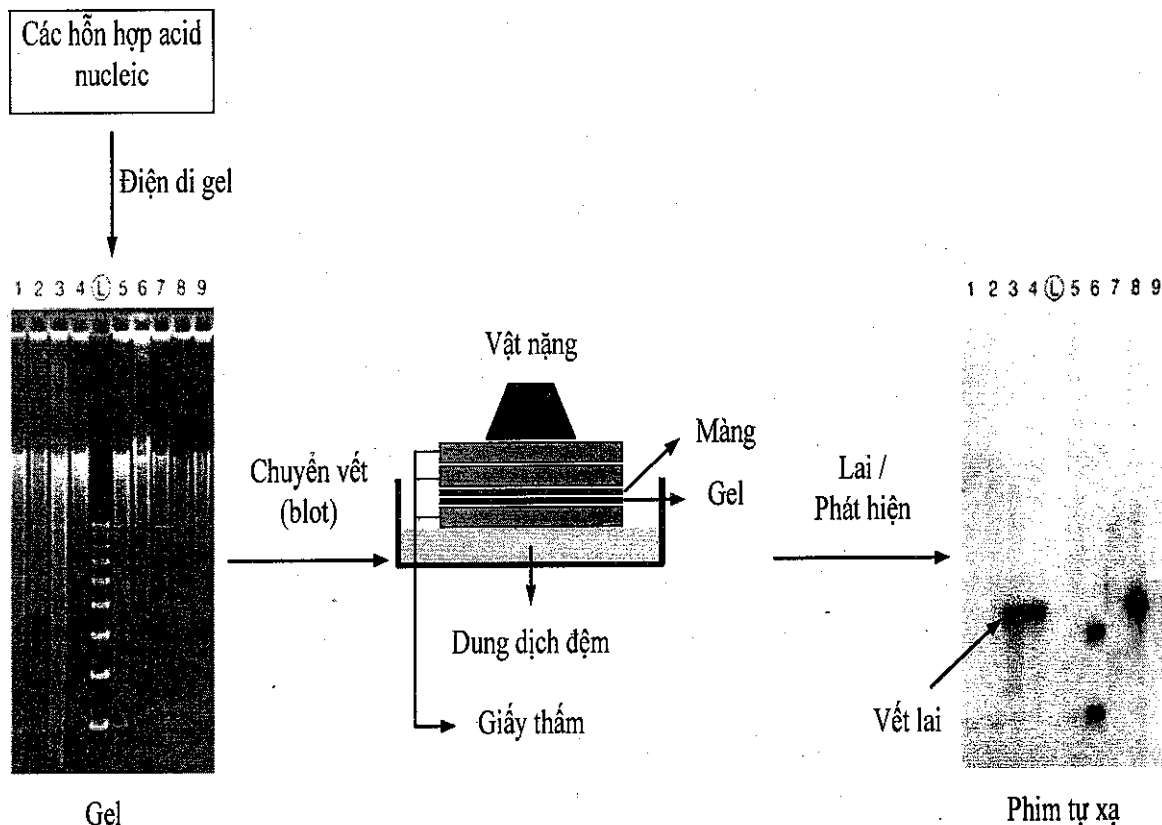
Lai trên màng rắn: Một trong các phân tử tham gia được cố định trên chất mang rắn, phân tử còn lại tự do trong dung dịch, khi sự lai diễn ra, duplex sẽ được cố định trên chất mang. Lai trên màng rắn là phương pháp phát hiện acid nucleic trong kỹ thuật blot, trong đó màng nitrocellulose thường được sử dụng. Các biến thể của lai trên màng:

– Southern blot: lai được thực hiện với đích là ADN sau khi đã được điện di trong gel rồi được chuyển và cố định trên màng.

– Northern blot: lai được thực hiện với đích là ARN sau khi đã được điện di trong gel rồi được chuyển và cố định trên màng.

– Dot blot: lai được thực hiện với hỗn hợp acid nucleic đích được chấm trực tiếp lên màng.

Southern và northern blot cho phép phát hiện và xác định kích thước của acid nucleic đích, trong khi dot blot chỉ phát hiện. Độ đậm của vết lai tương quan với lượng acid nucleic đích.



Hình 9.5. Quy trình Southern / Northern blot

Lai tại chỗ: Là kỹ thuật lai dùng để định vị acid nucleic đích trong mô hay tế bào.

Lai trong dung dịch: Phản ứng được tiến hành giữa phân tử đích và đoạn dò tự do trong dung dịch, duplex sau đó được tách ra khỏi dung dịch để phát hiện. Kỹ thuật này chủ yếu áp dụng trong chẩn đoán phát hiện dựa trên acid nucleic.

9.2.3. Kỹ thuật PCR (polymerase chain reaction)

9.2.3.1. Nguyên lý

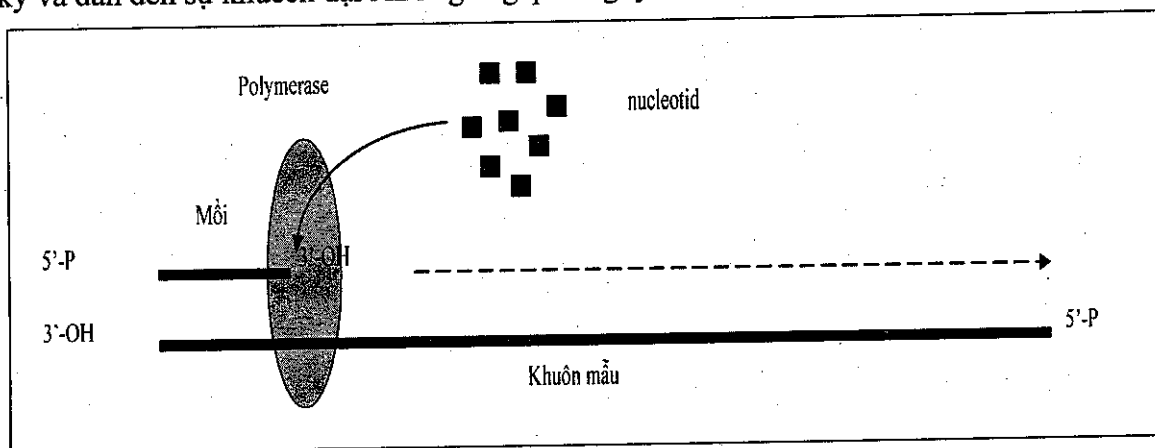
Kỹ thuật PCR dựa trên đặc điểm của quá trình sao chép gen trong tế bào: khi có sự hiện diện của các nucleotid, enzym ADN polymerase sẽ tổng hợp mạch ADN bổ sung với mạch kia nếu đầu 3'-OH còn trống (xem Hình 9.6). Kỹ thuật này do Karl Mullis và cộng sự phát minh năm 1985.

Kỹ thuật PCR bao gồm sự lặp lại của một tiến trình gồm 3 bước (xem Hình 9.6 và Hình 9.7):

1. **Biến tính:** Biến tính ADN sợi đôi thành các sợi đơn khuôn mẫu.
2. **Gắn mồi:** Ủ ADN sợi đơn với các đoạn mồi (oligonucleotid gắn đặc hiệu với ADN khuôn mẫu), sao cho một mồi gắn với sợi 5' (mồi xuôi) và mồi kia gắn với sợi 3' (mồi ngược).

3. Kéo dài: Enzym polymerase kéo dài các đoạn mồi theo nguyên tắc bổ sung với ADN sợi đơn khuôn mẫu tạo thành 2 sợi đôi mới.

Tiến trình trên được lặp lại nhiều lần. Từ chu kỳ thứ 2, đoạn ADN gốc và các sợi ADN mới tổng hợp sẽ là khuôn mẫu. Như vậy, sau mỗi chu kỳ PCR số phân tử khuôn mẫu sẽ tăng gấp đôi, tạo thành hiệu ứng khuếch đại. Một phản ứng điển hình có từ 20 đến 40 chu kỳ và dẫn đến sự khuếch đại ADN gốc gấp hàng tỷ lần.



Hình 9.6. Quá trình sao chép ADN ở tế bào

9.2.3.2. Các yếu tố kỹ thuật

Chu kỳ nhiệt:

Do có sự thay đổi nhiệt độ lặp lại giữa các bước, phản ứng được tiến hành trong máy luân nhiệt tự động, có khả năng thay đổi nhiệt độ một cách nhanh chóng theo chương trình. Một chu kỳ nhiệt PCR điển hình như sau:

– Biến tính khởi đầu: 95°C trong 5-10 phút (ở nhiệt độ này tất cả các ADN sợi đôi, bất kể chiều dài, đều bị biến tính).

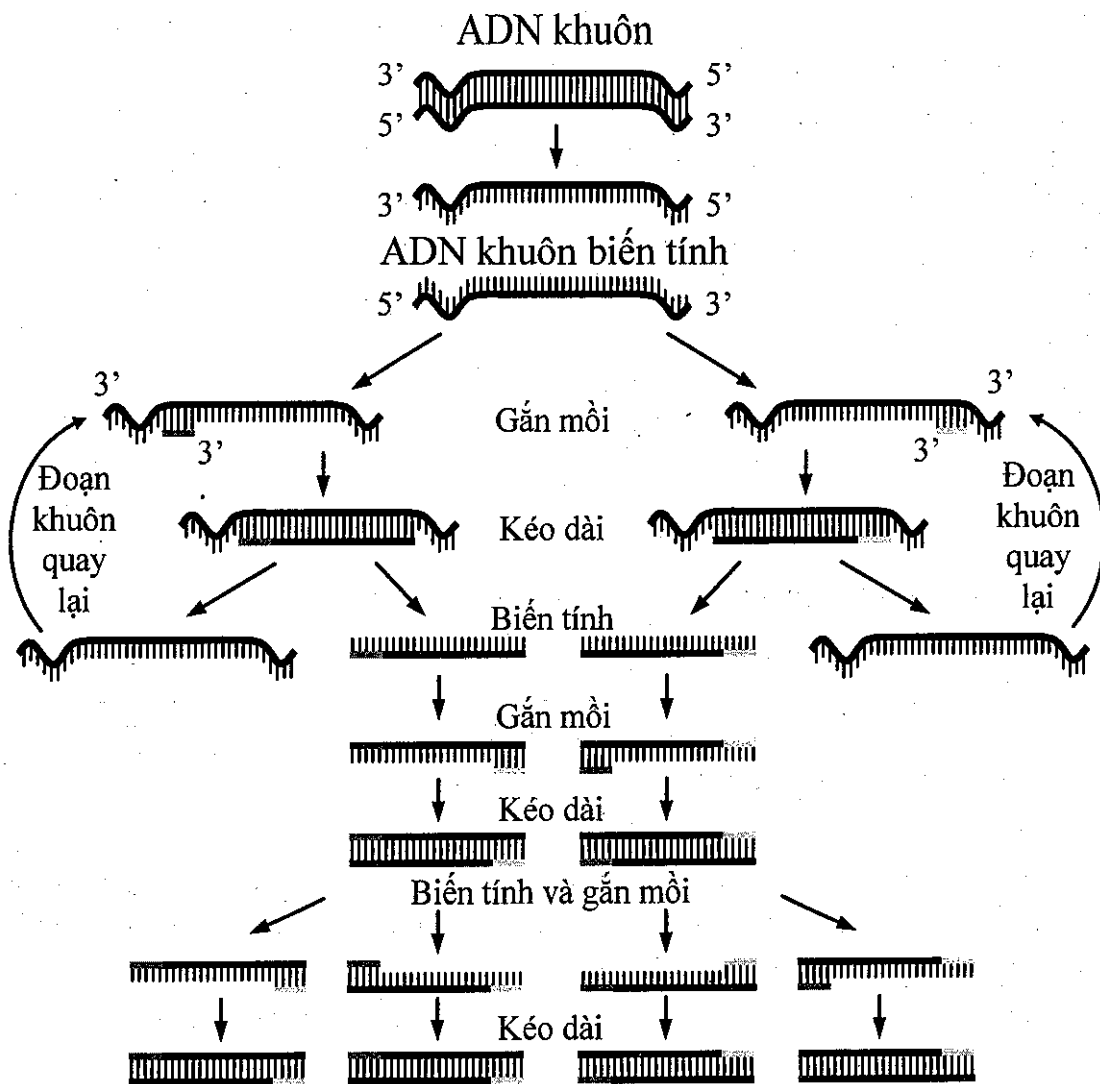
– Chu kỳ nhiệt: gồm 3 bước lặp lại 20-40 lần.

+ Biến tính: 94-95°C trong 30 giây đến vài phút.

+ Gắn mồi: 40-70°C (< T_m của mồi) trong 30 - 60 giây.

+ Kéo dài: được thực hiện ở nhiệt độ gần với nhiệt độ tối thích của enzym (thường là 65-74°C đối với enzym polymerase chịu nhiệt), thời gian tùy theo độ dài khuôn (thường 1 phút cho mỗi kb).

– Kéo dài kết thúc: ở nhiệt độ tối thích của enzym (thường là 72-74°C đối với enzym chịu nhiệt) trong thời gian gấp vài lần thời gian kéo dài.



Tiếp tục biến tính, gắn mồi và kéo dài trong 20-40 chu kỳ

Hình 9.7. Nguyên lý PCR

Thành phần phản ứng:

Phản ứng với thể tích 20-100 μl được tiến hành trong ống nhựa thành mỏng với thành phần phản ứng như sau:

- Khuôn mẫu: 1 pg - 1 μg , nồng độ cao sẽ ức chế phản ứng.
- Mồi: 0,1 - 1 μM , nồng độ quá cao sẽ dẫn đến bắt cặp sai và làm giảm tính đặc hiệu.
- Đệm: cung cấp pH thích hợp cho phản ứng và có nồng độ Mg^{2+} thay đổi (cần tối ưu hóa), thường trong khoảng 0,5-5 mM.
- Hỗn hợp đồng lượng các dNTP: A, T, G và C, nồng độ 200-250 μM .

- ADN polymerase chịu nhiệt: 0,5 - 2,5 đơn vị cho mỗi phản ứng.
- Nước vừa đủ.

Mồi PCR:

Gồm mồi xuôi và mồi ngược, là các oligonucleotid, được thiết kế theo các nguyên tắc:

- Chiều dài 18-30 (tối thiểu 15, nhưng có thể dài hơn 30) nucleotid, có trình tự bổ sung đặc hiệu với khuôn mẫu.
- Vị trí bắt cặp trên khuôn mẫu của hai mồi sẽ quyết định vùng ADN được khuếch đại trong PCR. Khoảng cách các vị trí bắt cặp trên khuôn mẫu giữa hai mồi không quá 10 kb (tốt nhất là < 3kb, nhưng một số polymerase đặc biệt có thể kéo dài đoạn ADN đến 40 kb).
- Tỷ lệ GC chiếm 40-60%. Các nucleotid AT và GC được phân bố đều, tránh dồn vào một chỗ. Đầu 3' của mồi nên có các nucleotid G hoặc C. Trình tự của mồi được lựa chọn sao cho không tạo cấu trúc kẹp tóc và hai mồi không bắt cặp với nhau.
- Nhiệt độ chảy của hai mồi phải gần nhau và trong khoảng 40-65°C, nhưng tốt nhất là 52-58°C.

Cách tính nhiệt độ gắn mồi:

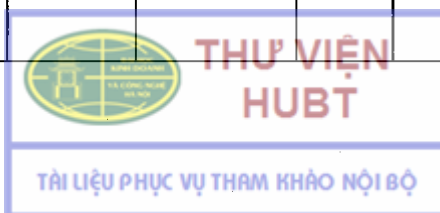
Nhiệt độ gắn mồi thường thấp hơn nhiệt độ chảy (melting temperature-T_m) khoảng 3-5°C. Nhiệt độ chảy phụ thuộc chiều dài, tỷ lệ GC và điều kiện môi trường như nồng độ muối, nồng độ ADN và được tính toán bằng nhiều cách khác nhau, trong đó đơn giản nhất là công thức sau: $T_m (^{\circ}C) = 4(GC) + 2(AT)$. Trên thực tế nhiệt độ gắn mồi cần được tối ưu hóa bằng thực nghiệm với xuất phát điểm là T_m lý thuyết của mồi.

Polymerase chịu nhiệt:

Do phản ứng sử dụng nhiệt độ cao để biến tính ADN khuôn mẫu, PCR ngày nay được tiến hành với các ADN polymerase chịu nhiệt, có nguồn gốc từ các vi sinh vật ưa nhiệt, ví dụ *Taq* polymerase của vi khuẩn *Thermus aquaticus*. Các polymerase khác nhau có độ bền nhiệt, tốc độ xử lý, khả năng xử lý ADN dài và tần suất lỗi khác nhau và sản phẩm có thể có đầu tù hay so le.

Bảng 9.2. Đặc tính một số polymerase chịu nhiệt

Enzym	Nguồn gốc	Nhiệt độ tối ưu (°C)	Hoạt tính exonuclease	Độ chính xác	Độ bền nhiệt (phút / °C)	Ghi chú
<i>Taq</i>	<i>T. aquaticus</i>	75-80	5'-3'	Thấp	9 / 97,5	Được sử dụng phổ biến nhất
<i>rTth</i>	<i>T. thermophilus</i>	75-80	5'-3'	Thấp	20 / 95	Có cả hoạt tính reverse transcriptase



<i>Tfl</i>	<i>T. flavus</i>	70	Không	Thấp	120 / 70	Dùng trong khuếch đại ADN dài
<i>Pwo</i>	<i>Pyrococcus woesei</i>	60-65	3'-5'	Cao	> 2 h / 100	Sản phẩm là ADN đầu tù
<i>Tli</i>	<i>Thermococcus litoralis</i>	70-80	3'-5'	Thấp	1,6 h / 100	Sản phẩm là ADN đầu tù
<i>Vent</i>	<i>Thermococcus litoralis</i>	70-80	3'-5'	Cao	6,7 h / 95	Sản phẩm là ADN đầu tù
<i>Vent exo-</i>	Tái tổ hợp	70-80	Không	Thấp	6,7 h / 95	Dùng trong giải trình tự
<i>DeepVent</i>	Tái tổ hợp	70-80	3'-5'	Cao	23 h / 95	Có độ chính xác và độ bền cao Sản phẩm là ADN đầu tù
<i>Pfu</i>	<i>Pyrococcus furiosus</i>	72-78	3'-5'	Cao	240 / 95	Có độ chính xác và độ bền cao
<i>UITma</i>	<i>Thermotoga maritima</i>	75-80	3'-5'	Cao	50 / 95	

9.2.3.3. Các hạn chế của kỹ thuật PCR

PCR có một số nhược điểm: Phản ứng dễ bị ảnh hưởng bởi sự nhiễm tạp của các đoạn ADN ngoại lai và chúng có thể được khuếch đại cùng với mẫu. Điều quan trọng là các ADN lạ này có thể được mang sang từ các sự khuếch đại trước đó (amplicon) hoặc từ các nguồn khác. Để loại bỏ tình trạng này, quy trình thực hiện phải được tuân thủ chặt chẽ, kiểm tra chất lượng enzym một cách nghiêm ngặt và sử dụng bộ pipet riêng và các thành phần phản ứng nên được phân liều trước, bên cạnh đó khu vực chiết tách ADN, thực hiện phản ứng và phát hiện sản phẩm nên tách biệt và thường xuyên được chiếu UV để làm giảm acid nucleic trong không khí.

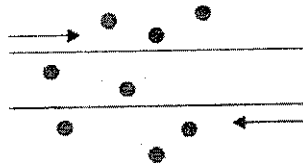
PCR đòi hỏi nhiều trang thiết bị thí nghiệm đặc biệt. Mặc dù, PCR là bán tự động, kỹ thuật vẫn đòi hỏi một lượng lao động đáng kể để chuẩn bị mẫu và phát hiện sản phẩm ADN sau đó.

1. SYBR Green I

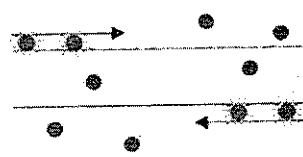
2. Phân hủy đoạn dò

3. Lai hóa đoạn dò

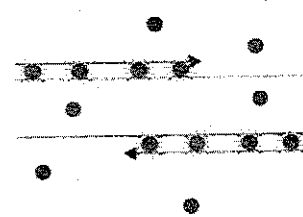
Gắn mồi / đoạn dò



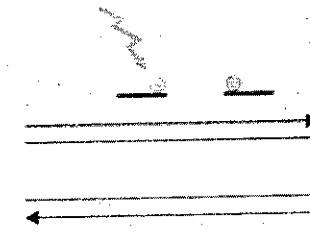
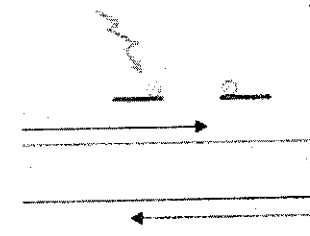
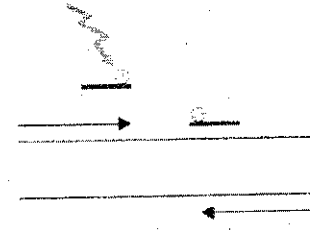
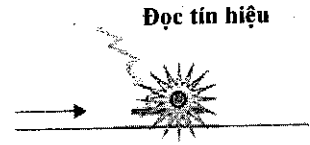
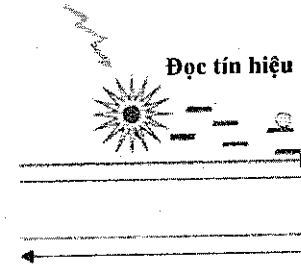
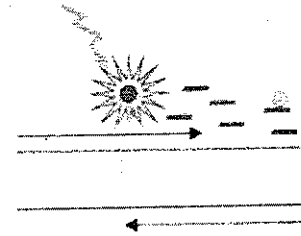
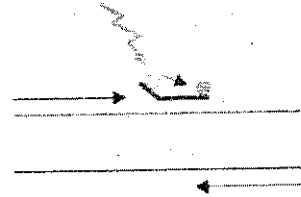
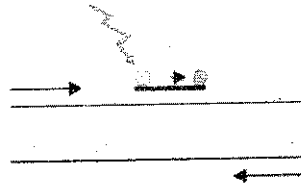
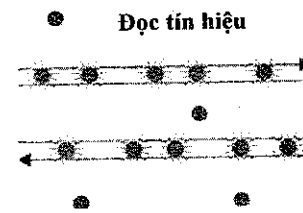
Pha kéo dài 1



Pha kéo dài 2



Kết thúc chu kỳ PCR



Hình 9.8. Các nguyên lý real time PCR

9.2.3.4. Các cải tiến

Hiện thời có nhiều kỹ thuật PCR khác nhau như PCR sử dụng các đoạn mồi tổ (nested), PCR ngược (reverse transcriptase, RT-PCR), PCR nghịch đảo (inverse), PCR neo (anchored)... Một tiến bộ quan trọng gần đây trong PCR là real time PCR cho phép phát hiện và định lượng sản phẩm PCR bằng huỳnh quang theo thời gian thực trong tube kín. Realtime PCR được ứng dụng nhiều trong chẩn đoán, trong sinh học phân tử kỹ thuật này được sử dụng để nghiên cứu sự biểu hiện gen. Việc phát hiện sản phẩm được thực hiện nhờ một trong các nguyên tắc sau (xem Hình 9.8):

1. Nhuộm xen giữa: Phản ứng PCR được bổ sung chất nhuộm xen giữa như SYBR Green I, chất này nếu tự do trong dung dịch thì không phát huỳnh quang, nhưng nếu được chèn vào sợi đôi ADN, ví dụ sản phẩm PCR, thì sẽ phát huỳnh quang. Tín hiệu được đọc khi chu kỳ PCR kết thúc.

2. Phân hủy đoạn dò: 1 đoạn dò đặc hiệu với trình tự sản phẩm được thiết kế sao cho 1 đầu được gắn phân tử phát huỳnh quang, 1 đầu gắn phân tử dập, do đó khi còn nguyên vẹn hoặc gắn trên trình tự đích thì không thể phát huỳnh quang được. Nhưng trong pha kéo dài của PCR, *Taq* polymerase sẽ sử dụng hoạt tính 5' exonuclease để phân hủy đoạn dò, dẫn đến phân tử phát huỳnh quang và phân tử dập không còn ở gần nhau nữa nên huỳnh quang sẽ được giải phóng. Tín hiệu được đọc khi chu kỳ PCR kết thúc.

3. Lai hóa đoạn dò: hai đoạn dò có gắn phân tử phát huỳnh quang được thiết kế sao cho đều gắn đặc hiệu với trình tự đích (sản phẩm PCR) theo kiểu đuôi-đầu. Do đó, khi cùng lai hóa trên trình tự đích huỳnh quang của phân tử thứ nhất sẽ kích thích sự phát quang của phân tử thứ 2 tạo ra tín hiệu, nhưng trong pha kéo dài, *Taq* polymerase sẽ tách 2 đoạn dò ra thành dạng tự do nên tín hiệu sẽ mất đi. Tín hiệu được đọc khi bắt đầu chu kỳ PCR sau bước gắn mới.

9.2.3.5. Ứng dụng của PCR

Kỹ thuật PCR cho phép sao chép một cách đặc hiệu đoạn ADN được giới hạn giữa hai môi lên hàng tỷ lần, do đó ứng dụng trực tiếp của nó là để thu lượng lớn đoạn ADN đích để dùng trong các phân tử và nghiên cứu sinh học phân tử, ví dụ để làm ADN đích trong tạo dòng gen.

Bằng cách sử dụng đoạn môi chứa nucleotid đột biến so với trình tự nguyên thủy, PCR là một công cụ hữu hiệu để tạo các đột biến điểm định hướng chính xác trong trình tự đích.

Tuy nhiên, ứng dụng được biết đến nhiều nhất của nó là để phát hiện ADN trong các phương pháp chẩn đoán dựa trên acid nucleic. Bằng cách thiết kế các đoạn môi đặc hiệu cho gen cần phát hiện, nó cho phép chẩn đoán (vi sinh vật hay gen gây bệnh) với độ nhạy và độ đặc hiệu cao.

Bên cạnh đó, PCR cũng được ứng dụng trong xác định quan hệ huyết thống hay trong pháp y.

9.2.4. Kỹ thuật tái tổ hợp ADN và tạo dòng gen

Để cô lập đoạn gen ở dạng tinh khiết, với số lượng lớn và nghiên cứu cấu trúc, chức năng của nó ta cần đưa đoạn gen đó vào một tế bào chủ sao cho nó có thể nhân bản được - nghĩa là tạo dòng gen. Điều này hoàn toàn khác với kỹ thuật PCR - khuếch đại đoạn gen *in vitro* tạo ra số lượng bản sao lớn của gen với độ tinh khiết đủ để thao tác - đòi hỏi đoạn gen đã được nghiên cứu xong về cấu trúc (trình tự), do vậy phải tạo dòng gen trước.

Bằng kỹ thuật tạo dòng gen, hiện nay người ta đã xây dựng các thư viện hay ngân hàng gen của nhiều loại sinh vật - với thành tựu nổi bật nhất là giải mã bộ gen người trong dự án Giải mã Bộ gen người (Human Genome Project) - trong đó mỗi gen nằm trong một dòng tế bào chủ.

Kỹ thuật tạo dòng gen gồm bốn phần cơ bản: tạo đoạn ADN, nối đoạn ADN cần tạo dòng vào vector để tạo phân tử tái tổ hợp, đưa ADN tái tổ hợp vào tế bào chủ và chọn lọc dòng tái tổ hợp.

9.2.4.1. Tạo đoạn gen mong muốn

Dùng RE:

Một trong những cách để tạo ra các đoạn ADN cho việc tạo dòng là cắt phân tử ADN lớn (vd: ADN nhiễm sắc thể) thành các đoạn nhỏ bằng RE. Điều này thường được sử dụng khi xây dựng các thư viện gen (genomic library).

Việc cắt được thực hiện tương đối đơn giản, tuy nhiên, việc phân tích và tìm ra đoạn gen thích hợp sau đó khá phức tạp nhất là đối với các bộ gen lớn vì số đoạn tạo ra lên đến hàng trăm nghìn. Để cắt các bộ gen người ta thường sử dụng 2 loại RE khác nhau để các đoạn tạo ra có kích thước vừa phải cho việc tạo dòng. Đồng thời với việc cắt bằng 2 RE khác nhau ta có thể phân tích các đoạn thu được và lập bản đồ các vị trí cắt của bộ gen.

RE được chọn để cắt cần lưu ý vị trí cắt sẽ tạo ra đầu tù (blunt end) hay đầu dính (sticky end) để sau đó khi duỗi thẳng vector ta cũng phải chọn RE tạo loại đầu tương tự thì việc nối đoạn gen với vector mới thực hiện được dễ dàng đặc biệt đối với các đầu dính.

Nếu ADN nguyên thủy có các điểm nhận diện duy nhất biết trước của RE (ví dụ trong trường hợp kiểm tra kết quả của việc tạo dòng) thì ta dễ dàng dùng các RE thích hợp để cắt lấy đoạn ADN mong muốn.

Kỹ thuật PCR:

Nếu cấu trúc đoạn gen đã được biết, ta dễ dàng thiết kế cặp mồi đặc hiệu để thực hiện phản ứng khuếch đại đoạn gen đó dùng cho việc tạo dòng.

Thông thường đoạn ADN thu được bằng PCR có các đầu bị so le (lệch) không xác định về trình tự. Do vậy, trước khi nối người ta phải làm tà đầu sản phẩm PCR và sau đó có thể gắn linker hoặc adaptor nếu cần thiết để có các đầu thích hợp cho việc nối với vector trong bước tiếp theo. Một cách khác để khắc phục vấn đề này là thiết kế mồi có chứa trình tự nhận diện cắt của RE, sản phẩm PCR thu được sẽ mang trình tự này và được cắt với RE tương ứng để tạo đầu cắt thích hợp cho việc nối vào vector.

9.2.4.2. Nối ADN

Vector tạo dòng cũng được cắt để duỗi thẳng bằng RE thích hợp (thường là cùng loại với RE dùng để cắt đoạn gen).



Để nối các ADN trong thực nghiệm người ta thường dùng ligase của T₄ (được tinh chế từ *E. coli* bị nhiễm phage T₄). Khi trộn vector đã đuôi thẳng và đoạn ADN với ligase trong điều kiện thích hợp, phản ứng diễn ra tạo liên kết phosphodiester và nối các ADN lại. Trong trường hợp tạo vòng, ADN tái tổ hợp phải được nối hoàn chỉnh với các yếu tố của vector để sẵn sàng được đưa vào và hoạt động trong tế bào.

Mặc dù ngày nay kỹ thuật nối (ligation) đã tiến bộ và cho phép nối các ADN có đầu tù tương đối dễ dàng, nhưng nếu lượng ADN thu được quá nhỏ thì việc nối bằng đầu dính sẽ dễ thực hiện hơn. Mặt khác, việc nối các đầu dính sẽ đặc hiệu hơn. Nhưng không phải lúc nào cũng chọn được enzym thích hợp cho việc cắt đầu dính hoặc các sản phẩm PCR thường có đầu không xác định được. Trong điều kiện như vậy ADN cần được biến đổi để có đầu dính.

Linker (đoạn nối):

Đây là một oligonucleotid sợi đôi tổng hợp ngắn có đầu tù và chứa trình tự cắt (đầu dính) của một RE. Ligase sẽ nối linker với đoạn ADN có đầu tù, việc này có thể thực hiện dễ dàng mặc dù là đầu tù vì ta có thể sử dụng một lượng lớn linker để nối. Sau đó dùng RE cắt linker để tạo ra đầu dính.

Adaptor:

Việc sử dụng linker có nhược điểm là nếu trong đoạn ADN cần tạo vòng cũng có vị trí cắt của RE giống linker thì trong bước tiếp theo RE sẽ cắt luôn đoạn ADN mong muốn và làm hỏng nó. Để khắc phục người ta sử dụng một adaptor - đây là một oligonucleotid sợi đôi tổng hợp có sẵn một đầu dính (giống với đầu tạo ra do RE cắt đầu dính). Ligase sẽ nối adaptor với đoạn ADN có đầu tù và tạo thành đoạn mới có đầu dính.

Cần lưu ý cấu trúc phía 5' của đầu dính trên adaptor bị thay nhóm phosphate bằng nhóm hydroxyl để tránh các adaptor tự nối với nhau. Do đó, sau khi nối adaptor xong ta phải dùng polynucleotid kinase để trả lại nhóm phosphate.

Tạo đuôi homopolymer:

Sử dụng enzym doxynucleotidyl transferase để thêm một loạt các nucleotid (cùng loại) vào đầu 3'-OH của sợi ADN tạo thành đuôi homopolymer. Trên vector đã đuôi thẳng cũng thực hiện như vậy nhưng với nucleotid bổ sung với các nucleotid trên sợi ADN. Kết quả cũng sẽ tạo ra 2 đầu dính dễ dàng nối với nhau bằng ligase.

9.2.4.3. Chuyển ADN vào tế bào chủ

Các thao tác cắt nối trên về thực chất chính là kỹ thuật tái tổ hợp ADN. Công việc tạo vòng gen nằm ở chỗ đưa được phân tử tái tổ hợp vào trong tế bào chủ. Tùy theo vector được sử dụng ta có các phương pháp đưa gen vào tế bào khác nhau.

Biến nạp:

Biến nạp (transformation) là sự đánh bắt ADN từ môi trường vào tế bào bởi vi khuẩn. Thông thường ADN sau khi đánh bắt sẽ bị phân hủy nhanh chóng trong tế bào hoặc bị pha loãng khi tế bào phân chia qua các thế hệ. Tuy nhiên, nếu ADN ngoại lai tích hợp được

trên nhiễm sắc thể hoặc có khả năng tái bản được (có Ori) thì nó có thể tồn tại được trong tế bào qua các thế hệ. Đó chính là trường hợp của tạo dòng gen.

Trong tự nhiên, chỉ một số ít sinh vật có khả năng đánh bắt ADN từ môi trường. Mặt khác, tần suất diễn ra rất thấp và mang tính ngẫu nhiên. Do đó, để biến nạp thành công các tế bào có khả năng cần được xử lý thêm để gia tăng tần suất. Các tế bào như vậy gọi là tế bào khả nạp (competent cell).

– **Chuẩn bị tế bào khả nạp:** Người ta nhận thấy nếu tế bào vi khuẩn được xử lý với CaCl_2 (thường sử dụng nồng độ 50 - 100 mM) thì khả năng đánh bắt ADN ngoại lai tăng lên rất nhiều. Cơ chế của hiện tượng này đến nay chưa rõ, có thể CaCl_2 làm ADN kết tủa trên mặt ngoài tế bào hoặc gây ra một biến đổi nào đó làm tăng khả năng kết dính của ADN trên tế bào. Tại thời điểm sau khi xử lý, ADN chỉ dính ở mặt ngoài tế bào chứ chưa vào trong. Tế bào khả nạp phải được bảo quản lạnh và khi sử dụng để ở nhiệt độ thấp (4°C).

– **Đưa ADN vào tế bào chất:** Sau khi trộn ADN tái tổ hợp với tế bào khả nạp, việc đánh bắt ADN thực sự diễn ra khi tế bào bị gây sốc. Cơ chế hiện nay vẫn chưa xác định, có thể khi đỡ màng tế bào bị thủng làm cho ADN lọt vào trong. Sốc nhiệt: môi trường tế bào được nâng nhiệt độ lên 42°C trong thời gian ngắn (khoảng 2 phút) rồi làm lạnh nhanh.

Chuyển nhiễm (transfection):

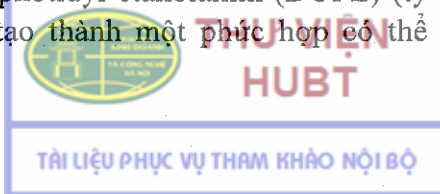
Là tập hợp một số phương pháp chuyển gen vào tế bào sinh vật nhân thật, đặc biệt là tế bào động vật có vú. Cơ chế chung cũng giống như biến nạp, tế bào được xử lý đặc biệt để có thể nhận gen ngoại lai khi bị sốc, nhưng khác với biến nạp các phương pháp này có thể áp dụng được cho các tế bào sinh vật không phải là vi khuẩn.

Chuyển nhiễm hiểu theo một nghĩa hẹp là sự chuyển gen bằng cách sử dụng phage gây nhiễm đối với sinh vật nhạy cảm.

– **Chuyển nhiễm nhờ phosphate calci:** Đây là một phương pháp tương đối đơn giản để đưa plasmid tái tổ hợp vào tế bào động vật có vú. Bằng phương pháp này có thể đưa cùng lúc 2 plasmid (co-transfection) vào tế bào. ADN được trộn trong kết tủa của phosphate calci và được tế bào đánh bắt bằng cơ chế không rõ. Khi thực hiện phương pháp này cần lưu ý độc tính với tế bào của kết tủa. Có thể sử dụng sốc glycerol (15%) để gia tăng khả năng đánh bắt ADN nếu tế bào chịu đựng được.

– **Chuyển nhiễm qua trong gian DEAE dextran:** Phương pháp này thích hợp cho việc tạo dòng không bền (transient transfection) hơn. Tế bào được treo trong Diethylaminoethyl (DEAE) dextran (phân tử lượng 500.000). Thời gian được xác định thực nghiệm tùy theo loại tế bào và có ảnh hưởng lớn đến hiệu quả.

– **Qua trung gian liposom:** Phương pháp này nhanh, đơn giản và hiệu quả, có thể chuyển đồng thời 2 plasmid vào tế bào với tần suất cao. Liposom được điều chế từ các chất béo cation như: N-(1-[2,3-di-o-leyloxy]propyl)-N,N,N-trimetylamonium cholrid (DOTMA) và dioleoyl phosphotidyl etanolamin (DOPE) (tỷ lệ 1:1). Các lipid này được cho tiếp xúc với ADN để tạo thành một phức hợp có thể được đánh bắt bởi tế bào.



Cần lưu ý ADN dùng cho phương pháp này phải thật tinh khiết thì hiệu quả mới cao. Và phải sử dụng vật chứa bằng polystyren thay vì polypropylen do liposom sẽ tạo phức không ổn định với chất sau.

– **Sử dụng Lipofectamine:** Tác nhân này là một dạng cationic lipid được điều chế thích hợp, được sử dụng nhiều trong thời gian gần đây vì cho hiệu quả chuyển nhiễm cao với nhiều loại tế bào và cách sử dụng rất đơn giản, chỉ cần trộn với ADN để tạo phức và cho vào môi trường nuôi cấy tế bào. Tác nhân này đặc biệt hiệu quả khi cần chuyển nhiễm ARNi vào tế bào.

Thâm điện (electroporation):

Đây là một phương pháp rất hiệu quả để đưa ADN vào tế bào sinh vật gồm cả tế bào động vật có vú, thực vật, nấm men và vi khuẩn. Trong phương pháp này, một xung điện cao thế (500 - 1500 V/cm) được truyền qua tế bào được treo giữa hai điện cực trong một thời gian cực ngắn (2 - 20 msec). Xung điện có lẽ làm màng tế bào bị thủng vào để cho ADN lọt vào.

Nhiệt tạo ra, sự thay đổi tính thấm của màng tế bào chất, và các hiệu ứng khác có thể làm tổn thương tế bào, thậm chí chết. Do vậy, điều kiện tiến hành phải được xác định bằng thực nghiệm đối với các loại tế bào khác nhau. ADN dùng cho phương pháp này phải thật tinh khiết thì hiệu quả mới cao.

Đạn sinh học:

Hay “súng bắn gen” để chuyển gen vào từng tế bào đơn lẻ bằng cách sử dụng ADN được hấp phụ trên hạt vàng (đường kính 0,6 – 2 μm). Để chuyển nhiễm các hạt này được đặt trong một dụng cụ gọi là súng bắn gen và được phóng ra nhờ helium áp suất cao xuyên qua da vào tế bào. Hiệu suất chuyển cao hơn hàng trăm lần so với tiêm bắp vì gen được chuyển thẳng vào tế bào nên không bị phân hủy.

Liposom:

Liposom cấu tạo là một màng kép phospholipid bao quanh một hạt nước, nó có thể chứa nhiều loại chất, kể cả ADN plasmid. Liposom cũng có thể được gắn kháng thể, protein hoặc đường lên bề mặt và nhờ đó định hướng đến vai trò đặc hiệu. Khi tiếp xúc với tế bào các màng phospholipid sẽ dung hợp và phóng thích nội dung của liposom vào tế bào.

Lipoplex và polyplex:

Là các phức hợp ADN với một phân tử lipid hay polyme tích điện dương. Sau khi kết hợp ADN, các hạt này tích điện dương và được thực bào. Lợi thế của hệ thống này là:

- Bảo vệ plasmid, giảm lượng ADN cần;
- Khi tương tác với tế bào bằng cơ chế thực bào, các hạt này có thể kích thích đáp ứng miễn dịch tốt hơn và do đó thích hợp cho ADN vaccin;
- Có thể sử dụng qua nhiều đường như miệng, mũi và phở;
- Dễ dùng;
- Ổn định khi bảo quản dưới dạng đông khô.

9.2.4.4. Chọn lọc dòng tái tổ hợp

Thông thường, việc nối ADN vào vector không bao giờ thành công hoàn toàn và đặc hiệu nên hỗn hợp nối sẽ có nhiều kết quả nối khác nhau (ví dụ: ADN nối lẫn nhau, vector tự đóng vòng mà không mang đoạn chèn, ...). Do đó, sau khi đưa ADN vào tế bào, tiến hành nuôi cấy và phân lập các dòng thu được để phát hiện dòng mang ADN tái tổ hợp.

Việc chọn lọc dòng tái tổ hợp dựa vào dấu hiệu chọn lọc (selectable marker) của vector, vị trí chèn và tế bào chủ. Trên các vector thường mang các gen đề kháng kháng sinh, do đó môi trường phân lập thường chứa kháng sinh tương ứng và chỉ dòng tế bào nào nhận được vector (có thể tái tổ hợp hoặc không) mới mọc được. Trong số các khóm thu được (có mang vector), dựa vào vị trí chèn và kỹ thuật đánh dấu sử dụng để phát hiện dòng mang vector tái tổ hợp. Nhìn chung có một số phương pháp chính như sau:

– Sử dụng kỹ thuật bổ sung alpha: dựa trên sự hiện diện của hoạt tính β -galactosidase (chịu trách nhiệm thủy phân lactose thành galactose). Enzym này là một tetramer, trong đó mỗi monomer có 2 mảnh (alpha và omega). Bộ gen của tế bào chủ bị đột biến mất hoặc bất hoạt phần mã hóa cho mảnh alpha, do đó chỉ sản xuất được mảnh omega của enzym, không đủ để có hoạt tính (kiểu hình *lac*⁻, không thủy phân được lactose). Trong khi đó, vector lại mang trình tự mã hóa cho đoạn alpha và vị trí chèn gen ngoại lai sẽ nằm trong trình tự này. Do đó, nếu tế bào nhận vector mang chưa tái tổ hợp thì sẽ được bổ sung mảnh alpha và có kiểu hình *lac*⁺, ngược lại nếu vector bị tái tổ hợp với gen lạ trong trình tự mã hóa mảnh alpha thì mảnh alpha sẽ bị bất hoạt và tế bào sẽ có kiểu hình *lac*⁻.

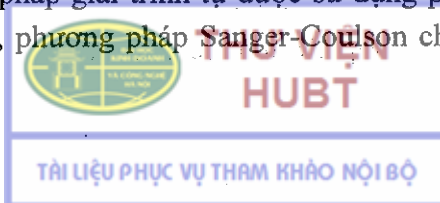
– Sử dụng sự đề kháng với kháng sinh: vector mang 2 gen kháng kháng sinh, vị trí chèn nằm trong 1 gen kháng kháng sinh. Nếu vector mang ADN tái tổ hợp dòng thu được sẽ mất khả năng đề kháng với 1 trong 2 kháng sinh. Do đó ta có thể sử dụng phương pháp in dấu để phát hiện.

– Chọn lọc trực tiếp: đây là cách tốt nhất nhưng khó thực hiện trên thực tế. Trong phương pháp này, đoạn chèn sẽ mang đến cho tế bào nhận được nó các đặc điểm kiểu hình mà các dòng khác không có và ta dễ dàng nhận ra khi phân lập. Ví dụ, đoạn chèn mang gen mã hóa sắc tố hay enzym.

– Lai khóm (colony hybridization): áp dụng cho trường hợp đoạn chèn đã biết trình tự. Các khóm sau khi phân lập được in dấu lên màng nitrocellulose hay nylon, tế bào trên màng bị phá hủy để bộ lộ vật liệu di truyền. Màng lọc lúc này được cho lai với đoạn dò được đánh dấu tương ứng của đoạn chèn. Nếu khóm nào mang đoạn chèn sẽ có tín hiệu, từ vị trí tín hiệu trên màng lọc suy ra khóm tương ứng trên hộp thạch ban đầu.

9.2.5. Xác định trình tự của acid nucleic

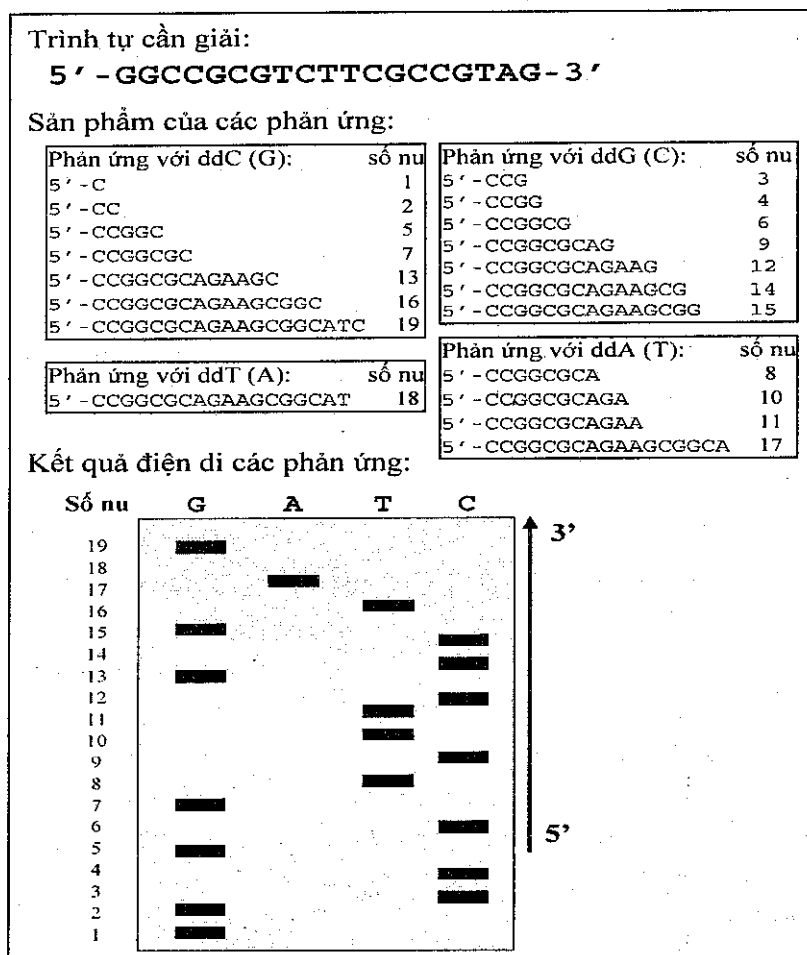
Hiện nay có hai phương pháp giải trình tự được sử dụng phổ biến là Sanger-Coulson và Maxam-Gilbert. Trong đó, phương pháp Sanger-Coulson cho hiệu quả cao, ít độc hại



hơn và là phương pháp được sử dụng trong hầu hết trường hợp, tuy nhiên cần phải biết trước một phần trình tự để thiết kế mỗi giải trình tự. Phương pháp Maxam-Gilbert cho hiệu quả kém, sử dụng nhiều hóa chất độc hại và chất phóng xạ nhưng cho phép giải *de novo* vì không cần môi.

9.2.5.1. Phương pháp Sanger-Coulson

Phương pháp này do các nhà hóa học người Anh là F. Sanger và A.R. Coulson phát minh năm 1977. Phương pháp này dựa trên sự tổng hợp chuỗi ADN bổ sung với chuỗi cần giải trình tự bằng enzym polymerase, trong đó khi nucleotid được thêm vào mạch là didesoxynucleotid thì sự tổng hợp bị chấm dứt (vì ở C₃-OH của đường desoxyribose bị lấy đi OH thay bằng H, khiến các didesoxynucleotid không còn khả năng hình thành nối phosphodiester với nucleotid tiếp theo). Do đó, phương pháp này còn gọi là phương pháp nucleotid ngừng chuỗi (chain termination) hay là phương pháp enzym. Phương pháp nguyên thủy được thực hiện trên ADN sợi đơn, tuy nhiên ngày nay nó được thực hiện với ADN sợi đôi nhờ biến tính với nhiệt tương tự PCR nhưng chỉ với 1 môi.



Hình 9.9. Giải trình tự bằng phương pháp Sanger



Quy trình:

Thực hiện phản ứng tổng hợp ADN bằng polymerase với nguyên liệu là dNTP có bổ sung khoảng 1% các ddNTP khác nhau, do sự kéo dài mạch sẽ ngừng lại khi gặp ddNTP, sản phẩm sẽ là hỗn hợp các ADN bị kết thúc ở các nucleotid tương ứng với ddNTP của phản ứng đó. Cụ thể:

– Thực hiện 4 phản ứng tổng hợp ADN riêng biệt với thành phần gồm:

+ Phần chung (giống nhau) gồm:

* ADN cần giải trình tự (khuôn mẫu).

* Môi: là oligonucleotid bắt cặp bổ sung với phần đầu trình tự cần giải.

* Các thành phần khác của phản ứng tổng hợp ADN: đệm, ion Mg^{2+} , dNTP, polymerase.

+ 4 phản ứng khác nhau do có thêm 1% ddNTP khác nhau:

* T: có ddATP

* A: có ddTTP

* C: có ddGTP

* G: có ddCTP.

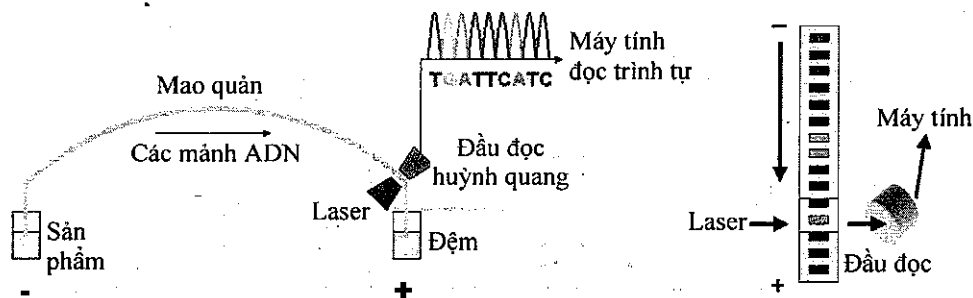
Mỗi hoặc dNTP hoặc ddNTP sẽ được đánh dấu bằng phóng xạ để hiện hình sản phẩm bằng tự xạ ký.

– Mỗi ống sản phẩm được nạp vào một giếng điện di.

– Kết quả điện di sẽ được hiện hình nhờ kỹ thuật tự xạ ký và được đọc thành trình tự ADN cần giải (xem Hình 9.9).

Giải trình tự tự động:

Gần đây, người ta cải tiến phương pháp Sanger-Coulson để có thể tự động hóa hoàn toàn (xem Hình 9.10) bằng cách gắn các dấu huỳnh quang khác nhau vào từng loại dideoxynucleotid. Ví dụ màu xanh cho Adenin, màu đỏ cho Cytosin. Do các chuỗi ngừng đối với từng loại ddNTP khác nhau được đánh dấu huỳnh quang với màu khác nhau, người ta không cần phải tách thành 4 ống phản ứng mà thực hiện tất cả trong cùng một ống. Sản phẩm được điện di bằng mao quản và thứ tự màu tương ứng với ddNTP sẽ được đọc trực tiếp thành thứ tự các base nhờ một đầu đọc laser đặt ở cuối mao quản.



Hình 9.10. Giải trình tự tự động



THƯ VIỆN
HUBT

9.2.5.2. Phương pháp Maxam và Gilbert

Phương pháp này dựa vào sự thủy giải hóa học tại các nucleotid đặc hiệu của phân tử ADN cần xác định trình tự, do đó còn gọi là phương pháp hóa học, do Allan Maxam và Water Gilbert (Mỹ) phát minh năm 1977.

Quy trình:

Việc giải trình tự bằng phương pháp hóa học được tiến hành qua các bước:

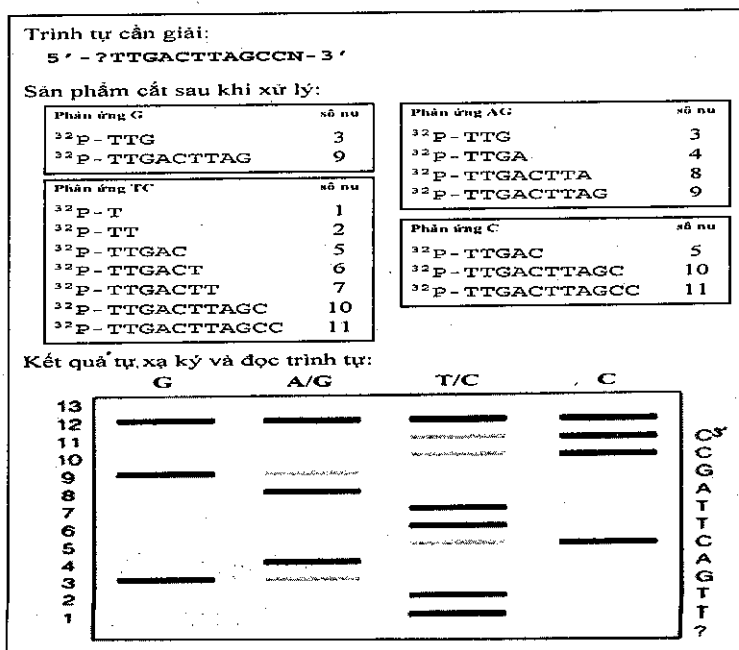
- Tạo ADN sợi đơn từ đoạn ADN cần giải trình tự.
- Đánh dấu các phân tử ADN ở một đầu bằng ^{32}P .
- Thực hiện 4 (hoặc 5) ống phản ứng riêng biệt bằng cách xử lý ADN với:
 - + G: Dimethyl sulphate làm cho Guanin bị methyl hóa
 - + AG: Formic acid làm cho Adenin và Guanin bị methyl hóa
 - + TC: Hydrazine gây biến đổi ở Thymin và Cytosin
 - + C: Hydrazine + 1,5 M NaCl gây biến đổi ở Cytosin

Xử lý 4 ống trên với piperidine 1M/90°C để cắt ADN ở liên kết phosphat kế nucleotid bị biến đổi. Ngoài ra, người ta có thể thực hiện thêm phản ứng:

+ A>C: NaOH 1,2N/90°C cắt ADN ở Adenin và một số ở Cytosin.

- Mỗi ống được nạp vào một giếng điện di và hiện kết quả bằng kỹ thuật tự xạ ký, trong đó chỉ mảnh ADN bị cắt được đánh dấu mới hiện hình.

- Kết quả tự xạ ký được đọc và biện luận thành trình tự (xem Hình 9.11).



Hình 9.11. Giải trình tự bằng phương pháp hóa học



CÂU HỎI:

1. Cách nào **không** dùng để tinh chế AND?
 - a) Sắc ký ái lực
 - b) Sắc ký lọc gel
 - c) Sắc ký trao đổi ion trên vi cột
 - d) Sắc ký lỏng hiệu năng cao
 - e) Sắc ký khí
2. Enzym cắt giới hạn loại được ứng dụng nhiều trong kỹ thuật tái tổ hợp di truyền:
 - a) I
 - b) II
 - c) III
 - d) I và III
 - e) II và III
3. Đoạn Kelenow exo - không có hoạt tính nào?
 - a) Exonuclease 3' - 5'
 - b) Exonuclease 5' - 3'
 - c) Polymerase 5' - 3'
 - d) Cả A và B
4. Phương pháp thu hồi AND bằng hấp phụ trên silica **không** có ưu điểm nào?
 - a) Hiệu quả với AND nhỏ
 - b) Dễ thực hiện
 - c) Nhanh
 - d) Sạch
5. Yếu tố ảnh hưởng đến sự lai hóa:
 - a) Nồng độ AND
 - b) Nhiệt độ và thời gian phản ứng
 - c) Độ dài của các trình tự
 - d) Lực ion
 - e) Tất cả
6. Môi trong phản ứng PCR:
 - a) Đoạn AND ngắn, mạch hơn
 - b) Có trình tự bổ sung với AND khuôn tại điểm đầu sao chép
 - c) Dài từ 6-30 nucleotid
 - d) Là oligonucleotid
 - e) Tất cả
7. Tính nhiệt độ "chảy" của đoạn môi nhằm xác định nhiệt độ thích hợp để:
 - a) Biến tính môi
 - b) Môi gắn vào khuôn
 - c) Tổng hợp từ khuôn trên gel
 - d) Môi không gắn bổ sung vào nhau
 - e) Môi gắn vào các đoạn khác nhau

8. Chọn chu kỳ nhiệt PCR dựa vào 2 yếu tố là kích thước của khuôn và:
- a) Độ tinh khiết của khuôn
 - b) Nồng độ của khuôn
 - c) Kích thước môi
 - d) Trình tự môi
 - e) c và d
9. Realtime PCR sử dụng SYBR Green I, đọc tín hiệu sau giai đoạn nào?
- a) Biến tính
 - b) Gắn môi
 - c) Kéo dài
 - d) Kết thúc phản ứng
 - e) Lúc nào cũng được
10. Đặc điểm **không** thuộc phương pháp định trình tự của Sanger:
- a) Xử lý hóa chuyên biệt làm biến đổi đặc trưng một loại nucleotid
 - b) Sử dụng AND polymerase
 - c) Nucleotid được đánh dấu
 - d) Phản ứng tiến hành trong bốn phân đoạn
 - e) Có sử dụng dideoxynucleotid.

ĐÁP ÁN

Bài 1. Nhập môn sinh học phân tử

1. b 2. e 3. c 4. e 5. a

Bài 2. Sao chép ADN

1. d 2. b 3. d 4. e 5. e
6. e 7. a 8. b 9. b 10. b

Bài 3. Các loại ARN

1. e 2. a 3. a 4. d 5. b
6. b 7. e 8. c 9. d 10. e

Bài 4. Sự phiên mã và mã di truyền

1. e 2. e 3. e 4. c 5. e
6. b 7. c 8. c 9. e 10. a

Bài 5. Sinh tổng hợp protein

1. e 2. d 3. a 4. a 5. b
6. d 7. e 8. e 9. d 10. c

Bài 6. Điều hoà hoạt động gen

1. e 2. e 3. d 4. b 5. a
6. e 7. d 8. e 9. b 10. b

Bài 7. Bộ gen tế bào nhân thật

1. b 2. d 3. e 4. e 5. e
6. e 7. e 8. e 9. e 10. e

Bài 8. Đột biến gen

1. c 2. e 3. e 4. d 5. d
6. e 7. e 8. d 9. c 10. a

Bài 9. Một số phương pháp và kỹ thuật cơ bản trong sinh học phân tử

1. e 2. c 3. b 4. e 5. a
6. e 7. e 8. b 9. e 10. e



TÀI LIỆU THAM KHẢO

1. Benjamin Lewin, *Genes VIII. Prentice – Hall*, 2004.
2. Hồ Huỳnh Thuỳ Dương, *Sinh học phân tử*. Nxb. Giáo Dục, 1997.
3. James Swarbrick, *Pharmaceutical Gene Delivery Systems*. Marcel Dekker Inc. New York, 2003.
4. Lê Đình Lương. *Nguyên lý kỹ thuật di truyền*. Nxb. Khoa học Kỹ thuật, 2001.
5. Phạm Thành Hồ. *Di truyền học*. Nxb. Giáo Dục, 1998.
6. Richard A. Harvey, Pamela C. Champe. *Biochemistry*. 3rd ed. Lippincott Williams and Wilkins, Baltimore, 2005.
7. Richter, J. D. in *Translational Control of Gene Expression* (Hershey, J. W. B. , Mathews, M. B. , and Sonenberg, N., eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2000.
8. Sambrook, J., Fritsch, E., và Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2 Ed., Vol. 3, Cold Spring Harbor Press, Cold Spring Harbor, New York, 1989.
9. Smith C. A. and Wood E. J. *Molecular biology and Biotechnology*. Chapman & Hall, 1991.
10. Tom Strachan and Andrew P. Read. *Human Molecular Genetics*. Bios Scientific Publishers Ltd, 1996.

Chịu trách nhiệm xuất bản:

Chủ tịch Hội đồng Thành viên kiêm Tổng Giám đốc NGƯT NGÔ TRẦN ÁI
Phó Tổng Giám đốc kiêm Tổng biên tập GS.TS VŨ VĂN HÙNG

Tổ chức bản thảo và chịu trách nhiệm nội dung:

Phó Tổng biên tập NGUYỄN VĂN TƯ
Giám đốc Công ty CP Sách ĐH-ĐN NGÔ THỊ THANH BÌNH

Biên tập nội dung:

TRẦN NGỌC OANH

Biên tập tái bản và sửa bản in:

VŨ BÁ SƠN

Trình bày bìa:

BÙI QUANG TUẤN

Chế bản:

THÁI SƠN



© Bản quyền thuộc Bộ Y tế (Cục Khoa học công nghệ và Đào tạo – Bộ Y tế)

SINH HỌC PHÂN TỬ
(Dùng cho đào tạo dược sĩ Đại học)

Mã số: 7K721y4-DAI

Số đăng kí KHXB : 453-2014/CXB/ 19- 280/GD.

In 500 cuốn (QĐ in số : 14), khổ 19 x 27 cm.

In tại Công ty CP In Phúc Yên.

In xong và nộp lưu chiểu tháng 03 năm 2014.

