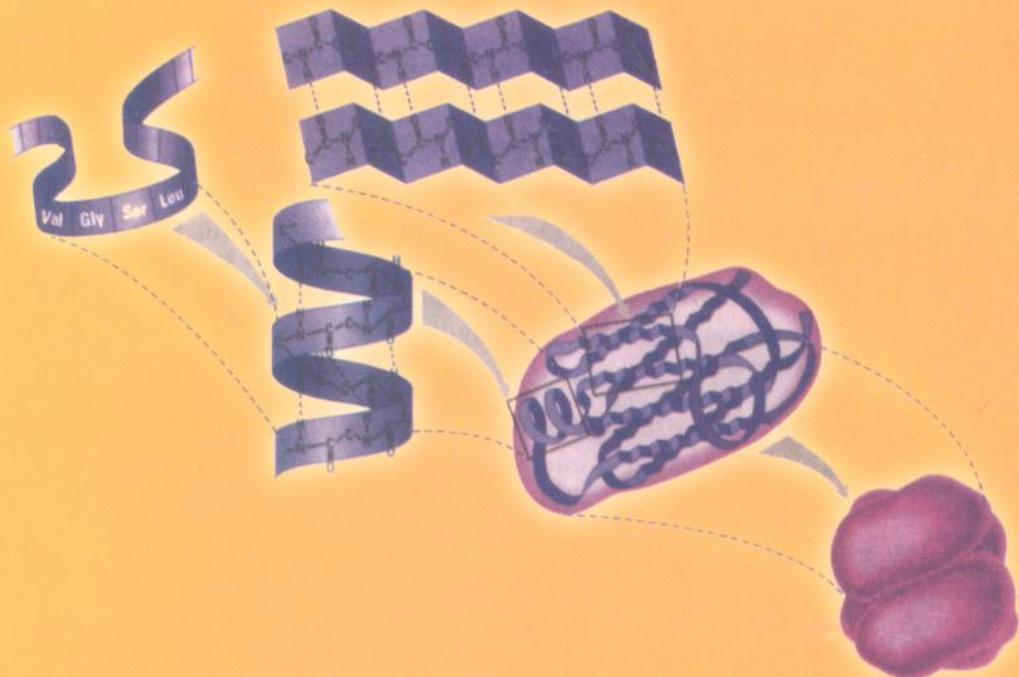


BỘ Y TẾ

SINH HỌC PHÂN TỬ

(DÙNG CHO ĐÀO TẠO DƯỢC SĨ ĐẠI HỌC)

GS. TS. NGUYỄN VĂN THANH (Chủ biên)



NHÀ XUẤT BẢN GIÁO DỤC

BỘ Y TẾ

SINH HỌC PHÂN TỬ[?]

(DÙNG CHO ĐÀO TẠO DƯỢC SĨ ĐẠI HỌC)

Mã số: Đ.20.X.06

NHÀ XUẤT BẢN GIÁO DỤC
HÀ NỘI – 2007

Chỉ đạo biên soạn:

VỤ KHOA HỌC VÀ ĐÀO TẠO – BỘ Y TẾ

Chủ biên:

GS.TS. NGUYỄN VĂN THANH

Những người biên soạn:

GS.TS. NGUYỄN VĂN THANH

TS. TRẦN THU HOA

TS. TRẦN CÁT ĐÔNG

ThS. HỒ THỊ YẾN LINH

Tham gia tổ chức bắn thảo:

ThS. PHÍ VĂN THÂM

TS. NGUYỄN MẠNH PHA

© Bản quyền thuộc Bộ Y tế (Vụ Khoa học và Đào tạo)

Lời giới thiệu

Thực hiện một số điều của Luật Giáo dục, Bộ Giáo dục & Đào tạo và Bộ Y tế đã ban hành chương trình khung đào tạo *Dược sĩ đại học*. Bộ Y tế tổ chức biên soạn tài liệu dạy - học các môn cơ sở và chuyên môn theo chương trình trên nhằm từng bước xây dựng bộ sách đạt chuẩn chuyên môn trong công tác đào tạo nhân lực y tế.

Sách *SINH HỌC PHÂN TỬ* được biên soạn dựa trên chương trình giáo dục của Trường Đại học Y Dược Thành phố Hồ Chí Minh trên cơ sở chương trình khung đã được phê duyệt. Sách được các tác giả GS.TS. Nguyễn Văn Thanh, TS. Trần Thu Hoa, TS. Trần Cát Đông, ThS. Hồ Thị Yến Linh biên soạn theo phương châm: Kiến thức cơ bản, hệ thống; nội dung chính xác, khoa học; cập nhật các tiến bộ khoa học, kỹ thuật hiện đại và thực tiễn ở Việt Nam.

Sách *SINH HỌC PHÂN TỬ* đã được Hội đồng chuyên môn thẩm định sách và tài liệu dạy - học chuyên ngành *Dược sĩ đại học* của Bộ Y tế thẩm định vào năm 2007. Bộ Y tế quyết định ban hành là tài liệu dạy - học đạt chuẩn chuyên môn của ngành trong giai đoạn hiện nay. Trong thời gian từ 3 đến 5 năm, sách phải được chỉnh lý, bổ sung và cập nhật.

Bộ Y tế xin chân thành cảm ơn các tác giả và Hội đồng chuyên môn thẩm định đã giúp hoàn thành cuốn sách; Cảm ơn PGS.TS. Nguyễn Văn Ty, PGS.TS. Đinh Hữu Dung đã đọc và phản biện, để cuốn sách hoàn thành kịp thời phục vụ cho công tác đào tạo nhân lực y tế.

Lần đầu xuất bản, chúng tôi mong nhận được ý kiến đóng góp của đồng nghiệp, các bạn sinh viên và các độc giả để lần xuất bản sau được hoàn thiện hơn.

VỤ KHOA HỌC VÀ ĐÀO TẠO – BỘ Y TẾ

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" (Elemente der exakten Erblichkeitslehre), 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 – 50000 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 "Sinh học phân tử" nhằm giúp cho sinh viên Dược khoa hiểu được cấu trúc cơ bản và chức năng của gen. Sách gồm các bài: Nhập môn; Sao chép ADN; Các loại ARN; Sự phiên mã và mã di truyền; Sinh tổng hợp protein; Điều hòa hoạt động gen; Bộ gen tế bào nhân thật; Đột biến gen – Các phương pháp phân tích ADN.

Sách xuất bản lần đầu nên khó tránh khỏi thiếu sót. Các tác giả rất mong nhận được sự góp ý của độc giả.

CÁC TÁC GIẢ

MỤC LỤC

	<i>Trang</i>
LỜI GIỚI THIỆU	3
LỜI NÓI ĐẦU	5
CÁC TỪ VIẾT TẮT	9
BẢNG ĐỔI CHIẾU THUẬT NGỮ VIỆT – ANH	11
BẢNG ĐỔI CHIẾU THUẬT NGỮ ANH – VIỆT	17
Bài 1 NHẬP MÔN SINH HỌC PHÂN TỬ	23
1.1. Lược sử	23
1.2. Đối tượng và phương pháp nghiên cứu	28
1.3. Những đóng góp lớn của sinh học phân tử hiện nay	30
Câu hỏi.....	38
Bài 2 SAO CHÉP ADN	40
2.1. Khái niệm	40
2.2. Sự sao chép của ADN	41
2.3. Sửa sai trong sao chép và khi không sao chép.....	57
Câu hỏi.....	58
Bài 3 CÁC LOẠI ARN	60
3.1. Khái niệm	60
3.2. Các ARN và vai trò của chúng.....	61
Câu hỏi.....	80
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 thuỷ	84
4.4. Sự phiên mã ở tế bào nhân thật.....	91
4.5. Phiên mã ngược ở retrovirus.....	97
4.6. Mã di truyền	98
Câu hỏi.....	100
Bài 5 SINH TỔNG HỢP PROTEIN	102
5.1. Mở đầu	102
5.2. Các yếu tố cần thiết cho sự tổng hợp protein.....	102
5.3. Diễn biến dịch mã ở ribosom (chu trình ribosom).....	109

5.4. Nhu cầu năng lượng cho quá trình sinh tổng hợp protein	118
5.5. Độ chính xác của quá trình dịch mã.....	120
5.6. Các yếu tố ức chế quá trình dịch mã.....	121
Câu hỏi.....	123
Bài 6 ĐIỀU HOÀ HOẠT ĐỘNG GEN.....	125
6.1. Mở đầu	125
6.2. Điều hòa quá trình sao chép.....	126
6.3. Điều hòa quá trình phiên mã.....	127
6.4. Kiểm soát sau dịch mã	138
Câu hỏi.....	140
Bài 7 BỘ GEN TẾ BÀO NHÂN THẬT.....	142
7.1. Mở đầu	142
7.2. Tổ chức bộ gen ở tế bào nhân thật	142
7.3. Các mức độ điều hòa biểu hiện gen	152
7.4. Điều hòa hoạt tính gen của tế bào nhân thật	155
7.5. Sự kiểm soát các chất thường gặp trong nhân.....	161
Câu hỏi.....	163
Bài 8 ĐỘT BIẾN GEN	165
8.1. Mở đầu	165
8.2. Các loại đột biến.....	167
8.3. Nguyên nhân đột biến	169
8.4. Các cơ chế chống lại đột biến.....	178
8.5. Các tính trạng đột biến và protein đột biến	183
8.6. Đột biến gen và ung thư	187
8.7. Các hệ thống chọn lọc đột biến ở vi sinh vật.....	189
Câu hỏi.....	192
Bài 9 CÁC PHƯƠNG PHÁP PHÂN TÍCH ADN.....	194
9.1. Chiết tách ADN.....	194
9.2. Các phương pháp phân tích định tính và định lượng sơ bộ acid nucleic	195
9.3. Kỹ thuật cắt, nối, lai ADN và ứng dụng	197
9.4. Các phương pháp xác định trình tự ADN.....	200
9.5. PCR	202
Câu hỏi.....	216
ĐÁP ÁN	218
TÀI LIỆU THAM KHẢO.....	219

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 Trascrption Factor
DMD	Duchenne muscular dystrophy
DNase	Deoxyribonuclease
dNTP	Desoxyribonucleozid triphosphat
dsADN	Double – stranded DNA
eEF	Eukaryote elongation factor
EF	Elongation factor
eIF	Eukaryote initiation factor
eRF	Eukaryote release factor
EGF	Epidermal growth factor
FRET	Fluorescence resonance energy transfer
G	Guanine
GMO	Genetic modified organism
Hfr	High frequency recombination
Hft	High frequency transduction
HIV	Human immunodeficiency virus
hnARN	Heterogenous nuclear RNA
Icr	Insensitive to catabolite repression
IF	Initiation factor
IGS	Internal guide sequence
Kb	Kilobase
LINE	Long interspersed repetitive element
MALDI-TOF	Matrix – assisted lazer desorption ionization time – of – flight
mARN	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 monophosphate
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
SPR	Surface plasmon resonance
SRP	Signal recognition particle
SSB	Single – stranded binding
ssADN	Single – stranded DNA
T	Thymine
tRNA	Transfer RNA
THE	Transposable human element
THR	Thyroid Hormone Receptor
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 vê 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 lặng	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 mutagensis
Gen có thể cảm ứng	Inducible gene
Gen có thể ức chế	Repressible gene
Gen điều hòa	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	Proteomics
Heterogenous nuclear ARN	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 hòa	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 hòa	Regulatory protein
Protein gắn chớp	Cap – binding protein

Tiếng Việt	Tiếng Anh
Protein hiệu ứng	Effector protein
Proteinức 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 hỗ biến	Tautomerization
Sự phiên mã	Transcription
Sự sao chép	Replication
Sự tự cắt nối	Self – splicing
Sự tự điều hòa	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 thuỷ	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

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 ĐỔI 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 hòa
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	Dầu tù
Cap	Chóp
Cap – binding protein	Protein gắn chóp
Capsid	Vỏ virus
Catena	Lồng 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 hóa

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	Dị 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 vệt

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 hòa 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 hóa
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 hòa 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 thuỷ
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 hòa
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 hiện đại do sinh học phân tử đem lại.

1.1. LUỢC SỬ

Có thể phân chia sự phát triển của sinh học phân tử 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 thuỷ 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 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 ≠ 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 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 hòa 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. Người ta xem enzym này 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 tuỳ 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 (SHPT) đã thúc đẩy sự ra đời của một ngành công nghệ mới là công nghệ di truyền.

Năm 1985, R.K. Saiki và K. B. 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, biến 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 nhiễm 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).
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.

Thời gian	Sự kiện
1947	McClintock khám phá yếu tố di truyền vận động "gen nhảy" ở bắp ngô.
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.
	Tìm 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.
	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.

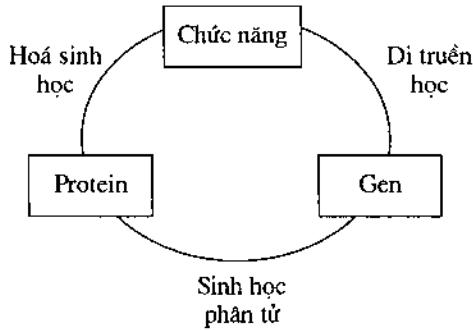
Thời gian	Sự kiện
1980	Trao bằng sáng chế kỹ thuật tạo dòng cho Cohen và Boyer. Phát triển máy tổng hợp gen đầu tiên. Giải Nobel Hóa học cho kỹ thuật tái tổ hợp phân tử: Berg, Gilbert, Sanger.
1982	Applied Biosystems, Inc., giới thiệu máy định trình tự protein pha khí đầu tiên, giảm thiểu lượng protein 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 hóa. Phát hiện ribozyme và retinoblastoma.
1990	Human Genome Project được khởi động.
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.

Thời gian	Sự kiện
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.
2003	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.
2004	Công bố phiên bản thô của bộ gen người hoàn chỉnh. 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 hóa 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 tinh chế protein người ta có thể hiểu được cách thức điều hòa 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.

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

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

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 ngũ cành tổng quát hơn. Như vậy, genomics chính là một nhánh khoa học giải thông tin mật mã trong ADN và khám phá các thông tin như số lượng gen, tổ chức và nội dung của bộ gen. Thông tin này có áp dụng cực kỳ to lớn trong Y tế, Nông nghiệp, Sinh học tiến hoá và các lĩnh vực khoa học khác.

Có nhiều loại genomics tuỳ theo khía cạnh và lĩnh vực mà nó nghiên cứu áp dụng. Trong mỗi loại genomics cũng lại gồm nhiều lĩnh vực, ví dụ như plant genomics gồm công nghệ sinh học thực vật, sinh học phân tử và các kỹ thuật liên quan đến thực vật (nông nghiệp, lâm nghiệp, làm vườn).

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).

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/phản 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 đọc 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 pseudobscura* và gần đây của cây lúa nước, đã mở ra một ngành khoa học mới là "Genom học – Genomics" chuyên nghiên cứu về cấu trúc và chức năng của toàn bộ bộ gen sinh vật.

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

Proteomics là nghiên cứu ở quy mô lớn về protein nhằm thiết lập các dấu hiệu 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. Từ "proteomics" được đặt ra để tương đương với genomics. Proteome (bộ protein) của một sinh vật là tập hợp các protein bởi nó trong thời gian sống của nó. Proteome của một tế bào ở tập hợp các protein trong nó. Từ "proteome" xuất phát từ "proteins" và "genome", vì genome (bộ gen) mã hoá các protein.

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

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 khói phổ (mass spectrometry), và kết hợp các phương pháp thực nghiệm như trình dien phage và các phương pháp máy tính.

1.3.2.2. Nhận diện protein

Nghiên cứu proteomics hiện nay trước hết là protein phải được phân giải, đôi khi ở quy mô lớn. Có thể tách protein bằng điện di hai chiều trên gel (two – dimensional gel electrophoresis). Kỹ thuật này đầu tiên tách protein theo điểm đẳng điện và sau đó theo khối lượng phân tử. Các điểm protein trên gel có thể nhìn được bằng cách dùng các chất nhuộm hóa học khác nhau hoặc các dấu hiệu huỳnh quang. Thường thì các protein có thể định lượng dựa vào cường độ nhuộm màu của chúng. Các protein đã được nhận diện khi đã được tách riêng và định lượng. Các đốm riêng rẽ được cắt ra khỏi gel và cắt thành các peptid bằng các enzym thuỷ phân giải protein. Sau đó các peptid này được nhận diện bằng khói phổ, đặc biệt là khói phổ MALDI – TOF. Trong quy trình này, peptid được đặt trên matrix làm peptid tạo các tinh thể. Sau đó peptid được ion hoá với tia laser và sự tăng điện thế ở matrix được dùng để bắn các ion về phía đầu dò (detector) mà thời gian ion đến đầu dò tùy thuộc vào khối lượng của nó. Ion có khối lượng càng cao càng mất nhiều thời gian bay. Trong khói phổ MALDI – TOF, các ion cũng có thể bị lệch hướng với gương phản xạ tinh điện cũng làm tập trung chùm ion. Như vậy, có thể xác định với độ chính xác cao sự đến đầu dò thứ hai các khối của các ion và các khối này có thể phát hiện chính xác các thành phần của các peptid, do đó nhận diện chúng.

Các hỗn hợp protein có thể được phân tích không cần tách trước. Quy trình này bắt đầu bằng phân giải các protein trong phức hỗn hợp. Các peptid tạo thành thường được bơm vào cột sắc ký lỏng hiệu năng cao (HPLC) làm tách các peptid dựa vào tính không ưa nước. Các peptid tách rửa từ cột có thể được nhận diện bằng khôi phổi song song (MS/MS). Giai đoạn đầu tiên của khôi phổi song song là phân lập các ion peptid riêng rẽ, giai đoạn thứ hai sẽ phá vỡ các peptid thành các đoạn và sử dụng khuôn mẫu của đoạn (fragmentation pattern) để xác định trình tự acid amin của chúng. Có thể dùng đánh dấu với các nhãn đồng vị để so sánh định lượng nồng độ các protein trong hai hoặc nhiều hơn hai mẫu protein.

1.3.2.3. Các kỹ thuật được sử dụng trong proteomics

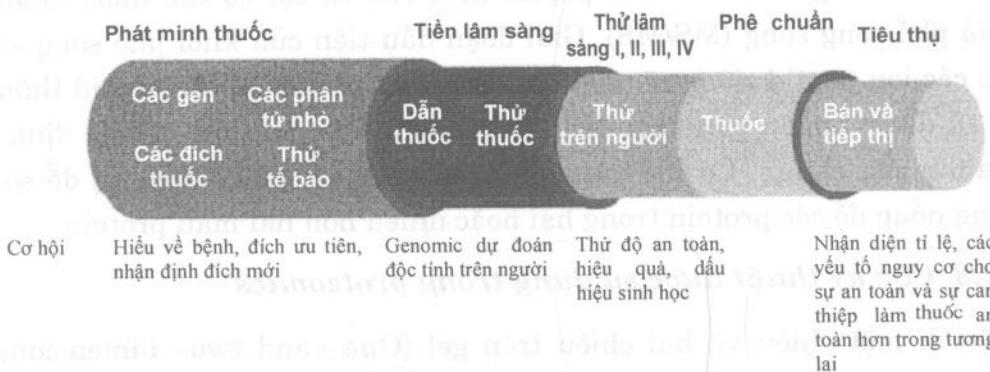
- Điện di một chiều và hai chiều trên gel (One – and two – dimensional gel electrophoresis).
- Tinh thể tia X và cộng hưởng từ hạt nhân (X – ray crystallography and nuclear magnetic resonance).
- Đa khôi phổi (Tandem mass spectrometry) kết hợp với sắc ký đảo pha (reverse phase chromatography) hoặc điện di hai chiều (2 – D electrophoresis).
- Khôi phổi (không song song), thường là MALDI – TOF.
- Sắc ký ái lực (Affinity chromatography), kỹ thuật lai đôi trên nấm men (yeast two hybrid techniques), chuyển năng lượng cộng hưởng huỳnh quang (FRET), và cộng hưởng plasmon bề mặt (SPR) được dùng để nhận định tương tác protein – protein và liên kết protein – ADN.
- X – ray Tomography.
- Phần mềm.

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 khôi phổi 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. Di truyền học có thể tác động đến từng pha của quá trình này. Tác động của di truyền học và bộ gen học lên xã hội sẽ là dự đoán di truyền về độ an toàn và hiệu quả của thuốc sớm nhất trong thương trường.



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. Từ này xuất phát từ pharmacology và genomics và như vậy nó là sự giao nhau của dược và di truyền.

Genomics tạo ra sự thay đổi cơ bản trong cách sử dụng thuốc và quá trình phát triển thuốc. Sự thay đổi quy mô này có lẽ sẽ xảy ra trong 5 đến 10 năm tới. Nguyên nhân của sự thay đổi là cuộc cách mạng khoa học và kỹ thuật trong lĩnh vực Bộ gen Người (Human genome, bản đồ đa hình đơn nucleotid (single – nucleotide polymorphism, SNP), định loại gen (genotyping) và tin sinh học (bioinformatics).

SNP là những dấu hiệu để đo lường và thử nghiệm tầm soát hiệu năng cao. Các tiến bộ trong đánh giá biểu hiện gen cũng là diệu kỳ. Trong vòng năm năm, số lượng các gen, độ nhạy của các thử nghiệm, độ lặp lại của các kết quả cũng đã tăng đáng kể khả năng phân tích các số liệu. Độ sự biểu hiện gen được dùng trong di truyền ung thư và dược học bộ gen ung thư (cancer pharmacogenomics). Sự biểu hiện gen giúp cách phân loại ung thư và liệu pháp thích hợp tốt hơn.

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. Có thể phát minh thuốc và cho các dấu hiệu thuốc (drug maker) để đưa ra liệu pháp có đích chuyên

biệt. Sự chính xác này không chỉ tối đa hoá hiệu quả trị liệu mà còn làm giảm sự phá huỷ các tế bào khỏe mạnh gần đó.

– Cho các thuốc tốt hơn, an toàn hơn ngay từ đầu. Thay vì phương pháp chuẩn 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 còn cho 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 khối 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. Các quá trình thích hợp làm dễ dàng thử nhầm trên nhóm dân số có di truyền chuyên biệt – cho mức độ thành công lớn hơn. Giá thành và nguy cơ của các thử nghiệm lâm sàng giảm do chỉ nhầm vào những người có thể đáp ứng với thuốc.

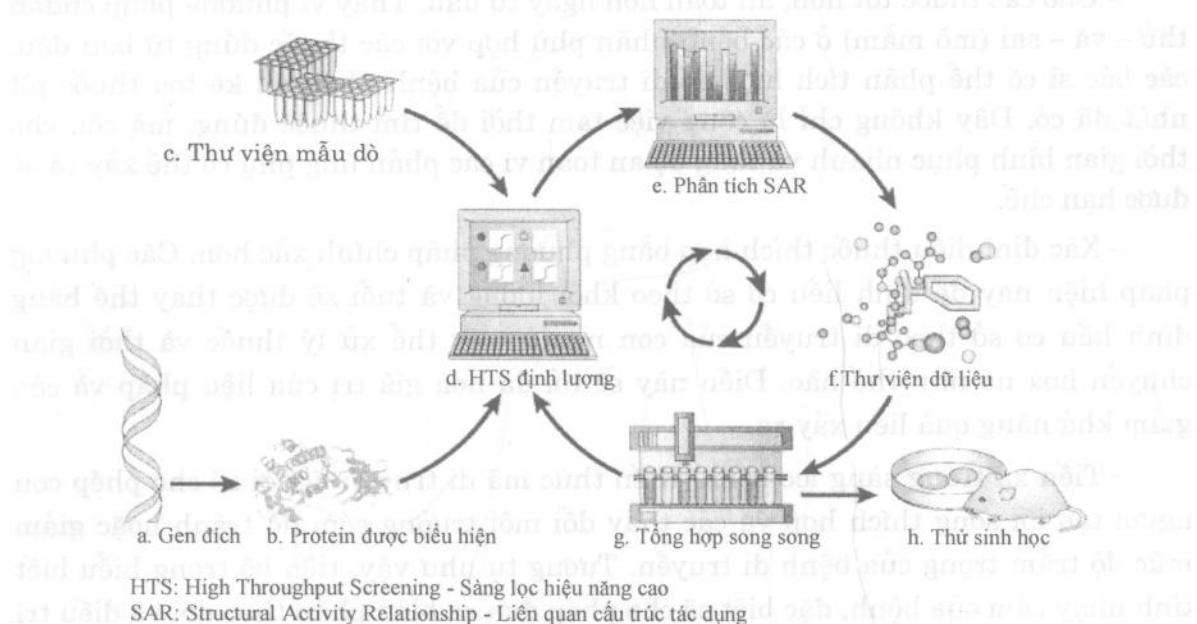
– Giảm chi phí trong chăm sóc sức khỏe. Giảm số lượng các phản ứng phụ của thuốc, số lượng thử thuốc thất bại, thời gian để thuốc được phê duyệt, thời gian bệnh nhân cần dùng thuốc, số lượng thuốc mà bệnh nhân phải dùng để điều trị hiệu quả, ảnh hưởng của bệnh lên cơ thể (do phát hiện sớm) và tăng khoảng cách đích có thể có của thuốc sẽ đẩy mạnh giảm mạng lưới giá thành trong chăm sóc sức khỏe.

1.3.3.2. Chiến lược chemogenomics để phát minh thuốc

Các thành tựu trong genomic 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.



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 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 hoá đa biến được thiết kế nhằm đảm bảo các hợp chất không chỉ được tối ưu hoá 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).

1.3.4. 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 peptide nhỏ), các chất dinh dưỡng (vitamin A trong hạt lúa),... 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 hóa 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 da 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 mổ 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. Bảng đồ gen người khi hoàn chỉnh cho thấy có bao nhiêu gen mã hoá cho protein?
 - a) 10 000 – 15 000
 - b) 20 000 – 25 000
 - c) 25 000 – 30 000
 - d) 15 000 – 20 000
 - 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 ADN có thể chuyển sang ARN.
 - c) Thông tin chỉ luân chuyển giữa các dạng acid nucleic khác nhau.
 - d) Sự sao chép, phiên mã, dịch mã là các quá trình chuyển thông tin trong tế bào.
 - 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

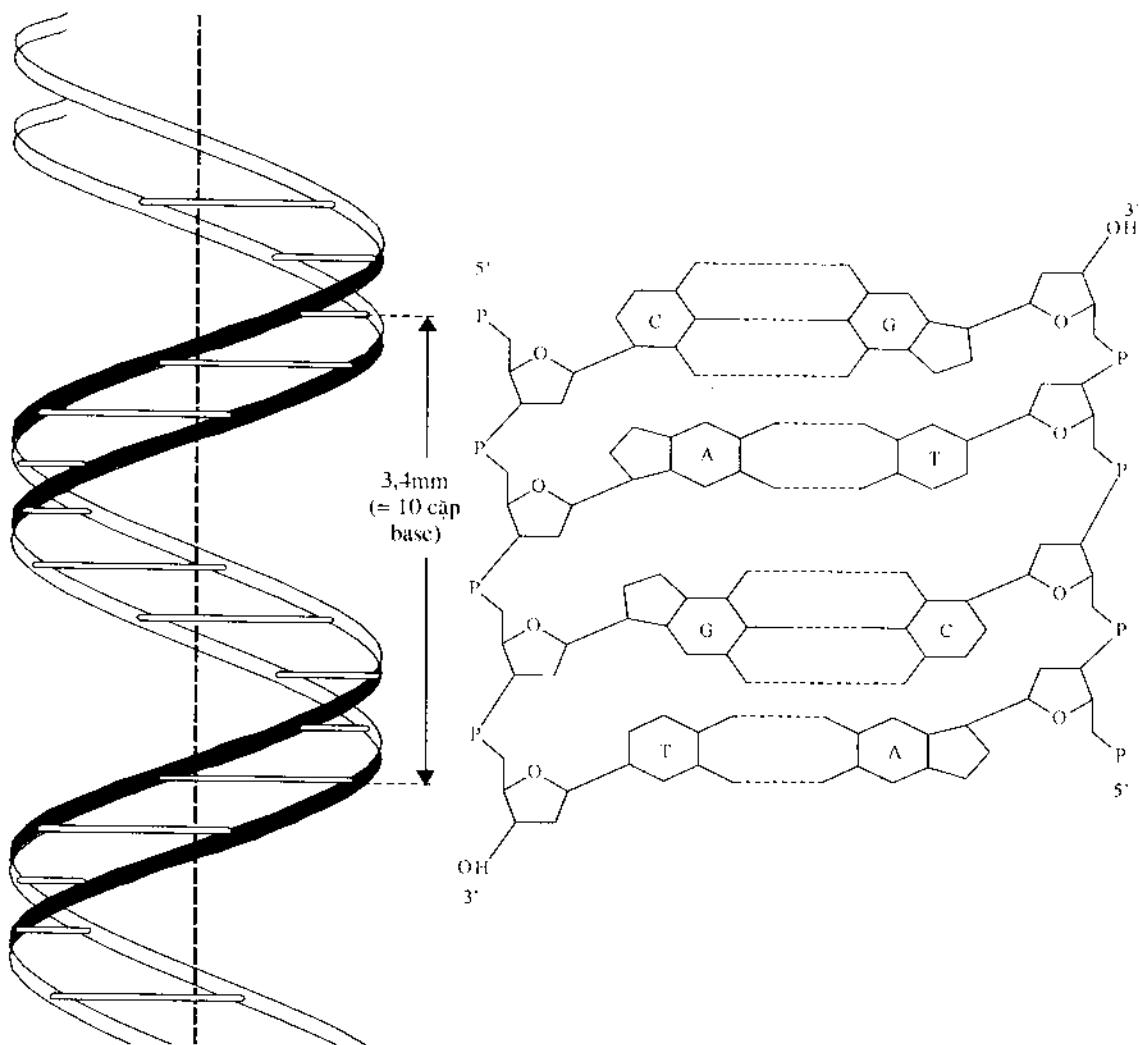
Một tế bào vi khuẩn điển hình có thể chứa khoảng 2000 loại protein, các tế bào nhân thật chứa khoảng 50000. Thông tin về các loại protein này được mã hoá trong các phân tử acid nucleic. Ở đa số loài, vật liệu di truyền chủ yếu là ADN, ở một số virus là ARN.

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 trực. Trình tự của các base nitơ tạo nên thông tin di truyền. Sự khác nhau trong trình 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.

Các base nối với nhau bằng liên kết hydro theo từng cặp bổ sung. Nhờ đó, khi chuỗi xoắn kép tách ra cho phép các base được sao chép tạo nên hai mạch mới, cơ bản giống với phân tử ADN gốc và đây là sự sao chép cơ bản trong hệ thống sinh học. Các enzym xúc tác quá trình sao chép của ADN là các ADN – polymerase. Một số enzym 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 huỷ ADN hay gây đột biến, sự hư hại này có thể được sửa chữa. Các yếu tố này còn có thể làm chết tế bào hoặc làm biến đổi bộ gen (genome). Tất cả những thay đổi này có thể di truyền được qua con đường sao chép ADN.



Hình 2.1. Cấu trúc phân tử xoắn kép của ADN

2.2. SỰ SAO CHÉP CỦA ADN

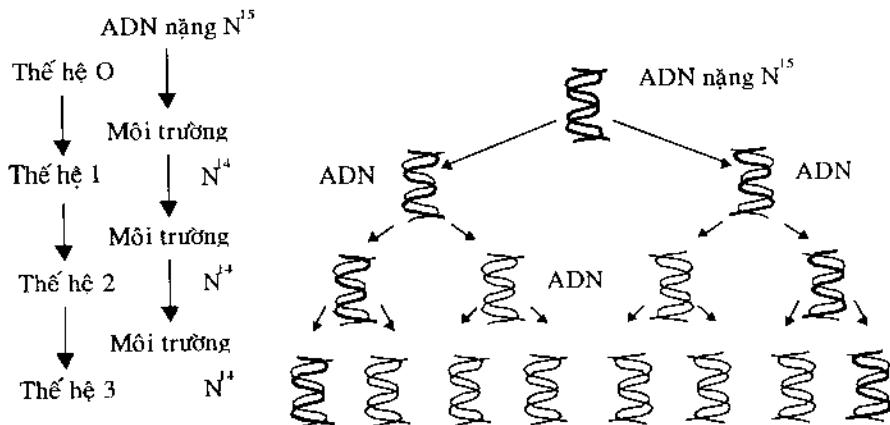
Người ta đưa ra hai cơ chế cơ bản của sự sao chép ADN: một cơ chế bảo tồn (conservative mechanism) trong đó phân tử ADN con gồm hai chuỗi hoàn toàn mới hoặc cơ chế bán bảo tồn (semiconservative mechanism) trong đó phân tử ADN con gồm một chuỗi mẹ kết hợp với một chuỗi mới được tổng hợp. Cơ chế sau được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

2.2.1. Thí nghiệm của Meselson và Stahl

Nếu tế bào *Escherichia coli* được cung cấp một nguồn nitơ 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 chuyển thành 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 ly tâm với thang nồng độ cesium clorid.

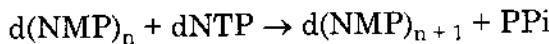
Trong thí nghiệm này, vi khuẩn được cho phát triển vài thế hệ trong môi trường N^{15} để đảm bảo cho tất cả ADN của vi khuẩn đều là loại nặng. Sau đó chuyển tế bào vào môi trường có chứa N^{14} , như là nguồn cung cấp nitơ duy nhất và để cho phân chia trong môi trường N^{14} , khi khối lượng ADN tăng 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 thuỷ (N^{15}) và một băng chứa N^{14} (hình 2.2). 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

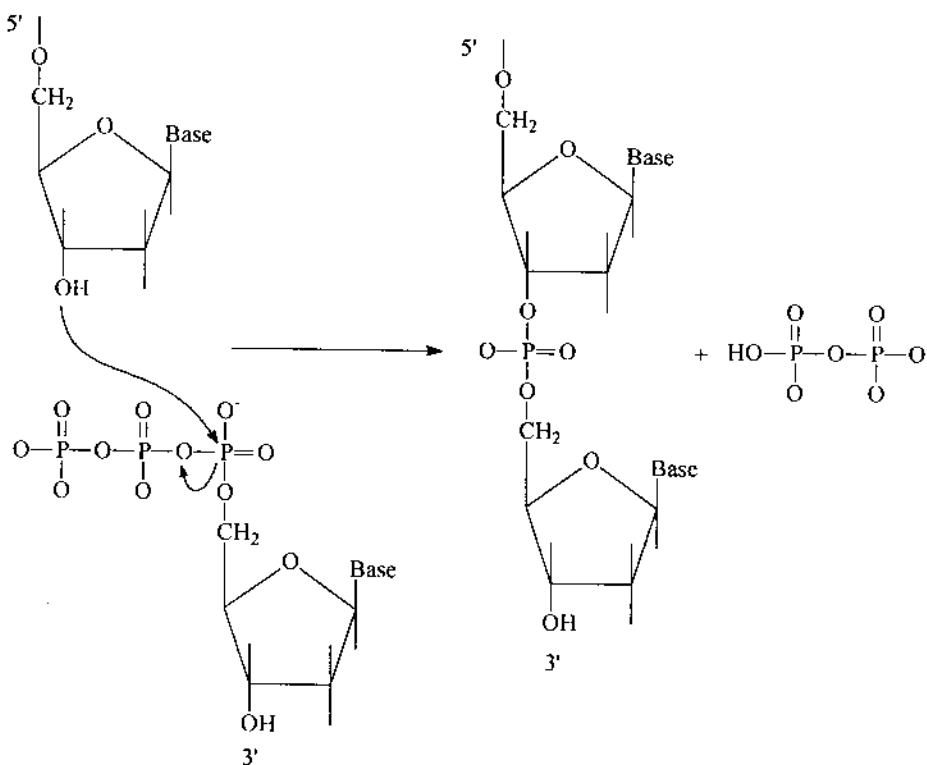
2.2.2. Các yếu tố cần thiết cho sự sao chép ADN

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), Mg^{2+} , 4 loại desoxyribonucleotid triphosphat (dNTP) và enzym ADN 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 thuỷ 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

ADN polymerase I, II, III được tìm thấy ở tế bào nhân nguyên thuỷ. Đặc tính của các ADN polymerase khác nhau được tóm tắt trong bảng 2.1.

ADN polymerase I của *E. coli* là một chuỗi polypeptid lớn có khối lượng phân tử 109000. Trong mỗi phân tử có chứa một nguyên tử kẽm. Enzym polymerase I chỉ có thể xúc tác sự polymer hoá theo hướng 5' → 3' do sự thêm một nucleotid từ một deoxyribonucleotid – 5' – triphosphat vào nhóm 3' – hydroxy của chuỗi ADN (hình 2.3).

Tất cả ADN polymerase của tế bào nhân nguyên thuỷ cũng có hoạt tính exonuclease, vì chúng thuỷ phân mạch đơn ADN của chuỗi từ đầu 3' theo hướng 3' → 5' (hoạt tính exonuclease 3' → 5'); hoặc chúng có thể thuỷ phân chuỗi ADN từ đầu 5' (hoạt tính exonuclease 5' → 3'). ADN polymerase II của *E. coli* rất giống ADN polymerase I nhưng không có hoạt tính exonuclease 5' → 3'. ADN

polymerase III được cấu tạo gồm 7 đơn vị nhỏ: α , β , γ , δ , ϵ , θ , và τ . Tất cả 7 đơn vị nhỏ này đều cần cho sự sao chép của ADN của *E. coli*. ADN polymerase III có thể là enzym sao chép chính của *E. coli* vì các đột biến với ADN polymerase III ưa nhiệt không thể sao chép được ở nhiệt độ 42°C. ADN polymerase I chịu trách nhiệm sửa chữa các ADN hư hỏng. Vai trò của ADN polymerase II chưa rõ.

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	3: α 140 000 ϵ 25 000 θ 10 000	2: α 65 000 β 95 000
Hoạt tính polymer hoá 5' → 3'	+	+	+	+
Hoạt tính exonuclease				
5' → 3'	+	-	+	-
3' → 5'	+	+	+	-

Ở tế bào 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 thuỷ. Polymerase α và δ liên quan tới sự sao chép của ADN nhiễm sắc thể. 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. Polymerase ϵ polymerase mới được tìm thấy gần đây, có một số điểm tương tự polymerase δ về hoạt tính.

ADN ϵ polymerase mới được tìm thấy gần đây, có một số điểm tương tự polymerase δ về hoạt tính.

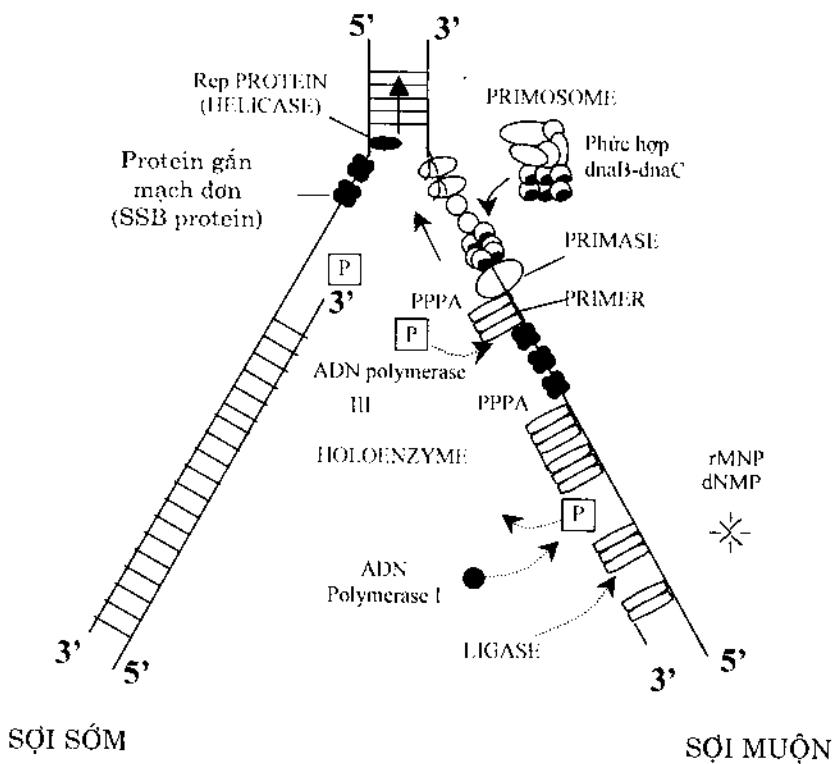
Bảng 2.2. Đặc điểm ADN polymerase của nhân thật

Polymerase	α	β	γ	δ	ϵ
Khối lượng phân tử (Mr)	300000	45000	140000	180000 – 240000	290000
Đơn vị nhỏ	4	1	4	2	2
Vị trí	Nhân	Nhân	Ty thể	Nhân	Nhân
Hoạt tính polymer hoá 5' → 3'	Sao chép ADN nhiễm sắc thể (sợi muộn)	Sửa chữa	Sao chép ADN ty thể	Sao chép ADN nhiễm sắc thể (sợi sớm)	Sao chép ADN nhiễm sắc thể

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 thuỷ 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 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. Dùng 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).

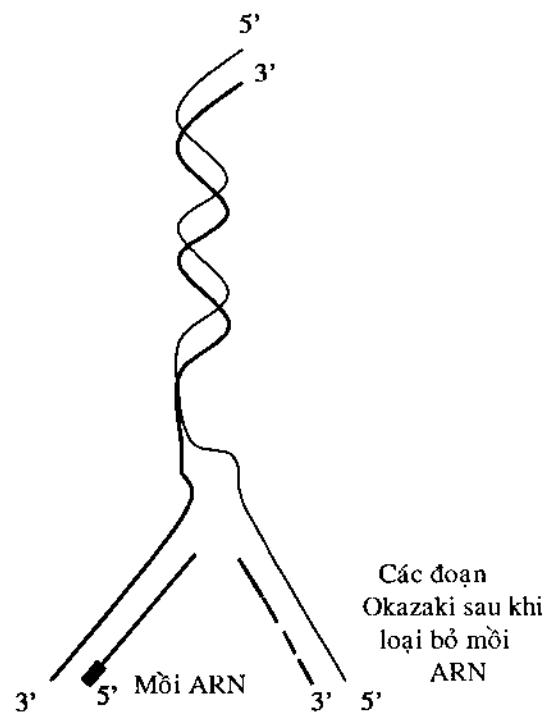
Các sợi ADN được tách ra và một sợi ADN có hướng $3' \rightarrow 5'$ sẽ được sao chép trực tiếp bởi ADN polymerase III. Sợi con bổ sung vừa tạo ra được gọi là sợi sói sóm (leading strand). Sợi gốc còn lại có hướng $5' \rightarrow 3'$, không được sao chép cho đến khi một phần của sợi gốc được tháo xoắn. 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 mồi 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 mồi ARN sẽ bị ADN polymerase I loại khỏi chuỗi ADN. Enzym này đồng thời làm nhiệm vụ lấp đầy các khoảng trống bằng các deoxyribonucleotid thích hợp do việc loại mồi. Các đoạn ADN được nối lại với nhau bởi ADN ligase tạo thành sợi ADN liên tục.



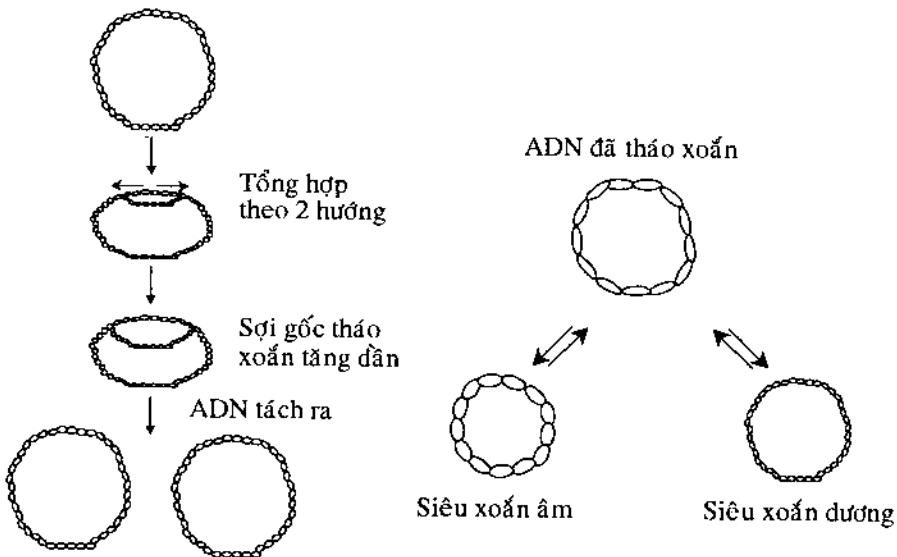
Hình 2.5. Sơ hình thành các đoạn Okazaki của sợi muộn ADN do sự sao chép không liên tục của sợi gốc

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).

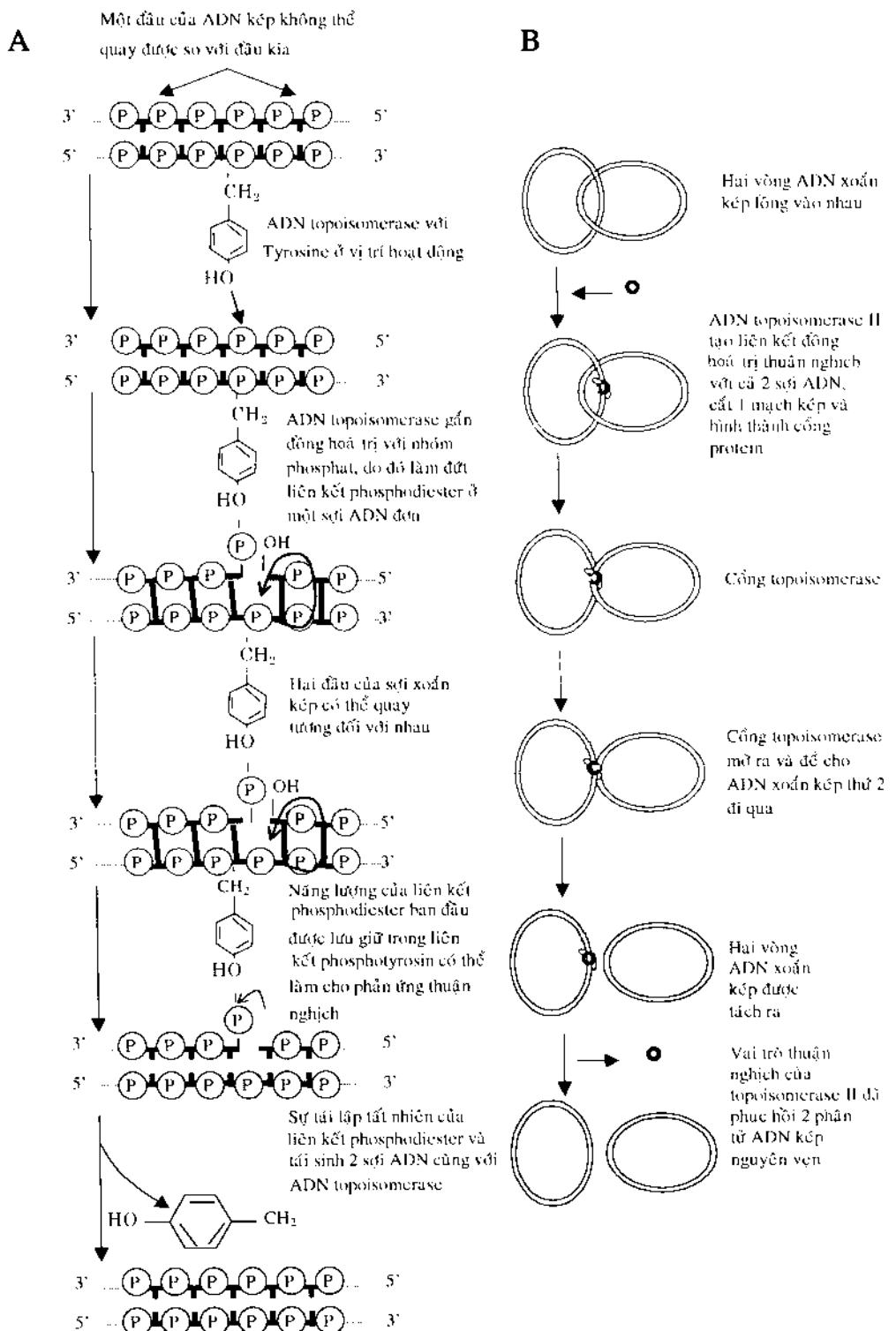


Hình 2.6. Sự hình thành cấu trúc theta do sự sao chép hai hướng của ADN kép

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).

Phân 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 sợi ADN và để phân tử ADN sợi kép đi xuyên qua điểm cắt, tạo phân tử ADN kép mới (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 phân 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 thuỷ phân ATP để di 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 muộn, enzym này tạo phức hợp với primase để tạo mồi 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 kép 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, thuỷ phân ATP để tạo năng lượng tự do cho helicase hoạt động. Tác động hiệp đồng của các protein này trong hình 2.4 và hình 2.9 và chức năng của chúng được tóm tắt trong bảng 2.3.

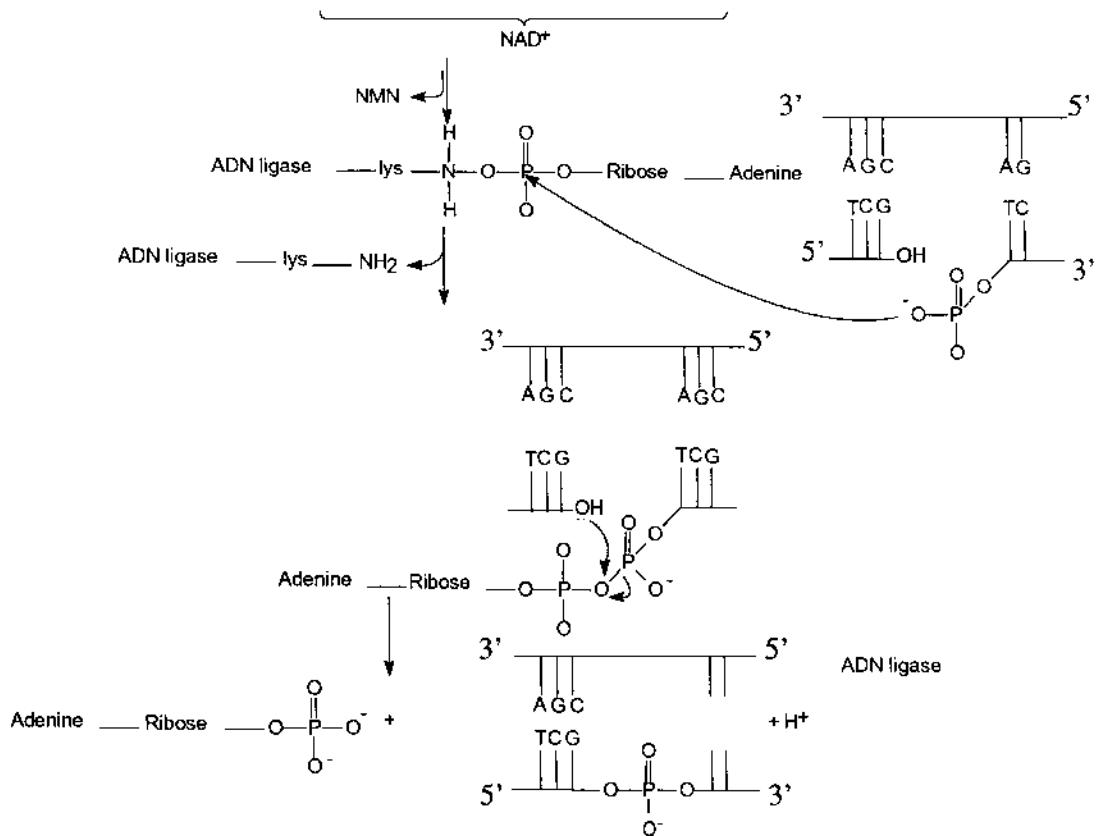
Bảng 2.3. Chức năng của các protein sao chép quan trọng của *E. coli*

Protein	Chức năng
Protein B	Nhận biết điểm Ori
N – protein	Cho phép primase hoạt động
Primase	Tổng hợp mồi ARN
Rep protein	Tháo xoắn sợi sớm
Helicase	Tháo xoắn sợi muộn
SSB – protein	Ôn định sợi AND đơn
ADN polymerase III	Tổng hợp AND
ADN topoisomerase I	Cắt khía một sợi đơn ADN
ADN topoisomerase II	Cắt khía 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 mồi 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 thực

2.2.5.1. Cơ chế sao chép

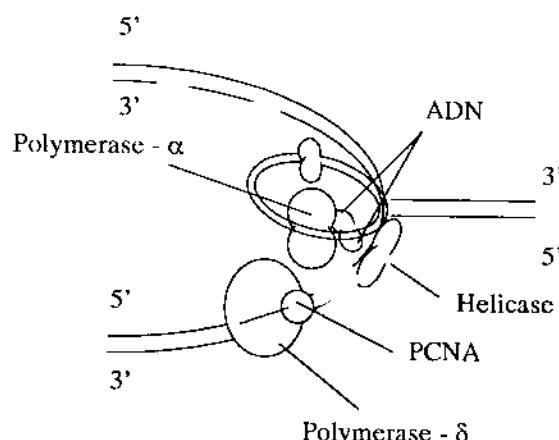
Cơ chế tổng hợp ADN ở tế bào nhân thực cũng tương tự như ở tế bào nhân nguyên thuỷ. Điểm khác biệt chủ yếu là ADN ở tế bào nhân thực đóng cuộn trong nhiều nhiễm sắc thể và dài hơn. Tốc độ di chuyển ADN polymerase ở tế bào nhân thực chậm hơn rất nhiều so với tế bào nhân nguyên thuỷ (khoảng 50 nucleotid/giây). Nhưng tế bào nhân thực chứa tới 20 000 phân tử enzym. Do vậy, ở nhiễm sắc thể tế bào nhân thực hình thành một lượng lớn chac 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 thực 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

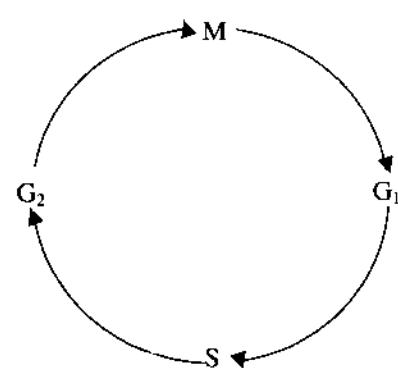
(PCNA = proliferating cell nuclear antigen = kháng nguyên tăng sinh nhân tế bào)

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 thuỷ.

2.2.5.2. Chu kỳ tế bào

Sự sao chép ADN diễn ra như một phần của tiến trình phôi hợp của sự phân chia tế bào. Chu kỳ tế bào nhân thật gồm 4 pha riêng biệt: pha M, G₁, S và G₂ (hình 2.10).

Pha G₁ là thời kỳ trước khi ADN bắt đầu tổng hợp, độ dài của G₁ thay đổi từ phút, giờ, tuần thậm chí đến hàng năm. Tế bào không bao giờ phân chia là tế bào trong đó G₁ bị ngưng trệ, thường được coi như là giai đoạn Go.

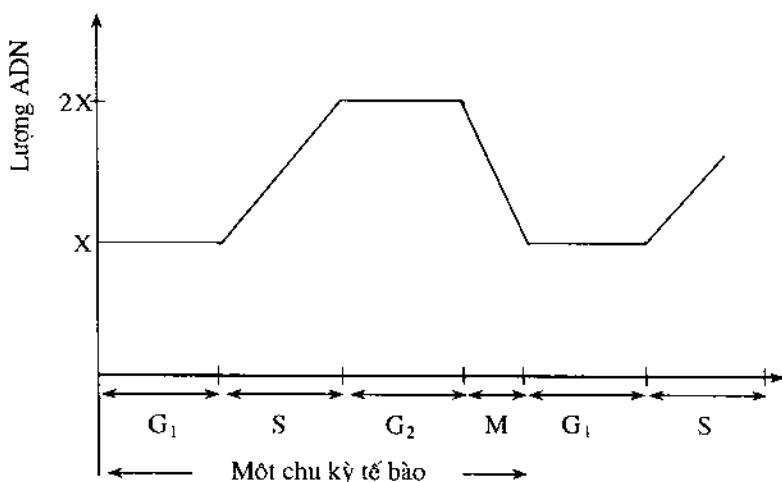


Hình 2.10. Chu kỳ tế bào nhân thật

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 G₂ ngắn hơn pha G₁, trong pha này người ta biết được ít những điều gì xảy ra ở tế bào. Pha G₂ 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.



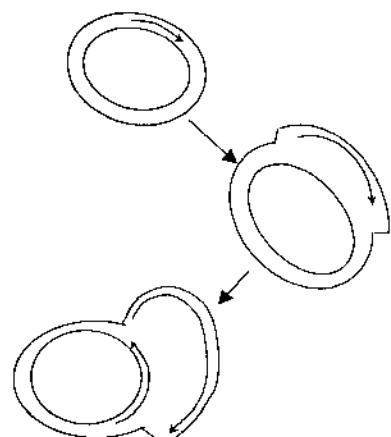
Hình 2.11. Lượng ADN được tạo ra trong chu kỳ tế bào

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

Ty thể và lạp thể có 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).

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.

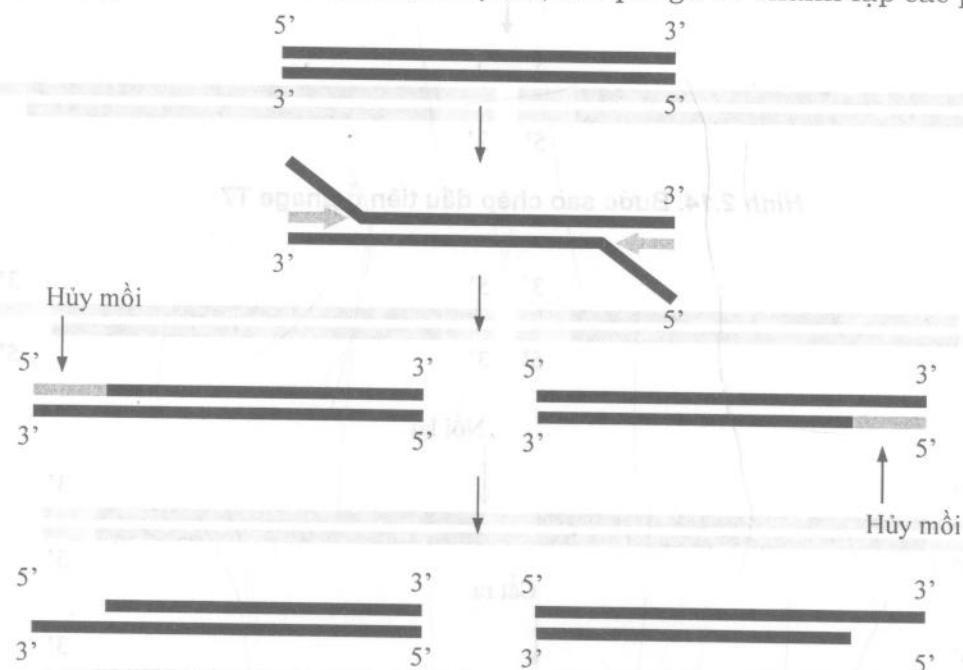


Hình 2.12. Sự hình thành cấu trúc D trong quá trình sao chép ADN ty thể

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 mồi ARN và chỉ hoạt động tổng hợp ADN theo chiều $5' \rightarrow 3'$ nên nó không thể tổng hợp chính xác các ADN thẳng (hình 2.13).

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.

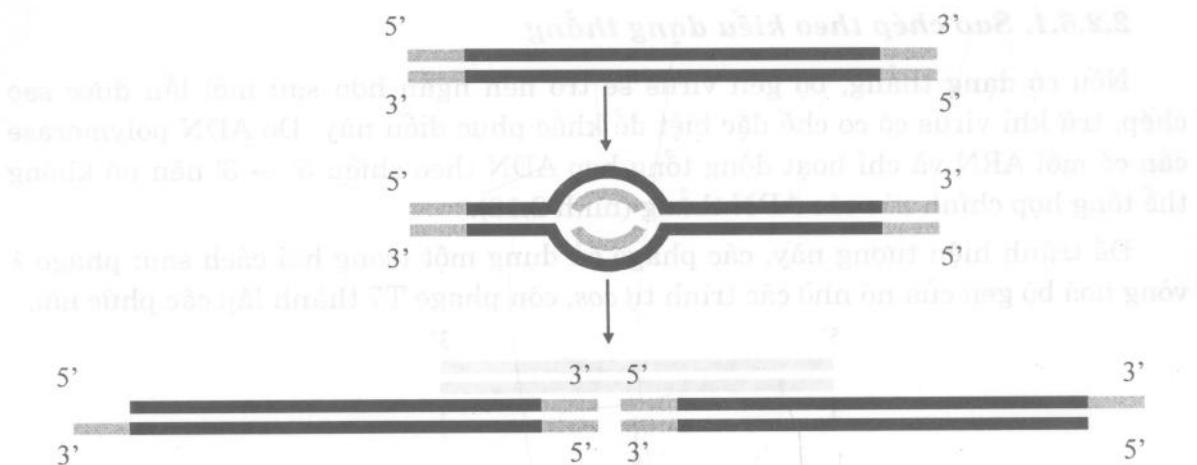


Hình 2.13. Sao chép ADN theo kiểu dạng thẳng

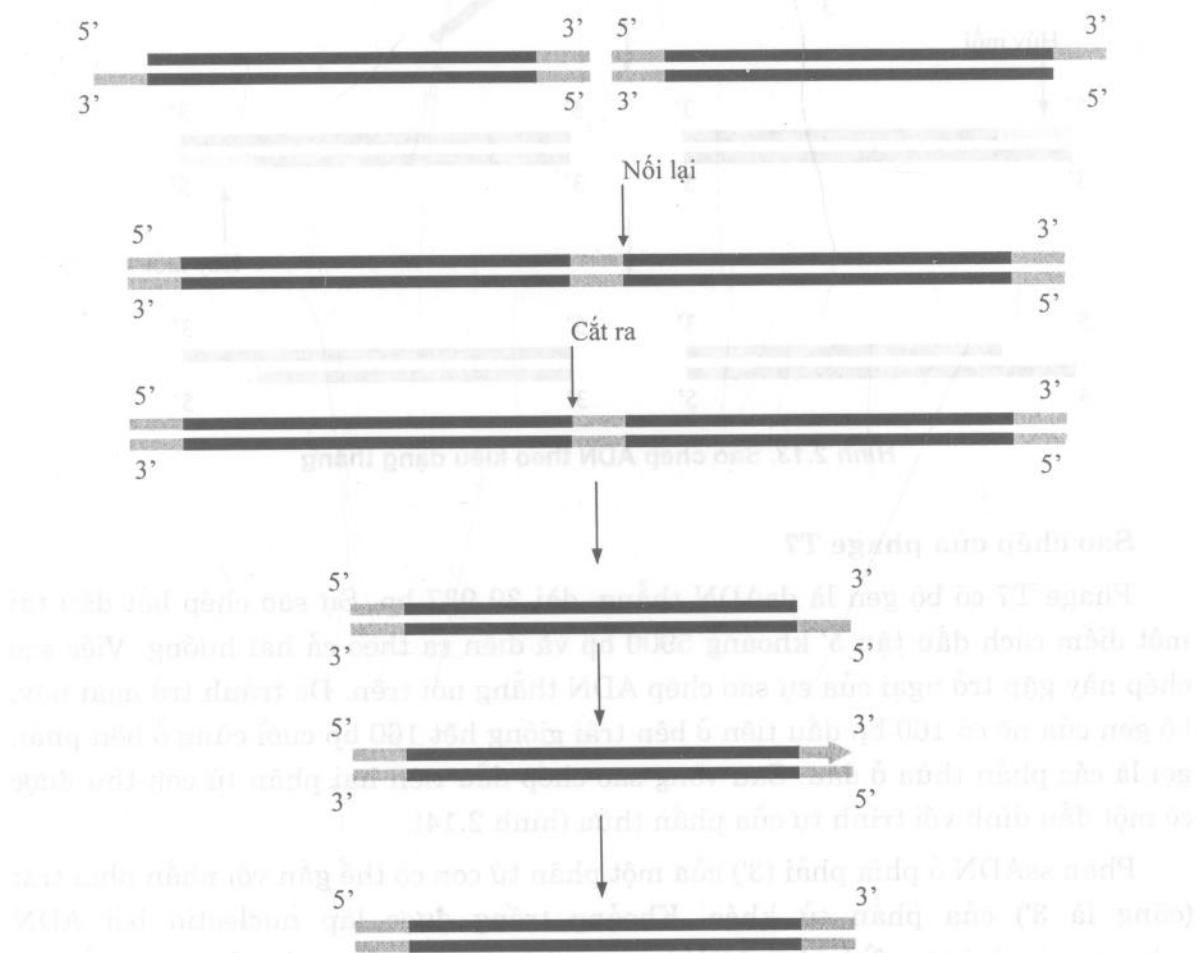
Sao chép của phage T7

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).

Phân ssADN ở phía phải (3') của một phân tử con có thể gắn với phân 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).



Hình 2.14. Bước sao chép đầu tiên ở phage T7



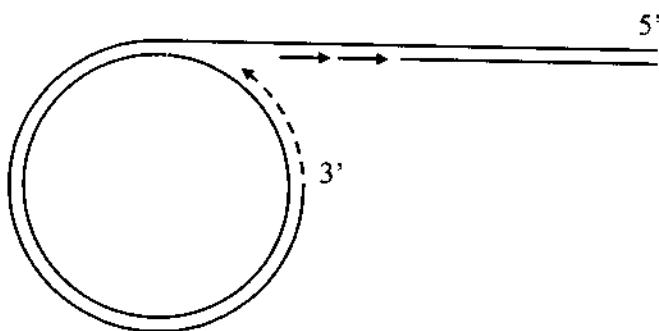
Hình 2.15. Phức nối được cắt ra và tái tạo bộ gen của phage T7

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

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 mồi 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 bọc bởi các SSB protein. Vì phage M13 không có gen mã hóa 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 mồi

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 thế 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 gọi là đầu dính, hai đoạn ở hai đầu có thể gắn bổ sung với nhau. Chúng dài 12 nucleotid và được gọi là vị trí cos.

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 tiêm 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. 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. Sản phẩm này giúp DnaB 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 thuỷ 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 thuỷ 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 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 hóa 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 hóa trị sai giữa base với các chất hóa 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 thuỷ. 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 thuỷ.

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 – Sao chép sợi muộn.
 - 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 thực
- a) I
 - b) II
 - c) α
 - d) β
 - e) b và d
4. Bong bóng sao chép còn được gọi là
- a) Nút sao chép
 - b) Chạc ba sao chép
 - c) Vị trí Origin
 - d) Vị trí Okazaki
 - e) Tất cả đều sai
5. Ý nào đúng với đoạn Okazaki ở tế bào nhân nguyên thuỷ?
- a) Gồm khoảng 1000 – 2000 nucleotid
 - b) Được nối lại bằng ADN ligase
 - c) Nối với nhau tạo thành sợi sầm
 - d) Còn gọi là loop
 - e) a và b
6. Vị trí Origin
- a) Điểm khởi đầu sự sao chép
 - b) Được nhận diện bởi protein B
 - c) Gồm 254 cặp base
 - d) a và b
 - e) a, b và c
7. Primase là enzym
- a) Tự bản thân không hoạt động được
 - b) Gồm nhiều N – protein gắn
 - c) Còn gọi là primosome
 - d) Lấp đầy các GAP bằng dNTP
 - e) a và c
8. Primase bắt đầu hoạt động khi
- a) N – protein được nhận diện
 - b) N – protein nhận diện được Ori
 - c) Protein – B nhận diện được Ori
 - d) Tạo phức hợp với các chuỗi polypeptid
 - e) Tạo phức hợp với N – protein
9. Bản sao ADN vòng mới được tổng hợp, tách ra khỏi bản gốc nhờ
- a) Topoisomerase I
 - b) Topoisomerase II
 - c) Tháo xoắn âm
 - d) Phân Tyrosin của Topoisomerase
 - e) Gốc phosphat tự do của Topoisomerase
10. Phage lambda sao chép bộ gen của nó theo kiểu
- a) Theta
 - b) Theta và lăn vòng
 - c) Sao chép ADN thẳng
 - d) Sao chép ADN vòng
 - e) 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 thuỷ 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 tuỳ 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.

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).

Tất cả ARN đều được tổng hợp từ các gen tương ứng trên ADN và đều bao gồm:

- Mạch đơn polyribonucleotid
- Đường ribose (5C)
- Các base nitơ Adenin, Guanin, Cytosin và Uracil (ở ADN là Thymin).

ARN 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.

3.2. CÁC ARN VÀ VAI TRÒ CỦA CHÚNG

3.2.1. ARN ribosom

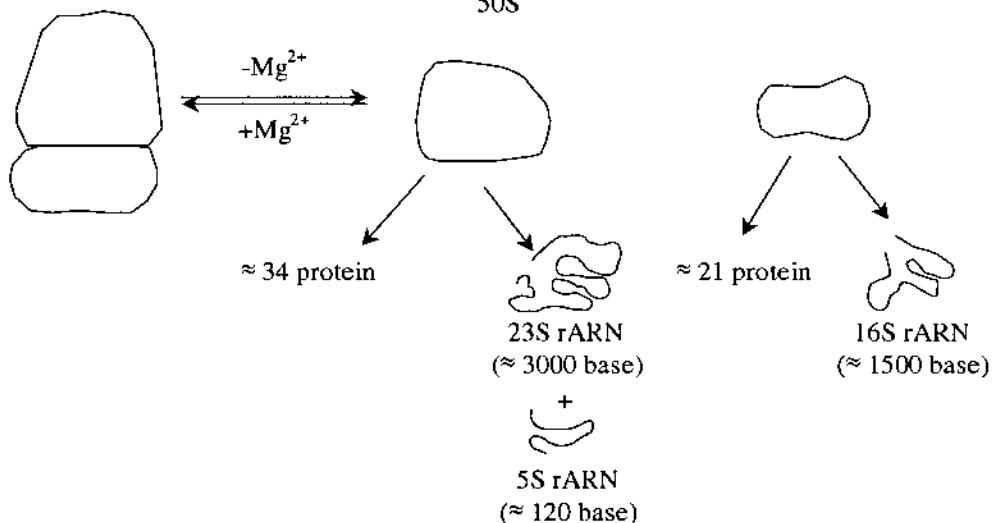
rARN là thành phần cấu tạo ribosom, chiếm phân nửa khối lượng của ribosom.

Trong dịch nuôi cấy *Escherichia coli* đang ở giai đoạn phân chia nhanh, ribosom chiếm tới 1/3 khối lượng tế bào. Khoảng 1/2 khối lượng ribosom là ARN nên rARN chiếm đến 75% của tổng số ARN.

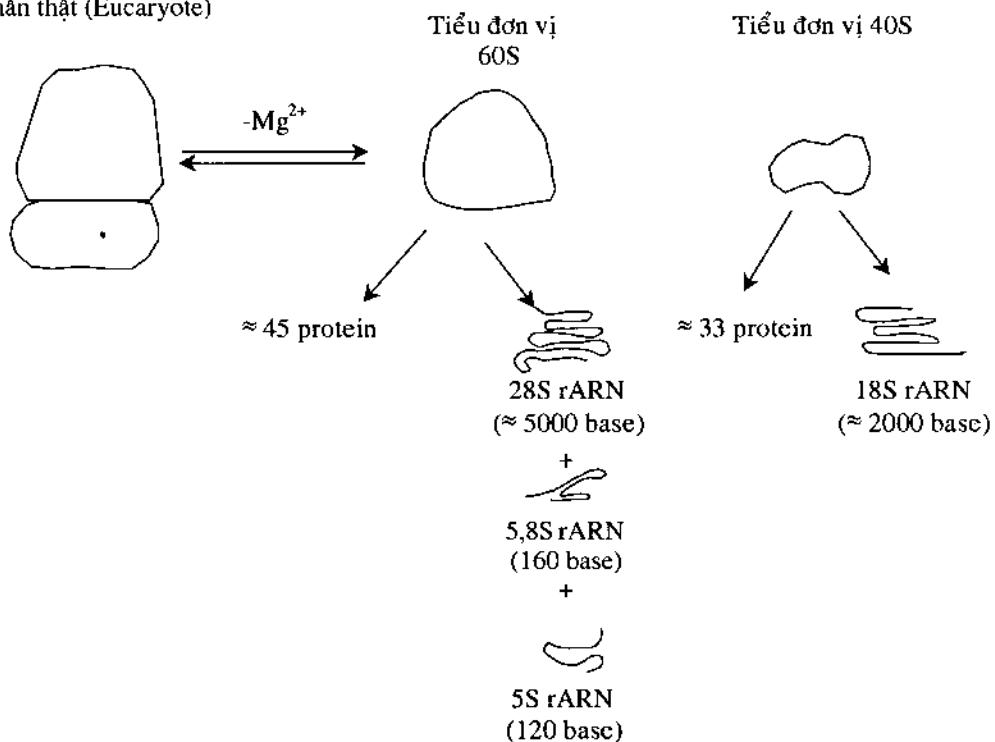
Mỗi một ribosom có $Mr = 2.5 - 4.5 \times 10^6$. Có 3 loại ribosom khác nhau, thường được phân biệt dựa vào hệ số lắng của chúng khi siêu ly tâm. Ribosom của *Eubacteria* và lục lạp có chỉ số lắng khoảng 70S. Ribosom của tế bào nhân thật là 80S, của ty thể động vật có vú khoảng 50S.

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 tuỳ theo nồng độ của Mg^{2+} (hình 3.1).

Ribosom 70S của tế bào nguyên thuỷ (Prokaryote)



Ribosom 80S của tế bào nhân thật (Eucaryote)



Hình 3.1. Thành phần của ribosom 70S (tế bào nhân nguyên thuỷ) và 80S (tế bào nhân thật)

Các ribosom của lục lạp và của tế bào nhân nguyên thuỷ có hệ số lắng khi ly tâm là 70S gồm 2 tiểu đơn vị:

- Đơn vị lớn 50S chứa 1 rARN 23S và 1 rARN 5S và 34 protein.
- Đơn vị nhỏ 30S chỉ chứa 1 rARN 16S và 21 – 23 protein.

Các ribosom của tế bào nhân thật có hệ số lắng khi ly tâm là 80S gồm 2 tiểu đơn vị:

- Đơn vị lớn 60S chứa 3 loại rARN khác nhau: 1 rARN 28S, 1 rARN 5,8S, 1 rARN 5S và 45 protein.
- Đơn vị nhỏ 40S chỉ có 1 rARN 18S và 33 protein.

3.2.1.1. Ribosom *E. coli*

Ribosom ở *E. coli* được nghiên cứu 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 50S được đặt tên từ L₁ – L₃₄. 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. Các protein L₇ và L₁₂ có những trình tự đã được xác định, L₇ là một dạng aminoacetyl hoá L₁₂ (sự acetyl hoá 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 protein. Các tiểu đơn vị ribosom có thể tái hợp lại 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 ái lực yếu không hoá trị giữa protein và các phân tử ARN, làm cho các protein trở nên không bền vững với ribosom. ARN và các protein không liên kết được tách ra khỏi nhau. Chúng 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.

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 lạp 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 3.1. 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

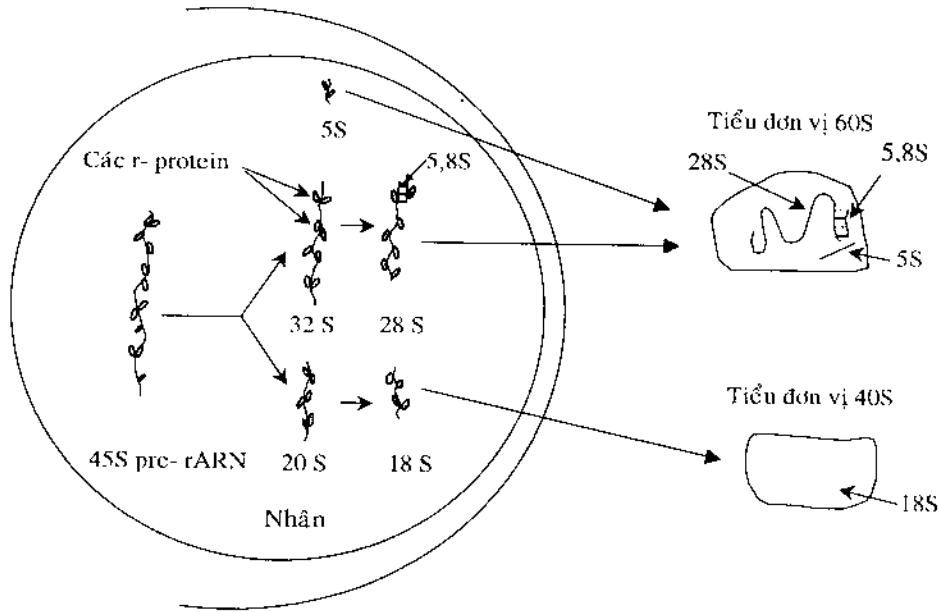
Cấu trúc bậc 2 chính xác của rARN có tầm quan trọng trong việc lắp ráp ribosom và chúng được lưu trữ tốt trong quá trình tiến hoá. Tất cả các protein của *E. coli* được gắn kết vào rARN cũng có cấu hình bậc hai như vậy và nếu rARN bị biến tính thì protein không gắn kết được.

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 – methylcytosine), còn rARN của tế bào chất tế bào nhân thật thường methyl hoá ở vị trí 2' của đường ribose (2' – O – methyladenosyl). Các vị trí của phân 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 tạo ra các ARN trưởng thành có khối lượng tương ứng với từng tiểu đơn vị.

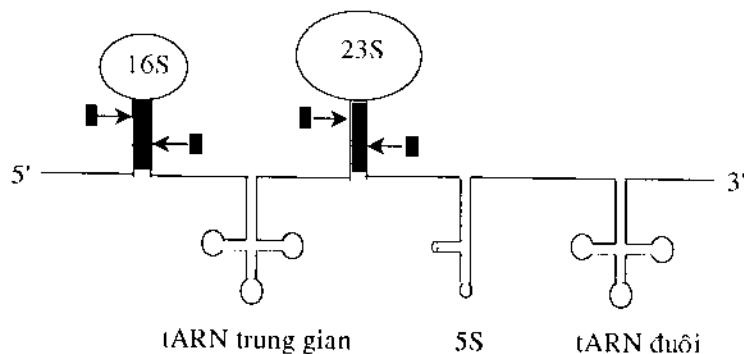
Cách biến đổi của rARN ở tế bào HeLa (người) được minh họa ở hình 3.2.



Hình 3.2. Quá trình sinh tổng hợp các tiểu đơn vị ribosom ở các tế bào HeLa

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 phân 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 thuỷ. Ở tế bào nhân nguyên thuỷ, 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 thuỷ 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 3.3).



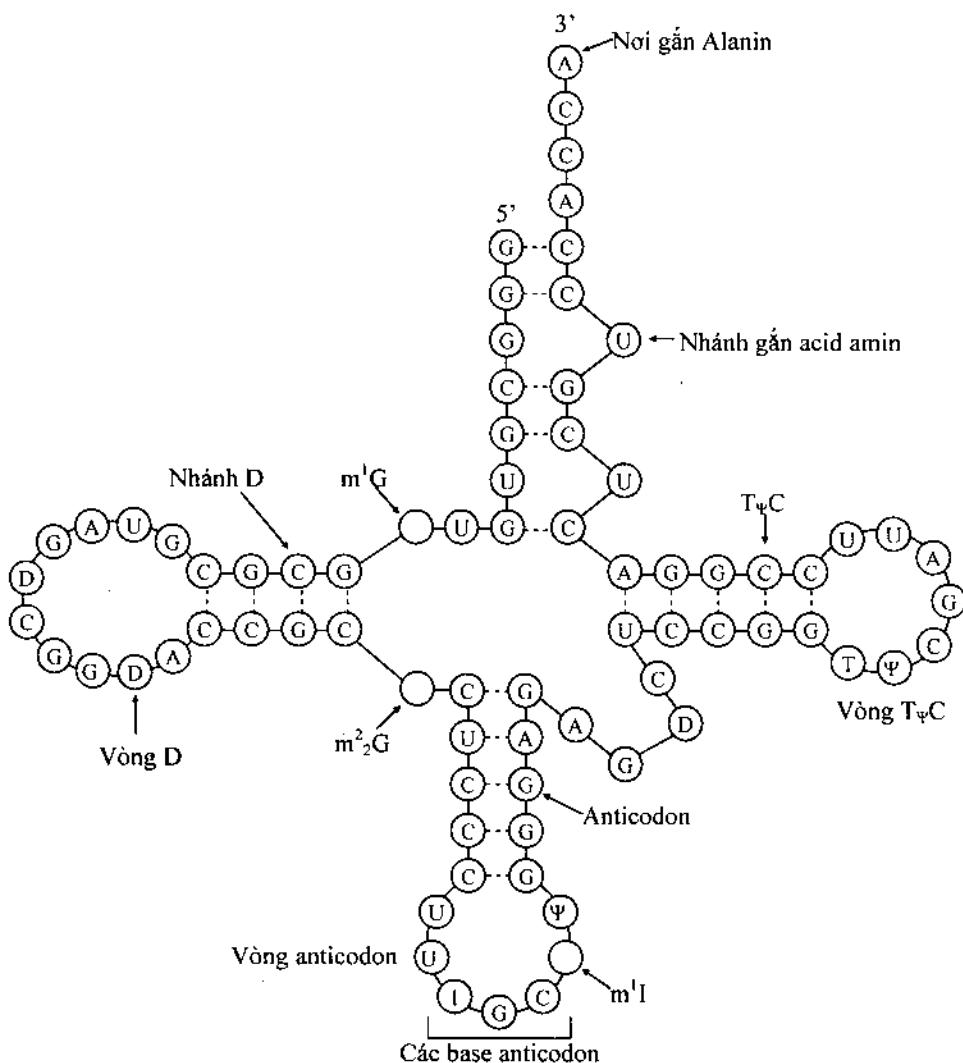
Hình 3.3. Quá trình cắt nối của pre – rARN ở *E. coli* dài khoảng 5000 nucleotid.

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 mARN. 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.

Ít nhất, mỗi loại tARN đặc hiệu cho một loại amino acid. Ví dụ: Arginine chỉ gắn đặc hiệu với ARN vận chuyển arginine. 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 chè ba (hình 3.4). Đầu mút 3' có trình tự kết thúc CCA.



Hình 3.4. Cấu trúc alanin – tARN

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

Các nucleotid thường gặp:	I = Inosine
	T = Ribothymidine
	ψ = pseudouridine
	m^1G = methyl guanosine
	m^2G = dimethylguanosine
	m^1I = methylinosine
	D = dihydrouridine

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

3.2.2.2. Biến đổi tARN

Ở tế bào nhân nguyên thuỷ 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 nội dài Tu.

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

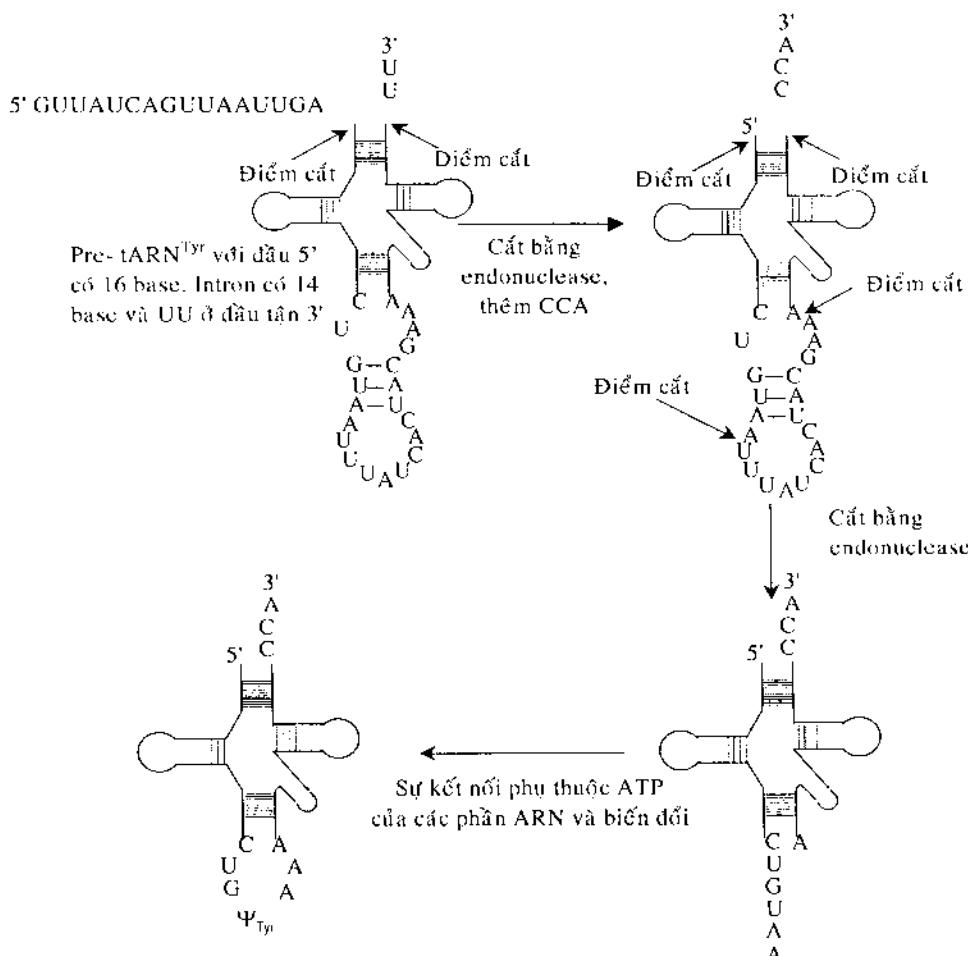
RNse P là một enzym, gồm 1 ARN có 375 nucleotid và một protein có trọng lượng phân tử 20 000. Trong các điều kiện sinh lý, cả 2 thành phần protein và ARN đều cần những yếu tố để hoạt động, nhưng khi nồng độ Mg⁺⁺ cao thì chỉ có thành phần ARN có hoạt tính xúc tác. Theo sau hoạt động của RNse P, một số 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 3.5).

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 thuỷ)
- Gắn những yếu tố nội dài
- Gắn ribosom
- Nhận diện codon – anticodon.



Hình 3.5. Quá trình biến đổi của tARN^{Tyr} ở nấm men

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 đằng (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 được tổng hợp khi enzym ARN polymerase II tiếp xúc với trình tự promoter trên phân tử ADN. mARN nguyên vẹn của tế bào vi khuẩn và của tế bào nhân thật chứa trình tự nucleotid nhiều hơn số được dùng để mã hoá protein.

3.2.3.1. Cấu trúc

Phân tử mARN gồm 3 phần:

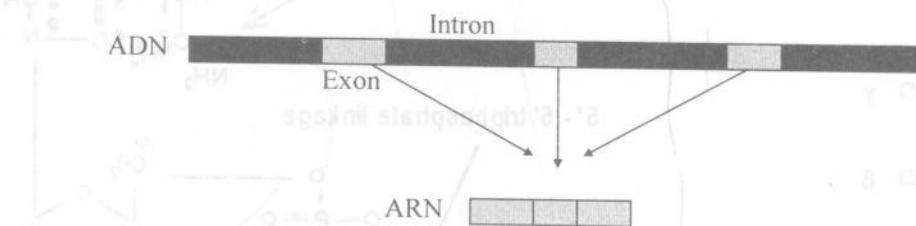
1. Đoạn 5' không mã hoá

2. Vùng mã hoá
3. Đoạn 3' không mã hoá

ADN polymerase khởi đầu sự phiên mã ở một đoạn nằm ngay trước vùng mã hoá cho phân tử protein được gọi là đoạn 5' không mã hoá. Do đó, mARN có đoạn đầu mang các tín hiệu cho ribosom nhận biết để gắn vào khi dịch mã.

Ở đuôi 3', sau dấu kết thúc có đoạn 3' không mã hoá là nơi gắn poly A (tế bào nhân thật).

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



Hình 3.6. Intron và exon ở tế bào nhân thật

Đoạn 5' và 3' là trình tự khởi đầu và kết thúc không được dịch mã, 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 hoá protein. Các trình tự này được phiên mã từ sợi ADN, ký hiệu 5'UTR và 3'UTR. 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 hoá, do vậy sẽ tổng hợp được nhiều protein hơn. Nếu không có các đoạn UTR, ARN sẽ bị thoái hoá 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.

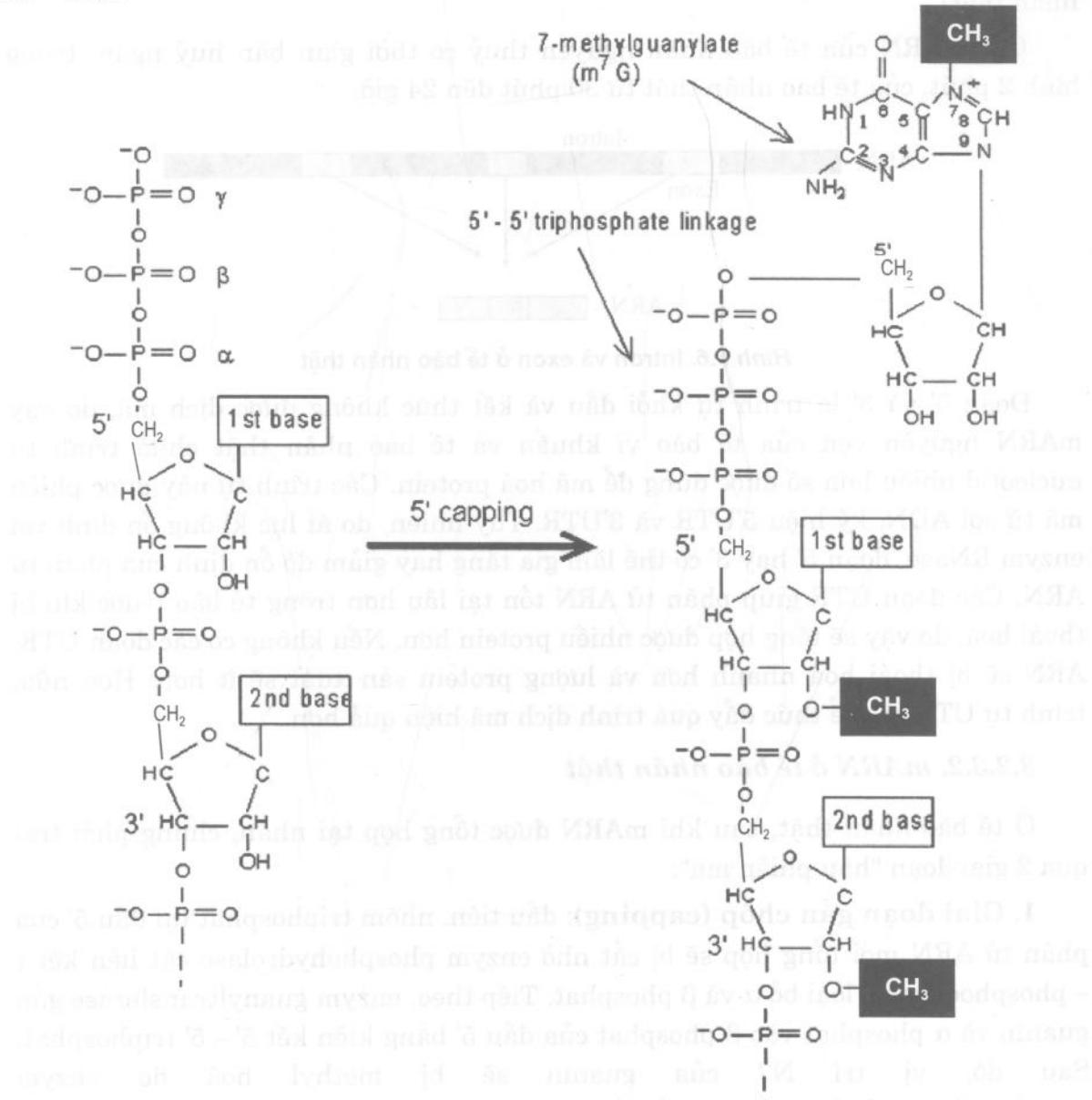
3.2.3.2. mARN ở tế bào nhân thật

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 chót (capping): Đầu tiên, nhóm triphosphat tại đầu 5' của phân tử ARN mới tổng hợp sẽ bị cắt nhở 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 kiêm 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 chót có thể xảy ra theo 1 trong 3 cách:

- a) Chớp 0: nhóm methyl gắn vào vị trí G7 do enzym guanin – 7 – methyl transferase xúc tác. Phản ứng này xảy ra ở tất cả mARN tế bào nhân thât.
- b) Chớp 1: nhóm methyl gắn vào vị trí 2' – OH đường ribose của nucleotid đầu tiên. Phản ứng do enzym 2' – O – methyl transferase xúc tác, xảy ra ở hầu hết mARN tế bào nhân thât.
- c) Chớp 2: phản ứng methyl hoá tương tự xảy ra ở nucleotid thứ 2 với tỉ lệ 10 – 15%.



Hình 3.7. Gắn chớp ở mARN của tế bào nhân thât

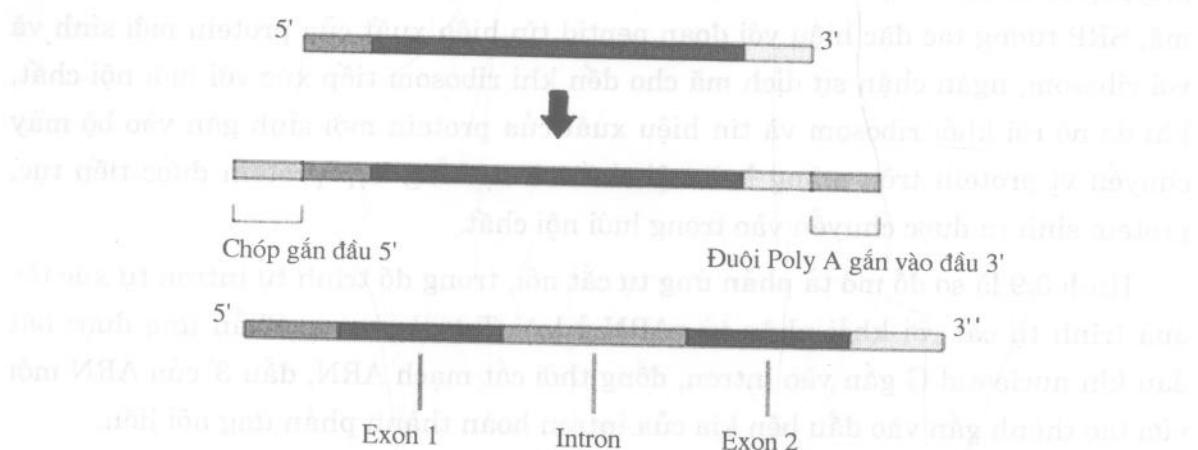
1st base: base ở cacbon vị trí số 1; 2nd base: base ở cacbon vị trí số 2; 5'capping:

gắn chớp ở đầu 5'; 5' – 5' triphosphaste linkage: liên kết triphosphat 5' – 5'.

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 chớp 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 chớp 5' trong việc bảo vệ phân tử ARN và gắn với ribosom. Do vậy, thời gian bán huỷ 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.



Hình 3.8. Cấu trúc mARNcủa tế bào nhân thật

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.

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

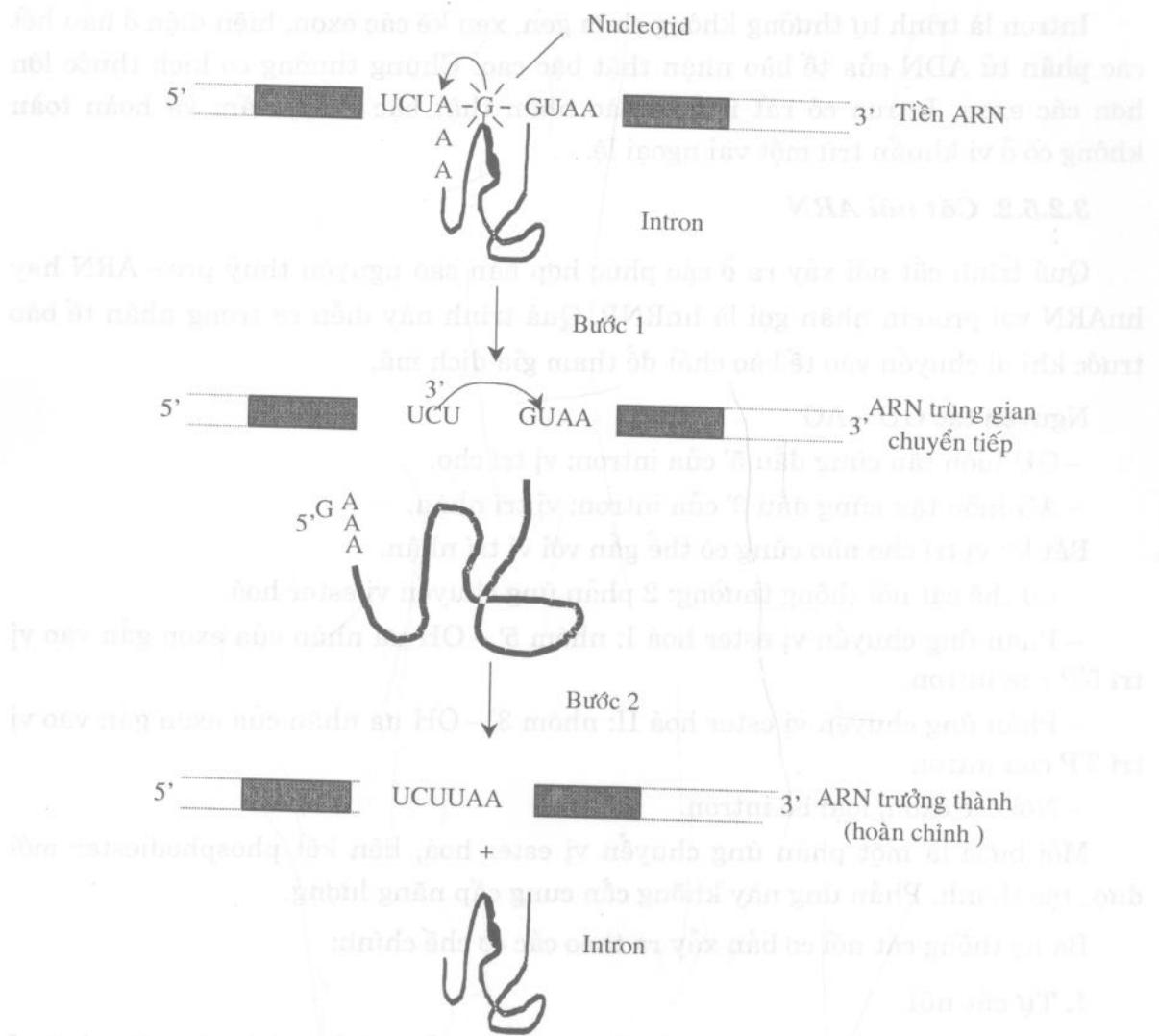
nên các hạt ribonucleoprotein (RNP) được gọi là snRNP hoặc scRNP tương ứng cho nhân hay cho tế bào chất. Chúng dài khoảng 90 – 300 nucleotid.

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 U₃ 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 U₁ – U₇ snRNP.

scRNP có vai trò biệt 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.

Hình 3.9 là sơ đồ mô tả phản ứng tự cắt nối, trong đó trình tự intron tự xúc tác quá trình tự cắt rời khỏi phân tử rARN ở loài *Tetrahymena*. Phản ứng được bắt đầu khi nucleotid G gắn vào intron, đồng thời cắt mạch ARN, đầu 3' của ARN mới vừa tạo thành gắn vào đầu bên kia của intron hoàn thành phản ứng nối liền.

Chỉ có một tỉ lệ nhỏ các ADN sinh vật tế bào nhân thật mã hoá protein. Các ADN dư thừa như vậy được phát hiện giữa các gen cấu trúc. Trong số những người đầu tiên phát hiện các đoạn ADN dư thừa này (intron) có Pierre Champon và đồng sự ở Đại học Strasbourg (Pháp) qua nghiên cứu sự di truyền trong tổng hợp ovalbumin lòng trắng trứng gà. Người ta thấy trong nhân tế bào chỉ có những phân tử lớn hơn mARN hoàn chỉnh trong tế bào chất rất nhiều và gọi chúng là những ARN nhân không đồng nhất, hnARN. So sánh ADN của gen và ADN phóng xạ được tổng hợp từ sao chép ngược của mARN trong tế bào chất, Champon nhận xét rằng các đoạn nucleotid dài có mặt trong ADN của gen ban đầu không thấy trong cADN phóng xạ của mARN tế bào chất. Các đoạn dài này không dịch mã, xuất hiện không chỉ ở đoạn đầu hoặc cuối của gen mà ở cả những đoạn giữa, làm gián đoạn các đoạn mã hoá protein (exon). Gen mã hoá ovalbumin bị gián đoạn bởi 7 intron.



Hình 3.9. Phản ứng tự cắt nối (self – splicing)
của rARN ở loài *Tetrahymena*

3.2.5. Tóm tắt các quá trình cắt nối ở ARN tế bào nhân thực

3.2.5.1. Khái niệm

Bộ gen tế bào nhân nguyên thuỷ là bộ gen gián đoạn. Trình tự không mã hoá xen kẽ các trình tự mã hoá. Quá trình loại bỏ các trình tự không mã hoá gọi là quá trình cắt nối ARN.

Exon là trình tự biểu hiện, có ở mARN hoàn chỉnh. Tuy nhiên, 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 hoá.

Quá trình loại bỏ các trình tự không mã hoá khỏi mARN để tạo ra mARN có khả năng biểu hiện protein được gọi là quá trình cắt nối ARN.

Intron là trình tự thường không chứa gen, xen kẽ các exon, hiện diện ở hầu hết các phân tử ADN của tế bào nhân thật bậc cao. Chúng thường có kích thước lớn hơn các exon. Intron có rất ít ở tế bào nhân thật bậc thấp, nấm và hoàn toàn không có ở vi khuẩn trừ một vài ngoại lệ.

3.2.5.2. Cắt nối ARN

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

Nguyên tắc GU – AG

- 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.

- Cơ chế cắt nối thông thường: 2 phản ứng chuyển vị ester hoá.
 - Phản ứng chuyển vị ester hoá I: nhóm 5' – OH ưa nhân của exon gắn vào vị trí 5'P của intron.
 - Phản ứng chuyển vị ester hoá II: nhóm 3' – OH ưa nhân của exon gắn vào vị trí 3'P của intron

– Nối các exon, loại bỏ intron.

Mỗi bước là một phản ứng chuyển vị ester hoá, liên kết phosphodiester mới được tạo thành. Phản ứng này không cần cung cấp năng lượng.

Ba hệ thống cắt nối cơ bản xảy ra theo các cơ chế chính:

1. Tự cắt nối

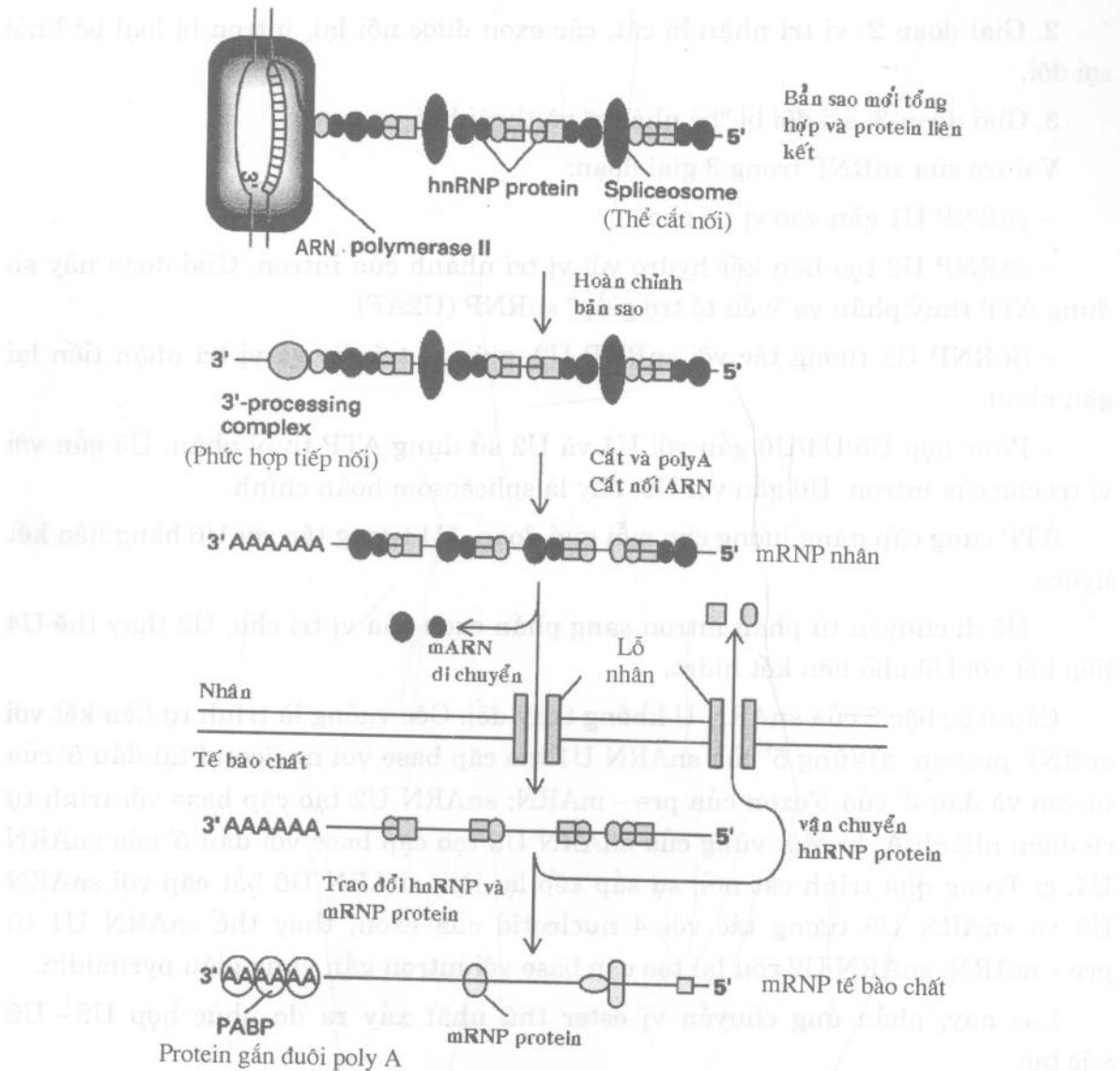
ARN đóng vai trò enzym, gọi là ribozym. Loại này có ở 2 nhóm intron: nhóm I và nhóm II

- Nhóm intron I: tự cắt nối ở vòng 9 của cấu trúc bậc 2. Quá trình này xảy ra ở gen rARN của tế bào nhân thật bậc thấp và nấm.
- Nhóm intron II: tương tự cơ chế cắt nối bằng spliceosom nhưng được xúc tác bởi các intron ARN.

2. Intron được loại bỏ khỏi tiền tARN nhân của nấm men tương tự quá trình xử lý tiền tARN.

3. Cắt nối nhờ spliceosom

Spliceosom là phức hợp chứa nhiều loại ribonucleoprotein. Chúng nhận diện các trình tự exon – intron để loại bỏ intron và nối các exon. Quá trình này thường xảy ra ở ARN nhân của tế bào nhân thật bậc cao



Hình 3.10. Các protein tham gia tổng hợp và tạo mARN hoàn chỉnh

3.2.5.3. Cắt nối mARN ở tế bào nhân thực

Spliceosom

Spliceosom cấu tạo bởi khoảng 40 protein quy định cắt nối và các snARN. Các phức hợp ribonucleoprotein này được gọi là snRNP hay snURP. 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 thực.

Phản ứng cắt nối: gồm 3 giai đoạn chính

1. Giai đoạn 1: vị trí cho bị cắt, đầu 5' gắn vào vị trí nhánh của intron bằng liên kết 5' – 2' phosphodiester tạo sợi đôi tạm thời. Vị trí nhánh gồm khoảng 30 nucleotid ở nhánh lên của vị trí nhận, có trình tự tương đồng.

2. Giai đoạn 2: vị trí nhận bị cắt, các exon được nối lại, intron bị loại bỏ khỏi sợi đôi.

3. Giai đoạn 3: sợi đôi bị "bẻ nhánh" và thoái hoá.

Vai trò của snRNP trong 3 giai đoạn:

- snRNP U1 gắn vào vị trí cho.

- snRNP U2 tạo liên kết hydro với vị trí nhánh của intron. Giai đoạn này sử dụng ATP thuỷ phân và "yếu tố trợ giúp" snRNP (U2AF).

- SnRNP U1 tương tác với snRNP U2, giúp vị trí cho và vị trí nhận tiến lại gần nhau.

- Phức hợp U5/U4/U6 gắn với U1 và U2 sử dụng ATP thuỷ phân; U5 gắn với vị trí cho của intron; U6 gắn với U2: đây là spliceosom hoàn chỉnh.

ATP cung cấp năng lượng cho mỗi giai đoạn. U4 tương tác với U6 bằng liên kết hydro.

- U5 di chuyển từ phần intron sang phần exon của vị trí cho, U2 thay thế U4 liên kết với U6 nhờ liên kết hidro.

Cấu trúc bậc 2 của snARN U không thay đổi. Góc vuông là trình tự liên kết với snRNP protein: a) Vùng 5' của snARN U1 tạo cặp base với nucleotid tại đầu 5' của intron và đầu 3' của 5'exon của pre – mARN; snARN U2 tạo cặp base với trình tự có điểm nhánh A. b) Một vùng của snARN U6 tạo cặp base với đầu 5' của snARN U4. c) Trong quá trình cắt nối, sự sắp xếp lại làm snARN U6 bắt cặp với snARN U2 và snARN U5 tương tác với 4 nucleotid của exon, thay thế snARN U1 từ pre – mARN. snARN U2 còn lại tạo cặp base với intron gần vùng giàu pyrimidin.

Lúc này, phản ứng chuyển vị ester thứ nhất xảy ra do phức hợp U2 – U6 xúc tác.

- snRNP U5 tương tác với phần intron của vị trí nhận và U2, xúc tác phản ứng cắt liên kết intron – exon và nối 2 exon. Phản ứng chuyển vị ester thứ 2 xảy ra.

- Spliceosom được tháo rời, giải phóng U1, U2, U4, U5, U6.

- Sợi đôi intron được tháo xoắn và thoái hoá.

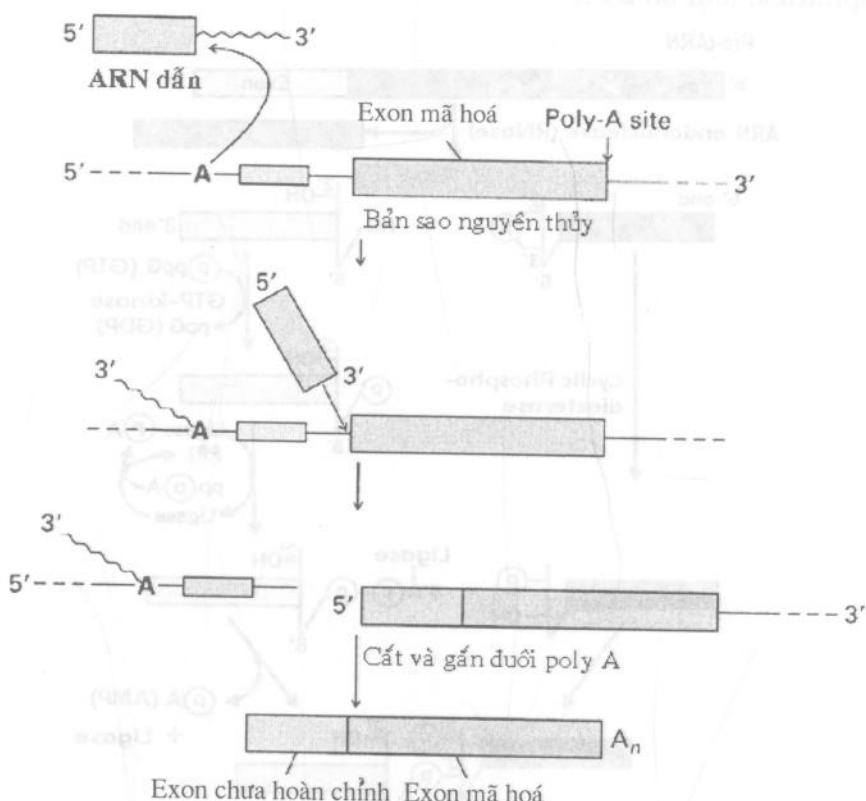
Intron nhóm II: ít gấp

Đây là phản ứng tự cắt nối, intron đóng vai trò như một enzym – gọi là ribozym.

Cắt nối chéo

Các phản ứng cắt nối trên xảy ra trên cùng một phân tử ARN, gọi là cắt nối kiểu cis; nếu xảy ra giữa các phân tử gọi là cắt nối kiểu trans.

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



Hình 3.11. Cắt nối chéo của ARN dẫn với exon mã hoá protein ở bản sao nguyên thuỷ polycistron loài Trypanosoma nhờ hai phản ứng chuyển vị ester hoá

3.2.5.4. Xử lý tARN

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

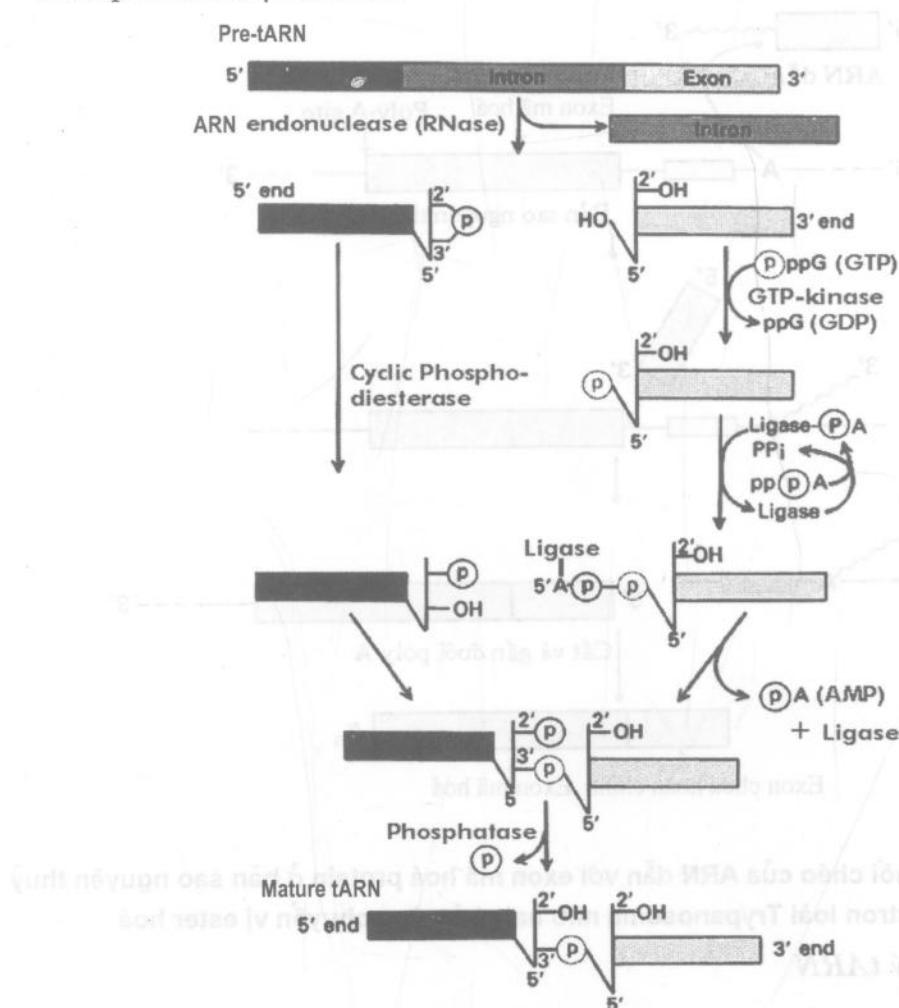
Xử lý đầu tận cùng 5' và 3'

- Xử lý đầu 5': enzym xúc tác là ribonuclease (RNseP)
- 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.

Cắt nối tARN ở nấm men

- ARN endonuclease (RNse) 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.



Hình 3.12. Cắt nối tARN ở nấm men

3.2.5.5. Xử lý rARN

Cắt nối theo cơ chế nhóm intron I ở tế bào nhân thực bậc thấp, nấm và 3 gen của thực khuẩn T4 ở *E.coli*. Có 2 tính chất chung:

- Tự cắt nối
- Cấu trúc bậc 2 của vòng 9

Cắt nối ở loài *tetrahymena*

Tiền rARN 35S được xử lý thành 18S, 5,8S, 26SrARN. Ở một vài chủng *Tetrahymena*, 26S có một intron gồm 400 base. Intron này có thể tự cắt nối nhưng đòi hỏi sự có mặt của các ion và nucleotid G (GTP, GDP, GMP...)

Phản ứng tự cắt nối: gồm 3 bước

1. G 3' – OH gắn đầu 5' của intron.
2. 3' – OH của exon A gắn đầu 5' của exon B; các exon được nối với nhau.
3. Intron mạch thẳng với đầu 5' tận cùng bằng G (414) sẽ được đóng vòng.

Đóng vòng intron mạch thẳng

4. Nhóm 3' – OH của intron tại G (414) gắn với pA (16) hay pU (20), tạo vòng intron và khởi đầu 5' 15 hay 19 nucleotid.
5. Nước thuỷ phân vòng intron tại cùng vị trí.
6. Nếu quá trình đóng vòng đầu tiên xảy ra tại pA (16) thì nhóm 3' – OH của intron gắn với pU (20), tạo vòng intron và cắt 3 nucleotid ở mạch thẳng.
7. Nước thuỷ phân vòng intron tại cùng vị trí, tạo L – 19ARN. L – 19ARN có hoạt tính enzym xúc tác phản ứng cắt và nối các oligonucleotid.

Vai trò cấu trúc bậc 2 của intron ARN, vòng 9

8. Nhánh P1 (exon 1 và IGS – InteARNI Guide Sequence) đóng vai trò quan trọng trong bước 1 của phản ứng cắt nối.
9. Nhánh P4 và P7 đóng vai trò quan trọng nhất trong bước 2 của phản ứng cắt nối, ngoài ra còn có P3, P6 và 2 base ở đầu 3' của intron.

Vai trò xúc tác của L – 19ARN

10. Gắn ARN với IGS – vị trí gắn kết.
11. Gắn G (414) hay G tự do vào vị trí gắn kết.

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 thymine
 - d) Được tổng hợp từ trong nhân
 - e) Có liên kết hydro giữa A = T
2. Cấu tạo từ 34 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 45 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) 34 phân tử protein + 1 rARN 23S, 1 rARN 5S
 - b) 21 phân tử protein + 1 rARN 16S
 - c) 45 phân tử protein + 1 rARN 28S, rARN 5,8S, rARN 5S
 - d) 33 phân tử protein + 1 rARN 18S
 - e) 45 phân tử protein + 1 rARN 23S + 1 rARN 5S
5. Tiểu đơn vị 30S của tế bào nhân nguyên thuỷ cấu tạo từ:
 - a) 34 phân tử protein + 1 rARN 23S, 1 rARN 5S
 - b) 21 phân tử protein + 1 rARN 16S
 - c) 45 phân tử protein + 1 rARN 28S, rARN 5.8S, rARN 5S

Bài 4

SỰ PHIÊN MÃ VÀ MÃ DI TRUYỀN

MỤC TIÊU

- Trình bày được quá trình phiên mã ở tế bào nhân nguyên thuỷ và tế bào nhân thật.
- Mô tả được sự phiên mã có điều hoà.
- Trình bày được sự phiên mã ngược, nếu 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 thuỷ, 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 thuỷ.

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.

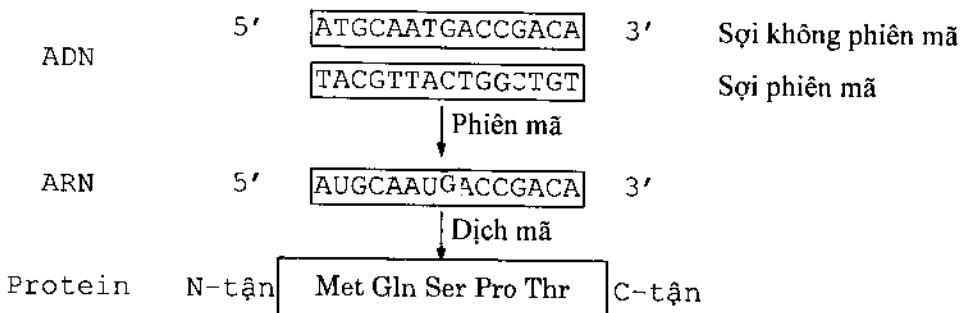
4.2. NGUYÊN TẮC CHUNG

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 di 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' → 5' trên ADN để cho "ARN" được tổng hợp theo hướng 5' → 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 thuỷ phân pyrophosphat vừa được giải phóng. Bằng cách này ARN được tiếp tục kéo dài và một tổ hợp lai 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'. Khi tổ hợp lai ADN – ARN đạt được chiều dài 12 base, đầu tận 5' của chuỗi ARN mới hình thành được tách ra từ mạch ADN bổ sung với nó và 2 mạch kép của ADN ban đầu phục hồi lại. ARN polymerase tiếp

tục di chuyển dọc theo mạch ADN khuôn mẫu đến khi gặp điểm kết thúc, đó là tín hiệu được mã hoá trong một trình tự ADN quy định chấm dứt sự nối dài của mạch. 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 THUỶ

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.

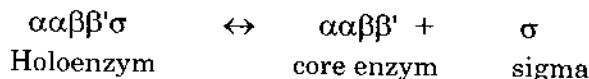
Hầu hết sự hiểu biết về quá trình phiên mã ở tế bào nhân nguyên thuỷ xuất phát từ sự nghiên cứu trên *E. coli* và thực khuẩn của nó. Cơ chế của sự phiên mã ở *E. coli* tương tự ở hầu hết tế bào nhân nguyên thuỷ, ngoại trừ *Archaeabacteria* (có intron trong ADN).

Nhiễm sắc thể đơn bội của *E. coli* là một vòng khép kín chứa $4 \cdot 10^6$ bp (cặp base) cùng với protein và ARN. 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 450000. Toàn bộ enzym được gọi là holoenzym chứa các tiểu đơn vị $\alpha\beta\beta'\sigma$.

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



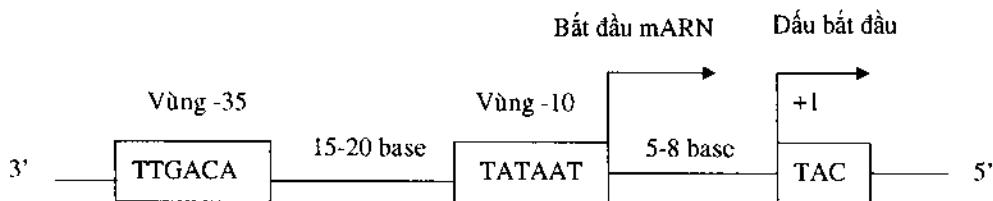
4.3.2. Promoter

Phần lớn promoter ở *E. coli* về cơ bản đều giống nhau. Nếu ký hiệu base đầu tiên phiên mã thành mARN (thường là Adenin) là + 1, thì các base theo hướng ngược lại với chiều phiên mã ký hiệu " - ".

Promoter gồm hai trình tự. Đó là trình tự TATAAT, tận cùng luôn luôn là base T, cách điểm xuất phát 6 – 8 base và có base trung tâm là –10. Trình tự này còn được gọi là hộp TATA hay hộp Pribnow; trình tự thứ 2 TTGACA có base trung tâm –35 (hình 4.2). Hai trình tự này chịu trách nhiệm cho phép ARN polymerase gắn vào vùng khởi động phiên mã. Vùng –35 tham gia vào việc gắn khởi đầu ARN polymerase (bảng 4.1). 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 (consensus).

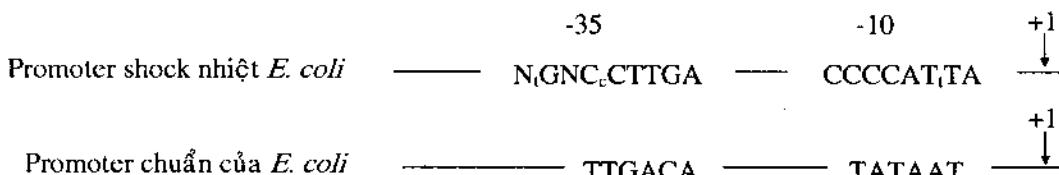
Bảng 4.1. Các trình tự chung –10 và –35 của một số promoter của *E. coli*

Promoter	–35	–10	Điểm xuất phát
Consensus	TTGACA	TATAAT	
<i>lac P</i>	TTTACA ← 18 bp →	TATGTT	← 6 bp → A
<i>Gal P</i> ₁ (Ức chế đồng hoá)		TATGGT	6 A
<i>Gal P</i> ₂ (Ức chế dị hoá)	GTCAC 16	TATGCT	6 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 hiệu ứng đột biến (một base được thay thế bởi một base khác) trong quá trình phiên mã từ promoter (hình 4.3). Nhiều đột biến như vậy ức chế tức khắc sự phiên mã của gen. Thật ra, vài vùng khởi động (promoters) ở *E. coli* có trình tự chung rất giống với trình tự chung lý tưởng. Do đó, để phân biệt được chúng có thể dựa vào cường độ (strength) hoặc tần suất khởi động phiên mã. Ví dụ, tế bào *E. coli* chứa khoảng 12 phân tử của lac repressor (thể 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 10000 ribosom phải được hình thành sau mỗi thế hệ tế bào (khoảng 20 phút). Nên mỗi gen – rARN được phiên mã khoảng 2 giây một lần. Sự khác biệt này do cường độ vùng khởi động.



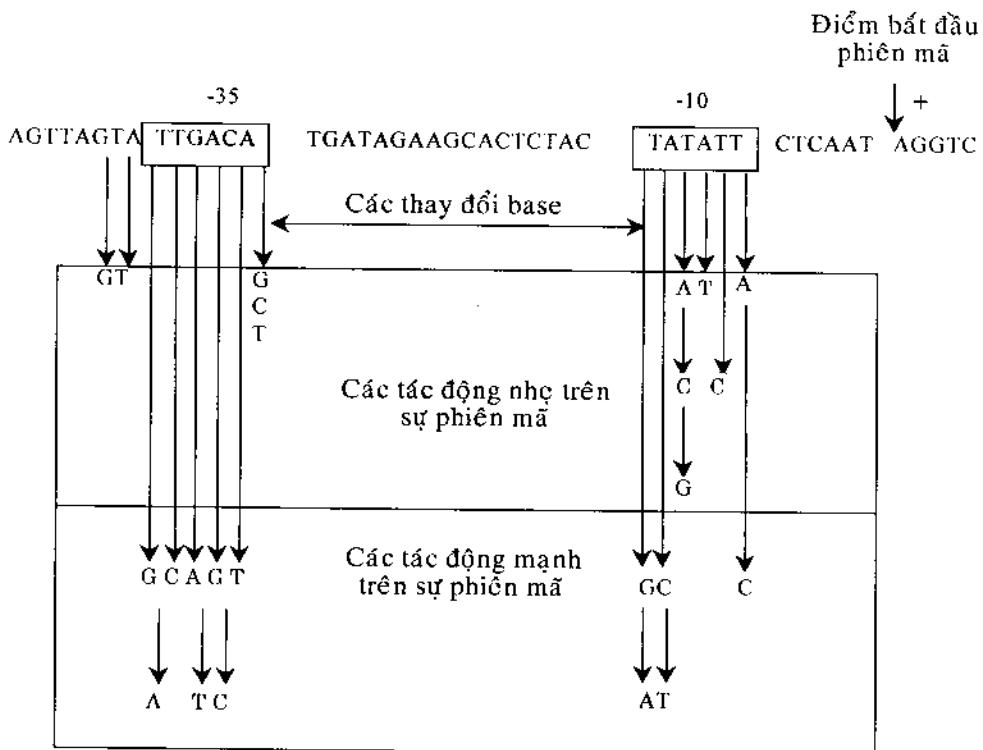
Hình 4.3. Các vị trí đột biến điểm tác động lên chức năng vùng khởi động ở *E. coli*

4.3.3. Các promoter shock nhiệt

Mãi cho tới gần đây, người ta vẫn cho rằng ARN polymerase ở *E. coli* có một loại holoenzym duy nhất có khả năng phiên mã tất cả các gen khác nhau. Tuy nhiên, những nghiên cứu gần đây trên sự đáp ứng shock nhiệt đã cho thấy rằng khả năng phiên mã các gen khác nhau có thể đạt được do sinh ra các tiểu đơn vị 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 thuỷ. 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 tạo ra bởi sản phẩm của gen *hptR*. Gen này đáp ứng với shock nhiệt và mã hoá cho

chuỗi polypeptid có Mr 32 000, đóng vai trò như một yếu tố σ biến đổi. Yếu tố này (được gọi σ³²) kết hợp với phần lõi bình thường của ARN polymerase để tạo nên holoenzym α₂ββ'σ³². 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.



Hình 4.4. Các trình tự chung của promoter sốc nhiệt ở *E. coli*

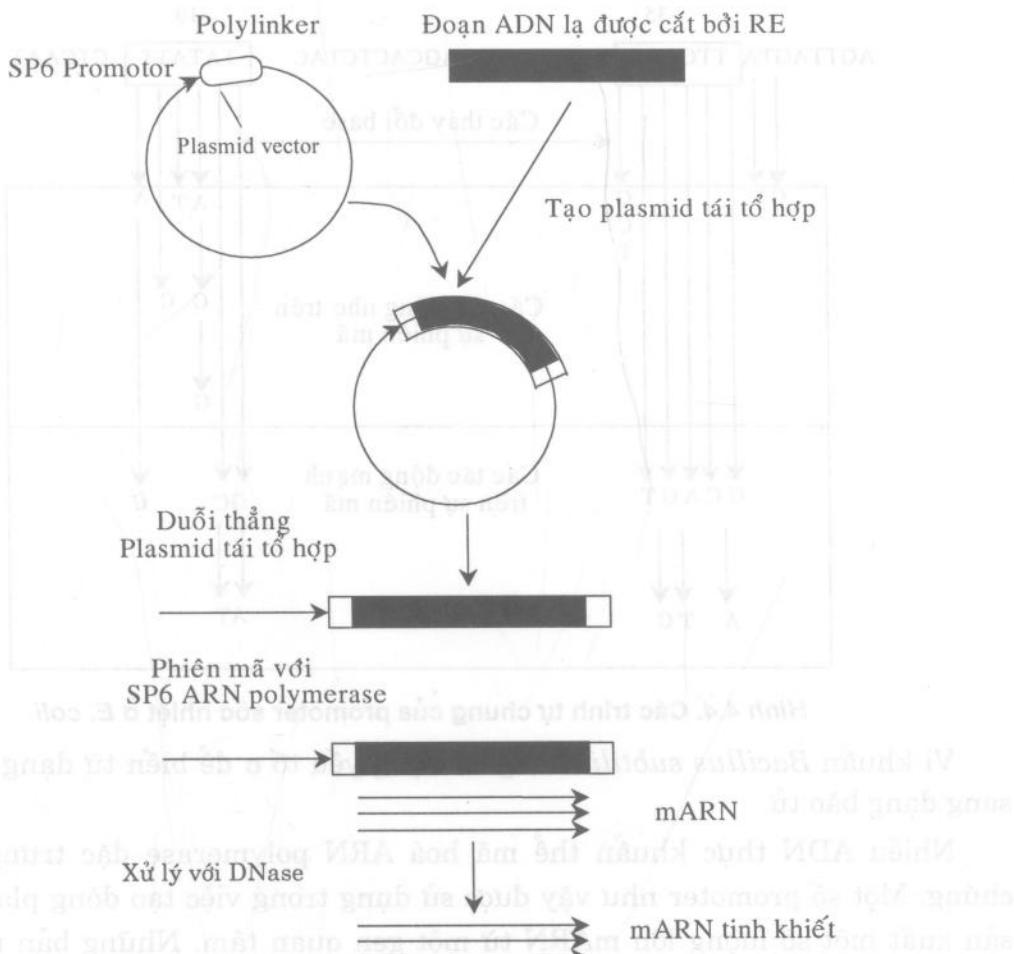
Vì khuẩn *Bacillus subtilis* cũng sử dụng yếu tố σ để biến từ dạng sinh dưỡng sang dạng bào tử.

Nhiều ADN thực khuẩn thể mã hoá ARN polymerase đặc trưng cho riêng chúng. Một số promoter như vậy được sử dụng trong việc tạo dòng plasmid nhằm sản xuất một số lượng lớn mARN 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.

Trình tự chung của promoter sốc nhiệt được sắp xếp cho 6 gen sốc nhiệt khác nhau. Những chữ cái ở dòng trên chỉ sự tương đồng mạnh, ở hàng dưới tương đồng yếu, n là nucleotide bất kỳ. Vị trí bắt đầu phiên mã ký hiệu +1.

Người ta thường dùng hệ thống promoter và ARN polymerase lấy từ thực khuẩn thể SP₆. SP₆ là một thực khuẩn thể của *Salmonella typhimurium*, mã hoá được một ARN – polymerase virus đặc hiệu. Các promoter của tế bào chủ và các

promoter có trong plasmid ly trich từ *E. coli* đều không được SP₆ polymerase nhận biết. Promoter SP₆ được chèn vào vector plasmid phía trước polylinker (polylinker là đoạn ADN nhân tạo chứa nhiều vị trí nhận biết duy nhất của các enzym cắt hạn chế). Đoạn ADN cần sao chép được chèn vào polylinker và plasmid tái tổ hợp được biến nạp vào *E. coli*. Plasmid chứa ADN lạ được "làm đuôi thẳng" bởi một enzym cắt hạn chế. Plasmid đuôi thẳng này được ủ với SP₆ polymerase và các NTP. Các plasmid – ADN được loại khỏi các bản sao mARN bằng cách ủ với DNase (hình 4.5).

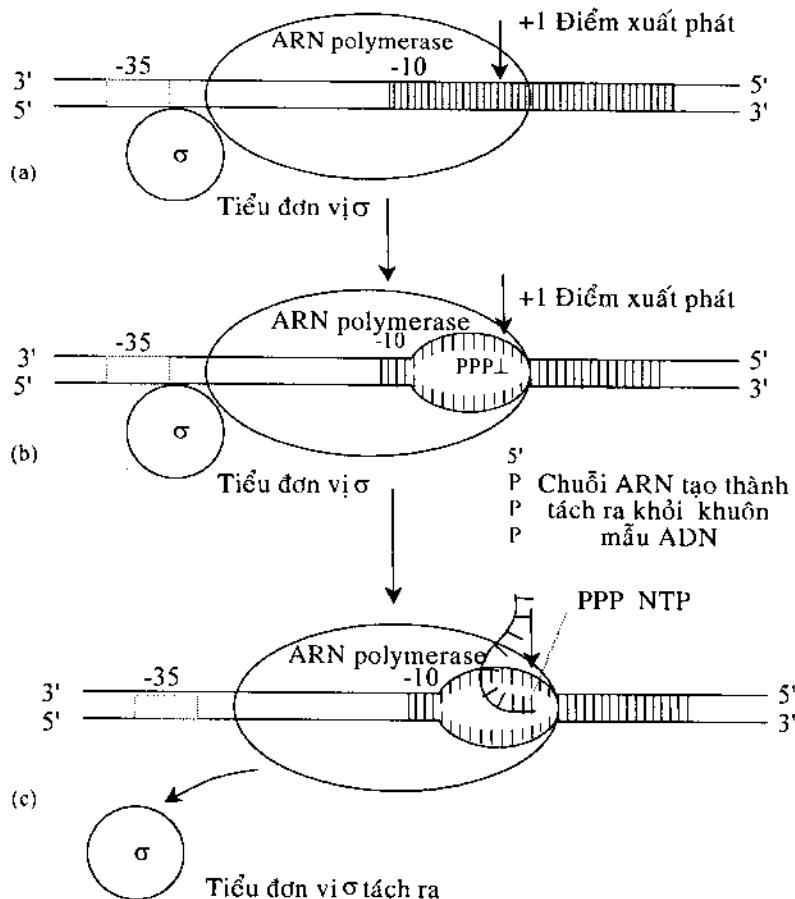


Hình 4.5. Dùng SP₆ promoter và ARN polymerase của nó để nhân bản sao mARN cần thiết

4.3.4. Khởi đầu và nối dài chuỗi ARN

Việc gắn của holoenzym ARN polymerase vào promoter tạo ra phức hợp khởi động đóng có cấu hình ADN xoắn kép (hình 4.6). Các điểm tiếp xúc chủ yếu giữa promoter và ARN polymerase nằm ở các vùng -10 và -35. Phức hợp đóng nhanh chóng biến thành phức hợp mở do các liên kết hydro bị enzym cắt trên một vùng

bao gồm 17 bp. Sợi ADN được tách ra nhờ sự tháo xoắn 1,6 vòng của xoắn kép B – ADN. 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. Như vậy, chuỗi mARN phát triển mang triphosphat ở đầu 5' và theo hướng 5' → 3'. Về điểm này, phiên mã giống với sao chép ADN.



Hình 4.6. 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) ARN polymerase làm chảy tại chỗ 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. Sự kéo dài chuỗi ARN bắt đầu khi các NTP polymer hóa vào đầu 3' – OH của đường ribose của chuỗi. Bong bóng phiên mã di chuyển 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 σ lại có thể liên kết với phần lõi tự do trong dịch bào tương để tham gia vào vòng khởi đầu mới.

Ở *E. coli*, hầu hết mARN có đời sống bán huỷ được tính bằng phút. Chúng nhanh chóng bị thoái hoá 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

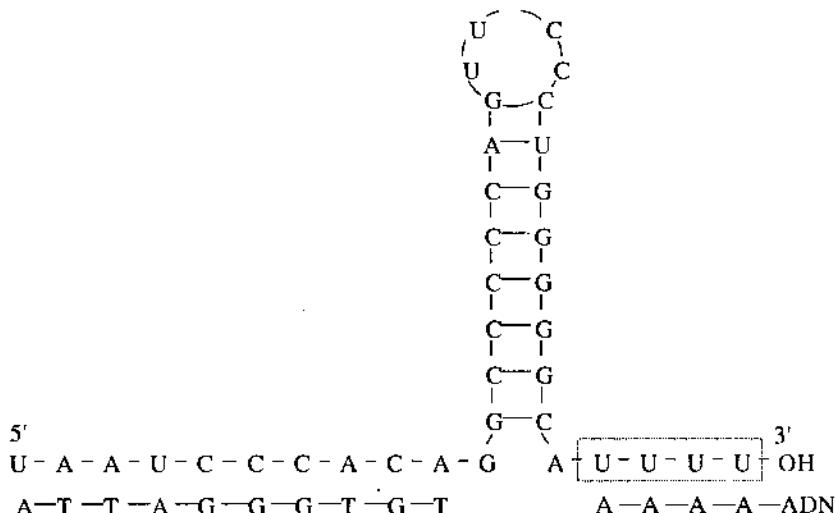
chuỗi mới sinh. Chúng gắn vào phức hợp ADN khuôn – ARN polymerase. Ribosom gắn vào vị trí gắn ribosom ở gần đầu 5' của mARN vừa mới rời phức hợp ADN khuôn – ARN polymerase. Do đó, sự phiên mã và sự dịch mã ở tế bào nhân nguyên thuỷ được thực hiện gần như song song với nhau. Chính điều này làm xuất hiện một dạng kiểm soát phiên mã được gọi là phiên mã suy giảm.

4.3.5. Kết thúc quá trình tổng hợp ARN

Chuỗi ARN 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 kèm được gọi là ρ (yếu tố rho) và một cơ chế khác không phụ thuộc vào các yếu tố này.

4.3.5.1. Kết thúc độc lập với yếu tố rho

Quá trình này đặc trưng bởi sự hiện diện của một trình tự giàu GC trên ADN. Trình tự này đối xứng 2 bên (dyad symmetry), có thêm 5 hoặc 6 adenin kèm theo. 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 4.7). Có thể sự hình thành xoắn kép ARN trong cấu trúc nút này, thích hợp cho việc hình thành ADN – ARN lai giữa sợi ADN phiên mã và chuỗi ARN mới sinh. Như vậy, phần dẫn 5' của nút kéo theo nửa đuôi 3' từ ADN, chỉ để lại trình tự oligo (U) tiếp tục nối dài theo ADN. Các dạng lai oligoA – oligoU cho thấy sự bắt cặp ít bền vững nhất và ARN có thể tách khỏi ADN. Sau đó, sợi kép ADN được hình thành trong "bong bóng phiên mã" và phần lõi của enzym polymerase, có ái lực thấp đối với ADN kép, được phóng thích ra.



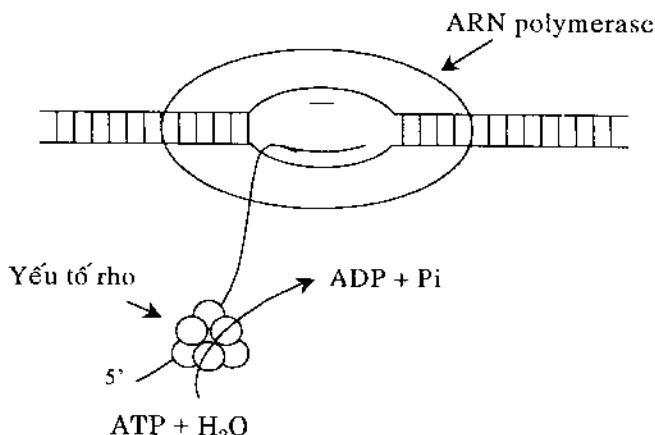
Hình 4.7. Cấu trúc bậc 2 được giả thiết của đầu 3' của bản sao ARN từ trp operon của *E. coli*

Nút vòng hay kẹp tóc và trình tự UUUU tạo nên điểm kết thúc

4.3.5.2. 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.

Khi bám dính vào ARN, yếu tố ρ thuỷ 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 4.8). Sau đó, yếu tố ρ tách đôi ADN – ARN lai bằng một cơ chế chưa rõ, phóng thích mARN vào tế bào chất.



Hình 4.8. 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.

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 mARN 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 ARN trưởng thành. Quá trình này được gọi là cắt nối (hình 4.9).

Khi mạch mARN 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.

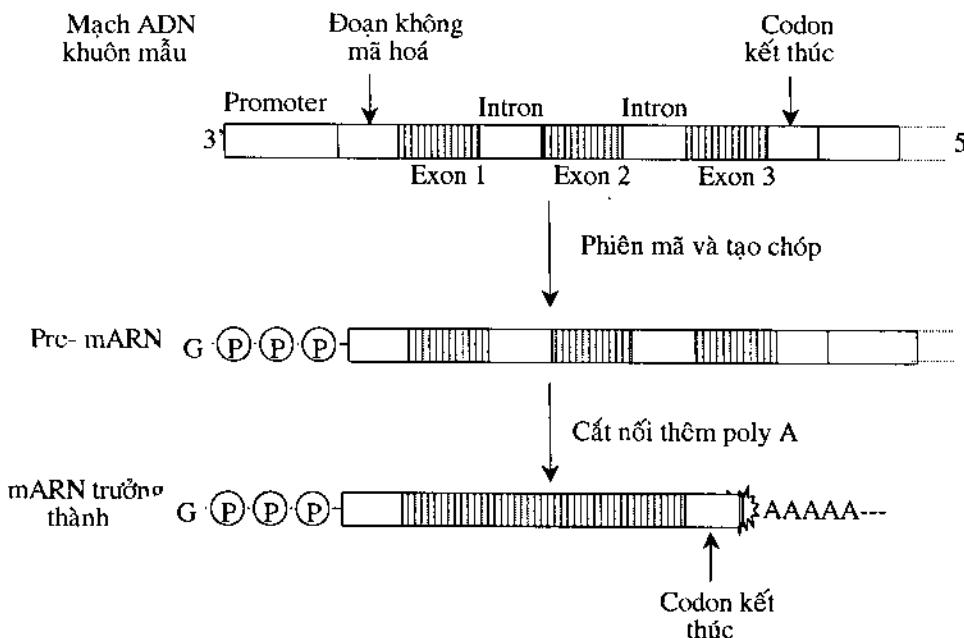
Một đoạn ngắn của đuôi 3' pre – mARN, tại phần không mã hoá sau bộ kết thúc, bị cắt để nối đuôi poly A.

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

Không phải mARN 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 trơ với sự phiên mã. Phiên mã chỉ xảy ra ở các nguyên nhiễm sắc xoắn lỏng lẻo.



Hình 4.9. Phiên mã gen gián đoạn ở tế bào nhân thật

4.4.2. ARN polymerase của tế bào nhân thực

Nhân của tế bào nhân thực có 3 loại ARN polymerase (I, II, III) (bảng 4.2).

Bảng 4.2. Các ARN polymerase nhân của tế bào nhân thực

Loại	Vị trí	Bản sao	Hiệu ứng á - 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

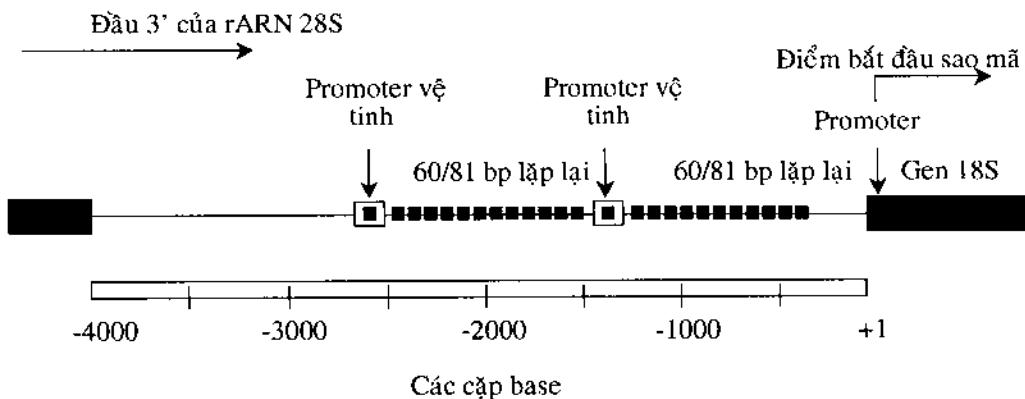
Polymerase I trong nhân, phiên mã các gen ARN ribosom. Bản sao chủ yếu là pre - rARN. Polymerase II tìm thấy trong dịch nhân, rất nhạy cảm với sự ức chế của α - amanitin (khoảng 50% ức chế ở nồng độ 10^{-8} M). Nó phiên mã các gen mã hoá protein và các bản sao sơ cấp không đồng nhất (hnARN), là tiền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

Cả 3 polymerase của tế bào nhân thực đều có phân tử lượng lớn với M vượt quá 500000. Chúng có 2 tiểu đơn vị lớn hơn 100000, mỗi phần là một polymerase chuyên biệt và có tới 12 tiểu đơn vị nhỏ hơn, trong đó có tới một vài phần không gắp ở cả 3 polymerase. Chức năng chính xác của các tiểu đơn vị khác nhau chưa được biết hết. Khác với các polymerase vi khuẩn, polymerase nhân tế bào nhân thực không có khả năng khởi đầu phiên mã ở các vùng tương ứng với các vị trí xuất phát *in vivo*, khi ủ 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 sao chép mạnh 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 sao chép các gen rARN được bảo lưu ở mức độ cao trong các chủng 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. Đó là một vùng không phiên mã nằm giữa các trình tự được phiên mã của các gen rARN của *Xenopus laevis* (cóc châu Phi) và các trình tự cần thiết cho sự phiên mã bởi polymerase I. Các trình tự cần thiết này được xác định bằng tác động xoá bỏ (hình 4.10).

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 – 1200 và – 2300 bp trong vùng vê tinh (vùng không phiên mã) và cài rất nhiều bản sao 60 hoặc 81 bp của promoter thực. Có lẽ điều này liên quan tới một yếu tố giới hạn số lượng. Các bản sao lặp lại 60/81 bp kích thích rất nhiều sự phiên mã của promoter thực. Các promoter vê tinh này có khả năng khởi động quá trình tổng hợp các bản mã sao rARN. Các phần tử polymerase "xếp hàng" trong vùng vê tinh trước khi chuyển động tới promoter thực.



Hình 4.10. Các phần tử promoter trong vùng vê tinh không phiên mã của các gen rARN *Xenopus laevis*

Ở vuông đen: các trình tự lặp lại 60 hoặc 80 bp có trong promoter thực. Các ô ở vị trí – 1200 và – 2300 là các bản sao hoàn chỉnh của các promoter thực ở vị trí + 6 đến – 142.

Ở người, 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 5,8 S. 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.

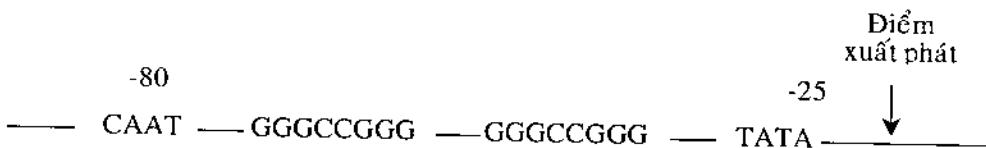
4.4.4. Sự phiên mã bởi polymerase II

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 hóa 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 hóa cho một loại protein.

Một số gen có bản sao thứ 2 trong quá trình tiến hóa. Cả 2 bản sao có thể chuyển đổi bổ trợ nhau, như trường hợp các gen α -globine. Các gen mã hóa 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 hóa galactose trong nấm men là một thí dụ.

Các promoter của ARN polymerase II nằm ở phía 5' của điểm bắt đầu phiên mã. Các trình tự tiếp theo rất có ý nghĩa đối với chức năng của promoter (hình 4.11).



Hình 4.11. Các nét chung của promoter của ARN polymerase II ở tế bào nhân thực bậc cao, số lượng và vị trí của các hộp GC rất khác nhau

Hộp TATA hay hộp Goldberg – Hogness có nét chung, nhưng không có nghĩa là phổ biến và có một trình tự chung TATAAT. Nó nằm cách khoảng 25 bp ngược chiều với vị trí bắt đầu sao mã ở các tế bào nhân thực bậc cao. Nhiều hộp TATA gắn kèm với các trình tự giàu GC. Các trình tự này đóng vai trò rất quan trọng trong việc lựa chọn vị trí bắt đầu phiên mã.

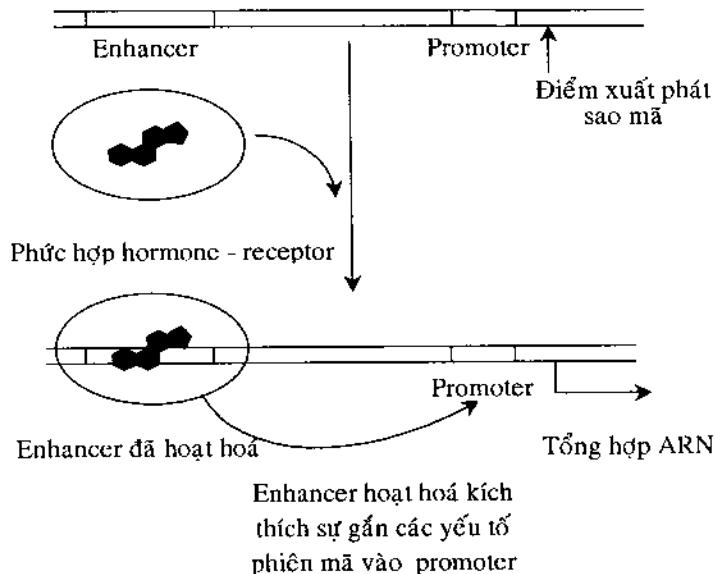
Hộp CAAT (nằm cách vị trí phiên mã 80 bp), có trình tự chung là GGCAATCT.

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 phần tử được đưa thêm vào, gọi là các trình tự tăng cường. Chúng không có khả năng khởi động chính mình, nhưng lại có khả năng kích thích phiên mã mạnh và có thể tác động ở các khoảng cách đáng kể (đến vài kb ADN). Trình tự tăng cường có hiệu quả như nhau đối với các gen cấu trúc do chúng điều hoà 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 hoà biểu hiện gen trong quá trình phát triển 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 hoà phiên mã gen đáp ứng với hormon steroid. Gắn steroid vào thụ thể protein tan trong nước, để cho phức hợp này lại gắn vào các trình tự tăng cường làm gia tăng hoạt động các gen đáp ứng steroid (hình 4.12).



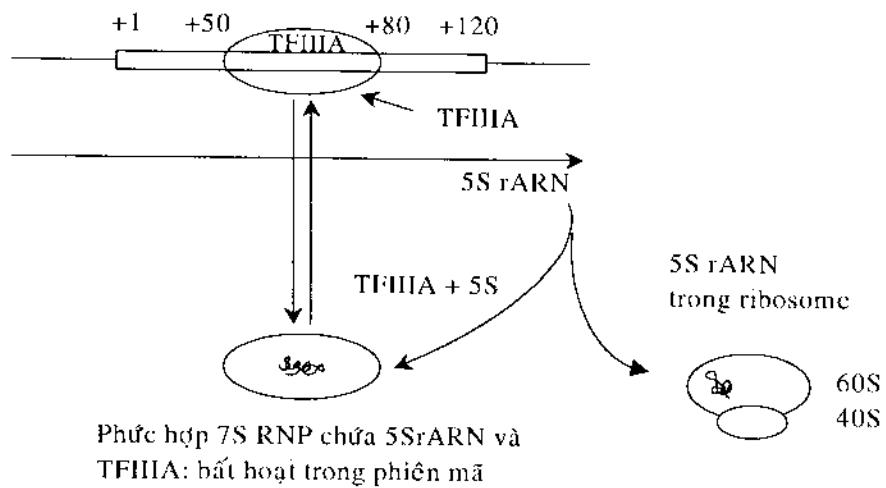
Hình 4.12. Hoạt hóa quá trình phiên mã của các gen đáp ứng hormon steroid (glucocorticoid)

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 trong ADN. 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 một enzym poly(A) polymerase nhằm gắn đuôi 3' ARN.

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.

Tham gia điều hoà tổng hợp các rARN 5S của noãn bào *X. laevis*, ngoài ARN polymerase III, còn có các yếu tố phiên mã khác là TFIIIA, TFIIIB và TFIIIC. Hai yếu tố TFIIIB và C cũng cần cho sự phiên mã các gen khác bởi ARN polymerase III (chẳng hạn các gen tARN), trong khi đó TFIIIA (một protein có Mr 40000) thì chuyên biệt cho noãn bào và gắn với vùng nội kiểm soát của gen rARN 5S. Sau đó, hướng ARN polymerase III tổng hợp khoảng 50 nucleotid (hình 4.13).



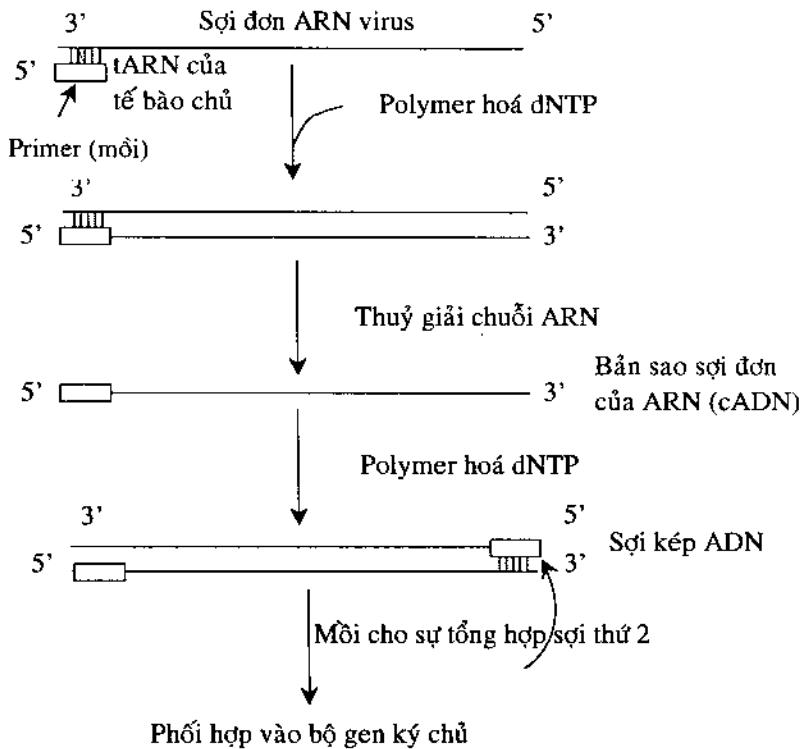
Hình 4.13. Kiểm soát quá trình tổng hợp rARN 5S noãn bào *X. laevis*

TFIIIA thường gắn bám 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 ribosome như các ARN 28S, 18 S và protein ribosom diễn ra sau 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, có thể xem sự tổng hợp rARN 5S noãn bào hoàn toàn là quá trình điều hòa.

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 4.14). Khác với ARN polymerase, enzym phiên mã ngược phải dùng tới mồi. Đoạn mồi 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 lai 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.



Hình 4.14. Sao chép ngược sợi đơn ARN của Retrovirus thành sợi kép ADN

4.6. MÃ DI TRUYỀN

Ta 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.

M. W Nirenberg và H. Matthaei (Mỹ) đã dùng enzym theo phương pháp của Ochoa tổng hợp ARN nhân tạo. Khi chỉ dùng một loại nucleotid là Uracil sẽ nhận được ARN là polyuracil, nếu chỉ Adenin sẽ nhận được polyadenin.

Năm 1961, khi dùng polyuracil thay cho mRNA tổng hợp protein trong hệ thống vô bào (có acid amin, enzym, nhưng không có ADN...), sản phẩm nhận được là mạch polypeptid polyphenylalanin. Điều đó chứng tỏ codon UUU mã hóa cho phenylalanin. Đây là codon đầu tiên được xác định. Nirenberg và Matthaei cũng chứng minh được AAA mã cho lysin, GGG cho Glycin, CCC cho prolin.

Năm 1964, H. G. Khorana tìm ra phương pháp tạo mRNA nhân tạo với trình tự lặp lại (như AAGAAGAAG...) và nhờ đó giải quyết các vấn đề chưa rõ.

Bảng mã di truyền (bảng 4.3) cho thấy trong 64 codon có 3 codon UAA, UAG, UGA không mã hoá 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 hoá, chỉ trừ methionin và tryptophan chỉ có một codon.

Các codon đồng nghĩa tức là mã hoá 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 hoá 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 hoá cho Met nội, vừa mã hoá cho acid amin mở đầu N – formyl Methionin trong tế bào tế bào nhân nguyên thuỷ hoặc methionin trong tế bào nhân thật.

Bảng 4.3. 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

Tóm tắt

Quá trình phiên mã trong tế bào nhân nguyên thuỷ và tế bào nhân thật có một vài chi tiết khác nhau đáng kể. Tế bào nhân nguyên thuỷ 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 thuỷ, 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

- Phiên mã đối xứng có thể xảy ra ở:
 - Vi khuẩn
 - Lạp thể
 - Ti thể của ruồi giấm
 - Ti thể côn trùng
 - Ti thể tế bào thú
 - Tiêu đơn vị σ tách ra khỏi phức hợp phiên mã khi ARN mới sinh đạt chiều dài:
 - 4 base
 - 6 base
 - 8 base
 - 10 base
 - 12 base
 - Vì sao ARN polymerase không cần có hoạt tính sửa sai?
 - Nucleotid kết hợp không đúng được thay thế ngay
 - Sai sót hiếm hoi không di truyền được
 - ARN không phải là nơi lưu trữ thông tin di truyền
 - ARN không tạo ra chính nó
 - a, b, c

4. ARN polymerase ở prokaryote là một holoenzym chứa các tiểu đơn vị:
- a) $\alpha\alpha\beta\beta\sigma$
 - b) $\alpha\alpha'\beta\beta\sigma$
 - c) $\alpha\alpha\beta\beta'\sigma$
 - d) $\alpha\alpha'\beta\beta'\sigma$
 - e) $\alpha\alpha'\beta\beta'\sigma'$
5. Ở *E. coli*, promoter gồm các vùng:
- a) Vùng TATAAT
 - b) Hộp TATA
 - c) Vùng TTGACA
 - d) Vùng -35
 - e) Tất cả
6. Chức năng quan trọng của hộp -10 và hộp -35 được phát hiện nhờ đột biến:
- a) Mất base
 - b) Thay base này bằng base khác
 - c) Thêm base
 - d) Đảo vị trí một cặp base
 - e) a, c
7. Sự kiện **không** đúng với hiện tượng phiên mã ngược:
- a) Cần mồi
 - b) Đoạn mồi là tARN của tế bào chủ
 - c) Đoạn mồi là ARN do primase tổng hợp
 - d) Đoạn mồi gắn vào đầu 3' của Retrovirus
 - e) ARN virus trong chuỗi lai bị phân huỷ bởi RNaseH
8. Đặc điểm nào **không** thuộc sự phiên mã ở tế bào nhân thật?
- a) mARN chứa thông tin một gen
 - b) Đầu 5' mARN có gắn chóp 7 - Methylguanosine
 - c) Bản phiên mã đầu tiên (pre - mARN) được sử dụng ngay cho việc tổng hợp protein
 - d) Có thêm đuôi polyA dài 100 – 200 nucleotid
 - e) Có 3 loại ARN polymerase I, II và III
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

SINH TỔNG HỢP PROTEIN

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 thuỷ.
- 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

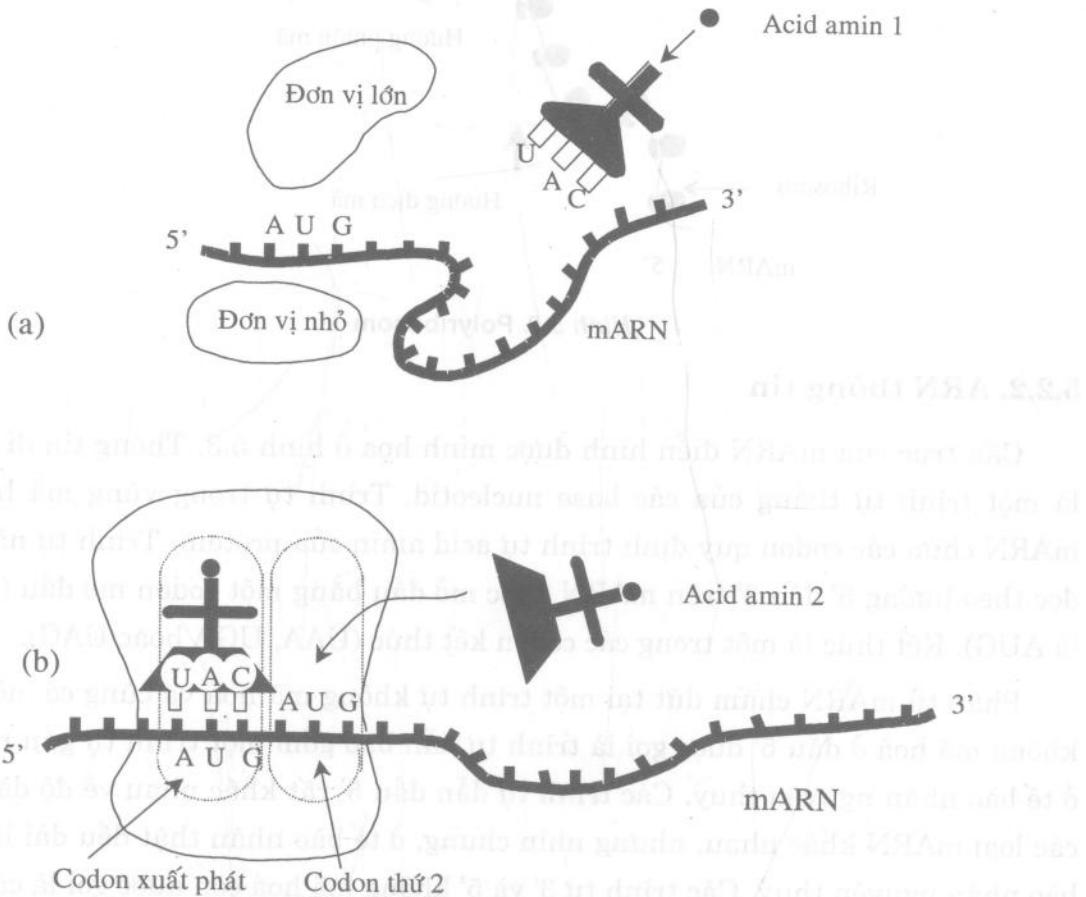
Giai đoạn cuối của sự biểu hiện thông tin di truyền xảy ra khi thông tin được mã hoá bằng trình tự nucleotid trên mARN được dịch mã thành trình tự acid amin tương ứng trong chuỗi polypeptid. Quá trình này được gọi là quá trình tổng hợp protein. Mặc dù, về cơ bản tiến trình dịch mã xảy ra như nhau ở mọi 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 hoá bởi chúng. Bộ máy dịch mã của các bào quan này giống tế bào nhân nguyên thuỷ hơn là tế bào nhân thật.

5.2. CÁC YẾU TỐ CẦN THIẾT CHO SỰ TỔNG HỢP PROTEIN

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 thuỷ 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.



Hình 5.1. Các tiểu đơn vị ribosom và sự dịch mã

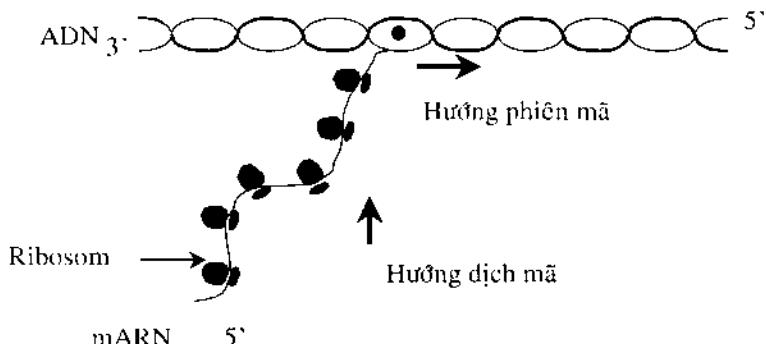
(a) 2 tiểu đơn vị ribosom với mARN; (b) sự dịch mã của ribosom

Ở tế bào nhân nguyên thuỷ, 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 thuỷ 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.

Ở cả tế bào nhân nguyên thuỷ lẫn tế bào nhân thật, khi ribosom đầu tiên dịch mã mARN được một đoạn, thì các ribosom khác có thể gắn vào phía trước để dịch

mã. Do đó, các ribosom xếp thành chuỗi trên mARN tạo nên cấu trúc polyribosom hay còn gọi là polysome (hình 5.2). Khoảng 15 – 20 ribosom có thể gắn cùng lúc trên mARN. Nhờ vậy, tốc độ tổng hợp protein tăng nhanh đáng kể.



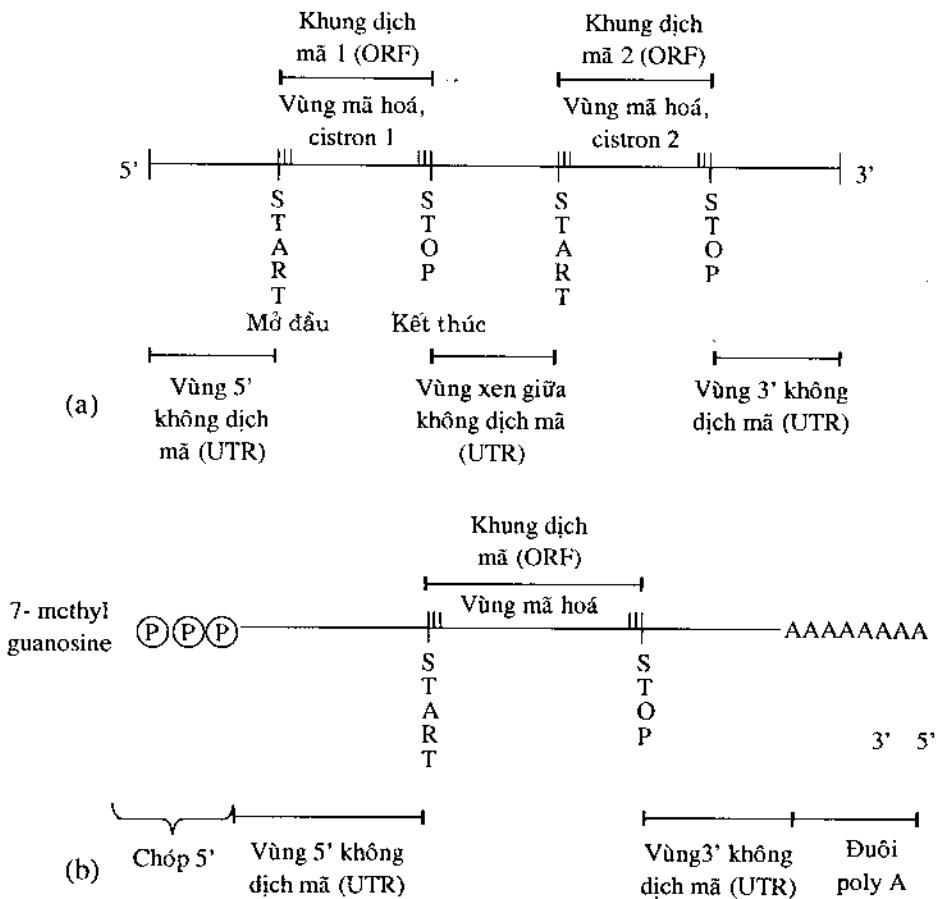
Hình 5.2. Polyribosom

5.2.2. ARN thông tin

Cấu trúc của mARN điển hình được minh họa ở hình 5.3. Thông tin di truyền là một trình tự thẳng của các base nucleotid. Trình tự trong vùng mã hoá của mARN chứa các codon quy định trình tự acid amin của protein. Trình tự này được đọc theo hướng 5' đến 3' trên mARN được mở đầu bằng một codon mở đầu (thường là AUG). Kết thúc là một trong các codon kết thúc (UAA, UGA hoặc UAG).

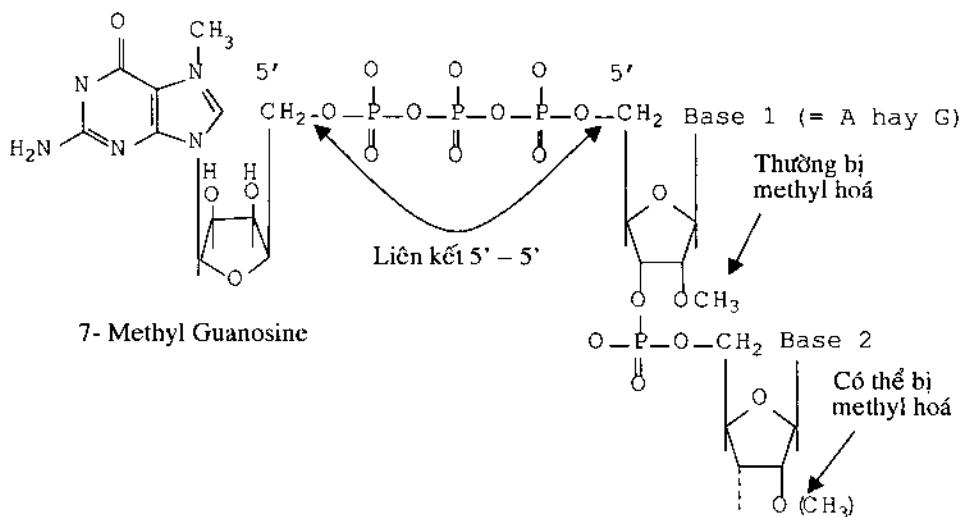
Phân tử mARN chấm dứt tại một trình tự không mã hoá và cũng có một vùng không mã hoá ở đầu 5' được gọi là trình tự dẫn bao gồm một trình tự gắn ribosom ở tế bào nhân nguyên thuỷ. Các trình tự dẫn đầu 5' rất khác nhau về độ dài trong các loại mARN khác nhau, nhưng nhìn chung, ở tế bào nhân thật đều dài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

Ngoài các trình tự không mã hoá ở đầu 5' và 3', các mARN 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 mARN và ribosom. Các trình tự không dịch mã ở đầu tận cùng 3' có chiều dài rất khác nhau. Hầu hết các phân tử mARN tế bào nhân thật ở đầu 3' có thêm đuôi 100 – 200 Adenin sau khi phiên mã. Chức năng của đuôi poly(A) chưa được rõ ràng. Chúng đóng vai trò quan trọng trong việc bảo vệ mARN khỏi bị thuỷ giải của exonuclease, nhưng lại không cần thiết cho sự dịch mã. Một vài loại mARN như histon – mARN và mARN vi khuẩn không có đuôi poly(A).



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 thực



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

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

Các ARN vận chuyển hoạt động như một chất kết nối giữa thông tin base – nucleotid của mARN với trình tự acid amin của phân tử protein tương ứng. Mỗi tARN đặc hiệu cho một loại acid amin riêng biệt, mặc dù nó có thể giải mã cho hơn một codon, số lượng các tARN biến động theo loài: tế bào nhân nguyên thuỷ 30 – 40, tế bào nhân thật 50. Do đó, các tARN phải được nhận biết chuyên biệt đồng thời bởi mARN và bởi enzym gắn acid amin vào tARN tương ứng (enzym tARN – aminoacyl 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 mARN thì nó sẽ cách xa với acid amin tương ứng với codon kế cận để hình thành liên kết peptid. Vì có tới 20 loại acid amin khác nhau cấu trúc nên bản dịch, do đó, cần phải có tới 20 loại aminoacyl – tARN synthetase khác nhau. Đó là Alanyl – tARN synthetase, Arginyl – tARN synthetase, Asparaginyl – tARN synthetase,... 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. Vấn đề một aminoacyl – tARN synthetase nhận diện được vài tARN khác nhau như thế nào hiện vẫn chưa rõ ràng. Dường như, sự nhận diện chính xác tARN của synthetase thông qua việc nhận diện anticodon. Enzym aminoacyl – tARN synthetase phải phù hợp và chọn đúng tARN: tARN chỉ đúng cho một synthetase thì được gọi là tARN duy nhất (cognate tARN). Các loại tARN liên quan đến cùng một acid amin thì được gọi là tARN đồng nhận (isoaccepting tARN).

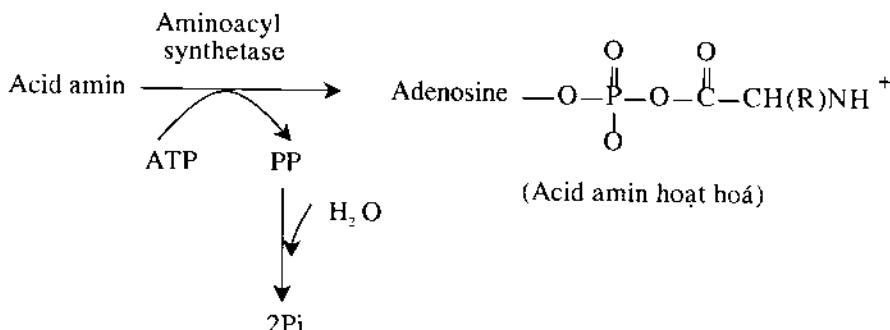
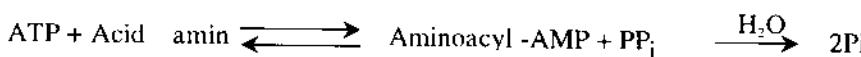
Vấn đề mấu chốt là các bước aminoacyl hoá phải chính xác. Bởi vì, nếu một tARN chấp nhận một acid amin sai dẫn đến việc gắn kết không đúng vào protein và sẽ không có "bước đọc sửa sai" theo sau sự aminoacyl hoá. Bước đọc sửa sai chỉ được thực hiện ở giai đoạn aminoacyl hoá bởi chính enzym aminoacyl – tARN synthetase. Quá trình aminoacyl hoá cùng lúc thực hiện 2 chức năng. Trước tiên, nó cung cấp sự liên hệ đầu tiên của mã di truyền (dựa vào acid nucleic) và cấu trúc protein (với các acid amin). Tiếp đến, nó hoạt hoá acid amin trước khi kết nối vào protein. Liên kết ester aminoacyl giữa acid amin và tARN dự trữ một năng lượng thuỷ giải tự do tương đối cao, năng lượng này rất cần thiết cho việc hình thành liên kết peptid trong quá trình dịch mã.

5.2.4. Aminoacyl hoá

Quá trình aminoacyl hoá xảy ra theo 2 bước:

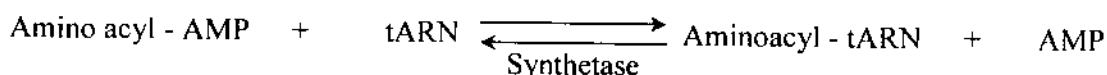
5.2.5.1. Hình thành aminoacyl – adenylate hoạt hoá

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



5.2.5.2. Hình thành aminoacyl – tARN

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



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 đobble 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, có một vài codon cho một acid amin được sử dụng và có một vài phân tử tARN cho mỗi codon. Hiện tượng này được gọi là sự dị biệt codon và đây là điểm khác nhau quan trọng nhất giữa loài này và loài khác.

5.2.5. Các tARN và các codon khởi đầu

Trình tự mã của phân tử mARN được khởi đầu với codon AUG, mặc dù ở vi khuẩn còn có codon GUG được sử dụng cho một số mARN (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.

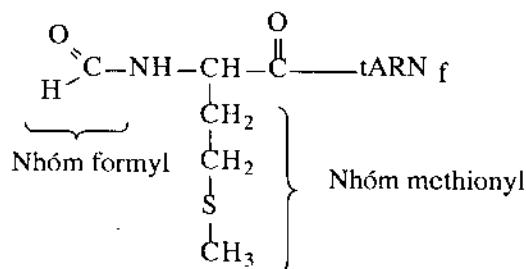
Dấu hiệu khởi đầu là codon AUG, cũng là codon mã hoá cho methionin (Met). Mặc dù, chỉ có một codon mã hoá cho Met, nhưng trong tế bào chất lại hiện diện 2

loại tARN có khả năng mang Met đến kết hợp với codon đó ở tế bào nhân nguyên thủy:

(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^{iMet} kết hợp với codon AUG khởi động, gắn Met mở đầu vào chuỗi polypeptid.

Ở vi khuẩn, tARN^{iMet} khởi đầu không chỉ vận chuyển methionin mà còn vận chuyển formyl – methionin: tARN này được gọi là tARN^{fMet} (viết tắt là fMet) (hình 5.5).



Hình 5.5. Formyl – methionyl – tARN

Sự tổng hợp tất cả protein ở vi khuẩn đều bắt đầu bằng formyl – methionin, mặc dù nhóm formyl, cũng như nhóm methionyl, 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}. Còn 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, 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

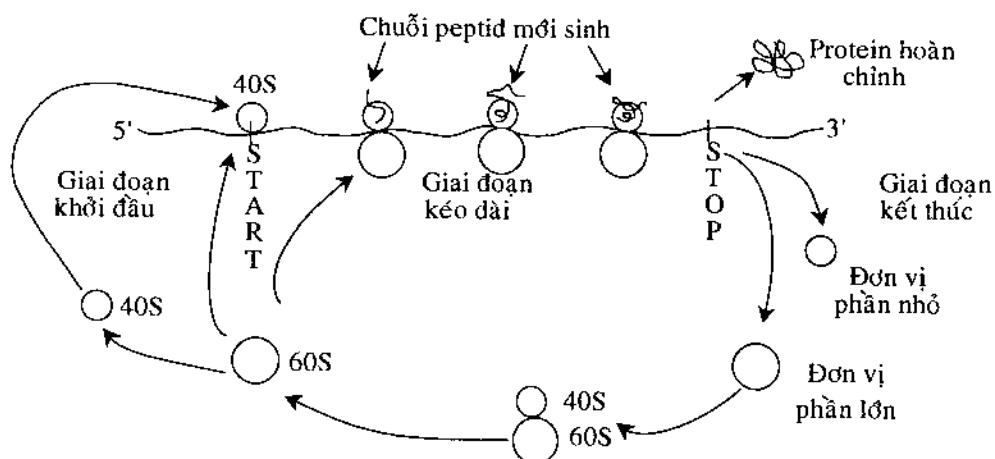
	Vị khuẩn	Tế bào nhân thật
Codon khởi đầu	AUG (GUG)	AUG
tARN khởi đầu	tARN ^{fMet}	tARN ^{iMet}
tARN mang	Formyl – methionin	Methionin

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

Quá trình dịch mã mRNA được mô tả trong hình 5.6, bao gồm các giai đoạn: khởi đầu, kéo dài và kết thúc.

5.3.1. Khởi đầu

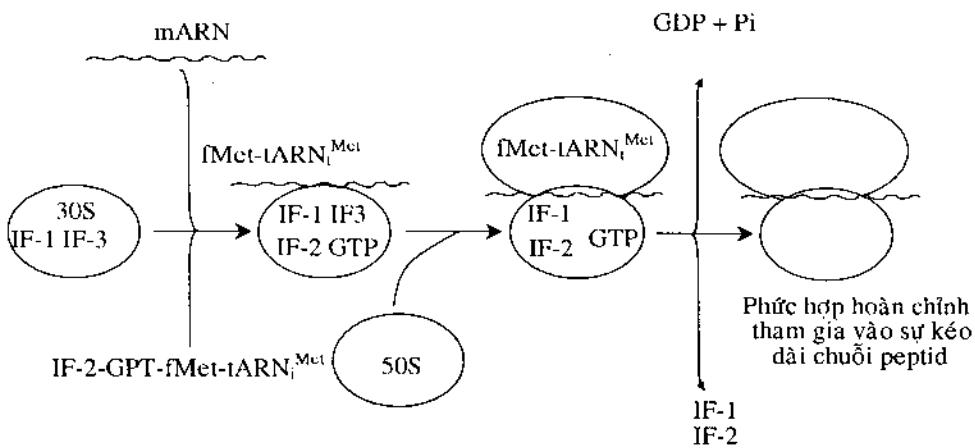
Giai đoạn khởi đầu, có sự tham gia của 2 tiểu đơn vị của ribosom, một phân tử mRNA (được sắp thẳng hàng với ribosom) và tRNA khởi đầu mang methionin hoặc formyl-methionin ở tế bào nhân nguyên thuỷ (bảng 5.2). Và các yếu tố khởi đầu như IF₁, IF₂, IF₃ ở vi khuẩn và 10 yếu tố eIF (IF của tế bào nhân thật)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



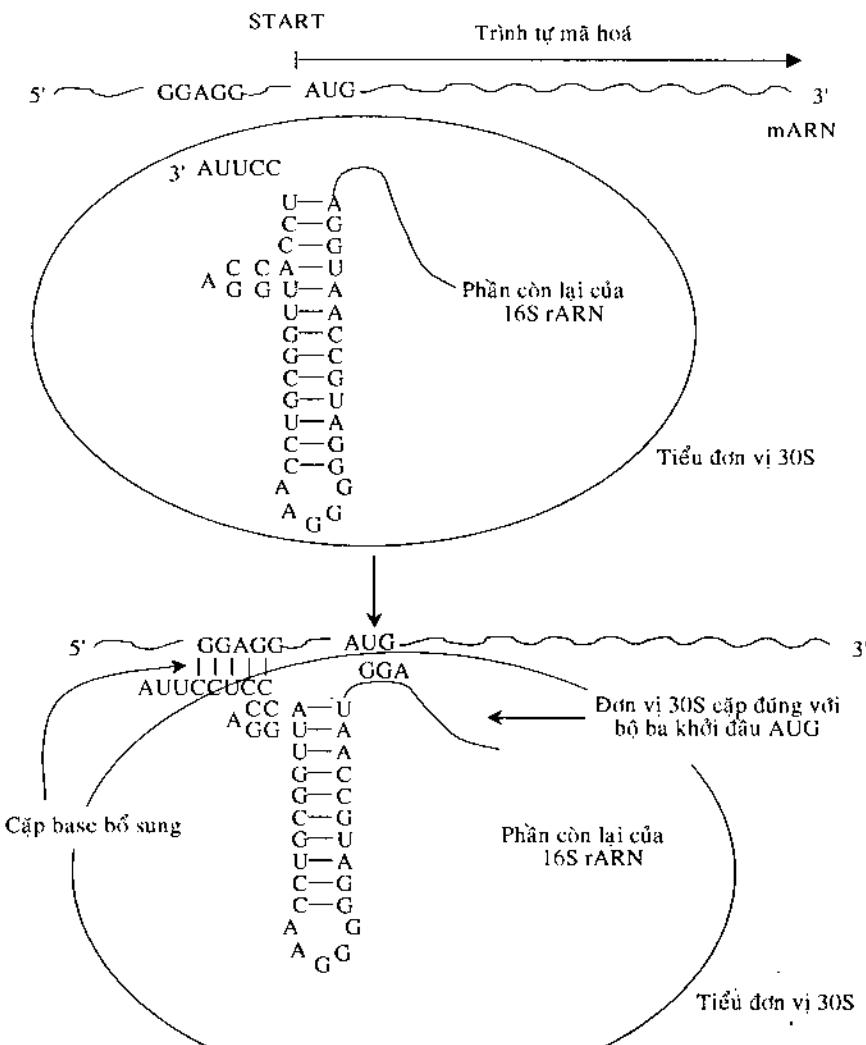
Hình 5.6. Chu trình ribosom tế bào nhân thật

5.3.1.1. Khởi đầu chuỗi peptid ở vi khuẩn

Đặc trưng của sự khởi đầu chuỗi peptid ở vi khuẩn là sự tương tác trực tiếp giữa mRNA và rRNA của tiểu đơn vị nhỏ 30S (hình 5.7). Sự tương tác này dẫn đến việc gắn mRNA vào ribosom và quan trọng hơn là nhận diện được codon khởi đầu của trình tự mRNA. Các mRNA vi khuẩn có chứa một trình tự dẫn ở đầu 5' khá ngắn. Nó bao gồm trình tự polypurin ngắn bổ sung với một trình tự giàu pyrimidin ở tận cùng 3' của rRNA 16S (hình 5.8). 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 mRNA. Do đó, người ta gọi trình tự ngắn này là trình tự Shine – Dalgarno.



Hình 5.7. Sự khởi đầu dịch mã ở vi khuẩn



Hình 5.8. Vai trò của trình tự Shine – Dalgarno

Vai trò IF – 2 là làm trung gian trong việc ghép tARN khởi đầu (fMet – tARNfMet) vào tiểu đơn vị 30S. Điều này phụ thuộc vào GTP. Bởi vì, IF – 2 tạo ra phức 3 yếu tố với GTP và với tARN khởi đầu. Đó là phức [IF – 2 – GTP – fMet – tARNfMet] mà tARN đã gắn với ribosom. Các phân tử tARN luôn luôn gắn với ribosom như những phức có chứa GTP và yếu tố protein. Chức năng chính xác của IF – 1 chưa được xác định, nhưng trong một số trường hợp cho thấy nó tạo điều kiện cho tác động của IF – 2. Có thể mARN hoặc phức hợp tARN gắn trước tiên vào tiểu đơn vị 30S. Người ta mới chỉ biết được rằng, cả 2 bước này đều xảy ra trước khi tiểu đơn vị 50S gắn vào để tạo ra phức hợp 70S cho việc nối dài. Vai trò chủ yếu của IF – 3 là giữ sự gắn kết mARN với ribosom.

Sự khởi đầu chuỗi peptid ở vi khuẩn bao gồm sự nhận diện trình tự Shine – Dalgarno, nằm gần điểm xuất phát của trình tự mã hoá mARN. Do đó, sự khởi đầu dịch mã không nhất thiết nằm gần với điểm khởi đầu 5' của mARN. Đúng vậy, các mARN tế bào nhân nguyên thuỷ có chứa tới vài vùng mã hoá, đúng hơn là vi khuẩn thường có polycistron mARN. Đặc biệt, rõ nhất là mARN chứa operon mã hoá cho 3 loại protein khác nhau ở *lac* operon.

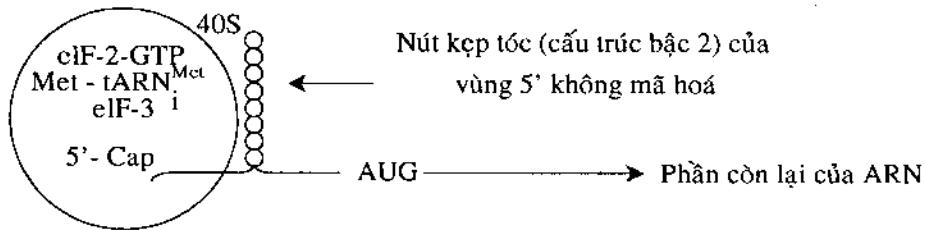
5.3.1.2. Khởi đầu chuỗi peptid ở tế bào nhân thật

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 thuỷ. Trong khi, ở *E. coli* chỉ có 3 yếu tố khởi đầu thì ở tế bào nhân thật có tới 10. Điều này được phản ánh trong các cách thức gắn kết mARN và xác định codon khởi đầu.

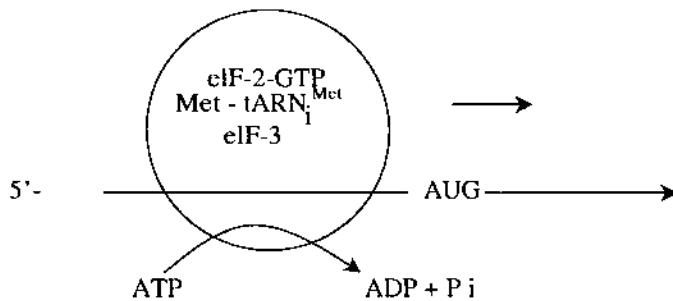
Trình tự dẫn mARN tế bào nhân thật không được xem như trình tự Shine – Dalgarno ở tế bào nhân nguyên thuỷ. Đầu tận cùng 3' của rARN 18S (tương đương với rARN 16S ở vi khuẩn) không có trình tự CCUCC. Và mARN ở tế bào nhân thật là monocistron, chúng chỉ chứa một trình tự mã hoá và khởi đầu, thường xảy ra ở codon khởi đầu AUG trong mARN. Những nhận xét trên giúp cho Kozak xây dựng mô hình khởi đầu sinh tổng hợp protein ở tế bào nhân thật (hình 5.9).

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 thuỷ phân ATP và sự tương tác khởi đầu với mARN có thể liên quan tới cấu trúc chỏp 5' (CAP). 95% các trường hợp khởi đầu tế bào nhân thật xảy ra tại codon khởi đầu AUG, hầu như không dùng tới GUG.

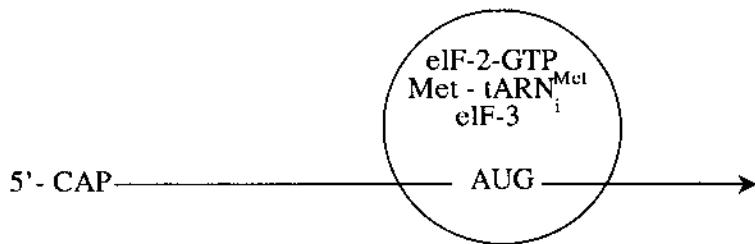
(a) Tiêu đơn vị nhỏ 40S gắn vào mARN ở chót 5' và qua chót (Cap) hình thành nên phức hợp chót – các yếu tố khởi đầu.



(b) Tiêu đơn vị 40S quét dọc theo mARN, tìm AUG khởi đầu. Quá trình này cần ATP và có thể tháo xoắn nút kẹp tóc.



(c) Codon khởi đầu, thường là AUG được lựa chọn.



Hình 5.9. Mô hình Kozak

Phức hợp mới hình thành bấy giờ đã có thể tương tác với tiêu đơn vị 60S để tham gia vào giai đoạn nối dài.

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 5.10.



Hình 5.10. Ngữ cảnh nhận diện codon khởi đầu

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

Tiêu 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.

Bảng 5.2. Các yếu tố khởi đầu của tế bào nhân nguyên thuỷ 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 – 1	Vai trò chưa rõ
IF – 2	eIF – 2	Gắn Met – tARN (hoặc fMet – tARN) vào ribosom thành phức hợp với GTP
IF – 3	eIF – 3	Ngăn cản sự tách đôi của các tiểu đơn vị ribosom
	eIF – 4A eIF – 4B eIF – 4E eIF – 4F	Nhóm yếu tố liên quan đáp ứng cho việc gắn chopy 5' vào đầu mARN và tháo xoắn cấu trúc bậc 2 tạo điều kiện cho ribosom đậu vào codon khởi đầu
	eIF – 5	Giải phóng eIF – 2 và eIF – 3 khỏi ribosom và giúp cho tiểu đơn vị 60S kết nối
	eIF – 6	Tham gia vào việc tách ribosom thành các tiểu đơn vị
	GEF	Yếu tố biến đổi nucleotid Guanin: tái sinh eIF – 2 bởi sự biến đổi trung gian GDP thành GTP trong quá trình tương tự để tái sinh EF – Tu bởi EF – Ts (hình 5.7)

Bảng 5.2 đã giới thiệu 10 yếu tố khởi đầu của tế bào nhân thật. Ở đây, eIF – 3 (giống IF – 3) cản trở các tiểu đơn vị ribosom tách đôi. eIF – 2 (giống IF – 2) là trung gian gắn kết tARN – khởi đầu với tiểu đơn vị nhỏ thành phức hợp với GTP. Tuy nhiên, ở tế bào nhân thật Met – tARNIMet phải được gắn kết trước vào mARN. Điều đó cho thấy rằng đối mã (CAU) trên tARN – khởi đầu đã nhận ra AUG trong khi tiểu đơn vị 40S đang quét trên mARN.

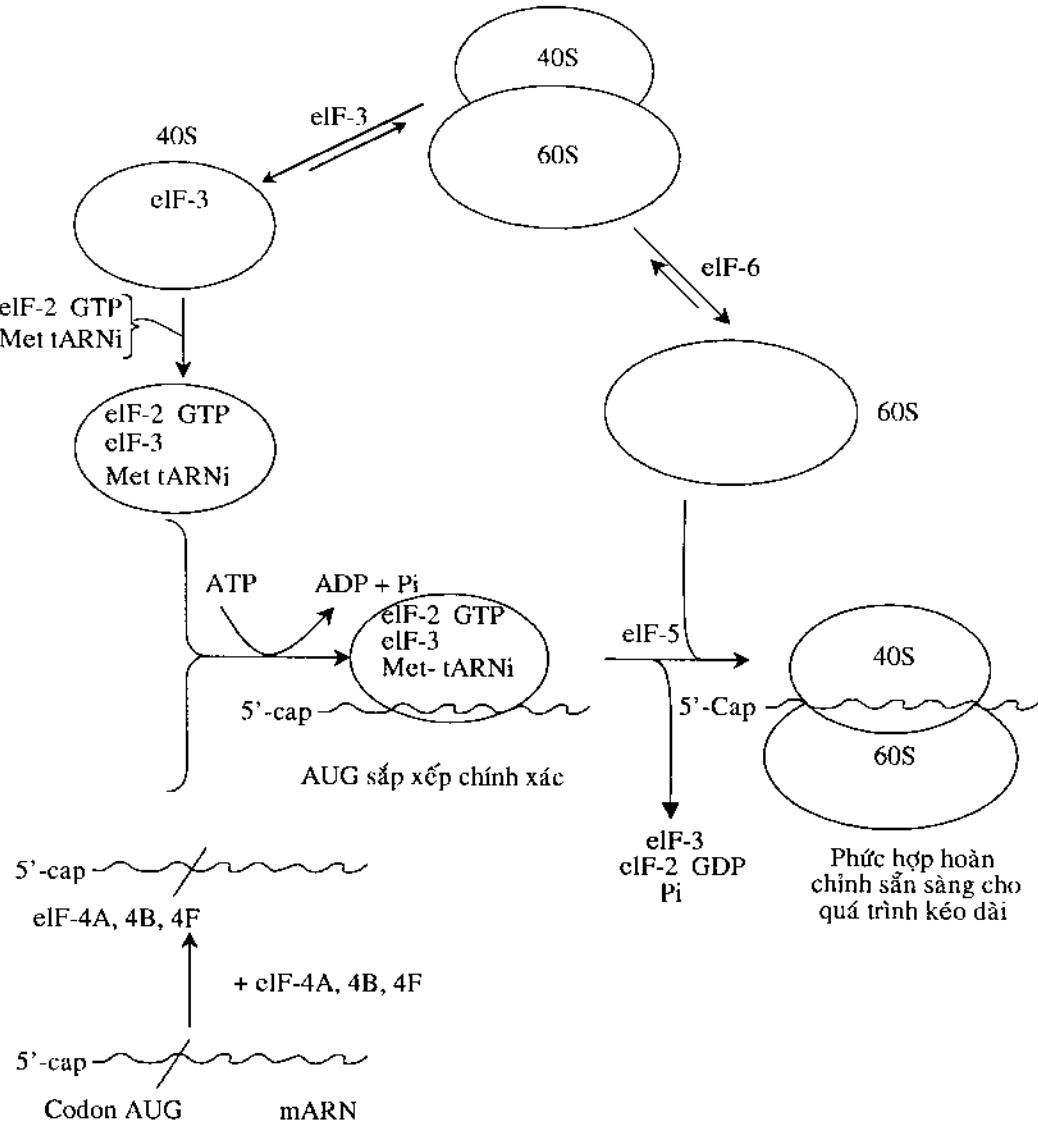
eIF – 6 tương tác với tiểu đơn vị lớn 60S và tách đôi được 2 tiểu đơn vị của ribosom. eIF – 5 giúp tiểu đơn vị 60S kết nối và giải phóng các yếu tố khởi đầu khác và các yếu tố khác liên quan tới sự gắn mARN như eIF – 4A, eIF – 4B, eIF – 4E, eIF – 4F.

Các yếu tố này thường là những protein gắn chopy do chúng gắn vào chopy 5'; chúng gắn vào các hợp phần của chopy (như f – methyl GTP chẳng hạn). eIF – 4F là một phức hợp protein chứa các thành phần tương tự eIF – 4E và eIF – 4A, vì cùng

là protein có phân tử lượng Mr 220000. Người ta cho rằng eIF – 4F thông qua eIF – 4E gắn trực tiếp vào chỏp trước và sau đó eIF – 4A và eIF – 4B mới gắn vào.

Các trình tự dẫn đầu 5' của mARN tế bào nhân thật thường dài đến vài trăm base và có thể chứa các vùng bổ sung, làm nhô lên các nút kẹp tóc. Đây là cấu trúc bậc 2 của mARN. Để cho ribosom có thể dịch mã được, điều cần thiết là phải tháo xoắn cấu trúc bậc 2 này. Bởi vì, có thể cấu trúc bậc 2 này ngăn cản sự nhận dạng của ribosom đối với mARN hoặc che khuất codon khởi đầu AUG.

eIF – 4A có khả năng tháo xoắn các cấu trúc bậc 2 (cần thuỷ phân ATP). Vai trò của eIF – 4B chưa rõ, nhưng có thể làm gia tăng tác động của eIF – 4A (hình 5.11).



Hình 5.11. Sự khởi đầu ở tế bào nhân thật

5.3.2. Nối dài

Quá trình dịch mã trên mARN được nối dài theo hướng 5' → 3' và chuỗi polypeptid được bắt đầu tổng hợp ở đầu tận cùng – N. Quá trình này giống nhau ở tế bào nhân nguyên thuỷ và tế bào nhân thật (hình 5.13).

Sự nối dài là một quá trình có chu kỳ, mà mỗi chu kỳ chịu trách nhiệm 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 – 1 α (tương ứng EF – Tu) và EF – 2 (tương đương EF – G), EF – 1 $\beta\gamma$ – Tu (tương đương EF – T_s) (bảng 5.3).

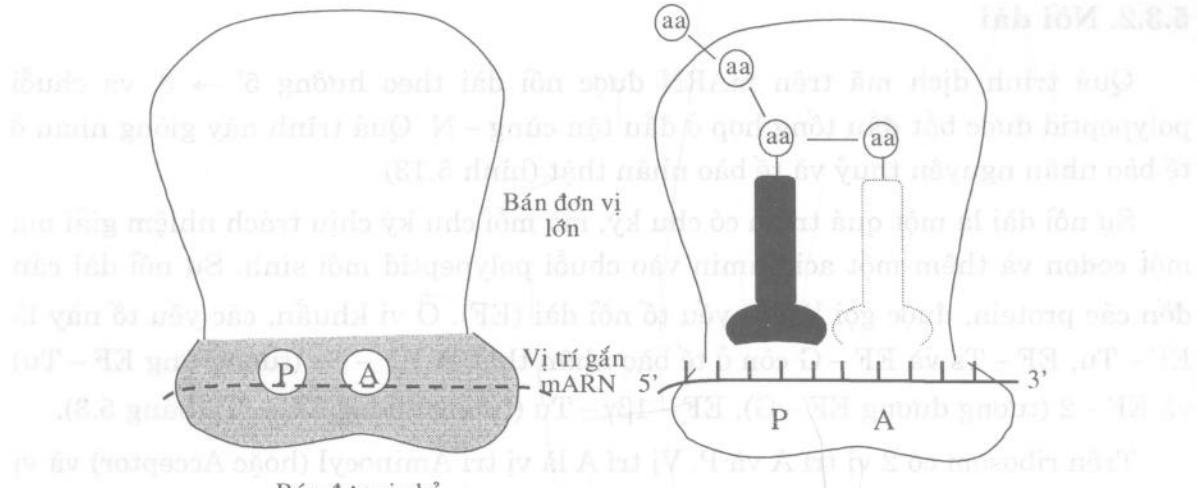
Trên ribosom có 2 vị trí A và P. Vị trí A là vị trí Aminocyl (hoặc Acceptor) và vị trí P là vị trí Peptidyl, nối giữ phức hợp peptidyl – tARN, tức là chuỗi polypeptid đang hình thành vẫn còn gắn với tARN trước đó.

Bảng 5.3. Các yếu tố nối dài

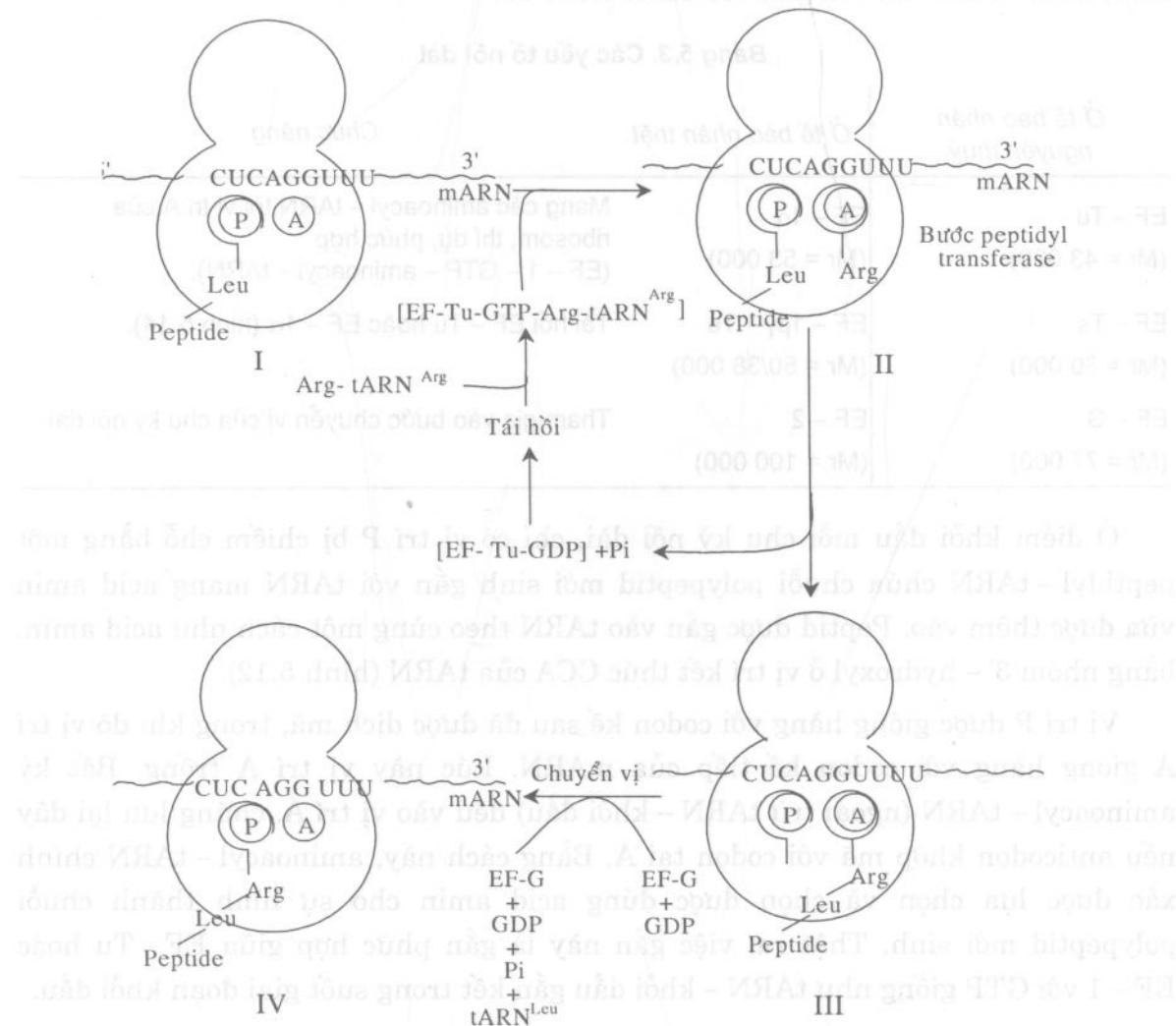
Ở tế bào nhân nguyên thuỷ	Ở tế bào nhân thật	Chức năng
EF – Tu (Mr = 43 000)	EF – 1 α (Mr = 53 000)	Mang các aminoacyl – tARN tới vị trí A của ribosom, thí dụ, phức hợp (EF – 1 – GTP – aminoacyl – tARN).
EF – Ts (Mr = 30 000)	EF – 1 $\beta\gamma$ – Tu (Mr = 50/38 000)	Tái hồi EF – Tu hoặc EF – 1 α (hình 5.14).
EF – G (Mr = 77 000)	EF – 2 (Mr = 100 000)	Tham gia vào bước chuyển vị của chu kỳ nối dài.

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 (hình 5.12).

Vị trí P được gióng hàng với codon kế sau đã được dịch mã, trong khi đó vị trí A gióng hàng với codon kế tiếp của mARN. Lúc này vị trí A trống. 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. Thật ra, việc gắn này là gắn phức hợp giữa EF – Tu hoặc EF – 1 với GTP giống như tARN – khởi đầu gắn kết trong suốt giai đoạn khởi đầu.



Hình 5.12. Cấu trúc ribosom



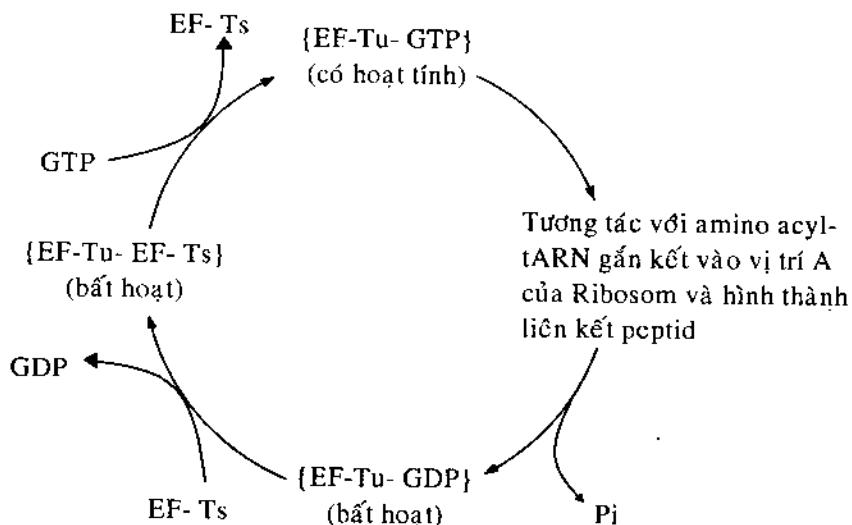
Hình 5.13. Quá trình nối dài gần như giống nhau giữa vi khuẩn và tế bào nhân thực

Theo sau sự gắn kết phức hợp vào vị trí A, là quá trình thuỷ phân GTP giải phóng GDP và Pi (hình 5.13). GDP được phóng thích khỏi ribosom sẽ gắn kết với yếu tố nối dài tạo phức hợp [EF – Tu – GDP] hoặc [EF – 1 – GDP]. Để yếu tố nối dài lại có thể tham gia vào quá trình gắn kết aminoacyl – tARN tiếp theo vào ribosom thì GDP phải được thay thế bởi GTP. Ở vi khuẩn, yếu tố EF – Ts đóng vai trò trung gian cho việc chuyển đổi GDP/GTP (hình 5.14), tương tự các yếu tố của protein – G liên quan tới sự tải nạp tín hiệu hormon. Giữa acid amin sau cùng của chuỗi mới sinh với acid amin trên tARN ở vị trí A, hình thành liên kết peptid. Phản ứng này được xúc tác bởi enzym peptidyl – transferase. Kết quả là tARN ở vị trí P lúc bấy giờ không còn mang acid amin hoặc là tách khỏi chuỗi polypeptid.

Bước chuyển vị (translocation step) cuối cùng của giai đoạn nối dài cần một yếu tố nối dài khác là EF – G hoặc EF – 2 và GTP. GTP thuỷ phân thành GDP và Pi. Sự chuyển vị có 3 hiện tượng xảy ra đồng thời:

- tARN vận chuyển xong được tách khỏi vị trí P
- Peptidyl – tARN di chuyển từ vị trí A sang vị trí P
- Ribosom chuyển dịch theo từng codon trên mARN, do đó, codon tiếp theo được đưa vào giống hệt với vị trí A.

Quá trình chuyển vị 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 thuỷ 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.



Hình 5.14. Sự tái hồi EF – Tu bởi EF – Ts và đồng phân của protein – G

EF – Tu được giải phóng khỏi ribosom sau mỗi chu kỳ nối dài thành phức hợp với GDP (bất hoạt), nghĩa là phức hợp này không thể gắn kết aminoacyl – tARN. Để cho EF – Tu tham gia vào vòng nối dài tiếp theo thì GDP cần phải được biến đổi thành GTP tạo phức hợp có hoạt tính [EF – Tu – GTP]. Sự biến đổi nucleotid Guanin cần có một yếu tố protein khác, đó là EF – Ts.

5.3.3. Kết thúc

Chu trình dịch mã thực hiện thêm khoảng 15 acid amin một giây vào mạch polypeptid. Chu trình được chấm dứt khi trải qua codon kết thúc là UAA, UAG và UGA.

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.

Ở vi khuẩn có 3 yếu tố phóng thích: RF – 1 chuyên biệt cho các codon UAA và UAG; RF – 2 liên quan tới UAA hoặc UGA và RF – 3 kích thích cả 2 yếu tố kia. Ở đây còn cần yếu tố nối dài EF – G và yếu tố phóng thích ribosom (RRF) chuyên biệt.

Ở tế bào nhân thật chỉ có một yếu tố phóng thích eRF đã được biết (bảng 5.4).

Bảng 5.4. Các yếu tố kết thúc

<i>Vi khuẩn (E. coli)</i>	<i>Tế bào nhân thật</i>
RF – 1 (cho UAA hoặc UAG)	
RF – 2 (cho UAA hoặc UGA)	eRF (phù hợp với cả 3 codon kết thúc)
RF – 3 (Kích thích RF – 2)	

Sự kết thúc liên quan đến việc phân giải liên kết giữa acid amin cuối cùng của chuỗi peptid mới sinh và tARN cuối cùng. Phản ứng này được xúc tác bởi enzym peptidyl transferase của ribosom và đòi hỏi GTP. Nhờ đó, polypeptid hoàn chỉnh được phóng thích. Sau đó, 2 tiểu đơn vị của ribosom được phóng thích cùng với mARN. Tất cả các thành phần tham gia chu trình đều được sử dụng lại.

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 thuỷ phân ATP và GTP (bảng 5.5).

Bảng 5.5. Nhu cầu năng lượng cho quá trình dịch mã

Giai đoạn	Tế bào nhân nguyên thuỷ	Tế bào nhân thật
Aminoacyl hoá tARN	ATP → AMP + PPi (2 phosphoanhydrid cho một gốc acid)	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 chớp và tháo xoắn 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	Không cần năng lượng	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.

Có thể tính năng lượng chuyển hoá cần dùng cho sự tổng hợp một chuỗi polypeptid điển hình ở vi khuẩn gồm 300 acid amin như sau:

- Để aminoacyl hoá tARN cần $300\text{ATP} \rightarrow 300\text{AMP} \approx 600$ liên kết phosphoanhydrid.
- Một phản ứng khởi đầu, $1\text{GTP} \rightarrow 1\text{GDP} \approx 1$ liên kết phosphoanhydrid

Cho 229 vòng nối dài:

$299\text{ GTP} \rightarrow 299\text{ GDP}$ ở giai đoạn EF – Tu ≈ 299 liên kết phosphoanhydrid

$299\text{ GTP} \rightarrow 299\text{ GDP}$ ở giai đoạn EF – G ≈ 299 liên kết phosphoanhydrid

Tổng năng lượng cần ≈ 1198 liên kết phosphoanhydrid

Tóm lại, tính ra chính xác có 4 liên kết phosphoanhydrid được sử dụng cho mỗi acid amin trong quá trình sinh tổng hợp protein.

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 xúe 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, công kênh 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, thuỷ 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 hoá (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 thuỷ phân (H), nên bị thuỷ 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 hoá tARN. Bước sửa sai được minh họa bằng hình 5.15.

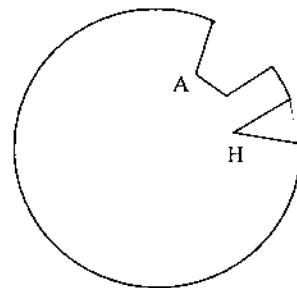
5.5.2. Tương tác codon – anticodon

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.

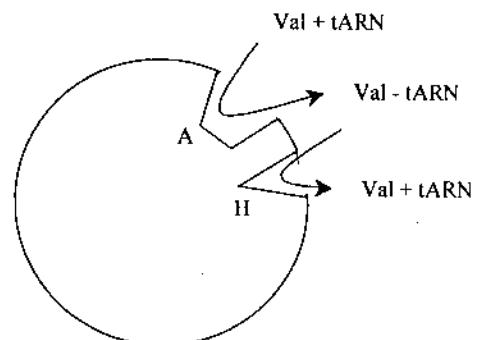
5.5.3. Kết thúc sớm, kết thúc muộn

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 hoá 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.

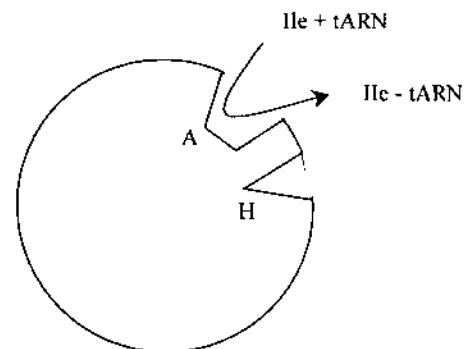
(a) Các acid amin lớn hơn isoleucin không vào được vị trí A (như phenylalanin). Kết quả: phản ứng không xảy ra



(b) Các acid amin nhỏ hơn isoleucin (như valin) vào vị trí A và tạo aminoacyl-tARN, nhưng cũng đi vào được vị trí H nên bị thuỷ phân thành 2 phần tARN và acid amin. Kết quả: phản ứng không xảy ra



c) Chỉ có isoleucin gắn vào vị trí A tạo isoleucyl-tARN nhưng không thể gắn vào vị trí H và do đó không bị thuỷ phân. Kết quả: Acyl hoá chính xác isoleucin



Hình 5.15. Sơ đồ bước sửa sai aminoacyl hoá tARN của enzym synthetase

A: vị trí acyl hoá (acylation); H: vị trí thuỷ phân (hydrolytic).

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 ỨC CHẾ 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 5.6). 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 5.7).

Bảng 5.6. Các chất ức chế dịch mã

Chất ức chế	Chọn lọc trên	Tác động ở giai đoạn
Chloramphenicol	Tế bào nhân nguyên thuỷ	Nối dài: ức chế peptidyl transferase
Cycloheximide	Tế bào nhân thật	Nối dài: chưa rõ cơ chế
Erythromycin	Tế bào nhân nguyên thuỷ	Nối dài: ức chế sự chuyển peptid hoá
Acid fusidic	Cả 2 giới (tế bào nhân nguyên thuỷ và tế bào nhân thật)	Nối dài: ức chế sự phóng thích của EF – G (hoặc EF – 2): phức hợp GDP
Kanamycin	Cả 2 giới	Nối dài: gây đọc sai
Neomycin	Cả 2 giới	Khởi đầu và nối dài: mARN 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 thuỷ	Nối dài: ức chế chuyển peptid hoá
Streptomycin	Tế bào nhân nguyên thuỷ	Nối dài: gắn vào 30S và ảnh hưởng tới tương tác codon – anticodon gây ra sự đọc nhầm của mARN
Tetracyclin	Tế bào nhân nguyên thuỷ	Nối dài: phong bế việc gắn aminoacyl – tARN vào vị trí A

Bảng 5.7. 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> để kháng penicillin
Kanamycin	Dùng cho nhiễm trùng gram (–) nặng
Spectinomycin	Điều trị lao
Nhóm tetracyclin	Kháng sinh có phổ kháng khuẩn rộng, nhưng nhiều vi khuẩn đã đề kháng với nhóm này.

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.

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 hoá dưới dạng codon, mỗi codon là một tổ hợp 3 nucleotid mã hoá 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 – tARNiMet), (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 thực 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

1. Năng lượng giải phóng được từ GTP thành GDP + Pi dùng để:
 - a) Ghép tARN khởi đầu với tiểu đơn vị 50S
 - b) Ghép tARN khởi đầu với tiểu đơn vị 30S
 - c) Dịch chuyển ribosome
 - d) Hoạt hoá acid amin
 - e) Gắn kết mARN với ribosome
2. Trong quá trình dịch mã:
 - a) Mỗi tARN có một tARN – aminoacyl synthetase tương ứng
 - b) Một tARN – aminoacyl synthetase chung cho tất cả acid amin
 - c) tARN – aminoacyl synthetase kéo dài chuỗi peptid
 - d) Một tARN – aminoacyl synthetase cho mỗi loại acid amin
 - e) tARN – aminoacyl synthetase xúc tác sự chuyển vị
3. Trình tự Shine – Dalgarno là:
 - a) Vị trí gắn ribosom
 - b) Yếu tố khởi lắp ráp rARN

- c) Yếu tố khởi đầu dịch mã d) Vị trí kết thúc dịch mã
e) Là trình tự codon khởi đầu
4. Acid amin khởi đầu chuỗi peptid ở tế bào nhân nguyên thuỷ:
a) Formyl – methionin b) Methyl – Methionin
c) Methionin d) AUG
e) GUG
5. Vai trò của eIF – 2 trong tổng hợp protein ở tế bào nhân thật:
a) Mang aminoacyl – tARN tới vị trí A
b) Gắn Met – tARN vào ribosom
c) Hoạt hoá acid amin cần để nối dài
d) Tái hồi EF – Tu
e) Tái hồi EF – 1á
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 hòa hoạt động gen ở tế bào nhân nguyên thuỷ.
- Mô tả được cơ chế điều hòa 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 hòa 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.

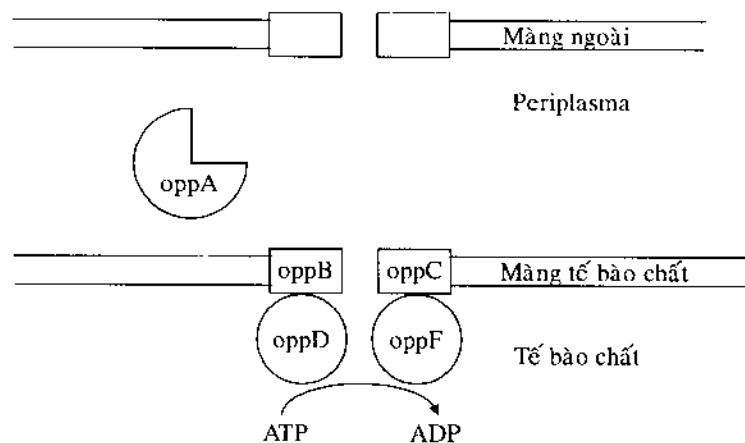
Ở vi khuẩn có ba loại enzym permease vận chuyển qua màng: permease dipeptid, permease tripeptid và oligopeptid permease. Oligopeptid permease có oppA, B, C, D, F lần lượt chuyển vật chất lẫn nhau và qua màng, trong đó oppA có tính đặc hiệu.

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, mắt xích di truyền cuối cùng của tế bào.



Hình 6.1. Các permease ở vi khuẩn

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 mồi (ARNo) 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 lợi 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 (ARNo) 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 dâng 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 sinh tổng hợp 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 cho chúng 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 sao 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.

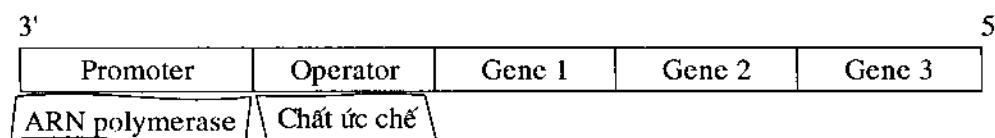
Hoạt động phiên mã của gen có thể không được điều hoà nếu sản phẩm của chúng không liên quan tới điều kiện môi trường. Có thể nói những sản phẩm không phụ thuộc vào môi trường như vậy được tổng hợp liên tục. Tuy nhiên số

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.

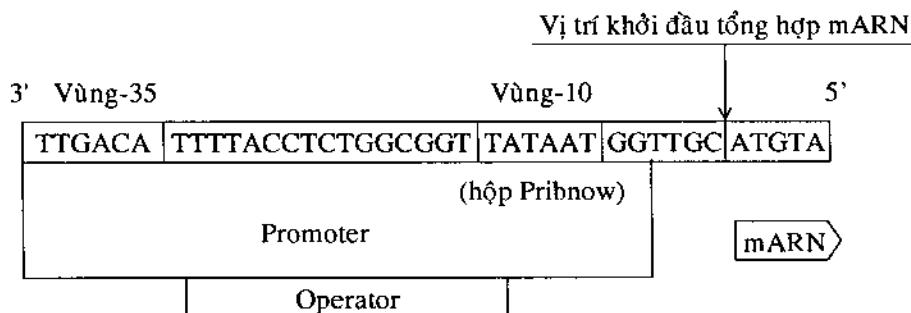
Đối với những protein đòi hỏi một điều kiện môi trường nhất định, thì sự hoạt động của các gen tạo ra chúng thường được điều hòa bởi một hay một số protein điều hòa.

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 khi 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 phần tử nhỏ (đường, acid amin...) đến những chất lớn hơn như các phức hợp steroid của hormon (tế bào nhân thực) 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 -galatosidase 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, mà trong đó 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 -galatosidase (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 mẫn cảm với **chất ức chế**, chất này ngăn chặn sự phiên mã phần còn lại của operon.
- Oc : allen điều khiển ổn định; nó không mẫn cảm với chất ức chế và để cho các cistron cận 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 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.
- is : 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 chồng 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 pi

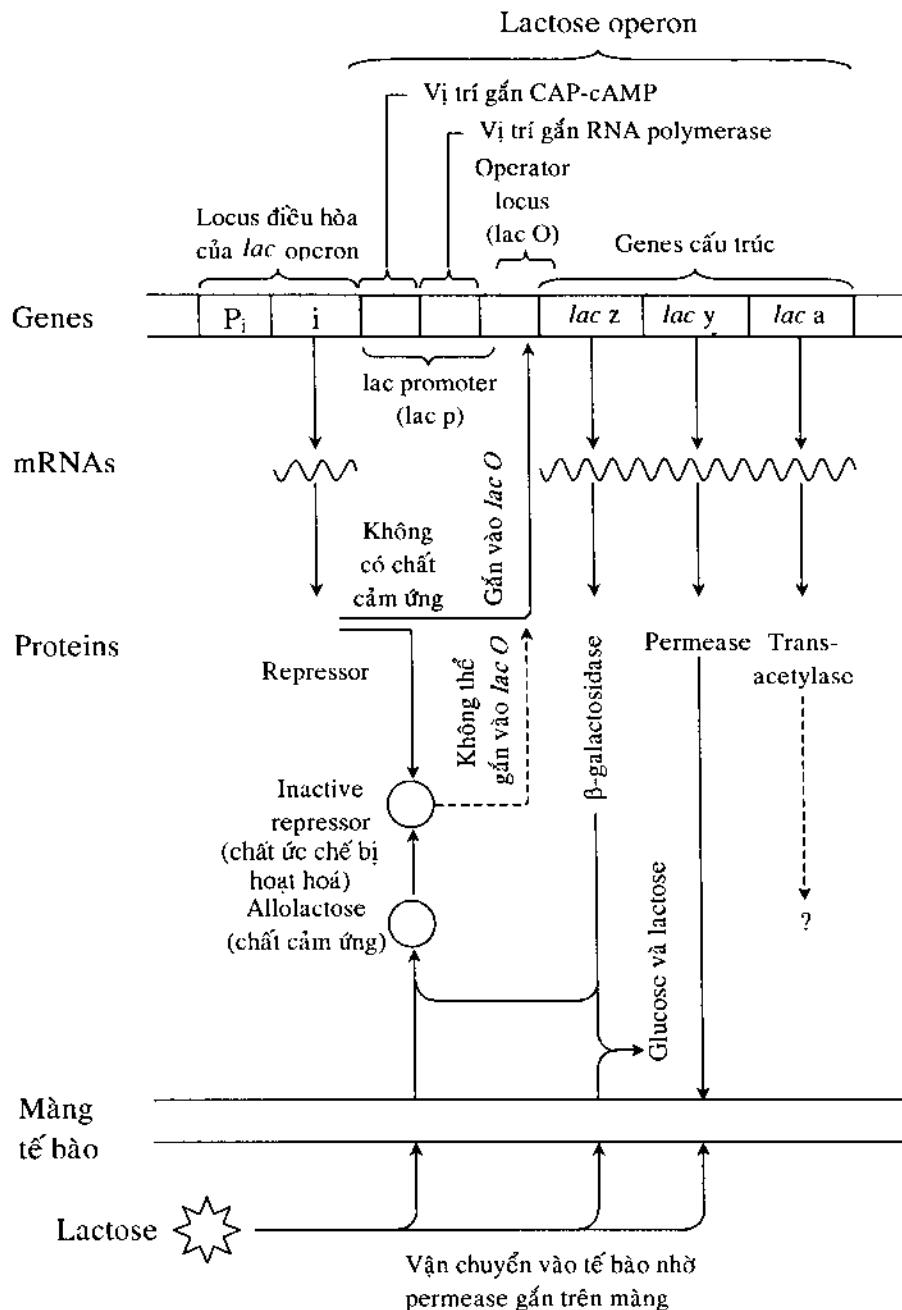
(inefficient promoter). Việc tổng hợp ra chất ức chế không thay đổi theo mức lactose trong tế bào. Ngược lại, promoter bình thường (P^+) của operon lactose (*lac* - p) lại gắn với ARN polymerase rất hiệu quả. Khi thiếu lactose (điều kiện không cảm ứng) thì protein ức chế có hoạt tính (sinh ra bởi i^+) 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. Từ 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 -galatosidase. 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 6.2). 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 huỷ bởi β -galatosidase. 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 lén 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 -galatosidase luôn luôn thường trực trong tế bào vi khuẩn.

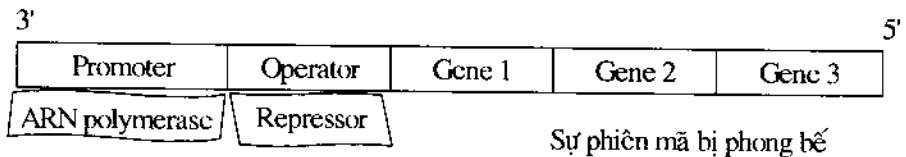
Các phân tử mARN vi khuẩn có thời gian bán huỷ rất ngắn (chỉ một vài phút), do đó sự tổng hợp protein ngừng lại rất nhanh sau khi phiên mã bị ức chế. Mặt

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.

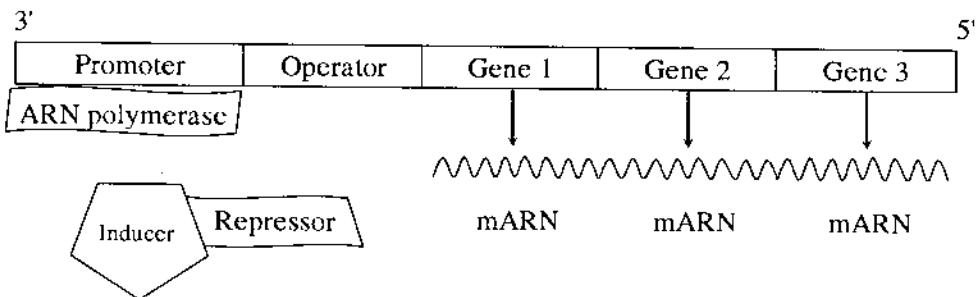


Hình 6.2. Sơ đồ các yếu tố chính kiểm soát lac operon

(a)



(b)

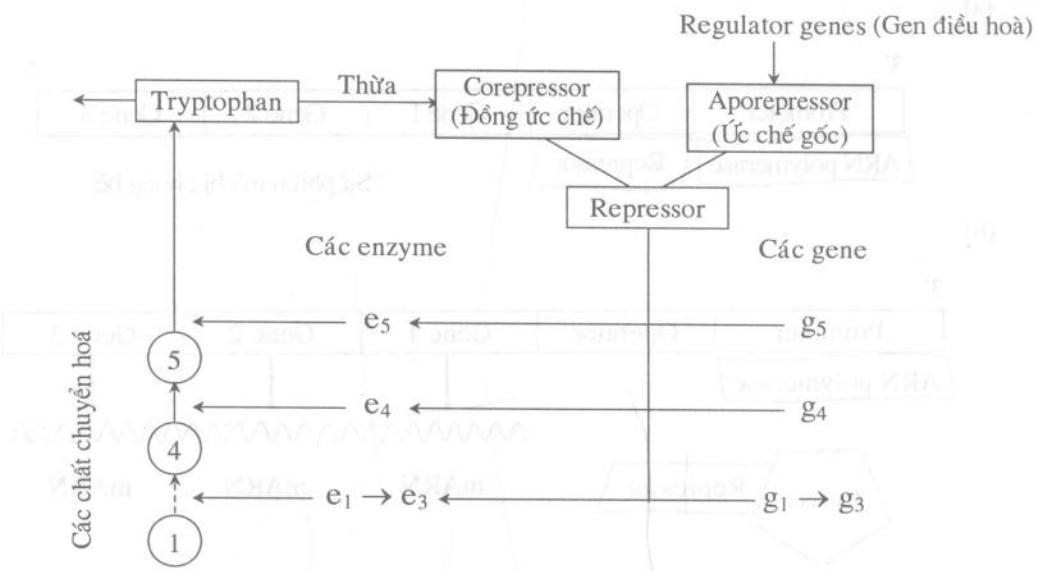
**Hình 6.3. Quá trình cảm ứng enzym**

- (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

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.

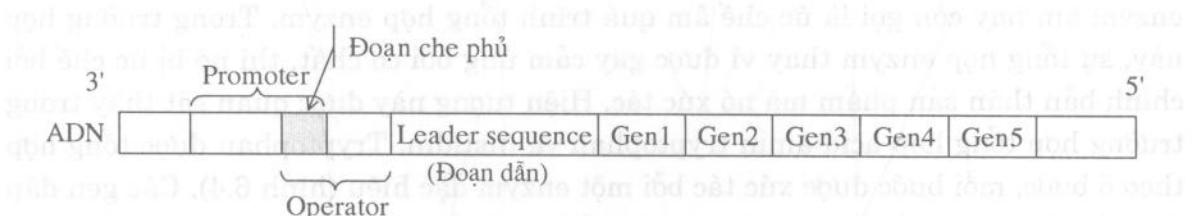
Gen điều hoà cho hệ thống này luôn luôn sản sinh ra một loại protein không chức năng, được gọi là chất **ức chế gốc**. Khi tryptophan sản xuất dư thừa, thì sự quá tải của tryptophan sẽ tác động như một chất **đồng ức chế**. Chất đồng ức chế gắn vào chất ức chế gốc hình thành nên một phức hợp ức chế chức 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 mARN đa cistron cho 5 enzym tổng hợp tryptophan.



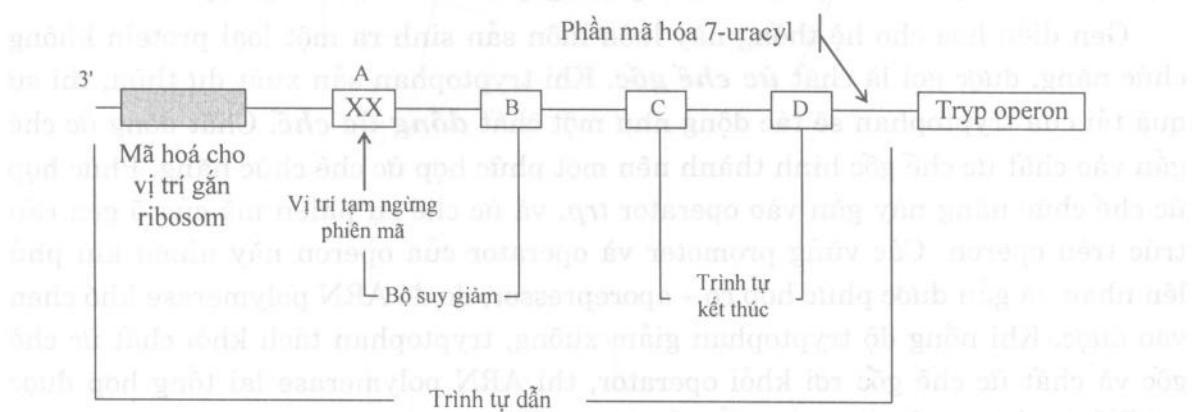
Hình 6.4. Kiểm soát ức chế âm trong hệ tryptophan của *E.coli*

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 mARN đa cistrone 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

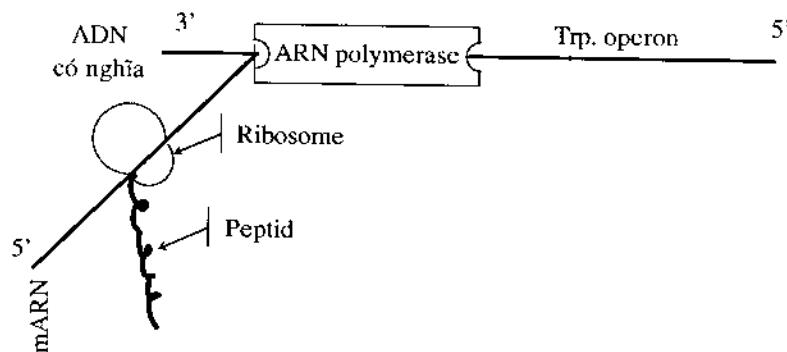


Hình 6.6. Sắp xếp các vùng nucleotid trong trình tự dẫn

Trình tự dẫn chứa một phần đóng vai trò bộ suy giảm, chính bộ suy giảm này mã hoá phân tử peptid dẫn chứa 14 acid amin báo hiệu cho sự phiên mã dừng lại.

Trình tự dẫn chứa 4 vùng, trong đó vùng A chứa hai codon mã hoá cho 2 codon mã sao cho tryptophan (UGGUGG) được ký hiệu XX. Phần D trước khi nối với Tryptophan operon có một trình tự mã hoá cho 7 – uracyl trên mARN.

ARN polymerase chuyển động đọc sợi có nghĩa ADN theo chiều 3' → 5' và phân tử mARN tương ứng sẽ được tổng hợp theo chiều 5' → 3'. Đầu 5' của mARN là vị trí gắn ribosom (hình 6.7).



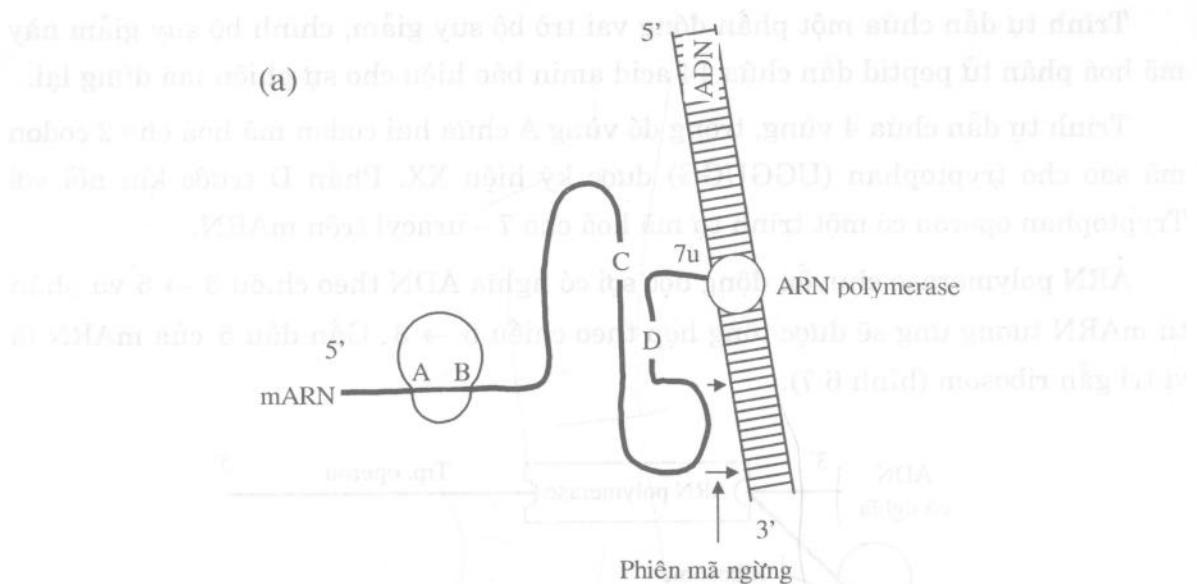
Hình 6.7. Sự chuyển dịch của ribosom trên mARN

Ribosom gần như chuyển dịch liên tiếp 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.

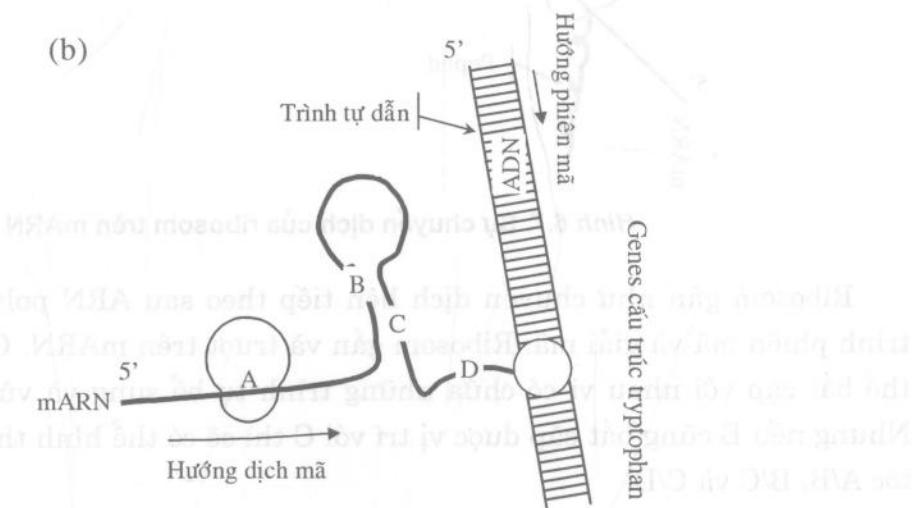
Trong điều kiện thừa tryptophan thì ribosom dịch mã ra sợi peptid dẫn, do đó vùng B và C của mARN không tạo nút được. Nhờ đó, vùng C và D bắt cặp được với nhau tạo thành nút kẹp tóc ngừng phong bế ARN polymerase. Khi đó Tryptophan operon không hoạt động. Nếu môi trường thiếu tryptophan thì ribosom sẽ 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 tổng hợp các enzym cho tryptophan (hình 6.8).

Sự phiên mã suy giảm của tryptophan operon cũng nhạy cảm với các nồng độ của một số acid amin khác ngoài tryptophan như histidin và leucin.



(b)



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

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

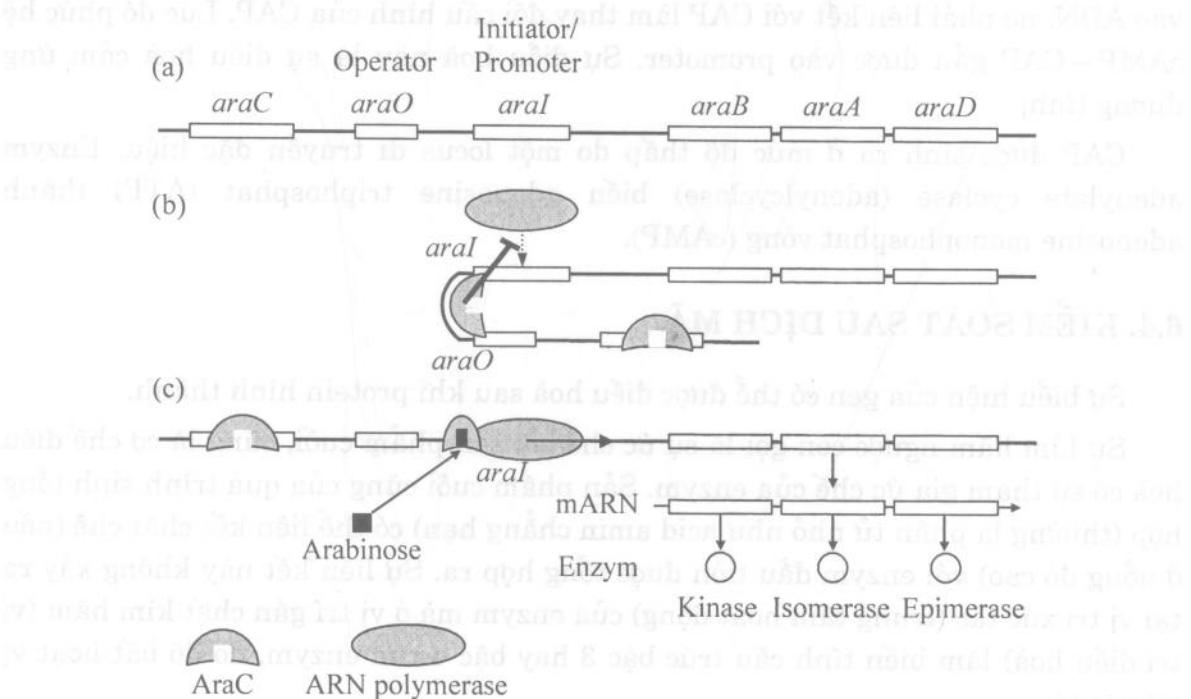
6.3.4. Kiểm soát cảm ứng dương

Đa số các hệ thống tổng hợp enzym cảm ứng được biết cho đến nay đều hoạt động nhờ sự điều hòa âm tính, giống như đã nêu trong ví dụ dị hoá 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 hoá của nó đòi hỏi 3 enzym được mã hoá 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 hoà bởi một gen *araC* nằm gần 3 gen trên nhưng cách chúng bởi phức hợp điều hoà (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*.

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 mARN 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.



Hình 6.9. Kiểm soát cảm ứng dương ở operon arabinose

Adenosine monophosphate 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 hợp 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.

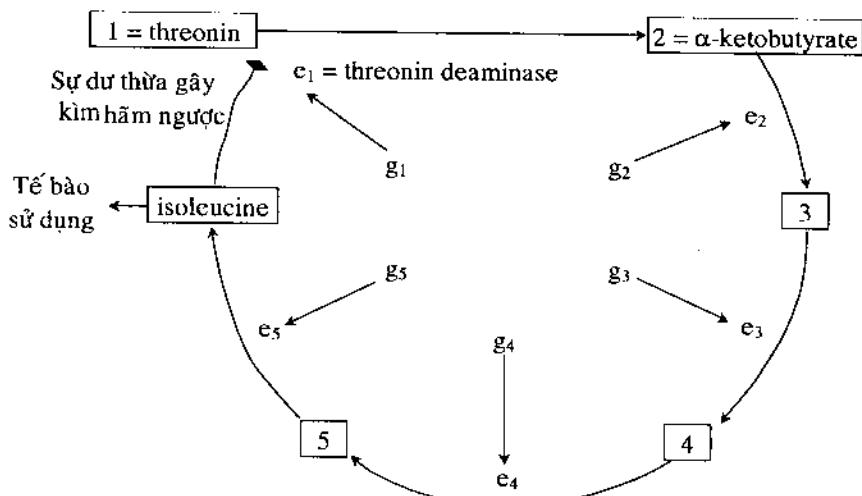
CAP được sinh ra ở mức độ thấp do một locus di truyền đặc hiệu. Enzym adenylate cyclase (adenylcyclase) biến adenosine triphosphat (ATP) thành adenosine monophosphat vòng (cAMP).

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 gần chất kìm 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.



Hình 6.10. Sơ đồ kìm hãm ngược trong *E. coli*

(e = enzym; g = gen; 1,2,3,4,5 = các chất 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 kìm hãm toàn bộ con đường này cho đến khi isoleucin trở lại mức bình thường 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 hóa 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 kìm 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.

Tóm tắt

Vi khuẩn sống trong một môi trường cạnh tranh gay gắt, chúng cần phải thay đổi nhanh chóng trong cơ chế chuyển hóa để sống sót. Nhiều vi khuẩn tạo ra các phương thức chuyển hóa để thích ứng với sự thay đổi này, các enzym trong quá trình chuyển hóa 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 hóa 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 hóa, 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 hòa 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 (Is)
 - b) Operon chứa promoter (Picr) và operator (Oc)
 - 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:
 - a) *E. coli* chuộng lactose hơn glucose
 - b) Promoter của operon này kém hiệu suất
 - c) Promoter của operon này gắn với ARN – polymerase
 - d) Chất ức chế được tổng hợp trong tế bào có thay đổi
 - e) a, b đúng
5. Chất ức chế gốc là:
 - a) Protein không chức năng sinh ra do gen điều hoà của hệ tryptophan
 - b) Tryptophan
 - c) 5 enzym tổng hợp tryptophan

- d) Trình tự dẫn
 - e) 5 chất chuyển hóa do 5 enzym tham gia tổng hợp tryptophan
6. Operon của arabinose được coi là operon nhạy cảm đối với glucose vì
- a) AMP tăng khi hàm lượng glucose tăng
 - b) AMP vòng có khả năng hoạt hóa promoter yếu
 - c) AMP muốn gắn vào promoter phải nhờ CAP
 - d) AMP gắn được vào promoter sẽ hoạt hóa ARN polymerase
 - e) b, c và d
7. Sự kìm hãm ngược **không** chấp nhận trường hợp:
- a) Sự ức chế bởi sản phẩm cuối cùng
 - b) Cơ chế điều hòa có sự tham gia ức chế của enzym
 - c) Sự liên kết giữa sản phẩm cuối cùng với enzym ở vị trí điều hòa của enzym làm bất hoạt vị trí xúc tác
 - d) Sự liên kết giữa sản phẩm cuối cùng với enzym ở vị trí xúc tác của enzym
 - e) Sự biến hình dị lập thể của enzym sẽ phong bế khả năng xúc tác của enzym
8. Operon gồm:
- a) Vùng khởi động (promoter)
 - b) Các gen cấu trúc
 - c) Vị trí điều hòa
 - d) Chóp GMP
 - e) a, b và c
9. Operator là:
- a) Đoạn mARN gắn được protein điều hòa
 - b) Đoạn ADN chuyên biệt gắn được protein điều hòa
 - 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 hóa 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 thuỷ.
- Mô tả được các loại trình tự chuỗi khác nhau tìm thấy 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 thuỷ. 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 thuỷ 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 thuỷ: sự dịch mã của phân tử mARN có thể bắt đầu trước khi mARN được tổng hợp xong.

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 biến 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 thuỷ (bảng 7.1).

Bảng 7.1. Bộ gen chứa acid nucleic

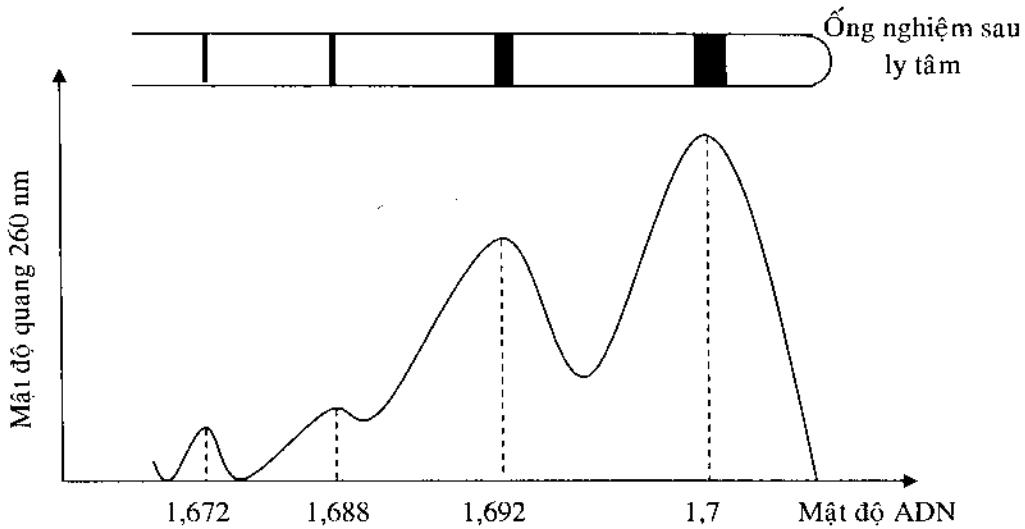
Bộ gen	Số lượng gen	Số cặp base
Thực vật	<50 000	$<10^{11}$
Người	25 000	$3 \cdot 10^9$
Giun	14 000	10^8
Ruồi	12 000	$1,6 \cdot 10^8$
Nấm	6 000	$1,3 \cdot 10^7$
Ví khuẩn	2 000 – 4 000	$<10^7$
Mycoplasma	500	$<10^6$
ssADN virus		
Parovirus	5	5 000
Phage fX174	11	5 387
dsARN virus		
Coronavirus	7	20 000
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 vẹt 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.

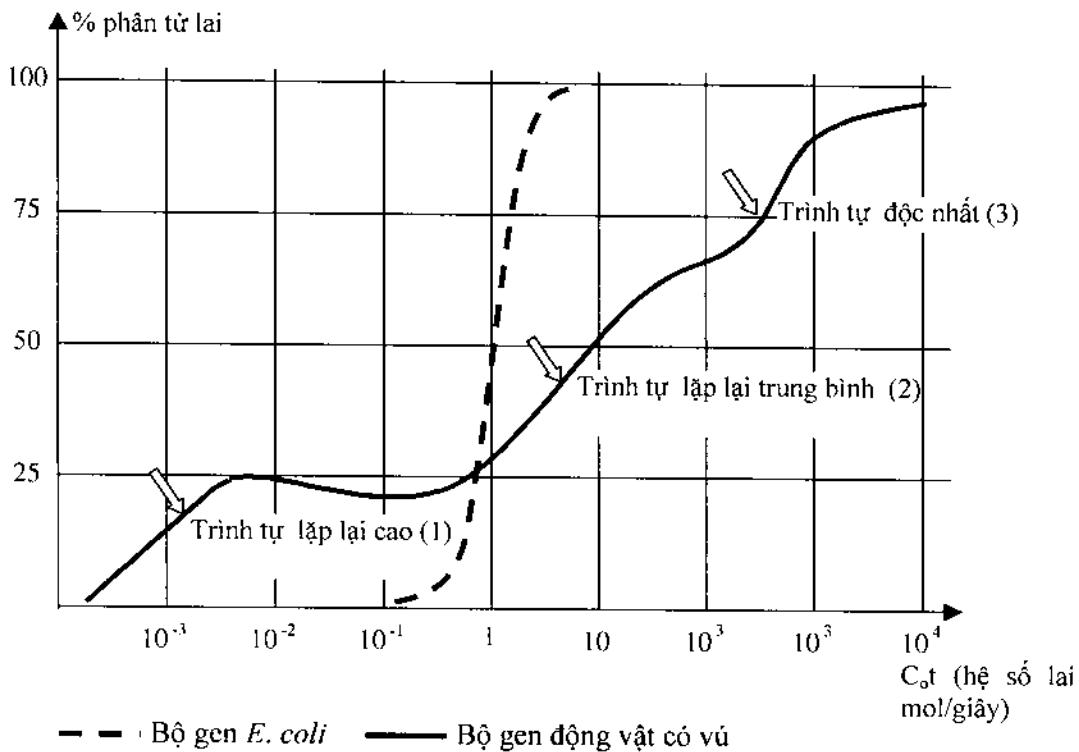
Động học tái hợp thành cặp của ADN tế bào nhân thật khác với tế bào nhân nguyên thuỷ (hình 7.2). Đường cong tái hồi của tế bào nhân nguyên thuỷ có dạng hình sigma điển hình, chứng tỏ sự đồng nhất của các trình tự trong tái hợp. Ở tế bào nhân thật đường cong phức tạp hơn, kéo dài một khoảng rộng các giá 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 $10^{-3} < C_0 t < 10$;
3. ADN bắt cặp rất chậm $10 < C_0 t < 10^3$

7.2.2. 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 các nồng độ gradient tỉ trọng cesium chlorid. 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 phân tử 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 di nhiễm sắc cơ cấu (constitutive heterochromatin).

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

- Họ ADN vi vệ tinh: gồm các đoạn nhỏ của trình tự lặp lại liên tục (tandem) 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 gấp, 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 gấp, 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 hoá, 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 vê 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 tandem 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 di nhiễm sắc như trình tự CEN và trình tự TEL.

+ Trình tự CEN: các trình tự CEN lặp lại cao là của các 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'TGATTTCGAA.

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 kìm 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 hoá.

Các chuỗi vê 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 ACAAACT, 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.

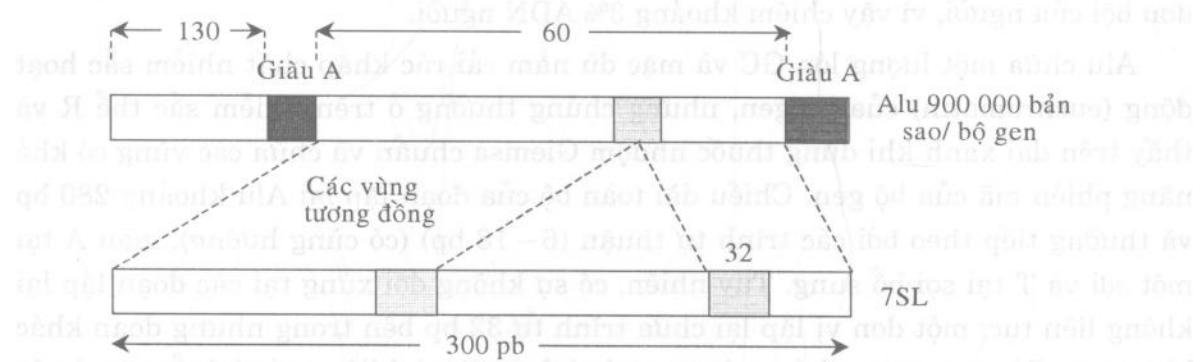
Alu chứa một lượng lớn GC và mặc dù nằm rải rác khắp chất nhiễm sắc hoạt động (euchromatin) của bộ gen, nhưng chúng thường ở trên nhiễm sắc thể R và thấy trên dải xanh khi dùng thuốc nhuộm Giemsa chuẩn và chứa các vùng có khả năng phiên mã của bộ gen. Chiều dài toàn bộ của đoạn lặp lại Alu khoảng 280 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 cüt 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 cüt 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 cự 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 gấp 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 biến hóa 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 mARN. *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.



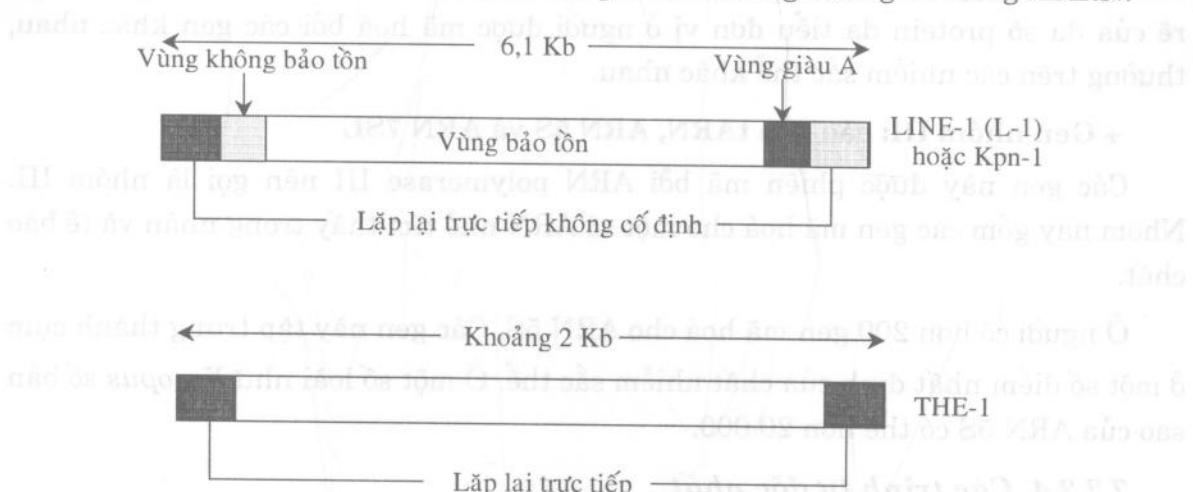
Hình 7.3. Các trình tự SINE

- 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 100000 bản sao từng phần rải khắp bộ gen người. Chúng là những trình tự lặp lại không mã hoá dài nhất và thường ở vùng giàu AT, giống như các trình tự Alu (hình 7.3). Các bản sao mã trình tự LINE gắn với protein tạo thành phức hợp ribonucleoprotein. Ở một dòng tế bào người bị ung thư vòm mạc (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.

ORF 1 ở gần cuối (đầu tận cùng 5') và mã hoá protein chưa rõ chức năng p40 (trọng lượng phân tử khoảng 40 kDa). ORF2 có những vùng giống với trình tự nucleotid mã hoá 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 3500 bản sao) và hầu hết các đoạn lặp lại bị cựt 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 dương) của nhiễm sắc thể kì giữa. Giống Alu, chúng thiếu các trình tự mã hoá nhưng có thể tìm thấy các trình tự không mã hoá 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 chép của 13 kb biểu hiện tiền 45S rARN sau đó trải qua nhiều phản ứng 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 hóa cho một số ARN nhỏ tìm thấy trong nhân và tế bào chất.

Ở người có hơn 200 gen mã hóa 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 hóa cho protein.
- Các gen giả.

Các gen mã hóa cho protein thuộc nhóm II vì nó được phiên mã do ARN polymerase II. Trừ các gen mã hóa 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 hóa 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 bổ 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.

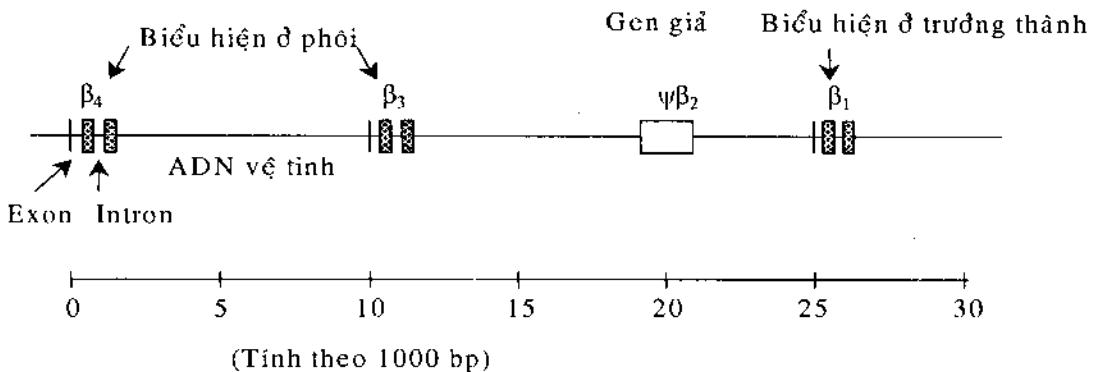
Bảng 7.2. Mạch Globine ở người trong quá trình phát triển cá thể

Peptide globine	Phôi	Thai	Người lớn
α - tương tự	ξ (xi)	α	α
β - tương tự	ϵ (epsilon)	γ , δ	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.

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 Hylap 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*.



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.

7.3. CÁC MỨC ĐỘ ĐIỀU HOÀ BIỂU HIỆN GEN

Ở tế bào nhân nguyên thuỷ, cũng như ở tế bào nhân thật, các cơ chế điều hòa sự biểu hiện của gen có thể tác động ở một hay nhiều mức độ khác nhau. Sự điều hòa 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).

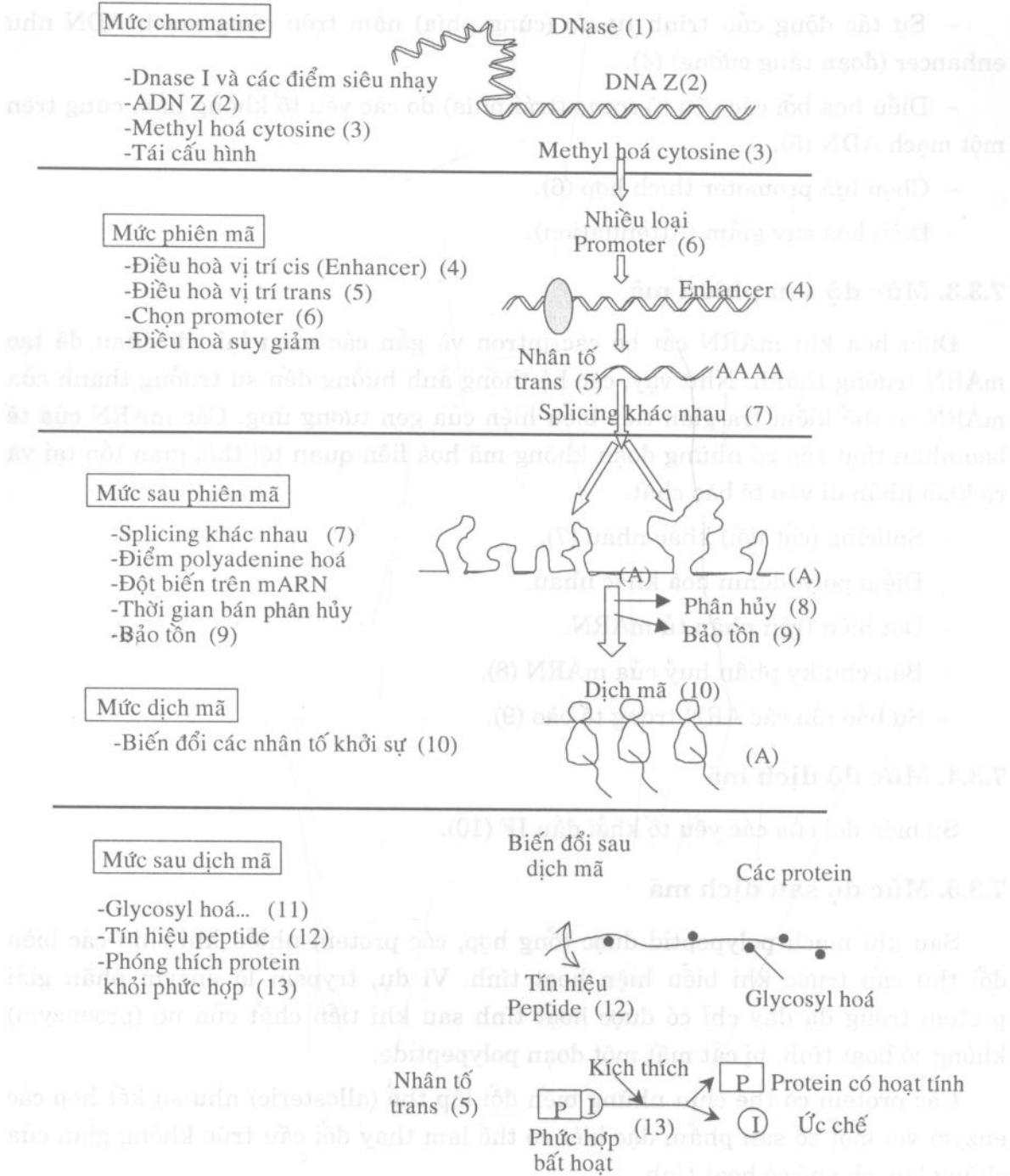
7.3.1. Mức độ chất nhiễm sắc

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 thuỷ xảy ra ở A và C, còn ở tế bào nhân thật ở C vị trí 5': Sự methyl hoá làm gen ngừng hoạt động. Ví dụ, nhiễm sắc thể X bất hoạt động ở người thuộc loại siêu methyl hoá.
- Sự thay đổi cấu hình có thể ảnh hưởng đến sự biểu hiện gen.



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.

- 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.

- 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 nhiều khi chịu các biến đổi thứ cấp trước khi biểu hiện 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 polypeptide.

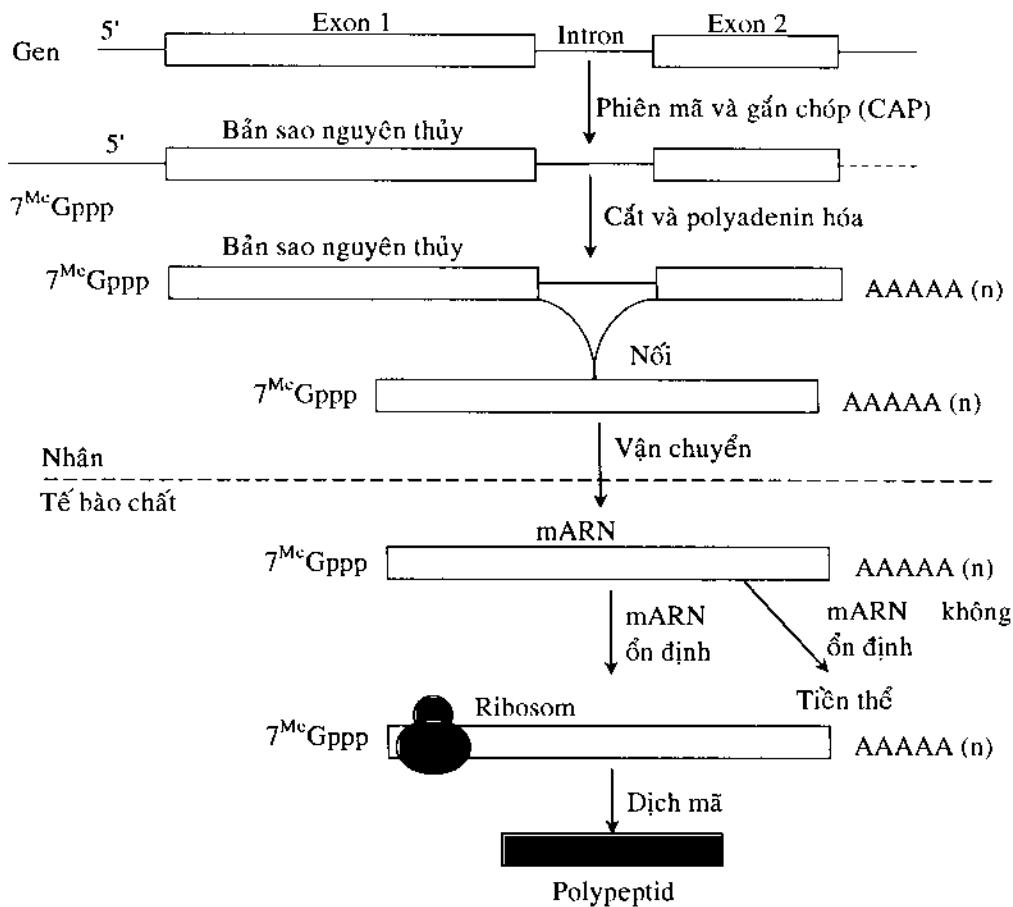
Các protein có thể chịu những biến đổi lập thể (allosteric) như sự kết hợp các enzym với một số sản phẩm đặc biệt có thể làm thay đổi cấu trúc không gian của chúng làm cho nó có hoạt tính.

- Glycosyl hoá, phosphoryl hoá... (11): gắn thêm các nhóm chất như đường, phosphate 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 lối 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

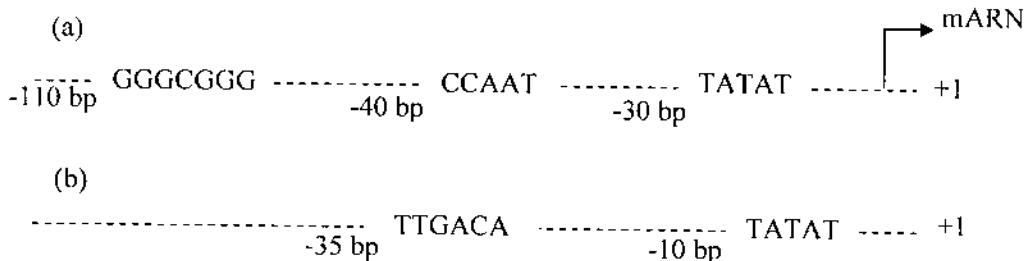
Một số đặc điểm của điều hoà hoạt động gen ở tế bào nhân thật cần được nhấn mạnh:

- Ở các operon của tế bào nhân nguyên thuỷ, 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 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 thực 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 phản hồi đáp lại tín hiệu ngoại sinh. Ngược lại, phản hồi sự điều hoà ở tế bào nhân thực 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 thực cũng nằm phía trước điểm xuất phát của mARN 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 thực; (b) Tế bào nhân nguyên thuỷ

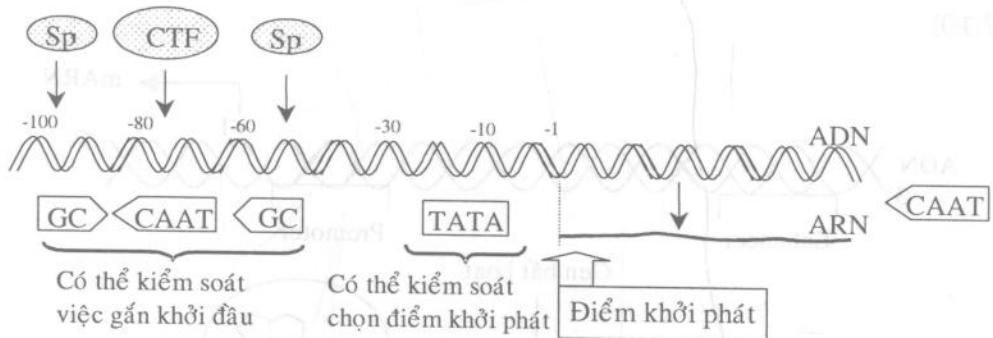
Sự thay đổi hộp TATA làm giảm tốc độ phiên mã. Sự giảm tốc độ phiên mã được đo bằng sự thay đổi của từng base trong promoter. Các thay đổi base ngoài hộp TATA và các trình tự phía trước không gây tác động đối với sự phiên mã. Khác với promoter của tế bào nhân nguyên thuỷ, các promoter của tế bào nhân thực 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 thực. Tuy nhiên, hầu hết các gen mã hoá 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 thực 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 thực, 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.

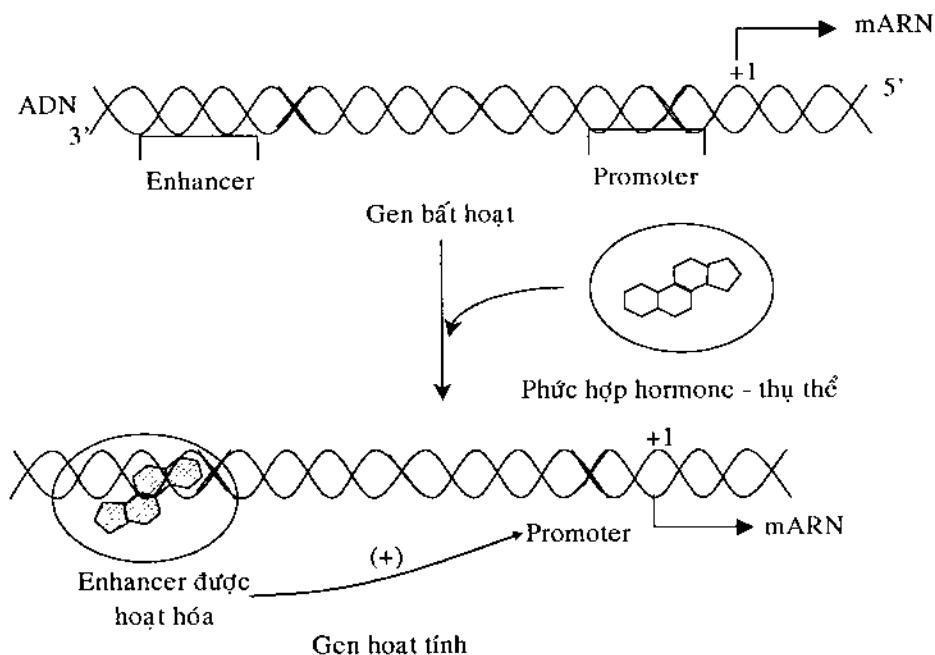
7.4.2. Các yếu tố tăng cường (Enhancer)

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 đặc 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 sao 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 thế 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 ở 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 của chuột. Khoảng cách giữa một trình tự tăng cường và một trình tự khởi động lớn nhất gấp ở gen albumine của 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 hòa 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 phân 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 dime (cấu tạo từ hai tiểu đơn vị).

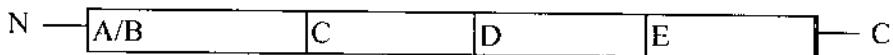
7.4.3. Các yếu tố protein tham gia vào quá trình điều hoà biểu hiện gen, các yếu tố *trans*

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.

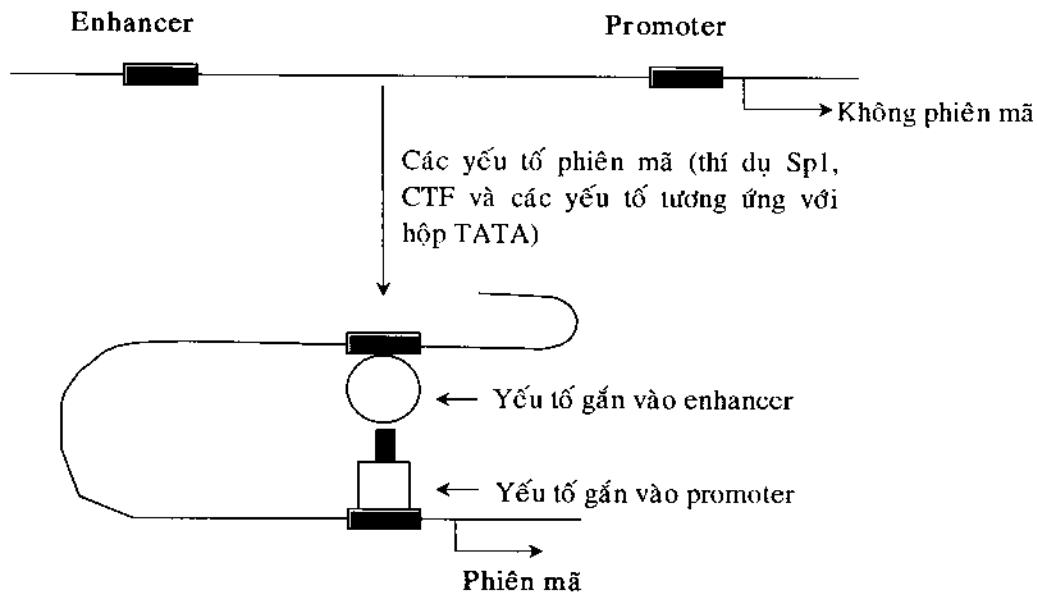
Hormon tuyến giáp khi gắn lên vùng E của thụ thể sẽ làm biến đổi cấu hình thụ thể. Lúc đó, vùng C của thụ thể có thể gắn lên vùng ADN điều hoà của các gen đích và kích thích sự phiên mã của các gen này.

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 bắt 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 polymesare 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 thực 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 đối đào 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 mARN dư. Ví dụ, globin trong tế bào lưỡi nội chất, sợi cơ và myosin trong các tế bào cơ. Vì vậy, các mARN dư nhiều thường là các bản sao của các gen chuyên biệt của tế bào.

Bảng 7.3. Các nhóm mARN dư ở vòi trứng gà và gan chuột nhắt

Loại tế bào	Nhóm dư	% của tổng mARN trong nhóm	Số lượng các loại mARN khác
Vòi 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

Một tế bào chứa vài trăm mARN dư trung bình, chiếm khoảng 33% tổng số mARN tế bào, nhưng mỗi loại mARN dư chỉ chiếm 0,1 – 0,5% tổng số mARN. Phần mARN 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 mARN.

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 *Drosophila*. 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 thuỷ. 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 dồi dào 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) Kiềm 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

Bài 8

ĐỘT BIẾN GEN

MỤC TIÊU

- Trình bày được đột biến xảy ra như thế nào, có thể có hậu quả gì với tế bào.
- Nêu 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 hóa.

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 đo 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 hâm 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
	$StrS \rightarrow StrR$ (nhạy → kháng Str)	4.10^{-10}	Đột biến / tế bào
<i>Neurospora crassa</i>	Adenin: $ade^- \rightarrow ade^+$	4.10^{-8}	Đột biến / bào tử vô tính
Bắp	$R \rightarrow r$ (đỏ → trắng)	$4.92.10^{-8}$	Đột biến / giao tử
	$C \rightarrow c$ (màu → không màu)		Đột biến / giao tử
	$Wx \rightarrow wx$ (tẻ → nếp)		Đột biến / giao tử
	$Y \rightarrow y$ (vàng → trắng)	$2.2.10^{-6}$	Đột biến / giao tử
<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 → 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 thuỷ 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 x 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 đâu 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

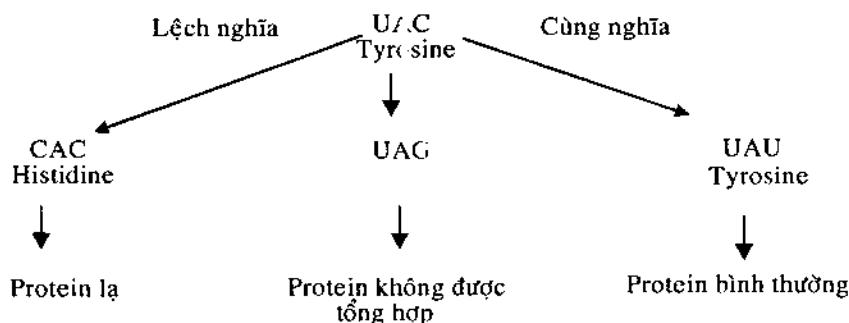
Phổ biến nhất là chuyển vị, pyrimidin được thay thế bởi pyrimidin khác hay purin bởi purin khác ($A \leftrightarrow G, C \leftrightarrow T$). Sự chuyển vị, bắt cặp sai, có thể gây bởi acid nitro hoặc các đồng đẳng base như 5 - bromo - 2 - deoxyuridin (BrdU). Cách đột biến điểm ít phổ biến hơn là đảo 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 thuỷ. 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 tuy 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 hòa 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 -thalassia 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 hóa 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 hóa. 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 mARN, 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 sao mã và dịch mã của bất kỳ gen nào chúng 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 bởi 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 hỗn biến, khử amin, chuyển vị, đảo chuyển, đột biến lệch khung, oxy hoá phá huỷ.

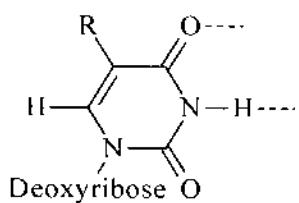
- Hỗn biến: 4 loại base trong ADN có thể tồn tại ở dạng hỗn biến, tức là có khả năng tồn tại hoán chuyển giữa hai dạng (Keto ↔ Enol và Amino ↔ Imino). Base pyrimidin bình thường có dạng keto (C=O), nhưng đôi khi chúng cũng có dạng enol (C – OH). Ở dạng enol, thymine có thể bắt cặp với guanine và enol cytosine có khả năng bắt cặp với adenine (hình 8.2a) tương tự base purin bình thường tồn tại dưới dạng amino (NH_2) nhưng cũng có dạng hỗn 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 hỗn biến chỉ gây các đột biến chuyển vị.

- Khử amin: Các phản ứng khử amin thông thường là phản ứng thuỷ phân cytosine thành uracil, quá trình này phóng thích ammonia và khử amin 5-methylcytosine thành thymine và ammonia (hình 8.2b). Trong ADN, sự khử amin cytosine tự phát được chỉnh lại bằng cách loại bỏ uracil (vì uracil không có trong ADN) và thay thế bằng cytosine, còn sự khử amin 5-methylcytosine không được chỉnh vì cơ chế sửa chữa không nhận diện Thymine là lỗi. Trong ADN bộ gen cytosine tại các trình tự CG (CpG) thường bị methyl hoá, 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.

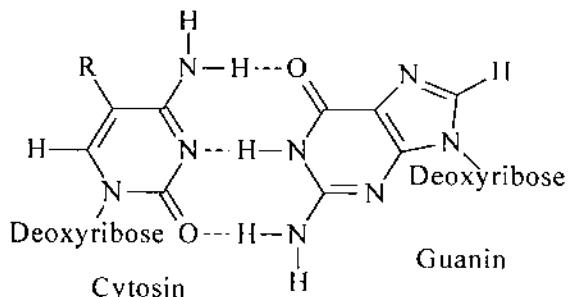
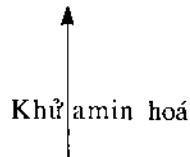
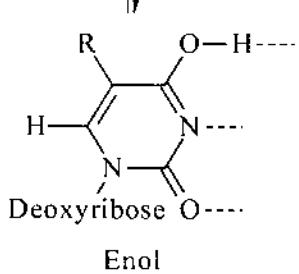
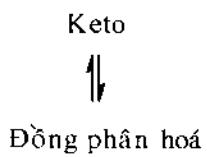
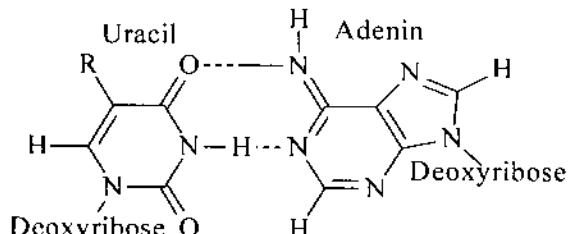
- Đột biến lệch khung (chèn hoặc mất trên một sợi), thường là lỗi của polymerase khi sao chép các đoạn lặp lại của một nucleotid. Mỗi polymerase có độ chính xác khác nhau. Yếu tố chính ảnh hưởng đến độ chính xác của polymerase là có exonuclease đọc sửa 3' – 5', làm loại bỏ các base bắt cặp không đúng chèn vào bởi polymerase.

– Oxy hoá phá huỷ do các gốc oxy. Các gốc oxy tăng trong tế bào do chuyển hoá oxy hoá (và cũng được tạo bởi các tác nhân vật lý như tia phóng xạ). Sản phẩm oxy hoá quan trọng là 8-hydroxy guanin bắt cặp sai với A, tạo đảo chuyển G – C thành T – A.

(a) Thymin ($R=CH_3$)
hay Bromouracil ($R=Brom$)



(b)



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 nitro (HNO_2), 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.

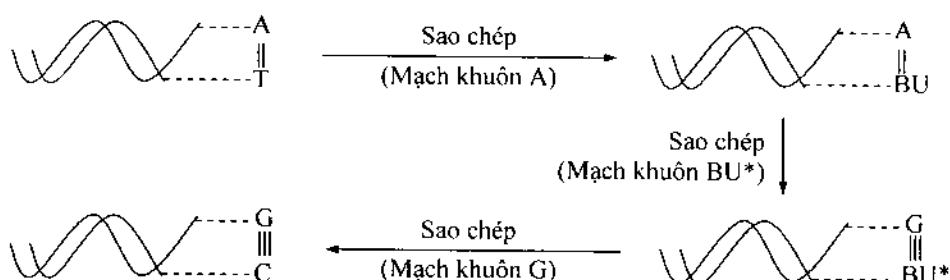
Tác nhân gây đột biến hóa học có thể được chia ra làm 4 nhóm:

- Nhóm base đồng đẳng.
- Các chất hóa 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.

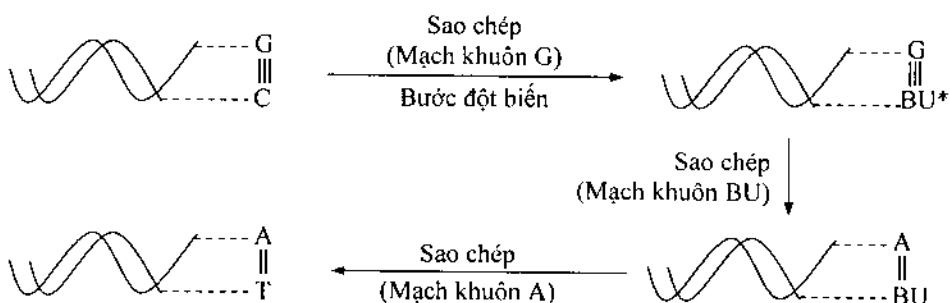
Base đồng đẳng:

Các chất hóa học này có cấu trúc hóa 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.

(a)



(b)



Hình 8.3. Tác động gây đột biến của bromouracil (BU)

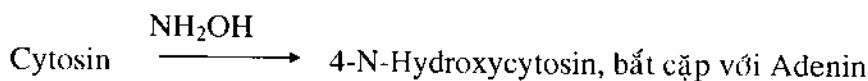
- **Bromouracil (BU)** là hợp chất nhân tạo dùng rộng rãi trong nghiên cứu. Giống với thymine (có nguyên tử Br thay cho nhóm methyl) và sẽ gắn vào ADN và bắt cặp với Adenine giống như Thymine. Có nhiều khả năng hô biến thành dạng enol (BU^*) và có khả năng bắt cặp với Guanine. Như vậy, BU có thể xâm nhập và bắt cặp bổ sung với Adenine thay cho Thymine tạo A - BU và ở vòng sao chép tiếp theo BU lại hô biến thành BU^* và bắt cặp với Guanine tạo $\text{BU}^* - \text{G}$, do đó A - T được thay thế thành G - C (hình 8.3a). Trong khi đó BU^* có thể xâm nhập sai, khi gắn

thay chỗ cho Cytosine tạo G – BU*, sau đó BU* lại hô 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 8.3b).

– **Aminopurin** là đồng đẳng của Adenin, có thể bắt cặp với Thymin hoặc (rất ít) với Cytosine, gây chuyển vị A – T thành G – C hoặc G – C thành A – T. Các đồng đẳng base gây chuyển vị như gây hô biến tự phát.

Các chất hóa học thay đổi cấu trúc và đặc tính bắt cặp của các base:

– **Các chất khử amin.** Ví dụ điển hình: acid nitro tạo bởi sự phân giải các nitrit (chất bảo quản) trong thực phẩm làm Cytosine thành Uracil, Methylcytosine thành Thymin, và Adenine thành Hypoxanthine khử amin, Hypoxanthine trong ADN bắt cặp với C gây chuyển vị. Methylcytosine có mặt trong các gen của sinh vật bậc cao có thể ảnh hưởng tới hoạt tính phiên mã. Các base ngoại như Xanthine, Hypoxanthine và Uracil có thể khởi động cơ chế sửa chữa. Thymin, dẫn xuất từ Methylcytosine nội sinh và nhiều sự chuyển vị có thể xảy ra theo cách này. Các gốc 5 – methylcytosine, sau đó trở thành "điểm nóng" của sự đột biến. Hydroxylamin (NH_2OH) chỉ khử amin của Cytosine, do đó có tác động rất giới hạn:



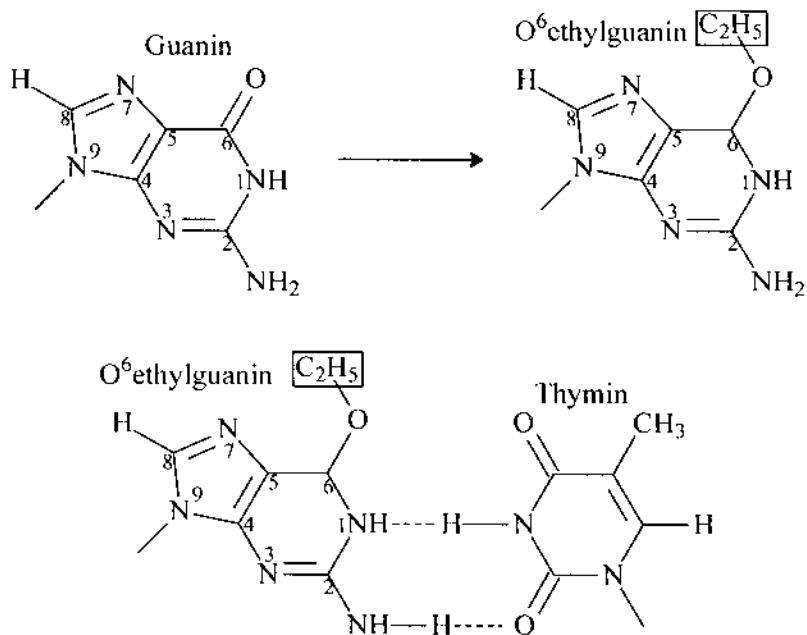
Các chất alkyl hoá. Ví dụ điển hình: Nitrosoguanidin (NTG), Methyl Methane Sulfonate (MMS), Ethyl Methane Sulfonate (EMS) là các tác nhân hoá học gây đột biến phản ứng với các base và thêm các nhóm Methyl hoặc Ethyl nên còn gọi là các tác nhân alkyl hoá (bảng 8.2). Tuỳ yếu tố bị tác động, chúng có thể gây đột biến ít nhất bằng 3 cách:

– Thêm nhóm Methyl ($-\text{CH}_3$) hay ethyl ($-\text{C}_2\text{H}_5$) vào Guanine tạo base đồng đẳng của Adenine dẫn đến sự bắt cặp bổ sung sai (hình 8.4).

– Mất purin do Guanine 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.

Guanine là base có thể dễ bị tấn công nhất, vòng pyrimidine, 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 guanine hay 4 – keto oxygen của Thymin và các vị trí cách xa vị trí ghép cặp như N – 7 của Guanine.



Sự bắt cặp của O⁶ ethylguanin với Thymin

**Hình 8.4. Hiệu quả dự đoán khi xử lý Guanin
với tác nhân alkyl hoá (EMS hay MMS)**

Bảng 8.2. Một số tác nhân alkyl hoá được sử dụng để gây đột biến cảm ứng

Tên hóa học	Viết tắt	Công thức hóa học
Methyl methan sulphonat	MMS	CH ₃ – SO ₂ – O – CH ₃
Ethyl methan sulphonat	EMS	CH ₃ – SO ₂ – O – CH ₂ – CH ₃
Dimethyl sulphonat	DMS	CH ₃ – O – SO ₂ – O – CH ₃
Diethyl sulphat	DES	CH ₃ – CH ₂ – O – SO ₂ – O – CH ₂ – 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. Tỉ 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.

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 càng) 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. Ba phát hiện này và những cái khác đã khai sinh Vật lý nguyên tử và kiến thức về bức xạ điện từ.

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. Mối tương quan giữa các đơn vị này là: rad \times RBE = 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 nên. 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 hóa 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 tuỳ 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:

- Dứt 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):

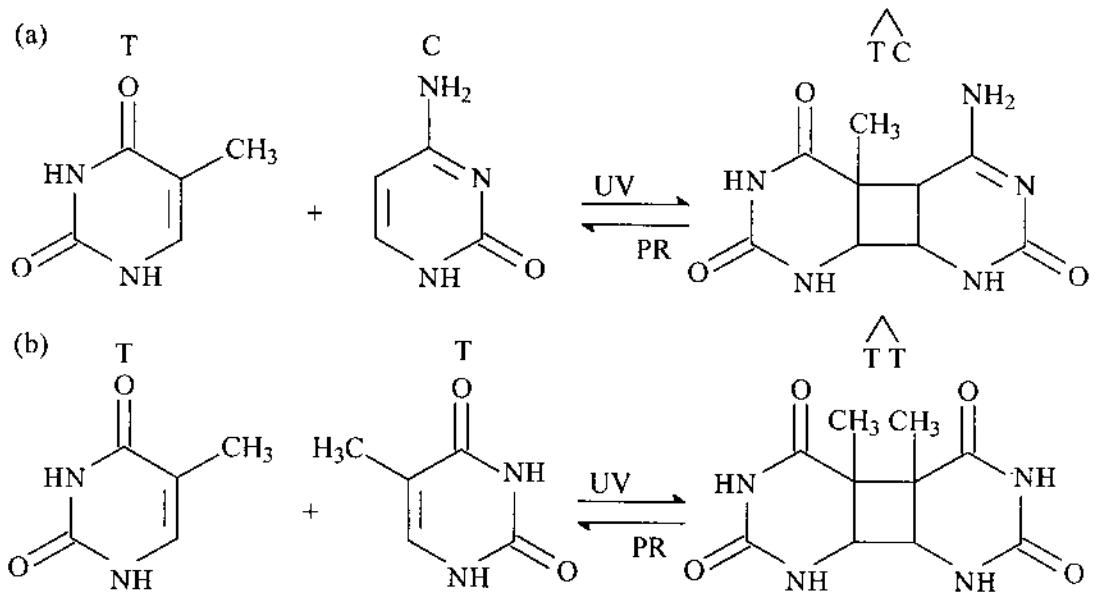
Bức xạ UV ít năng lượng và do đó không ion hoá, nhưng năng lượng của nó được ưu tiên hấp phụ bởi các base của ADN và bởi các acid amin thơm của các protein, vì thế nó cũng có ảnh hưởng sinh học và ảnh hưởng di truyền quan trọng.

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/dột biến chính của ánh sáng mặt trời; UV – A (320 nm – khả kiến) là "cận UV", cũng có tác dụng xoá mất (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 hoá 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 thu 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 thu tia tử ngoại mạnh nhất ở bước sóng 257 nm (\approx 260 nm), đây chính là bước sóng làm tăng tần số đột biến ở hạt phấn cây bắp (ngô). 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 phần 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:

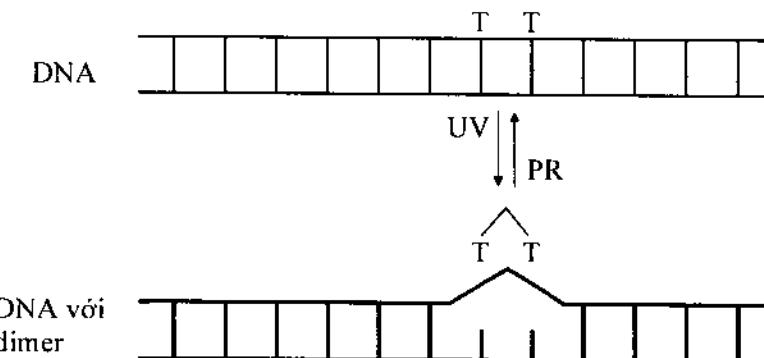
- Đảo nghịch sai hỏng: cơ chế này đơn giản nhất, enzym hoạt động phục hồi cấu trúc bình thường không có gãy khung.
- Loại bỏ sai hỏng: cắt bỏ và thay thế base hoặc phân nucleotid sai hỏng hoặc không thích hợp.
- Dung nạp sai hỏng: không sửa chữa thật sự, chép sai hỏng và tiếp tục sống.

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 hóa trị) khi có ánh sáng (hình 8.6).

Enzym photolyase xúc tác phản ứng này, nó có nhiều trong vi khuẩn, tế bào nhân thực 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.

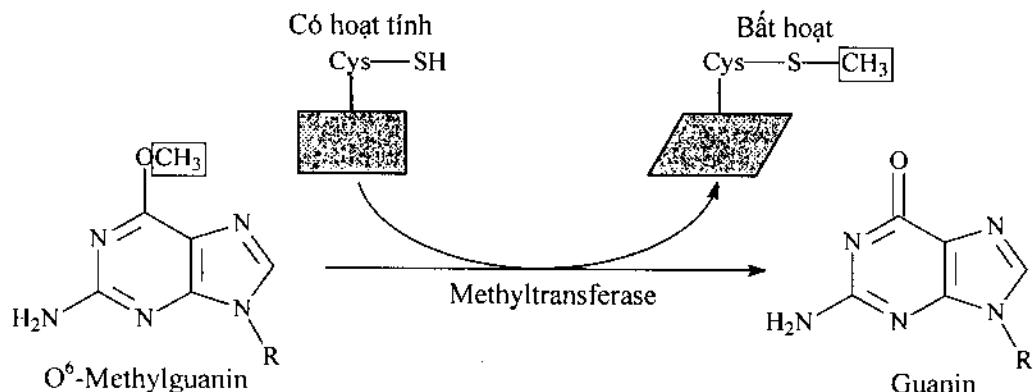


Hình 8.6. Quang phục hồi

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 – methylthymine, 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 O^6 – ethylguanin và O^6 – 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* (methyladenine 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 O^6 – MeGuanin và O^4 – MeThymin, để khôi phục base bình thường.



Hình 8.7. Cơ chế sửa sai bằng methyltransferase

8.4.1.3. Nỗi các chẽ đứt của sợi đơn

Tia X và một số hoá chất như peroxid có thể gây ra những chẽ đứt trên khung ADN. ADN ligase nhanh chóng sửa các chẽ 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 chẽ sai nhờ một hoặc vài enzym N – glycosylase. N – glycolase nhận biết base biến đổi và thuỷ giải liên kết N – glycosilic 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.

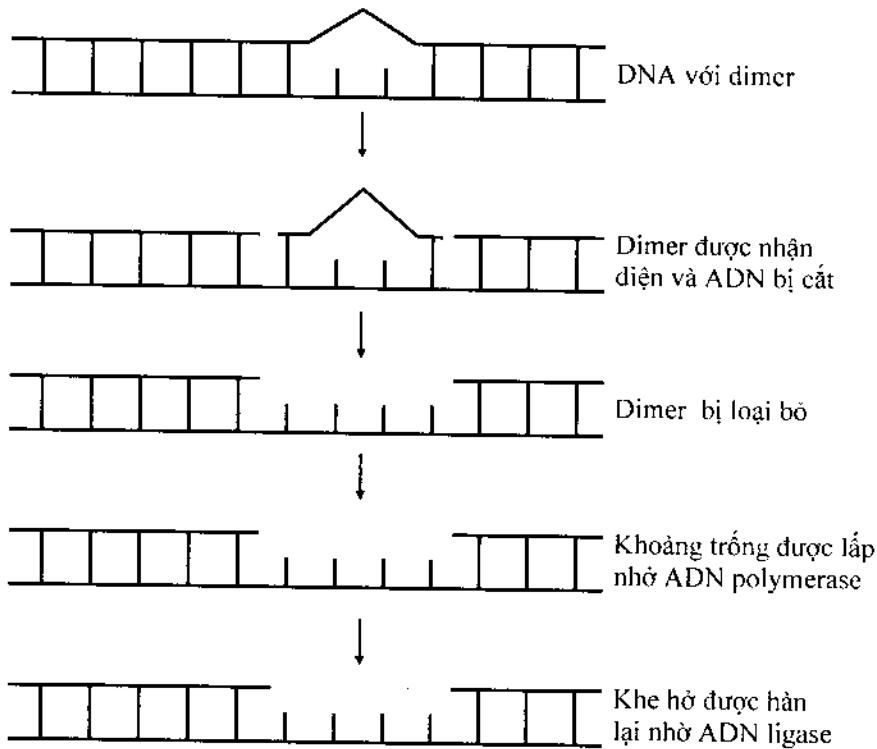
8.4.2.2. Sửa chữa lệch đôi

Quá trình này xảy ra sau sao chép ADN như là "đánh vân 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 bất 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 thuỷ 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).

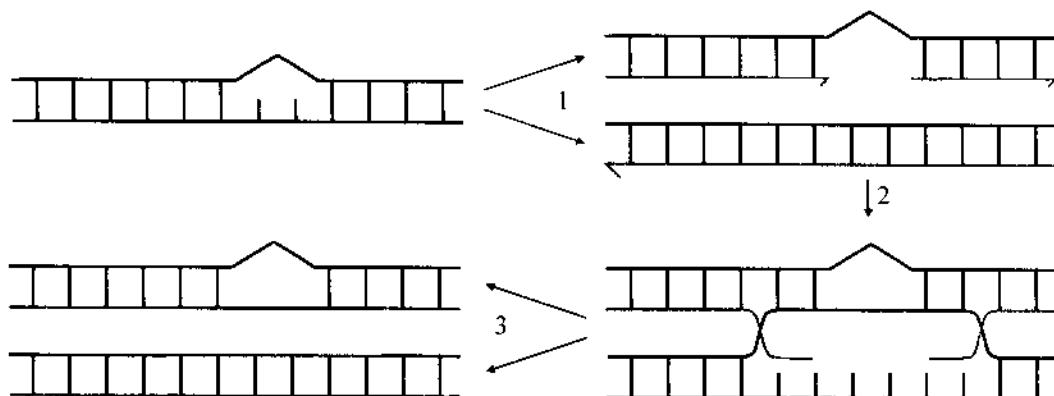
8.4.3. Dung nạp ADN sai hỏng

Không phải tất cả các ADN sai hỏng được loại bỏ ngay mà một số có thể tồn tại một thời gian. Nếu tại chac 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, chữa 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 đổi 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.

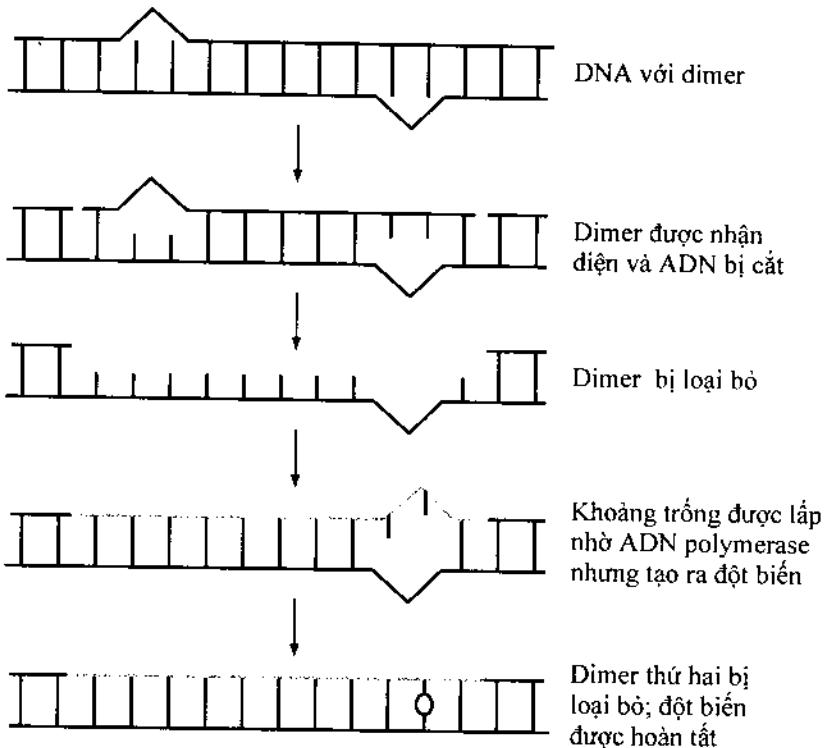


Hình 8.9. Các biến cố trong sửa chữa bằng tái tổ hợp

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 thực 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.

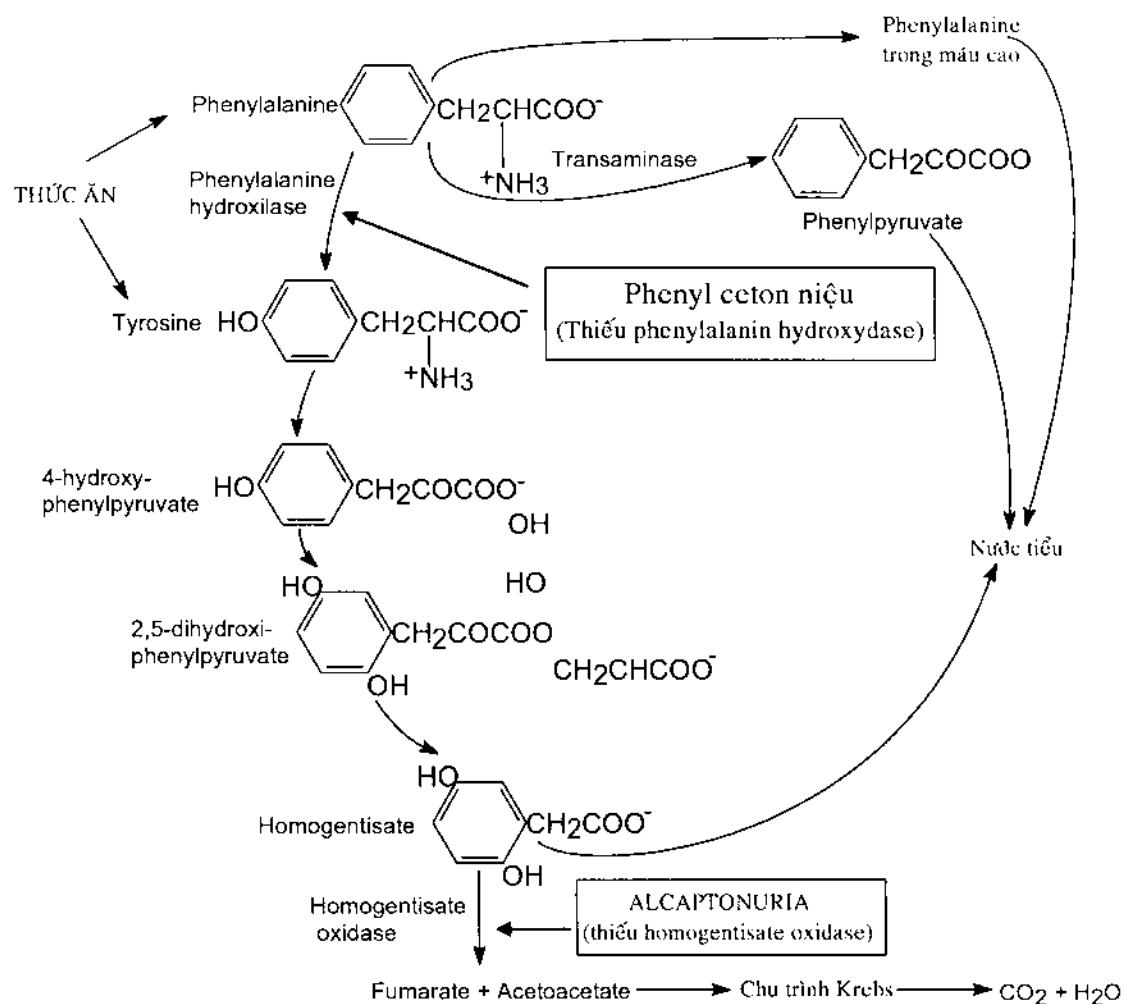
8.5. CÁC TÍNH TRẠNG ĐỘT BIẾN VÀ PROTEIN ĐỘT BIẾN

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 hóa học mang bản chất di truyền.

Các rối loạn chuyển hóa 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.



Hình 8.11. Chuỗi phản ứng sinh hóa của bệnh Alkapton niệu và Phenylketon niệu, hai sự rối loạn di truyền của sự thoái quá acid amin thơm

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.

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 phân 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

Loại rối loạn	Tên	Triệu chứng điển hình	Hoạt tính sinh hóa bị ảnh hưởng hay khiếm khuyết	Tần số tính trên 1000 người
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 (Pyruvate 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)
Chuyển hóa 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 hóa acid amin	Phenylketon niệu	Mức phenylalanine, mức phenylacetic acid trong huyết tương và nước tiểu cao; trì trệ chậm phát triển	Phenylalanine hydroxylase trong gan	0,1
Chuyển hóa purin	Hội chứng Lesh – Nyhan	Tổn thương não non thần kinh, thường có hành vi tự huỷ hoại	Hypoxanthine guanine phosphoribosyl transferase	0,1 (đàn ông)
Chuyển hóa Hem	Protoporphyrinia	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

<i>Loại rối loạn</i>	<i>Tên</i>	<i>Triệu chứng điển hình</i>	<i>Hoạt tính sinh hóa bị ảnh hưởng hay khiếm khuyết</i>	<i>Tần số tính trên 1000 người</i>
Vận chuyển lipoprotein	Cao cholesterol máu gia đình	Xơ vữa động mạch (Atherosclero – sis)	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 Glucosephingolipid	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. Allen biến đổi được gọi là oncogen (gen ung thư), còn allen 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
abl	Protein kinase (tyrosin)	Chuột, mèo	Pre – B – celleukemia
erb – B	Protein kinase (tyrosin); thụ thể của yếu tố tăng trưởng thượng bì (EGF)	Gà	Erythroleukemia
Sis	Yếu tố tăng trưởng từ tiểu cầu (PDGF)	Khỉ	Fibrosarcoma
Src	Protein kinase	Gà	Sarcoma, Carcinoma, Myelocytoma
			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.

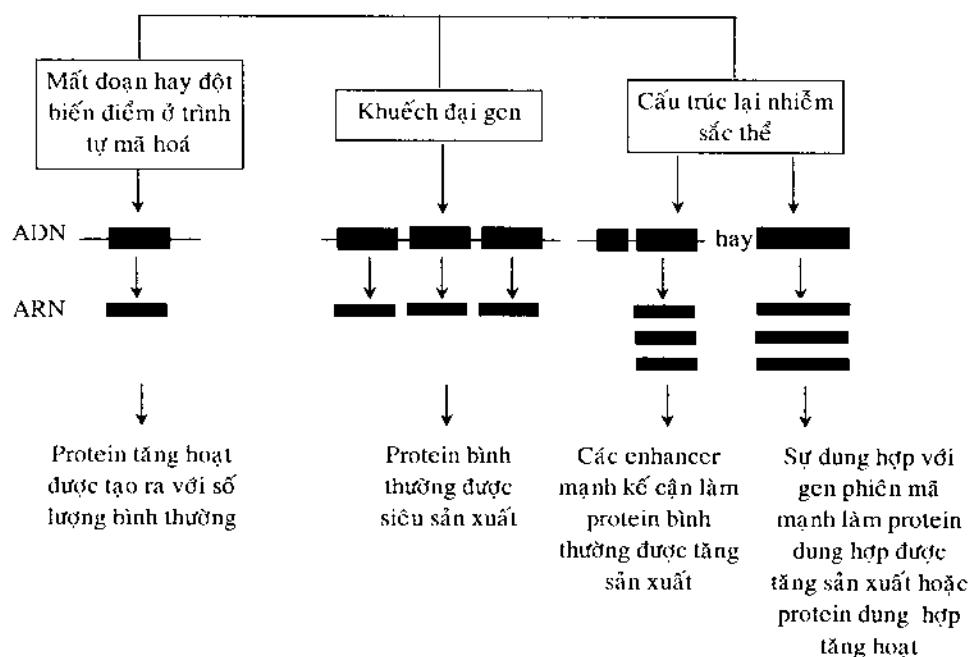
Các gen tiền ung thư hầu như liên quan đến các hệ thống tín hiệu của tế bào như các protein, các thụ thể xuyên màng, các gen điều hoà...

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 nghịch 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 retinoblastoma 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 cò 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 Đột biến	Đột biến có điều kiện		Đột biến khuyết dưỡng		Đột biến đề kháng	
Kiểu gen	Nhiệt độ thường	Nhiệt độ cao	Không bổ sung	Có bổ sung	Không có tác nhân	Có tác nhân
Hoang dai	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 rồi, 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

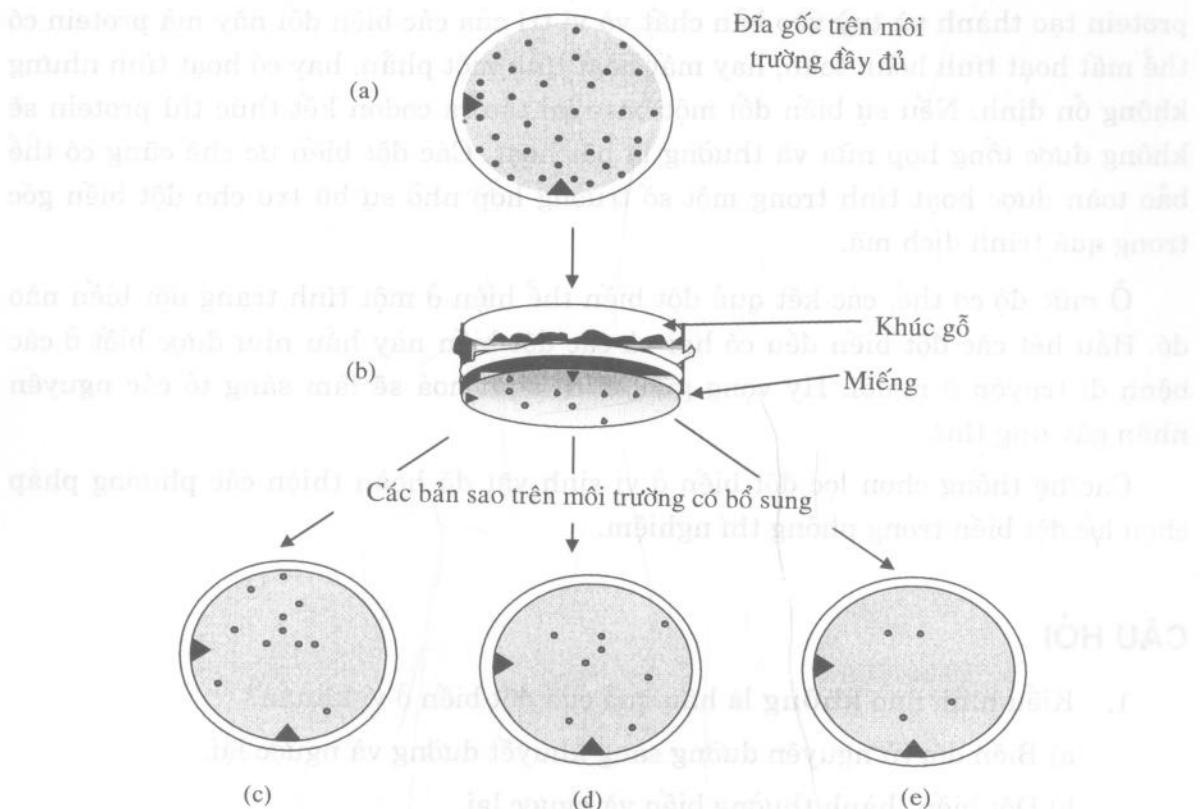
Penicillin có tác động diệt khuẩn dạng phân chia. Vì khuẩn được cho vào môi trường tối thiểu có penicillin. Các vi khuẩn đang tăng trưởng bị diệt, chỉ các tế bào đột biến không tăng trưởng còn sống sót. Sau đó, hỗn hợp được cấy lên môi trường không có penicillin thì các đột biến khuyết dưỡng mọc lên 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 thuỷ 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

Vì khuẩn được cấy để mọc rời 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 tuỳ 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 lai 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 tình 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.
2. Trong bao nhiêu bản mã sao của gen có một đột biến?
 - a) 10^5
 - b) 10^6
 - c) 10^7
 - d) 10^8
 - e) 10^9
3. Nhóm tác nhân gây đột biến nguy hại nhất:

a) Base đồng đẳng	b) Alkyl hoá
c) Deamin hoá	d) Ức chế tổng hợp base nitơ Chèn vào ADN
4. Cơ chế **không** đảo nghịch sai hỏng do đột biến:

a) Quang phục hồi	b) Làm mất nhóm alkyl
c) Nối lại bằng ligase	d) Sửa sai bằng glycosylase
e) b và c	

5. Cơ chế sửa chữa đột biến bằng cách loại bỏ sai hỏng trong and:
- a) Quang phục hồi
 - b) Làm mất nhóm alkyl
 - c) Nối lại bằng ligase
 - d) Sửa sai bằng glycosylase
 - e) Tái tổ hợp
6. Điểm khác biệt giữa kỹ thuật tái tổ hợp và kỹ thuật đột biến:
- a) Tái tổ hợp là kỹ thuật ghép gen
 - b) Tái tổ hợp phụ thuộc vào kỹ thuật di truyền
 - c) Đột biến không phụ thuộc vào kỹ thuật di truyền
 - d) a và b
 - e) a, b và c
7. Tần số đột biến là mức độ xuất hiện của đột biến trên:
- a) Một nhiễm sắc thể
 - b) Một tế bào
 - c) Một giao tử
 - d) Một lần sao chép
 - e) b, c, d
8. Đột biến tự phát **không** do
- a) Hỗn biến của base
 - b) Khử amin AP site
 - c) Khử amin tạo base đồng đẳng
 - d) Đột biến lệch khung khi polymerase sao chép các đoạn lặp lại của một nucleotid
 - 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

CÁC PHƯƠNG PHÁP PHÂN TÍCH ADN

MỤC TIÊU

- Nắm được nguyên tắc các bước tiến hành chính trong chiết tách ADN.
- Nắm được nguyên tắc và ứng dụng của kỹ thuật cắt, nối, lai ADN.
- Trình bày được cơ sở của phương pháp hoá học và phương pháp enzym học dùng xác định trình tự acid nucleic.
- Nêu được nguyên tắc của kỹ thuật PCR, các bước cơ bản để tiến hành và các ứng dụng của nó.

9.1. CHIẾT TÁCH ADN

ADN chứa thông tin di truyền của sinh vật, do đó muốn nghiên cứu ở mức độ sinh học phân tử cần chiết tách ADN. ADN gồm các loại: ADN bộ gen, ADN ty thể; ADN plasmid, ADN chloroplast và ADN virus. Tuỳ kích thước ADN, loại tế bào mà chọn phương pháp chiết khác nhau để đảm bảo chiết được ADN tinh khiết và có hiệu suất cao.

9.1.1. Phương pháp chiết tách ADN

ADN là phân tử có kích thước lớn, khi chiết tách cần lưu ý trong thao tác tránh mọi tác nhân cơ học hay hoá học quá mạnh có thể làm đứt gãy phân tử này.

Phương pháp chiết tách ADN gồm ba bước chính:

- Bước 1: Phá màng tế bào và màng nhân (tế bào nhân thật) bằng các tác nhân vật lý hay hoá học hoặc kết hợp nhiều tác nhân, giải phóng ADN ra môi trường.
- Bước 2: Loại bỏ các thành phần không mong muốn trong mẫu (protein, lipid, polysaccharid...). Thường dùng hỗn hợp phenol – chloroform – iso amyl alcôhl (25 : 24 : 1) làm biến tính protein, đồng thời không hoà tan acid nucleic; sau đó ly tâm: protein biến tính sẽ tách thành một lớp nằm giữa pha nước chứa acid nucleic và pha phenol – chloroform, thu nhận lại pha nước chứa acid nucleic.

– Bước 3: Kết tủa acid nucleic. Thường dùng cồn ethanol thể tích 2,5:1 và trong môi trường có lực ion cao (nồng độ muối cao) hoặc dùng isopropanol thể tích 1:1. Cặn acid nucleic thu được bằng ly tâm, sau đó hoà trong nước khử ion hoặc dung dịch đệm để bảo quản.

9.1.2. Tinh chế

9.1.2.1. Tinh chế acid nucleic bằng siêu ly tâm

– Siêu ly tâm trên gradient liên tục cesium chlorid (CsCl). Kỹ thuật này có thể đạt độ phân giải 1% đơn vị tỉ trọng, cho phép phân tách hai phân tử giống nhau trong đó một phân tử được đánh dấu bằng một đồng vị nặng.

– Siêu ly tâm trên đệm cesium chlorid (gradient không liên tục). Kỹ thuật này thường được áp dụng để tinh chế một lượng lớn acid nucleic.

– Siêu ly tâm trên gradient saccharose. Kỹ thuật này dùng để phân tách hỗn hợp các acid nucleic có kích thước chênh lệch nhau nhiều kb.

9.1.2.2. Tinh chế acid nucleic bằng sắc ký

– Sắc ký ái lực: sử dụng pha tĩnh là U – Sepharose hay oligodT – cellulose để tinh chế mARN.

– Sắc ký lọc gel để tách các nucleotid tự do sau khi tạo mẫu dò đánh dấu.

– Sắc ký trao đổi ion trên vi cột, áp dụng để thu hồi những lượng ADN rất nhỏ.

– Sắc ký lỏng hiệu năng cao: dùng để tinh chế các oligonucletid tổng hợp (độ phân giải là 1 nucleotid), plasmid, phân tách các đoạn ADN.

9.2. CÁC PHƯƠNG PHÁP PHÂN TÍCH ĐỊNH TÍNH VÀ ĐỊNH LƯỢNG SƠ BỘ ACID NUCLEIC

9.2.1. Quang phổ kế

Phương pháp này cho biết kết quả tương đối nồng độ acid nucleic trong mẫu.

Nguyên tắc phương pháp dựa vào sự hấp thụ mạnh ánh sáng tử ngoại ở bước sóng 260 nm của các base purin và pyrimidin. Giá trị mật độ quang ở bước sóng 260 nm ($OD_{260\text{nm}}$, Optical Density ở bước sóng 260 nm) cho phép xác định nồng độ acid nucleic trong mẫu dựa vào tương quan sau:

Một đơn vị $OD_{260\text{nm}}$ tương ứng với một nồng độ là:

50 µg/ml dung dịch ADN (hoặc ARN) sợi đôi

40 µg/ml dung dịch ARN (hoặc ADN) sợi đơn

30 µg/ml oligonucleotid (tối 70 base)

Ví dụ: giá trị OD_{260nm} = 0,9 sẽ tương ứng với:

Dung dịch có nồng độ ADN sợi đôi = $0,9 \times 50 = 45 \mu\text{g/ml}$

Dung dịch có nồng độ ARN sợi đơn = 36 µg/ml

Tuy nhiên cách tính này chỉ đúng với các dung dịch acid nucleic tinh khiết. Để kiểm tra độ tinh khiết của dung dịch, người ta đo thêm giá trị OD ở 280 nm (OD_{280nm}). 280 nm là bước sóng ở đó các protein có mức hấp thụ cao nhất, nhưng các protein cũng hấp thụ ánh sáng ở bước sóng 260 nm như các acid nucleic nên làm sai lệch giá trị thật của nồng độ acid nucleic. Một dung dịch acid nucleic được xem là tinh khiết (không tạp nhiễm protein) khi tỉ số OD_{260nm}/OD_{280nm} nằm trong khoảng 1,8 – 2.

9.2.2. Điện di phân tích ADN

Phương pháp này được áp dụng cả trong phân tích định tính lẫn trong việc thu nhận mẫu acid nucleic.

Nguyên tắc phương pháp dựa vào đặc tính cấu trúc của các acid nucleic. Đó là các đại phân tử tích điện âm đồng đều trên khắp bề mặt nên khi chịu tác động của một điện trường, chúng sẽ di chuyển về hướng cực dương. Sự di chuyển nhanh hay chậm tuỳ thuộc vào hình dạng và tỷ số điện tích: khối lượng, nhờ đó các chất khác nhau sẽ được tách ra thành từng vết riêng và được phát hiện bằng nhiều cách khác nhau. Tuy nhiên, hầu hết các phân tử ADN đều có hình dạng và tỷ số điện tích – khối lượng như nhau do đó không tách được bằng điện di thông thường, mà phải dùng kỹ thuật điện di trên gel. Gel được làm bằng agarose hay polyacrylamid có các lỗ do các dây polyme agarose hay polyacrylamid đan với nhau thành mạng lưới tạo ra, phân tử ADN khi đó phải chui qua các lỗ này để đến điện cực, do đó phân tử ADN nhỏ sẽ di chuyển nhanh hơn phân tử lớn, nhờ đó ta tách được ADN theo kích thước phân tử (cơ chế rây).

Tốc độ điện di của ADN qua gel tuỳ thuộc vào bốn thông số chính: (1) Kích thước phân tử ADN – là thông số quyết định nhất. Phân tử ADN sợi đôi duỗi thẳng di chuyển với tốc độ tỉ lệ nghịch với logarit của trọng lượng phân tử của chúng; (2) Cấu dạng của ADN, 3 dạng tồn tại của một ADN là dạng siêu xoắn, dạng vòng và dạng thẳng tuy có cùng trọng lượng phân tử và thành phần, nhưng di chuyển qua gel với tốc độ khác nhau; (3) Nồng độ gel xác định tốc độ di chuyển và hiệu quả tách các phân tử ADN, sự di chuyển của phân tử ADN tỉ lệ nghịch với

nồng độ của gel; và (4) Điện thế sử dụng – ở điện thế thấp, tốc độ di chuyển tỉ lệ với điện thế sử dụng.

Điện di trong trường xung được áp dụng để phân tách các phân tử ADN rất lớn (>50 kb). Nguyên tắc phương pháp dựa vào sự đổi hướng điện trường trong quá trình điện di. Mỗi lần hướng điện trường thay đổi thì phân tử ADN phải tự định hướng lại. Thời gian cần cho việc tự định hướng lại tuỳ thuộc kích thước của phân tử, phân tử càng dài thì thời gian định hướng lại càng lớn khiến cho nó di chuyển chậm hơn phân tử kích thước nhỏ.

Đối với ADN bộ gen nhiễm sắc thể của động vật có vú có kích thước lớn (khoảng 10^9 kDa), không thể phân giải bằng điện di agarose thường quy, ngay cả bằng điện di trong trường xung, phải thuỷ phân ADN với enzym giới hạn trước khi phân tích.

9.3. KỸ THUẬT CẮT, NỐI, LAI ADN VÀ ỨNG DỤNG

9.3.1. Cắt ADN bằng enzym cắt giới hạn

9.3.1.1. Khái niệm về enzym cắt giới hạn

Tế bào vi khuẩn có một hệ thống bảo vệ để chống lại sự tấn công của phage là các enzym cắt giới hạn (RE). Hệ thống này gồm hai loại enzym: (1) Các RE cắt ADN phage ở những vị trí chuyên biệt và (2) Các enzym methylase gắn nhóm methyl vào A hay C trên ADN tế bào vi khuẩn ở vị trí cắt của các RE nên ADN của vi khuẩn không bị chính các RE của chúng cắt.

Dựa vào khả năng nhận biết và cắt một trình tự xác định trên phân tử ADN, người ta chia RE thành ba loại: (1) Loại I – enzym cắt phân tử ADN tại điểm cách trình tự nhận diện khoảng 1000 – 5000 nucleotid và giải phóng vài chục nucleotid; (2) Loại II – enzym cắt ngay vị trí nhận biết trình tự; (3) Loại III – enzym cắt phân tử ADN tại điểm cách trình tự nhận diện khoảng 20 nucleotid. Trong kỹ thuật di truyền, các RE loại II được sử dụng rất nhiều để thao tác trên ADN nên sau đây chỉ trình bày các đặc tính của loại này.

9.3.1.2. Đặc điểm của RE loại II

Trình tự đặc trưng để RE nhận biết thường bao gồm 4 – 8 nucleotid. Mỗi enzym chỉ nhìn nhận một trật tự chuyên biệt có cấu trúc palindrome, nghĩa là hai mạch của trình tự hoàn toàn giống nhau khi chúng được đọc theo chiều 5' → 3', ví dụ trình tự GAATTC của enzym EcoRI.

Các enzym cắt ADN gồm hai loại: exonuclease và endonuclease. Exonuclease cắt ADN từ hai đầu mút, còn endonuclease cắt ADN ở giữa phân tử. Các RE cắt ADN ở giữa một cách đặc hiệu, được gọi là endonuclease giới hạn.

Từ năm 1962, khi Werner Arber lần đầu tiên phát hiện ra RE trong tế bào vi khuẩn, tới nay người ta đã phát hiện hơn 500 loại RE và hơn 100 loại đang được thương mại hóa.

Các RE loại II có hai kiểu cắt: (1) cắt tạo đầu tù do RE cắt hai mạch ADN tại cùng một điểm và (2) cắt tạo đầu so le hay đầu dính. Các enzym cắt tạo đầu dính được sử dụng rất nhiều trong tái tổ hợp di truyền *in vitro*, hai phân tử ADN có nguồn gốc khác nhau nhưng cùng được cắt bởi một RE sẽ có khả năng kết hợp thành một thông qua các đầu dính.

9.3.1.3. Ứng dụng

Việc sử dụng các RE có ý nghĩa quan trọng trong nghiên cứu di truyền phân tử. Chúng cho phép cắt nhỏ bộ gen khổng lồ của các sinh vật tế bào nhân thật, lập bản đồ giới hạn, so sánh bộ gen ở các loài khác nhau thông qua kỹ thuật RFLP.

9.3.2. Các ligase và ứng dụng để nối các đoạn acid nucleic

Các enzym ligase xúc tác sự hình thành liên kết nối hai đoạn ADN (ADN ligase) hay ARN (ARN ligase).

Các enzym này có thể được ly trich từ *E. coli* (*E. coli* ADN ligase) hoặc từ phage T4 xâm nhiễm *E. coli* (T4 ADN ligase và T4 ARN ligase).

E. coli ADN ligase có khả năng xúc tác phản ứng nối hai trình tự ADN có đầu dính; T4 ADN ligase cùng chức năng với *E. coli* ADN ligase nhưng đặc biệt còn có khả năng nối hai trình tự ADN có đầu bằng; và T4 ARN ligase xúc tác nối hai trình tự ARN bằng liên kết phosphodiester thường dùng để đánh dấu phóng xạ đầu 3' của các phân tử ARN, sản xuất đoạn dò.

9.3.3. Kỹ thuật lai nucleotid và ứng dụng

9.3.3.1. Khái niệm về lai phân tử

Khi nhiệt độ được nâng lên tới nhiệt độ "biến tính" (Tm), hai mạch phân tử ADN tách rời nhau, sự bắt cặp trở lại sẽ không xảy ra nếu nhiệt độ phản ứng hạ xuống đột ngột và phân tử ADN sẽ tồn tại trong môi trường ở dạng mạch đơn dưới một cấu hình không gian vô trật tự. Ngược lại, nếu sau khi hai mạch tách rời, nhiệt độ được làm giảm từ từ cộng với điều kiện thích hợp, hai mạch sẽ bắt cặp trở lại. Hiện tượng này được gọi là sự lai phân tử, nó có tính đặc hiệu tuyệt đối.

Các yếu tố ảnh hưởng đến sự lai phân tử:

bài tập lần 8.3.5.8.

- Nồng độ ADN và thời gian phản ứng.
- Nhiệt độ.
- Độ dài của các trình tự.
- Lực ion.

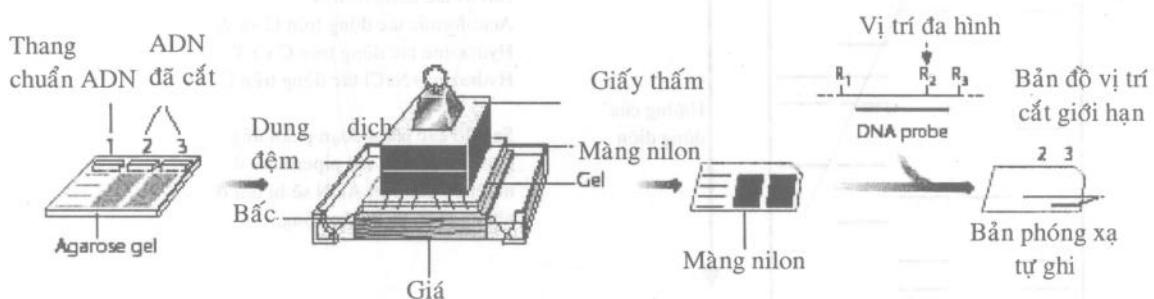
Các phương pháp lai phân tử rất đa dạng, trong bài này chỉ đề cập kiểu lai trên pha rắn và lai tại chỗ.

9.3.3.2. Lai trên pha rắn và ứng dụng

Nguyên tắc này là một trong hai trình tự bổ sung (các mạch đơn) nằm trong môi trường lỏng là một dung dịch đậm và trình tự kia (thường là trình tự đích, tức trình tự cần tìm) được cố định trên một giá thể rắn.

Các phương pháp lai trên pha rắn được ứng dụng nhiều nhất là Southern blot, Northern blot và dot blot.

Phương pháp Southern blot do Southern mô tả năm 1975, được sử dụng để định vị những trình tự đặc biệt trên ADN bộ gen, hay những ADN kích thước nhỏ như ADN plasmid, phage,... Phương pháp này đã được áp dụng để lập bản đồ giới hạn của một gen; phát hiện các đa dạng trình tự của cùng một gen ở các chủng hay các cá thể khác nhau qua sự so sánh bản đồ giới hạn của chúng; phát hiện các đột biến mất đoạn, đột biến điểm hay tái tổ hợp trên một gen vì chúng làm thay đổi bản đồ giới hạn.



Hình 9.1. Sơ đồ phương pháp Southern blot

Phương pháp Northern blot được sử dụng để xác định kích thước và hàm lượng của một mARN đặc trưng trong một hỗn hợp ARN.

Phương pháp dot blot cho phép xác định tương đối một ARN đặc trưng trong một hỗn hợp ARN mà không cần phải phân tách chúng ra. Phương pháp này có thể sử dụng cho ADN.

9.3.3.3. Lai tại chỗ

Nguyên tắc lai tại chỗ là một kiểu lai phân tử trong đó trình tự acid nucleic cần tìm (trình tự đích) nằm ngay trong tế bào hay trong mô. Lai tại chỗ cho phép nghiên cứu acid nucleic mà không cần phải tách chiết chúng ra khỏi mô, tế bào.

Các ứng dụng của kiểu lai này rất đa dạng đi từ kỹ thuật định vị gen trên nhiễm sắc thể, phát hiện dòng vi khuẩn tái tổ hợp trong phương pháp tạo dòng đến việc nghiên cứu một mARN chuyên biệt trong tế bào và mô.

9.3.3.4. Các loại mẫu dò

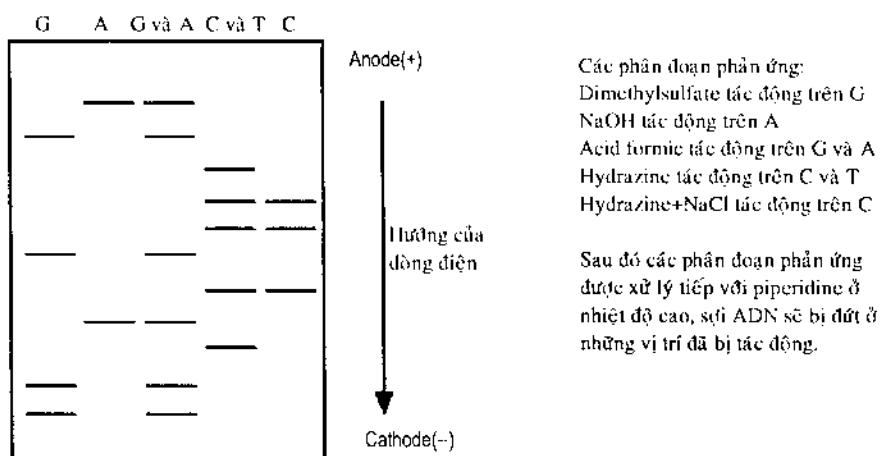
Mẫu dò là bản sao của một trình tự acid nucleic và được đánh dấu để dễ nhận biết. Mẫu dò có tính (1) chuyên biệt do chỉ lai với đúng trình tự bắt cặp bổ sung và (2) nhạy do được đánh dấu để phát hiện và định lượng. Tác nhân dùng để đánh dấu mẫu dò là các đồng vị phóng xạ: ^{32}P , ^{35}S , ^3H , hoặc hoá chất: hệ thống avidin (hay streptavidin) – biotin, hệ thống peroxydase – cơ chất.

9.4. CÁC PHƯƠNG PHÁP XÁC ĐỊNH TRÌNH TỰ ADN

9.4.1. Phương pháp Maxam và Gilbert – phương pháp hoá học

Phương pháp gồm các bước chính như sau:

- Đánh dấu các phân tử ADN bằng ^{32}P ở một đầu.



Hình 9.2. Sơ đồ phương pháp Maxam và Gilbert

- Chia hỗn hợp phản ứng thành năm phân đoạn, mỗi phân đoạn chịu một xử lý hoá học chuyên biệt làm biến đổi đặc trưng một loại nucleotid và sau đó cắt phân tử ADN ngay nucleotid đó. Năm phân đoạn tương ứng năm nhóm nucleotid bị tác động là: G, A, C, G + A, T + C. Sau khi xử lý, mỗi phân đoạn là một tập hợp

oligonucleotid có kích thước khác nhau nhưng đều chấm dứt tại cùng một loại nucleotid.

– Phân tách trên gel polyacrylamid, vị trí của các oligonucleotid trên gel tương ứng với kích thước của chúng và được phát hiện nhờ đầu có đánh dấu phóng xạ.

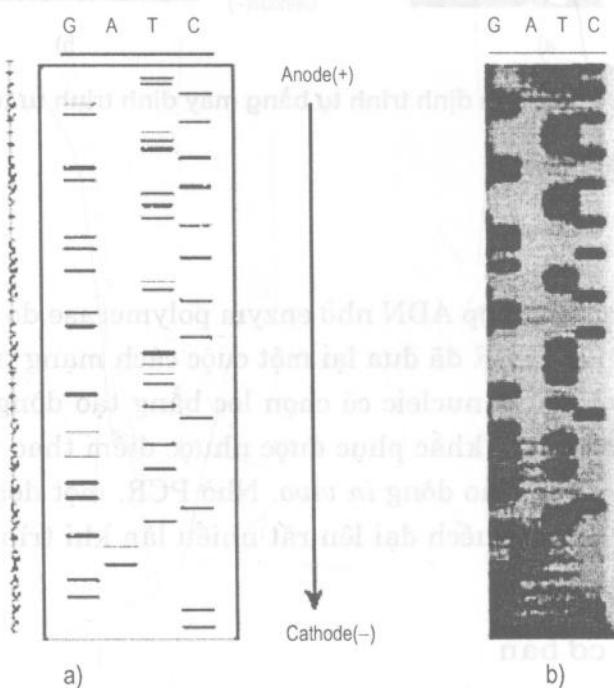
9.4.2. Phương pháp Sanger và cộng sự – phương pháp enzym

– Tạo dòng trình tự ADN cần xác định trong một vectơ mạch đơn (phage 13).

– Tổng hợp ADN trên khuôn ADN mạch đơn nhờ enzym ADN polymerase, với sự hiện diện của bốn loại nucleotid, trong đó có một loại được đánh dấu đồng vị phóng xạ (^{35}S).

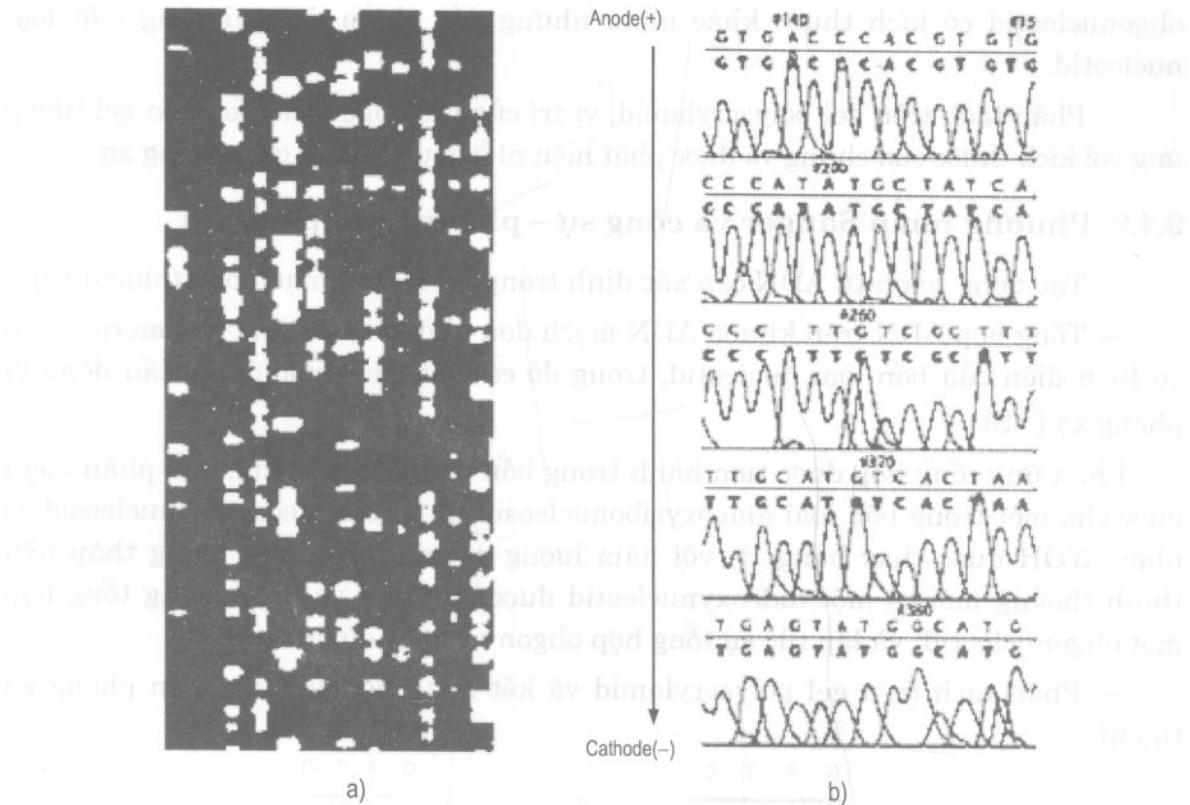
Phản ứng tổng hợp được tiến hành trong bốn phân đoạn riêng, mỗi phân đoạn được cho một trong bốn loại dideoxyribonucleosid (là những deoxyribonucleosid có nhóm 3'OH được thay bằng H) với hàm lượng rất nhỏ. Do hàm lượng thấp nên thỉnh thoảng mới có một dideoxynucleotid được sử dụng vào phản ứng tổng hợp một oligonucleotid, và lập tức sự tổng hợp oligonucleotid đó dừng lại.

– Phân tách trên gel polyacrylamid và kết quả được đọc trên bản phóng xạ tự ghi.



Hình 9.3. Kết quả định trình tự theo phương pháp Sanger

Có một số phương pháp khác được cải biến từ phương pháp Sanger, đơn giản hơn và rút gọn hơn, đó là phương pháp xác định trình tự bằng máy tự động và phương pháp PCR dùng trong xác định trình tự acid nucleic.



Hình 9.4. Kết quả định trình tự bằng máy định trình tự tự động

9.5. PCR

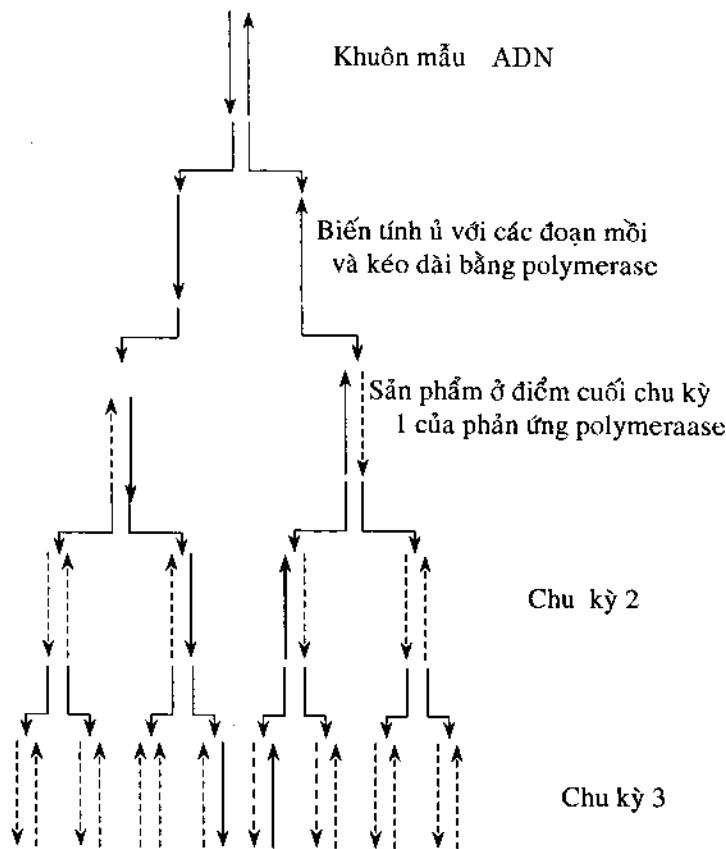
9.5.1. Mở đầu

PCR là phản ứng tổng hợp ADN nhờ enzym polymerase do Karl Mullis và cộng sự phát minh năm 1985. PCR đã đưa lại một cuộc cách mạng trong di truyền phân tử: khuếch đại *in vitro* acid nucleic có chọn lọc bằng tạo dòng *in vitro* mà không cần sự hiện diện của tế bào, khắc phục được nhược điểm thao tác phức tạp và cần thời gian dài của kỹ thuật tạo dòng *in vivo*. Nhờ PCR, một đoạn ADN ở một vùng bất kỳ trong bộ gen được khuếch đại lên rất nhiều lần khi trình tự nucleotid ở hai đầu đoạn đó đã biết.

9.5.2. Nguyên tắc cơ bản

PCR sử dụng các đặc điểm của quá trình sao chép ADN. ADN polymerase dùng các đoạn ADN mạch đơn làm khuôn để tổng hợp nên sợi mới bổ sung với sợi này. Các sợi ADN khuôn mạch đơn có thể tạo ra theo cách đơn giản là đun nóng dung dịch ADN mạch kép đến gần nhiệt độ sôi.

Tất cả các ADN polymerase khi hoạt động tổng hợp một mạch ADN mới từ mạch khuôn mẫu đều cần sự hiện diện của những mồi chuyên biệt. Mỗi là những đoạn ADN ngắn, có khả năng bắt cặp bổ sung với một đầu của mạch khuôn, và ADN polymerase sẽ nối dài mồi để hình thành mạch mới. Phản ứng chuỗi đã được hình thành dựa vào đặc tính đó của các ADN polymerase. Thật vậy, nếu ta cung cấp hai mồi chuyên biệt bắt cặp bổ sung với hai đầu của một trình tự ADN, ta sẽ chỉ tổng hợp đoạn ADN nằm giữa hai mồi. Điều đó có nghĩa là để khuếch đại một trình tự ADN xác định, ta phải có thông tin tối thiểu về trình tự đó đủ để tạo các mồi bổ sung chuyên biệt; các mồi này gồm một mồi xuôi và một mồi ngược (so với chiều phiên mã của gen).



Hình 9.5. Phản ứng chuỗi polymerase

Phản ứng chuỗi là một sự khuếch đại *in vitro* một trình tự ADN khuôn mẫu với xúc tác enzym. Phản ứng này là sự kéo dài đồng thời đoạn mồi của các mạch bổ sung ADN. Bằng cách lặp lại các chu kỳ gồm việc đưa nhiệt độ lên 95°C (khoảng 35 chu kỳ), hạ nhiệt độ xuống (37°C – 65°C) và kéo dài ở 72°C bằng cách sử dụng một loại ADN polymerase chịu nhiệt. Phản ứng được thực hiện trong một ống nhựa đặt trong một máy có thể điều chỉnh nhiệt độ gọi là máy chu kỳ nhiệt (máy

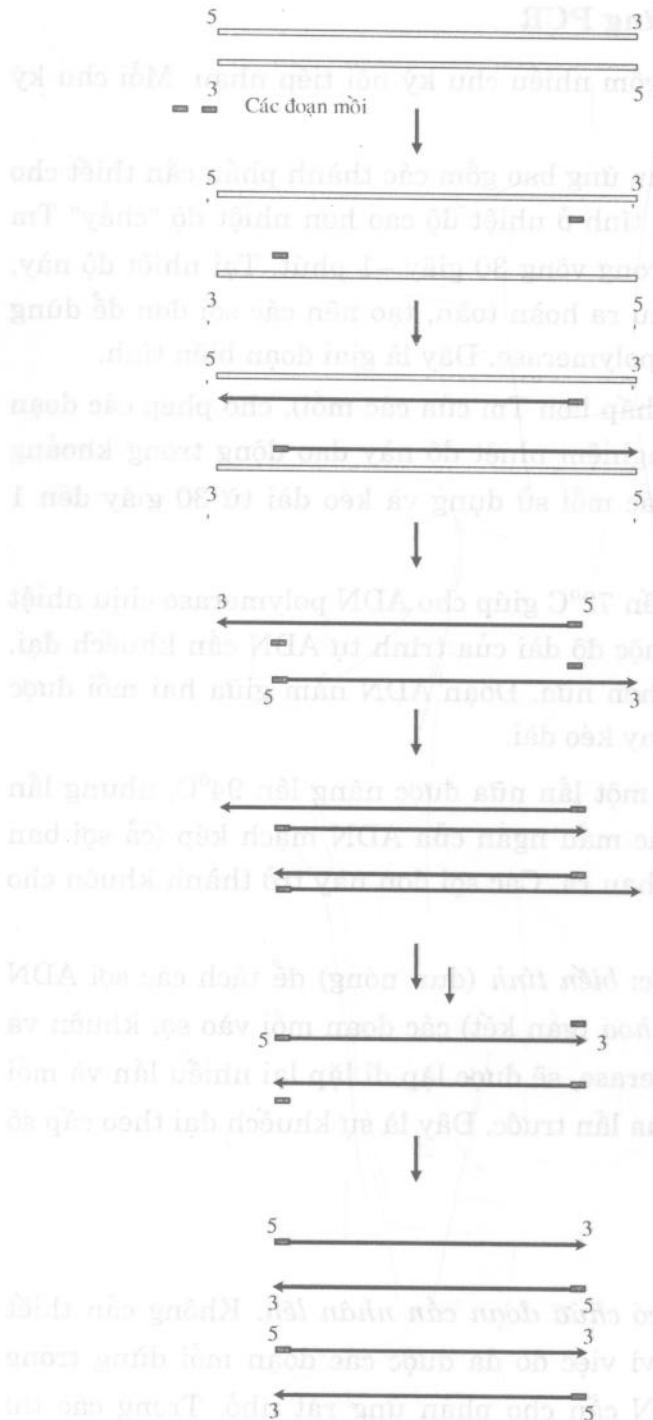
PCR), máy này có thể cài chương trình làm nóng và làm nguội theo chu kỳ cần thiết. Quá trình phản ứng sẽ làm gia tăng số bản sao một trình tự khuôn mẫu ban đầu (hình 9.5).

Cả hai sợi ADN đều được dùng để làm khuôn cho quá trình tổng hợp nếu các đoạn mồi được cung cấp cho cả hai sợi. Trong phản ứng PCR, các đoạn mồi được chọn để chặn hai đầu của đoạn ADN cần nhân lên, sao cho các sợi ADN tổng hợp mới được bắt đầu từ mỗi đoạn mồi và kéo dài về phía đoạn mồi nằm trên sợi kia (hình 9.6c). Như vậy, sau mỗi chu kỳ các điểm bám cho các đoạn mồi lại xuất hiện trên mỗi sợi ADN mới được tổng hợp.

Hỗn hợp phản ứng lại được đun nóng lên để tách sợi ban đầu khỏi sợi mới tổng hợp, các sợi này sau đó lại được dùng tiếp cho chu trình tiếp theo, bao gồm các bước gắn đoạn mồi, tổng hợp ADN và tách rời các đoạn. Kết quả cuối cùng của phản ứng PCR là sau n chu kỳ phản ứng, tính theo lý thuyết ta sẽ có 2^{n-2} bản sao các phân tử ADN mạch kép nằm giữa hai đoạn mồi (bảng 9.1).

Bảng 9.1. Hiệu quả nhân ADN của PCR

Số chu kỳ (n)	Số phân tử ADN mạch kép sinh ra	Số chu kỳ (n)	Số phân tử ADN mạch kép sinh ra
1	0	17	32768
2	0	18	65536
3	2	19	131072
4	4	20	262144
5	8	21	524288
6	16	22	1048567
7	32	23	2097152
8	64	24	4194304
9	128	25	8388608
10	256	26	16777216
11	512	27	33544432
12	1,024	28	67108864
13	2,048	29	134217728
14	4,096	30	268435456
15	8,192	31	536870912
16	19,384	32	1073741824



- a) Vật liệu khởi đầu là phân tử ADN mạch kép.
- b) Các sợi tách nhau ra khi gia nhiệt 94°C , sau đó hạ xuống $30\text{--}65^{\circ}\text{C}$, trong 30 giây để các đoạn mồi gắn vào hai vị trí tương hợp trên ADN.
- c) *Taq* polymerase tổng hợp các sợi ADN mới tương hợp với sợi khuôn, bắt đầu từ đoạn mồi kéo dài về phía đoạn mồi nằm bên sợi kia.
- d) Hỗn hợp phản ứng lại được gia nhiệt. Sợi ban đầu và sợi mới tổng hợp tách nhau ra. Trên hai sợi mới tổng hợp xuất hiện 2 điểm bám cho các đoạn mồi gắn vào.
- e) *Taq* polymerase tổng hợp các sợi ADN tương hợp mới, nhưng các chuỗi mới lần này chỉ tiến đến đúng vị trí đã được xác định bởi các cặp mồi mà các đoạn mồi đã lựa chọn.
- f) Quá trình được lặp lại, các đoạn mồi tiếp tục gắn vào các sợi mới tổng hợp.
- g) *Taq* polymerase tổng hợp các đoạn tương hợp sinh ra hai đoạn ADN mạch kép giống hệt đoạn cần tổng hợp. Quá trình cứ thế tiếp tục.

Hình 9.6. Sơ đồ nguyên lý phản ứng chuỗi polymerase

Để đơn giản hóa sơ đồ, từ phần (d) không vẽ hai sợi ADN khuôn ban đầu

9.5.3. Cách tiến hành một phản ứng PCR

PCR là một chuỗi các phản ứng gồm nhiều chu kỳ nối tiếp nhau. Mỗi chu kỳ gồm ba bước:

Bước 1: Trong một dung dịch phản ứng bao gồm các thành phần cần thiết cho sự sao chép, phân tử ADN được biến tính ở nhiệt độ cao hơn nhiệt độ "cháy" Tm của phân tử, thường là 94°C – 95°C trong vòng 30 giây – 1 phút. Tại nhiệt độ này, các phân tử ADN mạch kép tách nhau ra hoàn toàn, tạo nên các sợi đơn để dùng làm khuôn cho các đoạn mồi và ADN polymerase. Đây là giai đoạn biến tính.

Bước 2: Nhiệt độ được hạ thấp (thấp hơn Tm của các mồi), cho phép các đoạn mồi bắt cặp với khuôn. Trong thực nghiệm nhiệt độ này dao động trong khoảng 40°C – 70°C, tùy thuộc vào Tm của các mồi sử dụng và kéo dài từ 30 giây đến 1 phút. Đây là giai đoạn lai.

Bước 3: Nhiệt độ được tăng lên đến 72°C giúp cho ADN polymerase chịu nhiệt hoạt động tốt nhất. Thời gian tuỳ thuộc độ dài của trình tự ADN cần khuếch đại, thường kéo dài khoảng 30 giây hay hơn nữa. Đoạn ADN nằm giữa hai mồi được tổng hợp. Đây là giai đoạn tổng hợp hay kéo dài.

Vào cuối giai đoạn này, nhiệt độ một lần nữa được nâng lên 94°C, nhưng lần này chỉ trong vòng 30 giây để cho các mẫu ngắn của ADN mạch kép (cả sợi ban đầu và sợi mới được tổng hợp) tách nhau ra. Các sợi đơn này trở thành khuôn cho một chu kỳ tổng hợp ADN khác.

Tóm lại, một chu kỳ gồm ba bước: *biến tính* (đun nóng) để tách các sợi ADN kép thành sợi đơn để làm khuôn, *lai hoá* (gắn kết) các đoạn mồi vào sợi khuôn và *kéo dài* (tổng hợp) bằng ADN polymerase, sẽ được lặp đi lặp lại nhiều lần và mỗi lần lại làm tăng gấp đôi lượng mẫu của lần trước. Đây là sự khuếch đại theo cấp số nhân (hình 9.6).

9.5.4. Các thành phần phản ứng

Vật liệu gồm trước hết là *ADN có chứa đoạn cần nhân lên*. Không cần thiết phải tách riêng đoạn cần nhân lên vì việc đó đã được các đoạn mồi dùng trong phản ứng xác định. Hàm lượng ADN cần cho phản ứng rất nhỏ. Trong các thí nghiệm bình thường ở phòng thí nghiệm chỉ cần một microgram ADN bộ gen là đủ. Trong nhiều trường hợp, phản ứng PCR có thể nhân các đoạn ADN chỉ từ một phân tử riêng lẻ. *Hai đoạn mồi* để xác định điểm bắt đầu tổng hợp ADN, *ADN polymerase*, và hỗn hợp của 4 *deoxyribonucleotid* (ở dạng dNTP) và dung dịch đậm

có ion Mg^{2+} cần được đưa vào ống nghiệm có chứa ADN khuôn. Dung tích tổng số thường là 50 μ l hoặc 20 μ l. Hỗn hợp phản ứng được cho vào ống nhựa nhỏ chịu nhiệt, có nắp đậy.

Một phản ứng PCR có thể gồm các thành phần hay các vật liệu ban đầu: mache khuôn cho sự khuếch đại, các đoạn mồi oligonucleotide, dNTP, và ADN polymerase chịu nhiệt với dung dịch đậm phản ứng thích hợp và $MgCl_2$. Nồng độ của mỗi thành phần này tùy yêu cầu. Trừ ADN khuôn và các đoạn mồi, các thành phần cho PCR đều có sẵn từ các nhà cung cấp.

9.5.4.1. Các polymerase chịu nhiệt

Enzym ADN polymerase được sử dụng đầu tiên trong PCR là *đoạn Klenow* của ADN polymerase I. Đặc tính của đoạn này là xúc tác sự tổng hợp ADN theo chiều 5' – 3' và có hoạt tính exonuclease (thuỷ giải liên kết giữa các nucleotide từ đầu phân tử ADN) theo chiều 3' – 5'. Tuy nhiên enzym này không chịu được nhiệt nên thao tác phức tạp và hiệu quả thấp (phải thêm enzym mới vào phản ứng sau mỗi lần biến tính vì enzym cũ đã bị nhiệt phân huỷ, nhiệt độ lai thấp khiến sự khuếch đại không đặc hiệu vv...).

Hiện nay thường dùng ADN polymerase là *Taq polymerase* chịu nhiệt, được tách chiết từ một loài vi khuẩn suối nước nóng có tên *Thermus aquaticus*. Enzym này không bị phá huỷ ở nhiệt độ biến tính và xúc tác sự tổng hợp từ đầu đến cuối quá trình phản ứng.

Taq polymerase có sẵn trên thị trường từ các công ty công nghệ sinh học. Các polymerase chịu nhiệt khác có trên thị trường bao gồm Vent polymerase được phân lập từ *Thermococcus litoralis*, xúc tác tổng hợp nhiều bản sao chính xác từ mache khuôn hơn là *Taq* polymerase. Ngoài ra còn có *Tth* polymerase được phân lập từ *Thermus thermophilus*. Các polymerase chịu nhiệt là thành phần đắt nhất của PCR và thường được sử dụng 1 đơn vị cho một phản ứng. Hàm lượng này đủ để xúc tác cho sự khuếch đại khoảng 40 chu kỳ bằng máy PCR.

Trước đây, việc lựa chọn polymerase chịu nhiệt cho 1 thí nghiệm thì không mấy khó khăn vì chỉ có một loại ADN polymerase chịu nhiệt (*Taq*). Hiện nay có nhiều loại polymerase có nguồn gốc khác nhau, nhiều bản sao của các polymerase này được biến đổi di truyền và chúng có thể ưu việt hơn polymerase gốc trong phạm vi một vài ứng dụng (bảng 9.2). Sự chọn lựa một polymerase chịu nhiệt cho một thí nghiệm PCR phụ thuộc vào yêu cầu của từng thử nghiệm.

Bảng 9.2. Các loại ADN polymerase thường gặp

Enzym	Nguồn gốc	Nơi ở	Mức ổn định ở 95°C ($t_{1/2}$)	Hoạt động tối ưu	Cation cần thiết	3'→5' Exonuclease
Taq	<i>Thermus aquaticus</i>	Suối nước nóng	40 phút	50 – 75	MgCl ₂	–
AmpliTaq	Tag tái tổ hợp		40 phút	50 – 75	MgCl ₂	–
AmpliTaq Stoffel	AmpliTaq cắt bỏ 289 aa N – tân		80 phút	75 – 80	MgCl ₂	–
Tth	<i>Thermus thermophilus</i>	Suối nước nóng	20 phút	55 – 75	MnCl ₂ (RT), MgCl ₂ (PCR)	–
Ultima	<i>Thermotoga maritima</i>	Dòng chảy nóng ở biển sâu				+
Vent	<i>Thermococcus litoralis</i>	Dòng chảy nóng ở biển sâu	>95% hoạt động sau 1 h	80	MgSO ₄	+
Vent (Exo ⁻)	Vent tái tổ hợp				MgSO ₄	–
Deep Vent	<i>Pyrococcus sp.</i> Strain GB – D	Dòng chảy nóng ở biển sâu	>95% hoạt động sau 1 h	72 – 80	MgSO ₄	+
Pfu	<i>Pyrococcus furiosus</i>	Dòng chảy nóng ở biển sâu	>95% hoạt động sau 1 h	75	MgCl ₂	+

9.5.5.2. ADN khuôn mẫu

Phản ứng chuỗi cực nhạy, có thể chỉ cần dùng một phân tử ADN làm khuôn cũng thu được sản phẩm. Tuy nhiên, mức độ nhạy cảm cao là nguyên nhân chính gây ra sự dương tính giả do các ADN ngoại nhiễm.

Phản ứng khuếch đại tối ưu xảy ra trên ADN thật *tinh khiết* nhưng nhiều kỹ thuật chẩn đoán bằng PCR vẫn đạt kết quả tốt với ADN thu nhận trực tiếp từ dịch chiết mẫu thử (vết máu, vật liệu lưu trữ, ADN cũ và ngay cả vi khuẩn đã tiệt trùng). Tuy nhiên, trong nhiều trường hợp khác, các quá trình chuẩn bị mẫu phải được thực hiện để đảm bảo cho việc khuếch đại, vì các chất ức chế như heparin, EDTA, hay các acid có thể hiện diện trong mẫu.

Mạch ADN khuôn mẫu cũng có thể ảnh hưởng đến phản ứng chuỗi. Nếu ADN này có chứa nhiều GC thì sẽ khó khăn cho việc tách rời các chuỗi ADN ở nhiệt độ

cao. Trong các trường hợp này, việc thêm formamide hay dimethyl sulfoxide (DMSO) vào hỗn hợp phản ứng có thể giải quyết được vấn đề.

Nồng độ của ADN khuôn mẫu trong PCR sẽ ảnh hưởng đến mức độ khuếch đại. Với việc sử dụng các polymerase cho hiệu quả cao, lượng ADN mẫu sử dụng cũng có khuynh hướng giảm (1 microgram còn 100 nanogram). Mặc dù, các khuôn mẫu đơn được khuếch đại nhưng không phải mọi thí nghiệm đều đạt được sự nhạy cảm này. Do đó, nên thiết lập PCR với lượng ADN khuôn mẫu thích hợp. Lượng khuôn mẫu ADN dư thừa có thể ảnh hưởng xấu đến quá trình khuếch đại. Nếu lượng ADN khuôn mẫu quá cao thì sẽ có khuynh hướng tạo ra lượng lớn các sản phẩm được kéo dài từ đoạn mồi hơn là những sản phẩm nằm giữa các đoạn mồi. Việc lựa chọn mồi, quá trình ủ lặp lại không thích hợp và sự có mặt của các chất ức chế cũng có thể làm giảm sự khuếch đại.

ADN khuôn có thể là bộ gen tách ra từ tế bào, ADN từ các ngân hàng genome hoặc cADN (chỉ cần nồng độ dưới ng), các ADN bất kỳ, ARN toàn bộ hoặc ARNm.

9.5.4.3. Các đoạn mồi

Các đoạn mồi là thành phần quyết định sự khuếch đại đặc trưng và có hiệu quả cao của PCR. Việc chọn đoạn mồi phải tuân thủ một số nguyên tắc:

- Trình tự của mồi được chọn sao cho không thể có sự bắt cặp bổ sung giữa các mồi tạo cấu trúc bậc hai hay tạo dimer giữa các nucleotid ngay trong cùng đoạn mồi làm cho phản ứng giảm hiệu quả và ngăn cản sự tạo thành đoạn mong muốn.
- Các mồi chọn phải đặc trưng cho trình tự ADN cần khuếch đại, không trùng với các trình tự lặp lại trên gen.
- Kích thước một đoạn mồi tối ưu từ 18 – 25 nucleotid. Đoạn mồi với kích thước này đảm bảo cho việc gắn đặc hiệu ở nhiệt độ tối ưu và giá thành cho một đoạn mồi được giảm thiểu.
- Khoảng cách giữa hai mồi không dài quá 3 kb. Phản ứng PCR sẽ tối ưu trên những trình tự nhỏ hơn 1 kb. Tuy nhiên hiện nay có một số enzym polymerase chịu nhiệt có thể tổng hợp được đoạn ADN có kích thước khoảng 30kb.
- Tỷ lệ GC ở các mồi thường chiếm khoảng 40 – 75%. Trong thành phần các đoạn mồi tránh có quá nhiều cặp GC hay là một loạt các nucleotid G và C. Vì liên kết G: C mạnh hơn liên kết A: T (G nối với C bằng 3 liên kết hydro còn A với T chỉ có 2). Do vậy, việc bẻ gãy liên kết giữa G: C cần nhiều năng lượng hơn liên kết A: T. Người ta có thể tính được nhiệt độ tối ưu cho việc ủ các đoạn mồi trong PCR,

nhiệt độ này gọi là T_m –5°C là nhiệt độ "chảy" của đoạn mồi trừ đi 5. Kết quả này có thể tính đơn giản:

$$T_m = 2(A + T) + 4(G + C)$$

$$\text{Nhiệt độ } \bar{u} = T_m - 5^{\circ}\text{C}$$

Các đoạn mồi được tổng hợp hóa học trên một chất tải rắn trong máy tổng hợp oligonucleotide. Các máy này được tự động hóa cao, dễ dàng tổng hợp ra các đoạn mồi. Có nhiều quy tắc để thiết kế đoạn mồi và đã có một số phần mềm cho sự chọn lựa một cặp đoạn mồi tối ưu cho sự khuếch đại của một ADN khuôn mẫu có trình tự đã biết. Nồng độ mồi khoảng 0,1 μM đến 1 μM . Nồng độ thường được xác định dựa vào giá trị OD (optical density) đo ở 260 nm.

9.5.4.4. Magnesium chloride

Nồng độ MgCl_2 ảnh hưởng đến sự thành công của phản ứng, nếu không có mặt Mg^{+2} thì sự khuếch đại sẽ không xảy ra. MgCl_2 cần thiết cho hoạt động của ADN polymerase khi ủ ADN với các đoạn mồi và nó có một mồi tương quan với nucleotid triphosphat. Nồng độ MgCl_2 tối ưu cho PCR phải được xác định cho từng phản ứng bằng thực nghiệm. Có thể dùng phương pháp định phân bằng cách thiết lập một loạt phản ứng với các nồng độ MgCl_2 khác nhau từ 1 mM đến 5 mM và xác định nồng độ MgCl_2 tối ưu cho mẫu mong muốn.

9.5.4.5. Nucleotid triphosphat (dNTP)

Bốn loại nucleotide thường được sử dụng ở nồng độ là 20 – 200 μM / mỗi loại nucleotid. Nồng độ cao hơn dễ dẫn đến sự khuếch đại "ký sinh". Sự mất cân bằng trong thành phần các nucleotid làm tăng các lỗi sao chép của polymerase.

9.5.5. Chu kỳ nhiệt PCR

Phản ứng chuỗi khuếch đại ADN đạt được bằng các chu kỳ được lặp đi lặp lại. Mỗi chu kỳ gồm sự biến tính bằng cách nâng nhiệt độ, ủ bằng cách hạ nhiệt độ và sự kéo dài. Một chu kỳ điển hình cho việc khuếch đại đoạn ADN gồm 500 cặp base sẽ là biến tính ở 94°C trong 60 giây, ủ ở 50°C trong 60 giây và 72°C trong 60 giây. Mỗi PCR được thiết kế và được tối ưu hóa khác nhau và tuỳ thuộc một số yếu tố:

– Kích thước của đoạn được khuếch đại. Các sản phẩm khuếch đại lớn hơn có thể cần giai đoạn biến tính và kéo dài lâu hơn. Các đoạn ngắn hơn cần ít thời gian hơn ở giai đoạn kéo dài.

- *Trình tự* của các đoạn mồi và các điều kiện về ion sẽ quyết định nhiệt độ ủ tốt nhất cho một chu kỳ. Vì vậy, dựa vào các phản ứng đã được thực hiện, nhiệt độ ủ có thể thay đổi từ 37°C đến 65°C.

9.5.6. Ưu điểm và nhược điểm của phương pháp PCR

Việc ứng dụng rộng rãi kỹ thuật PCR có nhiều lý do:

- Rất nhanh chóng. Chỉ tốn khoảng 3 giờ là có thể khuếch đại một trình tự mong muốn đã biết so với các kỹ thuật công nghệ di truyền "cổ điển" phải mất một tuần hay hơn nữa.

- Đơn giản: PCR có thể được thực hiện trong một ống nhỏ với các thành phần tối thiểu và chỉ việc trộn chúng lại với nhau. Các phương pháp tạo dòng gen điển hình khác cần có các vật liệu mắc tiền như các màng và các nucleotid triphosphate được đánh dấu phóng xạ và các kỹ thuật đặc biệt. PCR có thể được thực hiện trên các mẫu có chứa ADN tương đối thô, ví dụ vết máu chưa được xử lý cho việc phân tích pháp y. Điều này ngược với các phương pháp về thao tác gen là cần phải có ADN cả khuôn mẫu lẫn vector tương đối tinh khiết. Các yếu tố trên cho thấy rằng PCR là một sự thay đổi đáng kể so với các phương pháp cổ điển cho việc khuếch đại các trình tự đặc biệt.

- Nhạy: Phản ứng PCR cực nhạy vì có thể chỉ cần dùng một phân tử ADN làm khuôn cũng thu được sản phẩm.

Tuy nhiên, cần phải nhấn mạnh rằng các ứng dụng rộng rãi của kỹ thuật PCR đòi hỏi phải biết rõ trình tự ADN khuôn mẫu. Đây là một điểm giới hạn của kỹ thuật này, nghĩa là nó vẫn phải dựa vào nền tảng của các kỹ thuật công nghệ di truyền chứ không hẳn là thay thế các kỹ thuật này. Do vậy người ta không sử dụng kỹ thuật này trong các kỹ thuật tạo dòng gen thiết kế. Trong một vài trường hợp phương pháp PCR không áp dụng được với các đoạn ADN kích thước lớn hơn 3 kb (thường 1 kb là tốt nhất). Ngoài ra khả năng ngoại nhiễm đối với phương pháp này là rất lớn.

9.5.7. Các phương pháp PCR cải tiến

Phản ứng PCR được áp dụng rộng rãi và có những biến đổi phù hợp với từng mục đích nghiên cứu riêng. Vì vậy chúng ta gặp nhiều phương pháp khác như PRC tổ (Nested – PCR), ADN nhánh (Branch – ADN), PCR đa thành phần (multiplex – PCR), PCR ngược (Reverse – PCR), RAPD – PCR... nhưng tất cả đều dựa trên nguyên tắc của phản ứng PCR.

9.5.7.1. PCR "tổ"

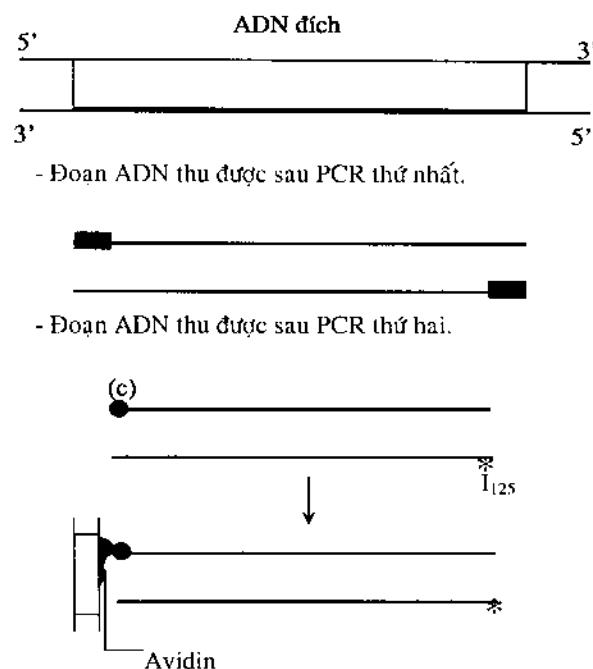
Để đạt được độ nhạy cao trong PCR, người ta có thể thực hiện 2 PCR liên tiếp. Trên một điện di đồ của các đoạn thu được sau PCR thứ nhất, người ta quan sát thấy ngoài một vệt chính còn có các vệt nhỏ khác. Điều này là do các mồi có thể bắt cặp với vùng khác của ADN đích bởi sự lai hoá không đặc hiệu.

PCR thứ 2 được thực hiện với các mồi giống PCR 1, khuếch đại các sản phẩm thu được từ PCR thứ nhất. PCR "tổ" có độ đặc hiệu cao do sử dụng 2 cặp mồi khác nhau.

Cặp mồi ngoại: Cặp mồi này tạo ra các đoạn ADN được khuếch đại giống với PCR cổ điển. Chúng đặc hiệu cho 2 trình tự mút của ADN cần khuếch đại. Các đoạn ADN thu được làm khuôn mẫu cho PCR thứ 2.

Cặp mồi nội: Cặp mồi này bổ sung với vùng nằm bên trong chuỗi nucleotid thu được với cặp mồi thứ nhất. Thuật ngữ "PCR tổ" là vì các đoạn mồi được tổ, đóng hộp trong lần PCR đầu.

Kỹ thuật này có độ nhạy và độ chuyên biệt cao. Người ta dùng kỹ thuật này để phát hiện virus gây bệnh viêm gan C, một loại virus có vật liệu di truyền là ARN. ARN virus không thể được phát hiện bằng sự lai hoá trực tiếp, nên phải thực hiện phản ứng chuỗi khuếch đại vật liệu di truyền của virus. PCR thứ nhất chuyển tất cả ARN virus thành cADN, đây chính là bước sao mă ngược nhờ enzym phiên mă ngược tạo cADN bằng kỹ thuật PCR "tổ". Trong PCR thứ 2, một mồi được biotin hoá, một mồi được gắn với iod 125. Các đoạn thu được sẽ được gắn trên một giá mang avidin (avidin là một chất có ái lực mạnh đối với biotin), và cuối cùng người ta đo hoạt tính phóng xạ để kết luận mẫu là dương tính hay âm tính (hình 9.7).



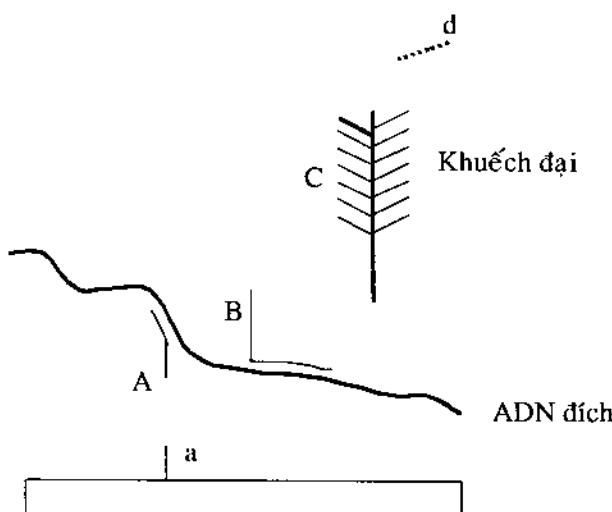
Hình 9.7. PCR "tổ" phát hiện HCV

PCR "tổ" có độ nhạy cao hơn PCR tiêu chuẩn là do lặp lại sự khuếch đại sản phẩm từ một PCR thứ 1 với một PCR thứ 2. Tuy nhiên vấn đề ngoại nhiễm cũng là vấn đề hạn chế của nó.

9.5.7.2. ADN "nhánh"

ADN đích được khuếch đại bằng PCR, sau đó cho lai với đoạn dò (là một phân đặc hiệu của ADN này). Tín hiệu được lai hoà đặc hiệu với đoạn dò theo cùng cách thức của ADN đích có gắn (chất huỳnh quang, chất dạ quang hay chất màu), tín hiệu sẽ được tạo ra nhờ chất đánh dấu hiện diện sau cùng ở máy khuếch đại. PCR "nhánh" có độ đặc hiệu cao.

Kỹ thuật ADN "nhánh" được dùng để phát hiện ADN của virus gây bệnh viêm gan B. Trong kỹ thuật này sử dụng 5 đoạn dò. Một đoạn dò thành phần là một trình tự cho phép sự lai hoà cùng một lúc trên 2 đoạn đích khác nhau.



Hình 9.8. Kỹ thuật ADN "nhánh" sử dụng 5 đoạn dò

9.5.7.3. PCR đa thành phần

Trong kỹ thuật này, 2 hay nhiều ADN đích được khuếch đại đồng thời. PCR đa thành phần có thể được sử dụng thường xuyên trong các xét nghiệm chẩn đoán, sử dụng 2 bộ mồi khác nhau: bộ thứ nhất để xác định sự đồng nhất của PCR và bộ thứ hai nhắm vào các trình tự ADN dư thừa. PCR đa thành phần cũng có thể được sử dụng trong phòng thí nghiệm vi sinh lâm sàng với cùng đoạn dò cho một mẫu xét nghiệm cho vài sinh vật khác nhau trong một ống PCR.

Khi tiến hành PCR đa thành phần cần phải xem xét một số yếu tố trong thiết kế đoạn mồi. Các mồi phải có nhiệt độ lai tương tự nhau. Sự khác biệt 10°C trong nhiệt độ lai giữa các bộ mồi có thể làm cho các sản phẩm khuếch đại rất khác nhau

hay tạo sự khuếch đại không thể phát hiện được của sản phẩm này hay sản phẩm khác. Các đoạn mồi được sử dụng cũng được thiết kế sao cho các sản phẩm khác nhau đủ về kích thước để mỗi sản phẩm có thể được xác định nhưng không quá khác nhau để cho sự khuếch đại của một ADN đích được ưu tiên hơn so với sự khuếch đại của ADN đích khác. PCR đa thành phần có các ADN đích khác nhau nhiều về kích thước thường ưu tiên khuếch đại các ADN đích ngắn trước các ADN đích dài, kết quả là có sự khác nhau về số lượng các sản phẩm được khuếch đại. Tuy nhiên, sự khác nhau đáng kể về nhiệt độ lai hay chiều dài ADN đích là không thể tránh khỏi. Người ta có thể điều chỉnh quy trình thực hiện để làm cân bằng phản ứng và đạt được số lượng sản phẩm tương đương nhau hoặc điều chỉnh nồng độ mồi cân bằng với một phản ứng đa thành phần.

9.5.7.4. PCR của ARN

ARN của retrovirus được ly trích và được chuyển đổi thành cADN bởi enzym sao mã ngược enzym phiên mã ngược, cADN mới được dùng làm khuôn mẫu cho PCR. Tuy nhiên, các phản ứng nhờ enzym phiên mã ngược khó thực hiện và việc sử dụng các enzym không chịu được nhiệt trên 42°C là một bất lợi vì xảy ra sự lai không nghiêm ngặt. Vì vậy, sự phiên mã ngược kém hiệu quả là một trở ngại làm cho sự phát hiện các ARN đích kém nhạy và kém đặc hiệu.

Người ta đã tìm ra ADN polymerase chịu nhiệt có hoạt tính enzym phiên mã ngược đó là ADN polymerase tái tổ hợp được trích từ vi khuẩn chịu nhiệt *Thermus thermophilus* (Tth pol). ADN polymerase chịu nhiệt này hoạt động hữu hiệu khi có sự hiện diện của ion Mn²⁺. Hỗn hợp phản ứng của PCR của ARN bao gồm: Tth pol, 2 loại mồi oligonucleotide (một được sử dụng cho sự tổng hợp cADN, và cả 2 được sử dụng cho PCR), ion mangan ở dạng MnCl₂, và tất cả các thành phần khác cần cho sự phiên mã ngược và PCR. Máy chu kỳ nhiệt được đun nóng trước đến 70°C, và những thành phần phiên mã ngược được ủ trong 15 phút. Sau phản ứng nhờ enzym phiên mã ngược, hỗn hợp phản ứng được nâng nhiệt độ lên 95°C để làm biến tính phức hợp ARN – ADN. PCR sau đó được bắt đầu với 2 chu kỳ ở 95°C trong 15 phút và 60°C trong 20 giây, tiếp theo là 38 chu kỳ ở 90°C trong 15 giây và 60°C khoảng 20 giây. Sự phiên mã ngược được thực hiện ở 70°C do vậy độ đặc hiệu và độ nhạy của việc phát hiện ARN đích cao.

9.5.8. Các lĩnh vực ứng dụng của PCR

Trong nghiên cứu khoa học, PCR giúp cho việc xác định trình tự nucleotid của các đoạn ADN được nhân lên; có thể sử dụng PCR để tách dòng những đoạn ADN đặc hiệu, mặc dù trong nhiều trường hợp tách dòng không cần đến PCR; nó giúp

phát hiện đột biến, nghiên cứu mARN hoặc tạo các đột biến gen, cho phép phân tích liên kết gen từ những tế bào riêng lẻ; giúp nghiên cứu quá trình tiến hoá ở mức phân tử. Thậm chí giúp phục hồi những gen trong các cổ vật đã tồn tại cách đây hàng chục triệu năm.

Trong chọn giống vật nuôi và cây trồng, đặc biệt đối với động vật và thực vật bậc cao, vốn trước đây việc lựa chọn cặp cha mẹ làm giống là rất khó khăn, cần nhiều năm, thì nay PCR cho phép thực hiện chỉ trong vài ngày, thậm chí vài giờ với hiệu quả rất cao.

Trong y học, PCR có thể chẩn đoán chính xác các bệnh nhiễm trùng từ virut đến vi khuẩn và các bệnh do nấm, kể cả HIV – AIDS, chẩn đoán sớm và theo dõi điều trị ung thư...

Trong tư vấn di truyền y học, PCR cho phép chẩn đoán nhanh, chính xác các bệnh di truyền kể cả chẩn đoán trước sinh giới tính và các dị tật bẩm sinh có thể có từ khi thai mới được 8 tuần tuổi và điều quan trọng là không cần chọc ối, chỉ cần lấy một giọt máu ngoại vi của mẹ để làm mẫu thí nghiệm.

Trong khoa học hình sự, PCR là không thể thiếu, nó giúp chẩn đoán nhanh, chính xác thủ phạm chỉ từ một vết máu khô, một sợi tóc hoặc một sinh phẩm nào đó mà thủ phạm để lại trên hiện trường. Ngoài ra kỹ thuật này còn cho phép xác định chính xác quan hệ huyết thống cha – con, ông – cháu v.v... cũng chỉ trong vài giờ.

Trong bảo vệ môi trường, việc xác định mức độ ô nhiễm sinh học có thể thực hiện rất hiệu quả và nhanh chóng bằng phản ứng PCR.

Những ứng dụng thực tiễn đa dạng của phản ứng chuỗi tổng hợp ADN thật vô cùng tận. Nếu như năm 1985 mới có 3 công trình được công bố về PCR thì 5 năm sau kỹ thuật này được sử dụng trong hàng nghìn phòng thí nghiệm trên khắp thế giới. Ngay ở nước ta, PCR cũng đã và đang được dùng trong nhiều trường đại học, viện nghiên cứu và ngày càng trở nên phổ biến. Phản ứng chuỗi tổng hợp ADN thật sự đã đưa lại một cuộc cách mạng trong lĩnh vực ứng dụng thực tế của sinh học phân tử. Ngày nay có thể nói mọi lĩnh vực nghiên cứu sinh học đều sử dụng PCR.

Tóm tắt

Chiết tách và tinh chế ADN đóng vai trò quan trọng trong các nghiên cứu sinh học phân tử. Các bước chính để chiết tách ADN là phá vỡ tế bào, loại bỏ protein, và tủa acid nucleic. Tinh chế ADN có thể áp dụng kỹ thuật siêu ly tâm hoặc sắc ký. Nồng độ ADN có thể đo bằng quang kế ở bước sóng 260 nm, kích thước tương đối

của ADN có thể xác định bằng kỹ thuật điện di trên gel đối chiếu với thang kích thước đã biết.

Các thao tác cắt, nối và lai ADN là các thao tác cơ bản của kỹ thuật di truyền phân tử. Trình tự ADN có thể xác định bằng phương pháp hoá học (Maxam và Gilbert) hoặc phương pháp enzym (Sanger và cộng sự). Phương pháp Sanger đã được cải tiến đơn giản hơn và có thể tự động hoá.

Phương pháp PCR sử dụng enzym ADN polymerase và máy luân nhiệt cho phép khuếch đại *in vitro* acid nucleic lên gấp tỉ lần sau 30 – 40 chu kỳ chỉ trong vài giờ. Kỹ thuật này có nhiều ưu điểm như đơn giản, nhanh và nhạy nên được ứng dụng rộng rãi trong nghiên cứu khoa học, chẩn đoán, khoa học hình sự, chọn giống...

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. Tốc độ điện di **không** phụ thuộc:
 - a) Kích thước phân tử ADN
 - b) Cấu dạng của ADN
 - c) Nồng độ ADN
 - d) Nồng độ gel
 - e) Điện thế sử dụng
3. 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
4. Yếu tố ảnh hưởng đến sự lai hoá:

a) Nồng độ ADN	b) Nhiệt độ và thời gian phản ứng
c) Độ dài của các trình tự	d) Lực ion
e) Tất cả	
5. Đặc điểm **không** thuộc phương pháp định trình tự của Sanger:
 - a) Xử lý hoá học chuyên biệt làm biến đổi đặc trưng một loại nucleotid
 - b) Sử dụng ADN polymerase

- c) Nucleotid được đánh dấu.
 - d) Phản ứng tiến hành trong bốn phân đoạn
 - e) Có sử dụng dideoxynucleotid
6. Chất làm giảm nhiệt độ biến tính của ADN trong PCR
- a) Formamid
 - b) MgCl₂
 - c) DMSO
 - d) EDTA
 - e) a và c
7. Mồi trong phản ứng PCR
- a) Đoạn ADN ngắn, mạch đơn
 - b) Có trình tự bổ sung với ADN khuôn tại điểm đầu sao chép
 - c) Dài từ 6 – 30 nucleotid
 - d) Là oligonucleotid
 - e) Tất cả
8. Tính nhiệt độ "chảy" của đoạn mồi nhằm xác định nhiệt độ thích hợp để
- a) Biến tính mồi
 - b) Mồi gắn vào khuôn
 - c) Tổng hợp từ khuôn trên gen
 - d) Mồi không gắn bổ sung vào nhau
 - e) Mồi gắn vào các đoạn khác nhau
9. Chọn chu kỳ nhiệt PCR dựa vào 2 yếu tố là kích thước của khuôn và
- a) Độ tinh khiết của khuôn
 - b) Năng độ của khuôn
 - c) Kích thước mồi
 - d) Trình tự mồi
 - e) c và d
10. PCR tổ:
- a) Dựa trên nguyên tắc của phản ứng PCR
 - b) Là hai PCR liên tiếp, sử dụng hai cặp mồi "ngoại" và "nội"
 - c) Có độ nhạy và chuyên biệt cao hơn PCR thường
 - d) Ứng dụng trong chẩn đoán
 - e) Tất cả đúng

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 hòa 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. Các phương pháp phân tích ADN

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 Thuỷ 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 Hổ. *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 HDQT kiêm Tổng Giám đốc NGÔ TRẦN ÁI
Phó Tổng Giám đốc kiêm Tổng biên tập NGUYỄN QUÝ THAO

Chịu trách nhiệm nội dung:

Chủ tịch HĐQT kiêm Giám đốc Công ty CP Sách ĐH – DN
TRẦN NHẬT TÂN

Biên tập và sửa bản in:

TRẦN NGỌC OANH

Trình bày bìa:

BÙI QUANG TUẤN

Chế bản:

THÁI SƠN

SINH HỌC PHÂN TỬ

Mã số: 7K721M7 – DAI

In 1.000 bản (QĐ 94), khổ 19 x 27 cm, tại Công ty CP In Anh Việt.

Địa chỉ : Số 74, ngõ 310 đường Nghi Tàm, Tây Hồ, Hà Nội.

Số ĐKKH xuất bản : 770 - 2007/CXB/3 – 1676/GD.

In xong và nộp lưu chiểu tháng 12 năm 2007.



CÔNG TY CỔ PHẦN SÁCH ĐẠI HỌC - DẠY NGHỀ
HEVOBCO
25 HÀN THUYỀN – HÀ NỘI
Website : www.hevobco.com.vn



VƯƠNG MIỆN KIM CƯƠNG
CHẤT LƯỢNG QUỐC TẾ

TÌM ĐỌC SÁCH THAM KHẢO Y HỌC CỦA NHÀ XUẤT BẢN GIÁO DỤC

- | | |
|-------------------------------------|--|
| 1. Sinh học phân tử | GS. TS. Nguyễn Văn Thành (Chủ biên) |
| 2. bào chế và sinh dược học – Tập 2 | PGS. TS. Lê Quan Nghiệm – TS. Huỳnh Văn Hoá (Đồng chủ biên) |
| 3. Thực vật dược | TS. Trương Thị Đẹp (Chủ biên) |
| 4. Ký sinh trùng | PGS. TS. Phạm Văn Thân (Chủ biên) |
| 5. Hóa đại cương | PGS. TSKH. Phan An (Chủ biên) |
| 6. Điều dưỡng cơ bản 1 | PGS. TS. Phạm Văn Linh – TS. Lê Văn An (Đồng Chủ biên) |
| 7. Điều dưỡng cơ bản 2 | PGS. TS. Hoàng Ngọc Chương – BSCKII. Trần Đức Thái (Đồng Chủ biên) |
| 8. Kiểm nghiệm thuốc | Trần Tich (Chủ biên) |
| 9. Nhãn khoa | PGS. TS. Hoàng Thị Phúc (Chủ biên) |
| 10. Sinh lý học | GS. TS. Phạm Thị Minh Đức (Chủ biên) |
| 11. Phẫu thuật miệng – Tập 1 | TS. BS. Lê Đức Lánh (Chủ biên) |
| 12. Hóa phân tích – Tập 1 | PGS. TS. Võ Thị Bạch Huệ (Chủ biên) |
| 13. Công nghệ bào chế dược phẩm | PGS. TS. Hoàng Minh Châu (Chủ biên) |
| 14. Dược lý học – Tập 1 | GS. TS. Đào Văn Phan (Chủ biên) |

Bạn đọc có thể mua tại các Công ty Sách - Thiết bị trường học ở các địa phương hoặc các Cửa hàng sách của Nhà xuất bản Giáo dục :

Tại Hà Nội : 25 Hàn Thuyên ; 187B Giảng Võ ; 232 Tây Sơn ; 23 Tràng Tiền ;

Tại Đà Nẵng : Số 15 Nguyễn Chí Thanh ; Số 62 Nguyễn Chí Thanh ;

Tại Thành phố Hồ Chí Minh : Cửa hàng 451B - 453, Hai Bà Trưng, Quận 3 ;
240 Trần Bình Trọng – Quận 5.

Tại Thành phố Cần Thơ : Số 5/5, đường 30/4 ;

Website : www.nxbgd.com.vn



8 934980 786352



Giá: 47.000 đ