2.1 Compoziţia chimică şi structura acizilor nucleici

2.1.1 Compoziţia chimică şi structura ADN

Acizii nucleici sunt macromolecule polinucleotidice care conţin informaţia genetică din celulă în formă codificată. Există două tipuri de acizi nucleici, diferenţiabili prin structura chimică şi funcţiile îndeplinite în celulă şi anume ADN sau acidul dezoxiribonucleic şi ARN sau acidul ribonucleic.

ADN este purtătorul informaţiei ereditare la organismele eucariote şi la unele organisme procariote (bacterii), dar şi la dezoxiribovirusuri.

Totalitatea materialului genetic al unei celule formează genomul celulei. La eucariote genomul nuclear este format din totalitatea cromozomilor cu structură liniară din nucleu. De exemplu, genomul nuclear haploid al unei celule somatice este distribuit în cei 46 cromozomi sub forma a 46 molecule de ADN. Cei 46 cromozomi sunt grupaţi în 23 perechi de omologi, care formează două seturi (2n). Fiecare set de cromozomi cuprinde aproximativ 3.200.000.000 pb ($\approx 3 \times 10^9$ perechi de baze). Genomul extranuclear din celula somatică animală este reprezentat de mitocondrii. O mitocondrie conţine o singură moleculă de ADN dublu catenar circular, care cuprinde 16.596 pb ($1,6 \times 10^4$). Cu excepţia hematiilor, toate celulele unui animal conţin propria copie sau copii ale genomului (Brown, T.A., 2002). În celulele eucariote vegetale găsim în plus, în citoplasmă şi genomul plastidial.

Procariotele (bacterii, alge unicelulare) nu au nucleu adevărat separat de membrană nucleară de restul celulei. Genomul nucleoidal de la bacterii constă în general dintr-un singur cromozom care conţine o singură moleculă de ADN circular. În citoplasma bacteriilor întâlnim material genetic accesoriu, molecule mici de ADN dublu-catenar circular, sau liniar la unele specii, denumite plasmide, care formează genomul extranucleoidal.

La virusuri nu întâlnim organizare celulară iar genomul este format dintr-o moleculă de ADN (dezoxiribovirusuri) sau de ARN (ribovirusuri), liniară sau circulară, dublucatenară sau monocatenară.

ADN (acidul dezoxiribonucleic) şi ARN (acidul ribonucleic) sunt moleculele de acizi nucleici în care informaţia genetică este codificată. Ambele sunt molecule polinucleotidice rezultate prin polimerizarea monomerilor (nucleotidelor) prin intermediul legăturilor fosfodiesterice. O catenă polinucleotidică este alcătuită din unităţi mai simple numite

nucleotide. Nucleotidele preiau denumirea acizilor nucleici în compoziţia cărora intră. Cele din compoziţia ADN se numesc dezoxiribonucleotide iar cele din ARN ribonucleotide.

ADN-ul are o structură primară şi una secundară. Structura primară a ADN se referă la elementele din care este compusă o catenă, iar structura secundară se referă la duplexul sau conformaţia bicatenară a moleculei de ADN. Aşadar, monomerul din care este alcătuit ADN-ul şi ARN-ul este nucleotida. Fiecare nucleotidă constă dintr-o pentoză (zahar), o bază azotată şi o grupare fosfat.

În ADN pentoza este dezoxiriboza, iar în ARN riboza. Cele două pentoze diferă între ele prin gruparea ataşată la carbonul numărul 2'. În cazul dezoxiribozei este vorba de un atom de hidrogen iar în cazul ribozei de gruparea hidroxil (Fig.2.5).

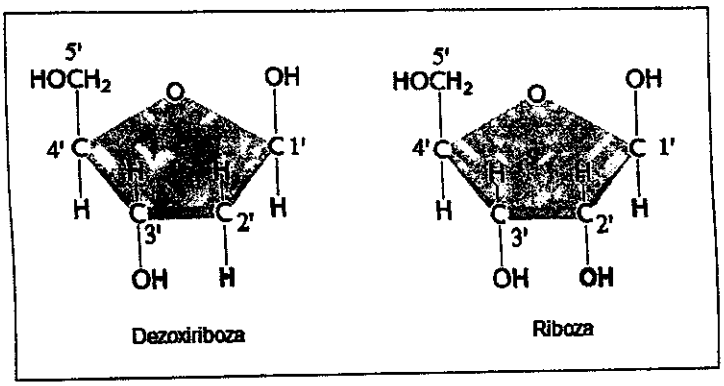


Fig.2.5 Structura pentozei-dezoxiriboza din ADN, comparativ cu riboza din ARN

Prin legarea unei baze azotate la o pentoză rezultă un nucleosid (Fig.2.7.a). Legăturile între acestea sunt covalente şi se realizează între carbonul C1 din pentoză şi azotul N1 dintr-o purină sau N9 dintr-o pirimidină. Nucleotida rezultă prin legarea unui radical fosfat la un nucleosid.

Există două tipuri de baze azotate care intră în structura ADN şi anume purinele, structuri cu două nuclee ciclice şi pirimidinele formate dintr-un singur nucleu ciclic (Fig.2.6). Adenina (A) şi guanina (G) sunt baze azotate purinice, iar timina şi citozina în ADN sunt bazele azotate pirimidinice. În ARN pirimidina timina este înlocuită cu altă bază, tot pirimidinică - uracilul (U).

Polimerizarea nucleotidelor de pe o catenă se face prin intermediul radicalului fosforic care leagă două nucleoside învecinate (o bază azotată şi pentoza), prin intermediul carbonilor 3' şi 5' din pentoze (Fig.2.7 b). Aceste legături sunt covalente,de tip fosfodiester, mult mai puternice decât legăturile de hidrogen dintre cele două catene (forţe Van der Waals).

Aşa cum am arătat anterior, elementele din care este alcătuită o singură catenă formează structura primară a ADN.

Conform modelului propus de Watson şi Crick, în celulă ADN-ul este bicatenar, structură rezultată prin legarea bazelor azotate complementare de pe cele două catene antiparalele, prin legături duble şi triple de hidrogen. Legăturile duble se stabilesc întotdeauna între bazele azotate A şi T, iar cele triple între C şi G, adică între o bază azotată purinică şi una pirimidinică. Această constatare a fost făcută prima dată de Cargaff, şi a dus în final la stabilirea structurii de dublu-helix a ADN. Cele două catene sunt antiparalele, cu direcţii contrare una faţă de cealaltă şi cu secvenţă complementară (Fig.2.8).

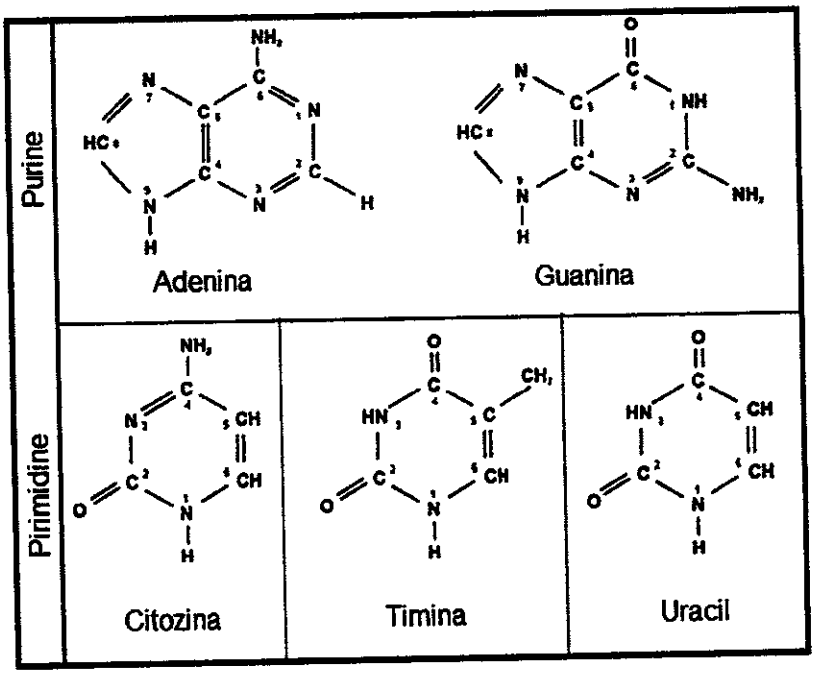


Fig.2.6 Bazele azotate purinice (A, G) şi pirimidinice (C,T,U) prezente în structura ADN şi ARN

Astfel, prin complementaritate una din catene dictează secvenţa nucleotidică a celeilalte catene. Catena 3'-5' este catena ascendentă, iar cealaltă, cu direcţia 5'-3', este descendentă. La replicarea ADN, catena ascendentă se mai numeşte conducătoare la procariote şi progresivă la eucariote deoarece sinteza ei se face continuu, iar catena descendentă se mai numeşte decalată (procariote) sau retrogradă (eucariote) deoarece sinteza ei se face pe fragmente Okazaki. Structura monocatenară a ADN apare numai în anumite etape din ciclul de viaţă al celulei, mai precis în momentul

replicării ADN în care cele două catene se desfac ca un fermoar formând ochiul de sinteză. Altă situaţie în care ADN-ul se află în celulă în stare monocatenară este în timpul sintezei proteinelor, când informaţia de pe una din catenele ADN este transcrisă în ARNm (ARN mesager). Prin urmare ADN-ul apare în celulă şi în stare monocatenară în timpul îndeplinirii celor două funcţii caracteristice: funcţia autocatalitică şi funcţia heterocatalitică.

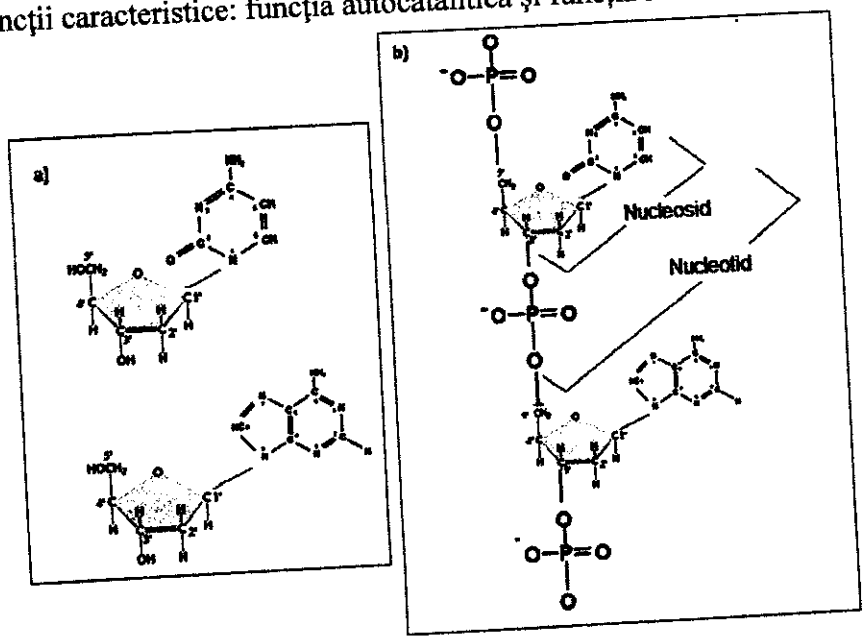


Fig.2.7 Structura unui nucleosid rezultat prin legarea C1 din pentoză la N1 de la o pirimidină sau la N9 de la o purină (a). Formarea unui nucleotid prin legarea grupării fosfat de nucleoside învecinate, legătură realizată prin intermediul C3' şi C5' din pentoză (b)

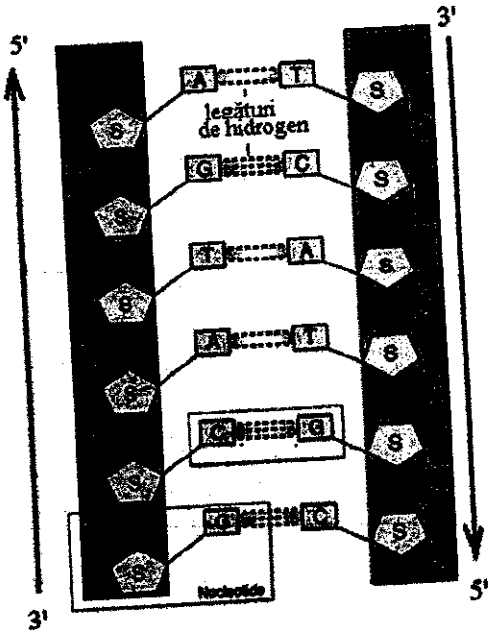


Fig.2.8 Structura bicatenară a ADN rezultă prin legarea celor două catene de direcţii contrare, prin punţi de hidrogen. A şi T se leagă prin legături duble iar C şi G prin legături triple. Direcţia ascendentă este notată 3'-5' iar cea descendentă 5'-3'

Funcţia autocatalitică face referire la replicarea ADN în timpul fazei sintetice S din interfază. În faza sintetică cromozomii sunt monocromatidici şi este iniţiată sinteza celei de-a doua cromatide. În fapt, în proces are loc biosinteza ADN după modelul semiconservativ prin care dintr-o moleculă parentală bicatenară vor rezulta două molecule fiice identice, fiecare cu câte o catenă veche provenită de la molecula mamă şi una nouă sintetizată pe bază de complementaritate cu matriţa.

Funcţia heterocatalitică a ADN este funcţia prin care informaţia codificată sub formă de codoni la nivel ADN este transcrisă în ARNm şi tradusă la nivelul ribozomilor într-o polipeptidă.

2.1.2 Compoziţia chimică şi structura ARN

Singurele entităţi care conţin ca material genetic ARN-ul sunt ribovirusurile. ARN-ul, la fel ca şi ADN-ul este tot o structură polinucleotidică, însă spre deosebire de ADN, ARN-ul este monocatenar. Într-o celulă se găsesc mai multe specii de ARN din care una singură este o moleculă informaţională - ARNm, restul îndeplinesc roluri funcţionale extrem de variate în celule.

ARN mesager (ARNm), ARN de transfer (ARNt) şi **ribozomal (ARNr)** sunt speciile clasice de ARN cunoscute încă cu mult timp în urmă. Alături de acestea, relativ recent au fost descoperite în celule şi alte specii cu rol reglator, grupate în categoria speciilor mici de ARN, denumite astfel datorită lungimii reduse a moleculelor.

Deşi toate tipurile de ARN sunt monocatenare, unele prezintă numeroase porţiuni fals bicatenare care apar datorită plierii catenei, când regiuni din ARN devin complementare între baze, între care se stabilesc legături de hidrogen. Această conformaţie la ARN este denumită structura secundară. ARN-ul în stare bicatenară poate să apară şi în cazul unor ribovirusuri la care este material genetic. La alte ribovirusuri însă, materialul genetic este reprezentat de ARN monocatenar. Structura bicatenară a ARN din genomul ribovirusurilor este similară structurii bicatenare a ADN adică se prezintă sub forma a două catene antiparalele.

Legăturile fosfodiesterice dintre bazele azotate din ARN sunt orientate la exteriorul helixului în timp ce bazele azotate sunt orientate spre interior. Ribonucleotida este monomerul din care este format lanţul de ARN, care spre deosebire de dezoxiribonucleotida din ADN deşi conţine aceleaşi trei elemente, prezintă unele diferenţe în structura chimică şi în compoziţie:

- pentoza sau zaharul este **riboza în ARN**
- bazele azotate sunt reprezentate de A, G (purine) şi C, **U** (pirimidine), timina din ADN fiind înlocuită de uracil (U).

Prin intermediul radicalului fosforic se leagă două ribonucleotide învecinate formând catena poliribonucletidică (Fig. 2.9) .

Dacă la ADN procentul de baze azotate complementare este acelaşi, la ARN această regulă nu se mai respectă datorită structurii monocatenare.

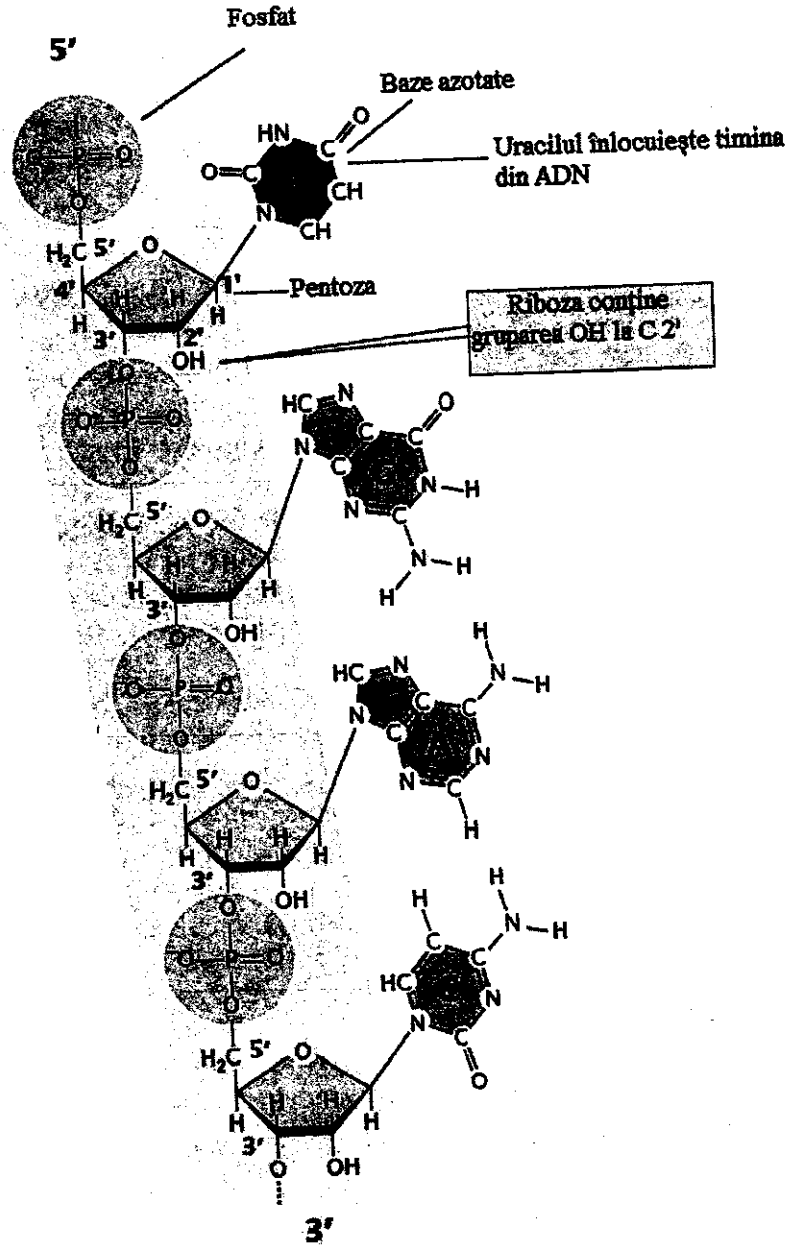


Fig.2.9 Structura poliribonucleotidică a ARN (modificată şi adaptată, sursa: Pierce şi colab., 2002)

Speciile de ARN din celule

ARN mesager (ARNm) se numeşte şi informaţional deoarece transportă la ribozomi mesajul genetic transcris din molecula de ADN.

La eucariote ARNm se sintetizează în nucleu şi se decodifică în citoplasmă, la ribozomi. La procariote nucleul nu este delimitat de membrană, deci şi sinteza ARNm se produce în acelaşi spaţiu. Lungimea sa este variabilă în funcţie de cantitatea de informaţie copiată (transcrisă) din ADN. Prin transcripţie informaţia genetică (succesiunea nucleotidelor) din catena matriţă a ADN (3'→5') este copiată pe bază de complementaritate în catena de ARNm cu direcţia 5'→3'. Prin urmare, mărimea ARNm este aceeaşi cu a ADN-ului matriţă pe baza căruia s-a sintetizat prin complementaritate.

La procariote informaţia genetică de pe mai multe gene structurale [redacted] este transcrisă într-o singură moleculă de ARNm denumit *ARNm policistronic*. Acest lucru se întâmplă deoarece genele la procariote sunt formate doar din unităţi informaţionale, iar ARNm conţine informaţia transcrisă de pe mai multe gene adiacente care formează operonul. În timp ce la procariote există o singură ARN polimerază care realizează transcripţia tuturor speciilor de ARN, inclusiv a ARNm, la eucariote sinteza ARNm se realizează cu ajutorul ARN polimerazei II, iar celelalte ARN polimeraze participă la sinteza altor specii de ARN.

La eucariote expresia genelor se realizează în etape. În prima etapă are loc transcripţia sau copierea informaţiei din ADN, în urma căreia se sintetizează o moleculă de *ARNm precursor* ce conţine copiată informaţia atât din exoni dar şi din unităţile noninformaţionale intronice. Acest ARNm precursor are dimensiunea catenei matriţă de ADN de pe care s-a sintetizat. Ulterior ARNm precursor suferă procese de prelucrare la capete prin **adiţia** unei **grupări 7-metil-guanozină** la capătul 5' şi a cozilor **poliadenilate** (poli-A) la capătul 3', rezultând astfel un ***ARN intermediar***. Acesta intră în procesul de prelucrare a informaţiei (*splicing*) prin care sunt eliminaţi pe rând intronii şi sunt asamblaţi exonii, rezultând în final *ARNm matur*. ARNm matur trece prin porii membranei nucleare în citoplasmă şi se asociază cu ribozomii, formând **poliribozomii**, unde va fi decodificat în secvenţe polipeptidice. Durata de viaţă a ARNm este relativ scurtă, acesta fiind degradat imediat după sinteza proteinei, însă sintetiza altor molecule de ARNm are loc continuu, pe măsura cerinţelor celulei. Procesul de degradare a moleculelor de ARNm se realizează cu ajutorul exonucleazelor

după ce iniţial s-a produs deadenilarea (îndepărtarea cozilor poli-A) şi a fos îndepărtată structura „cap" de la extremitatea 5'.

ARN ribozomal (ARNr) reprezintă 80-85% din cantitatea de AR din celulă, este monocatenar şi stabil şi nu se modifică în cursul sintez proteinelor. Monocatena de ARNr se poate contorsiona şi asocia în anumi regiuni rezultând conformaţia secundară şi terţiară. Porţiuni necomplementare din ARNr au rolul de legare a [redacted] şi a [redacted] la nivel ribozomilor, în timpul sintezei proteice (Coman, N., 2003).

În ribozomi - granule sferice dense cu diametrul cuprins între 150 şi 300 Å, ARNr este asociat cu proteinele ribozomale. ARNr ocupă 2/3 d ribozomii de la eucariote şi 1/3 din cei de la procariote.

Toţi ribozomii conţin două subunităţi - una mică şi alta mar diferenţiabile prin constanta de sedimentare Svedberg (S). O unita Svedberg este egală cu 1 x 10^{-13} secunde. Ribozomii 80S au fost evidenţia la eucariote; aceştia sunt formaţi dintr-o subunitate mare 60S şi o subunita mică 40S. Subunitatea mare conţine **ARN 5S, 5,8S şi 28S**. Subunitatea mic 40S conţine doar ARN 18S. Moleculele de ARNr conţin un număr variab de nucleotide, în funcţie de specia de ARNr. Astfel, ARNr 5S conţine 12 nucleotide (nt) , cel 5,8 S – 160 nucleotide, iar ARNr 28S - 4.718 nucleotic (Lodish şi colab., 2000). Această compoziţie poate fi exploatată c indicator molecular al gradului de înrudire dintre specii.

Ribozomii procariotelor au constanta de sedimentare Svedberg 7(Şi aceştia conţin tot două subunităţi: cea mare - 50S (conţine ARN 5 format din 120 nucleotide şi 23S format din 2904 nucleotide), iar cea mică 30S (conţine ARN 16S format din 1542 nucleotide). Spre deosebire d celelalte structuri celulare care îndeplinesc rol genetic, ribozomii nu a capacitate de autoreproducere deoarece nu conţin ADN propriu, însă conţi acizi nucleici (ARN) şi sunt sediul sintezei proteice. Sinteza lor începe i nucleoli ca protoribozomi şi se definitivează în citoplasmă.

La eucariote, genele care controlează sinteza ARNr se găsesc în sut chiar mii de copii. Genele pentru ARNr 5,8S, 18S şi 28S sunt prezente i tandem în regiunile organizatorilor nucleolari ale anumitor cromozomi, i timp ce genele pentru ARNr 5S se găsesc plasate pe alţi cromozomi care n prezintă organizatori nucleolari (Russell, P., 2010). La eucariote sinte ARNr se face de pe genele nucleare iar a ARNr specific mitocondriei plastidelor de la plante, se realizează de pe genele citoplasmatice.

Genele pentru sinteza ARNr **18S, 5,8S şi 28S** sunt plasate una dup alta într-o secvenţă repetată în tandem de 100-1000 ori, în funcţie d

rganism şi grupate în clustere. Transcripţia genelor are loc cu ajutorul nzimei **ARN polimeraza I** care va produce un precursor pre-ARNr, cu paţii între cele 3 gene şi două capete 5'-3' pentru fiecare dintre ele. Acest ranscript este ulterior prelucrat de către ribonucleaze prin eliminarea egiunilor spaţiatoare, rezultând cele trei subunităţi ale ARNr. Îndepărtarea egiunilor spaţiatore din transcriptul pre-ARNr duce la obţinerea a trei ranscripte separate, proces deosebit de cel de îndepărtare a intronilor din gene (*splicing*) în care se formează un singur ARNm cu exonii asamblaţi. Sinteza celui de-al patrulea ARNr, **5S,** se face cu ajutorul **ARN polimerazei III** de pe gene situate diferit în genom. (Lodish şi colab., 2003; Russel, P., 2010).

La procariote există 7 unităţi genice pentru sinteza ARNr. Fiecare nitate conţine informaţia pentru sinteza ARNr 16S, 23S şi 5S, aranjată în această ordine. Fiecare unitate codificatoare are un singur promotor ecunoscut de ARN polimeraza bacteriană care va transcrie unitar un precursor pre-ARNr format din cele trei specii de ARNr, fără spaţiere, în ordinea 5'-16S-23S-5S-3'. Proteinele ribozomale împreună cu pre-ARNr formează un complex ribonucleoproteic de dimensiune mare. Ulterior, prin procesarea transcriptului la nivelul acestui complex, vor rezulta cele trei specii de ARNr.

ARN de transport (ARNt) reprezintă 10-15% din totalul ARN celular şi are rol în procesul de biosinteză a proteinelor deoarece are proprietatea de a transporta aminoacizii activaţi la ribozomi, centrul sintezei proteinelor. Deoarece există cel puţin atâtea tipuri de ARNt câţi aminoacizi există, diferenţierea se face pe baza compoziţiei particulare în ribonucleotide. Diferenţele în structură explică de fapt abilitatea fiecărui tip de ARNt de a transporta un aminoacid specific. Aminoacidul este legat de ARNt specific cu ajutorul enzimei aminoacid - sintetaza ARNt; pentru fiecare aminocid existând o enzimă activatoare. Tipul aminoacidului transportat la ribozomi este dictat de codonul din ARNm, recunoscut prin bucla anticodon din ARNt. Pentru aminoacidul alanină spre exemplu, sunt 3 molecule de ARNt, iar pentru valină sunt 2 molecule de ARNt. Moleculele de ARNt specifice pentru acelaşi aminoacid, sunt identice între ele la toate speciile (Vlaic, A., 2011).

ARNt are greutate moleculară mică, fiind alcătuit din 70-96 ribonucleotide cu secvenţa diferită în funcţie de specia ARNt. Este monocatenar cu porţiuni bicatenare, având forma unei frunze de trifoi (treflă) (Fig.2.10).

La nivelul moleculei de ARNt se disting 4 regiuni de împerechere între baze azotate complementare şi 4 bucle care reprezintă regiuni neîmperecheate. Bucla numărul II conţine secvenţa **anticodon** formată din trei nucleotide, care recunoaşte specific secvenţa codon din ARNm. Câte un ARNt este adus specific la ribozomi iar specificitatea este dictată de codonul din ARNm şi nu de aminoacid.

Regiunea acceptoare are două extremităţi, 3' şi 5' libere. Toate speciile de ARNt au la extremitatea 3' tripletul CCA (citozină-citozină-adenină) care serveşte ca loc de ataşare pentru aminoacizi, iar extremitatea 5' este permanent guanilată, fiind de structură GGG.

Numerotarea buclelor în ARNt:

I. *Bucla dihidrouracil* (DHU), sau bucla D;

II. *Bucla anticodonului*, recunoaşte în mod specific un anumit codon din ARN mesager. Existenţa mai multor tipuri de ARNt atât la procariote cât şi la eucariote, arată că fiecare tip de ARNt trebuie să fie determinat de către o genă specifică.

III. *Braţul suplimentar* (extrabraţul) cu un număr variabil de nucleotide.

IV. *Bucla T Ψ C* G (ribotimină-pseudouracil-citozină-guanină) (T, pshi, C, G).

Sinteza ARNt la bacterii este asigurată de gene care se găsesc în una, până la câteva copii în cromozomul bacterian, iar la eucariote de gene care se repetă de mai multe ori în genom. ARNt bacterian este transcris de singura ARN polimerază bacteriană, în timp ce transcrierea genelor ARNt de la eucariote este asigurată de **ARN polimeraza III**. În urma transcripţiei rezultă un precursor pre – ARNt care este prelucrat. Prelucrările constau în eliminarea unui grup de nucleotide de la capătul 5' şi modificarea unor baze azotate din structură. Nucleotidele de la capătul 5' sunt eliminate de o ribonuclează în cursul procesului de prelucrare, iar la capătul 3' sunt modificate trei nucleotide şi transformate în secvenţa 3' –CCA care se găseşte în toate tipurile de ARNt. Celelalte prelucrări în secvenţa ARNt precursor constau în modificări ale unor baze azotate din structură. Experimental s-a demonstrat că circa 10% din bazele azotate sunt modificate în molecula matură de ARNt (Lodish şi colab., 2003).

Aşadar, enzimele implicate în transcripţia speciilor de ARN la eucariote sunt ARN polimeraza I, II şi III, în timp ce la procariote există o singură ARN polimerază pentru sinteza tuturor speciilor de ARN.

cât şi benzi corespunzătoare densităţii „uşoare" a ADN de la moleculele fiice. Acest lucru nu s-a întâmplat, cercetătorii demonstrând că, din punct de vedere biochimic, **replicarea ADN este semiconservativă.**

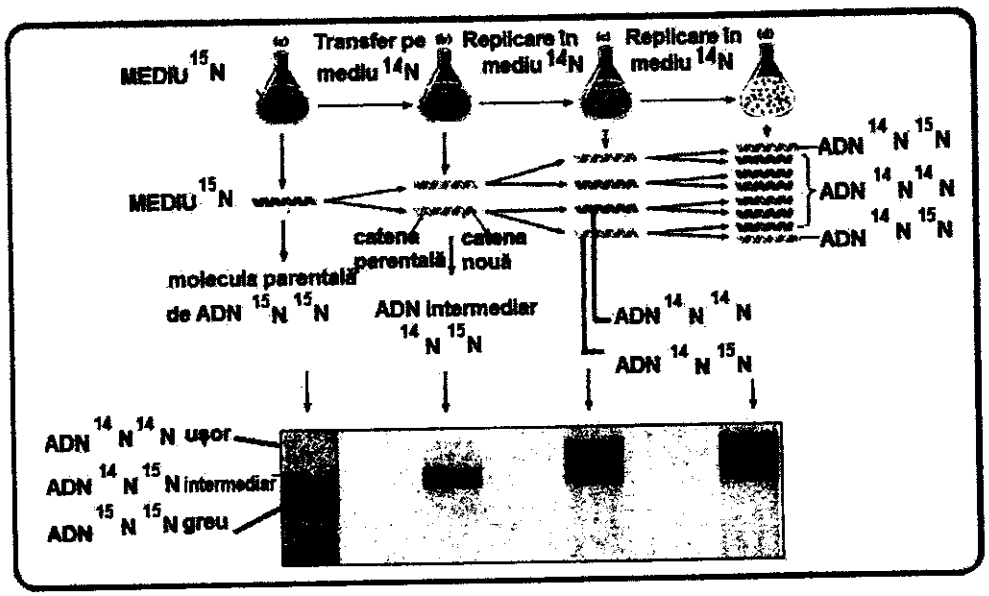


Fig.3.2 Experimentele lui Meselson şi Stahl. a) Celulele sunt crescute pe mediu cu ^{15}N-^{15}N. b) Celulele sunt trecute pe medii cu ^{14}N-^{14}N şi crescute mai multe generaţii (c,d). În partea inferioară a imaginii sunt redate profilele electroforetice obţinute după extracţia ADN din culturile celulare (a-d), în urma centrifugării în gradient de densitate cu CsCl2 (după Russel, P., 2010). În imagine se pot diferenţia densităţile diferite ale ADN corespunzător replicării semiconservative.

3.1 Mecanismul molecular al replicării ADN

3.1.1 Enzimele pentru replicarea ADN

Replicarea sau biosinteza ADN este procesul prin care se dublează cantitatea de material genetic din celulele somatice înaintea diviziunii mitotice. Procesul are loc în faza sintetică S a interfazei şi în acest fel se asigură continuitatea fizică a materialului genetic de la o generaţie la alta de celule. Ulterior, prin procesul de duplicare a materialului genetic care are loc în mitoză, se dublează şi numărul de cromozomi, asigurându-se astfel repartizarea egală a materialului genetic între cele două celule fiice rezultate.

Replicarea macromoleculei de ADN după modelul semiconservativ începe prin ruperea punţilor de hidrogen dintre cele două catene într-o

regiune specifică denumită ***originea replicării***, rezultând „ochiul de sinteză". Odată inițiată, replicarea ADN înaintează bidirecțional pe măsura desfacerii progresive a celor două catene, similar deschiderii unui fermoar. În acest fel ia naștere un model asemănător literei Y, denumit *furcă de replicare*. (Fig.3.3). Atât la cromozomii liniari, cât și la majoritatea celor circulari replicarea are loc bidirecțional, la nivelul a două furci de replicare, ambele inițiate în ochiul de sinteză. Există însă și situații în care sinteza are loc unidirecțional, așa cum este cazul **ADN mitocondrial cu structură circulară**.

Replicarea ADN are loc în trei etape: **inițierea replicării, elongația și terminarea replicării**, la care participă mai multe enzime ce intervin secvențial. După ordinea intrării în funcție acestea sunt: topoizomerazele, helicazele, proteinele de stabilizare a monocatenelor, primaza, ADN-polimerazele și ADN-ligazele.

Cea mai importantă enzimă de replicare este însă **ADN polimeraza** care catalizează biosinteza ADN prin formarea legăturilor fosfodiesterice între gruparea 3'-OH a ultimei nucleotide din secvența catenei fiice nou sintetizate și gruparea 5'- fosfat a precursorului nucleotidic liber din mediu, adăugat de enzimă în lanțul în creștere pe bază de complementaritate cu matrița.

În afară de aceste enzime, în replicare intervin și câteva proteine care acționează singure sau în complexe nucleoproteice cu ADN. Unele au rol de inițiator prin recunoașterea regiunii *originea replicării*, altele de a recruta alte enzimele implicate în replicare, altele de a stabiliza furca de replicare prevenind renaturarea ADN și altele de a glisa enzimele pe molecula de ADN.

Topoizomerazele de tip I și II sunt enzime care scindează reversibil una sau ambele catene de ADN, la nivelul legăturilor fosfodiesterice, ajutând la relaxarea helixului. Molecula de ADN este o structură plectonemială în care răsucirea celor două catene are loc spre dreapta (conformația B) și a cărei compactare prin asociere cu proteinele histonice, are loc prin pliere și superpliere.

În timpul înaintării furcilor de replicare molecula de ADN se superspiralizează pozitiv prin acțiunea helicazei și se tensionează tot mai puternic. În același timp, în fața furcii pot să apară și răsuciri negative la nivelul helixului. Pentru relaxarea moleculei în fața furcii, este nevoie de intervenția topoizomerazelor care să elimine răsucirile pozitive și negative, astfel încât ADN polimeraza să poată să înainteze.

Intervenţia topoizomerazelor este esenţială atât în procesul de replicare a ADN cât şi în transcripţia informaţiei genetice, deoarece în ambele situaţii intervine helicaza care rupe punţile de hidrogen pentru generarea monocatenelor, acţiune care tensionează molecula tot mai puternic.

La bacterii, topoizomeraza I poate înlătura doar suprarăsucirile negative, nu şi pe cele pozitive din helix. Suprarăsucirile pozitive sunt înlăturate de ADN giraza care, în acelaşi timp introduce în helix suprarăsuciri negative ce sunt înlăturate ulterior de topoizomeraza I.

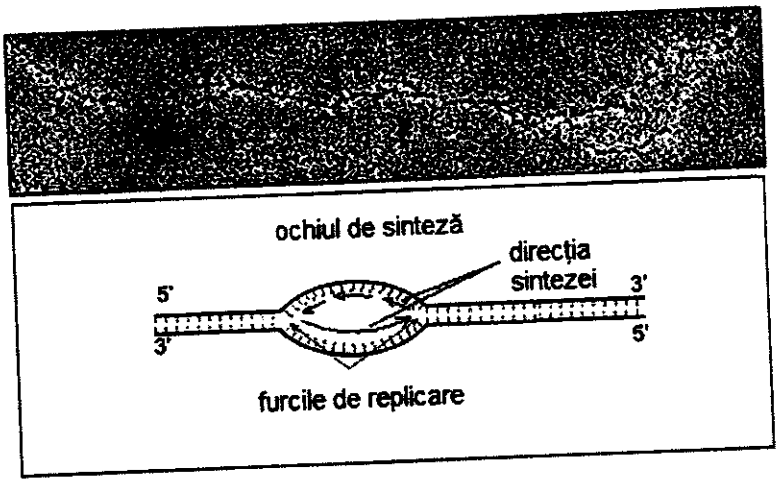


Fig.3.3 a. Imaginea electronomicroscopică a replicării ADN eucariot, în care sunt evidenţiate furcile de replicare la nivelul unor origini diferite.b) Producerea *ochiului de sinteză* şi a *furcii de replicare* ca urmare a ruperii punţilor de hidrogen şi iniţierea replicării

La eucariote topoizomeraza I poate înlătura din helix atât suprarăsucirile negative cât şi pe cele pozitive. Spre deosebire de giraza de la procariote, enzima nu introduce suprarăsuciri negative, deci intervine atât în relaxarea helixului cât şi în menţinerea noilor molecule de ADN, separate în timpul replicării (Fig.3.4).

Enzima topoizomeraza I de la eucariote recunoaşte regiunile tensionate şi acţionează printr-un mecanism de scindare – relaxare – refacere a ADN. Imediat după legare la ADN, enzima topoizomeraza I produce o tăietură pe una din catene şi simultan creează o legătură covalentă între gruparea 5' – fosfat liberă din catena de ADN clivată şi reziduul tirozină din structura proprie. Celălalt capăt 3'-OH produs în urma tăierii, trece printr-o stare liber – monocatenară de rotaţie, timp în care catena clivată suferă o singură dezrăsucire şi astfel este îndepărtată suprarăsucirea

pozitivă din helix. (Fig.3.4). Pe baza ultimelor cercetări, se pare că la eucariote acţionează cel puţin două topoizomeraze I (notate A şi B), diferenţa între acţiunile acestora fiind dată de legarea covalentă, fie de capătul 5', fie de cel 3' al ADN scindat.

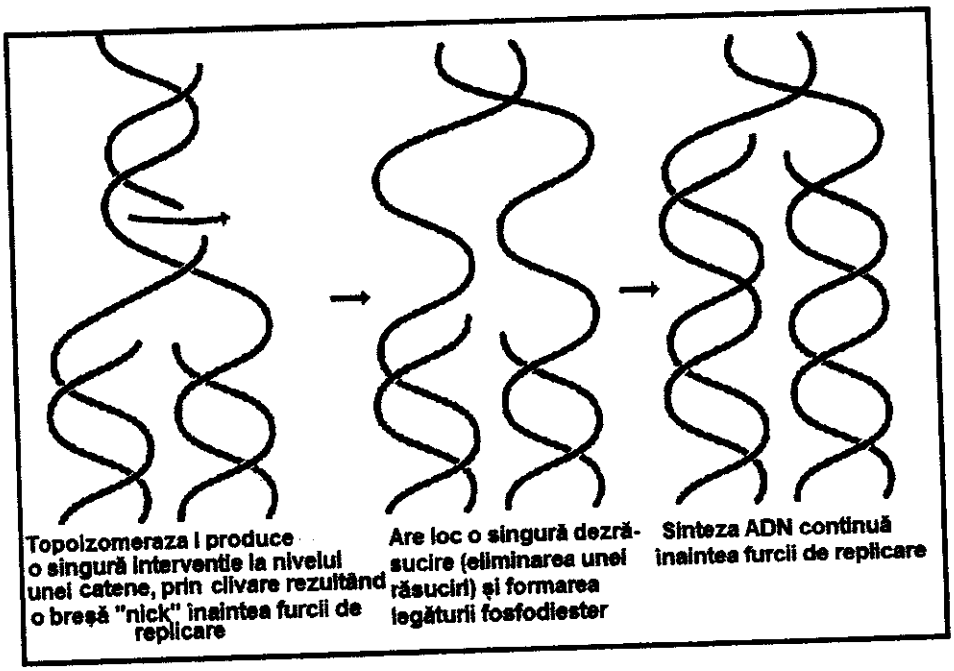


Fig.3.4. Acţiunea topoizomerazei I de la eucariote (sursa: Watson şi colab., 2003)

Imediat după relaxarea helixului, topoizomeraza I reuneşte capetele catenei de ADN, care astfel revine structural la starea iniţială iar enzima se eliberează de pe matriţă (Watson şi colab., 2003). Topoizomeraza II de la eucariote şi giraza la procariote scindează ambele catene de ADN, producând dezrăsucirea prin trecerea unui capăt al moleculei la nivelul tăieturii şi reformarea legăturilor fosfodiesterice (Fig.3.5).

În acţiunea topoizomerazei II un capăt scindat al moleculei de ADN se leagă covalent la enzimă, iar celălalt capăt execută mişcarea de rotaţie. În urma acestor acţiuni se produc două dezrăsuciri la nivel helixului.

În acţiunea topoizomerazelor prezenţa reziduului tirozină cu afinitate pentru nucleotide, constituie puntea legăturii covalente dintre nucleotidele capetelor scindate, permiţând astfel celuilalt capăt să se rotească liber şi să realizeze dezrăsucirea helixului, fie în sens pozitiv, fie negativ.

Dacă intervenţia topoizomerazei are loc pe o singură catenă se produce desfacerea unui pas al helixului (cazul topoizomerazelor tip I), în

timp ce clivarea ambelor catene va produce desfacerea helixului la nivelul a doi paşi din helix (cazul topoizomerazelor tip II).

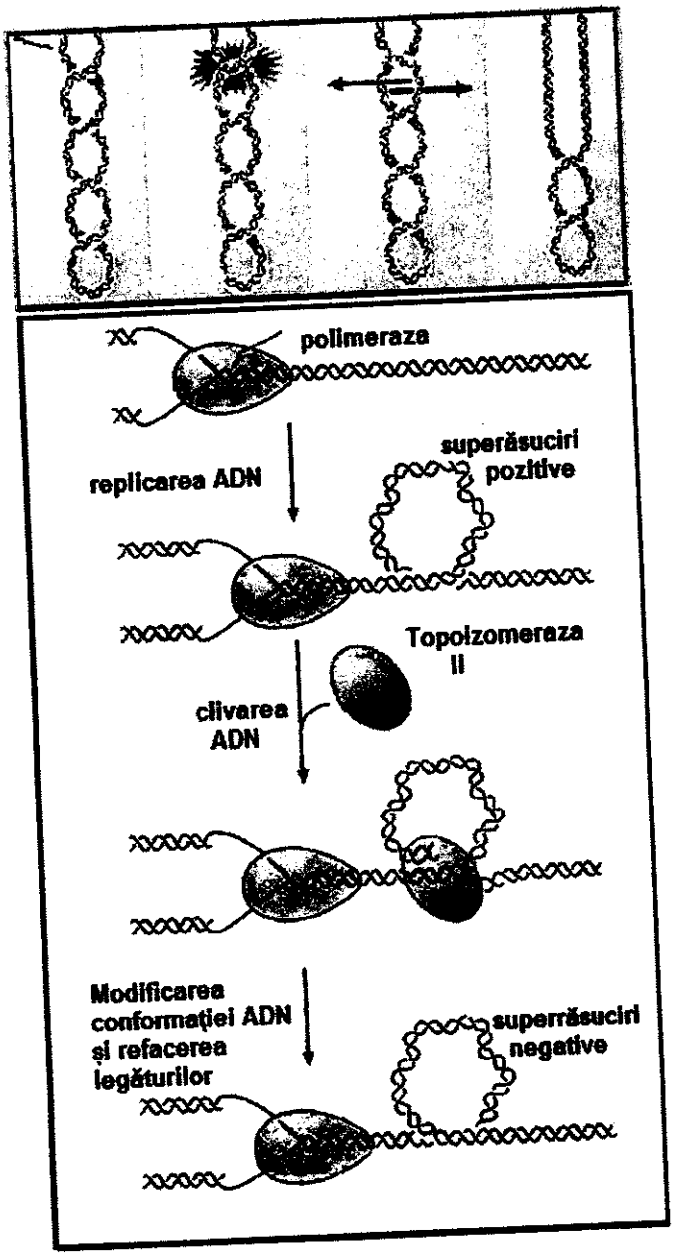


Fig.3.5. Acţiunea de eliminare a super-răsucirilor pozitive din helix a topizomerazei II prin producerea unei tăieturi tranzitorii la nivelul ambelor catene (după Watson şi colab., 2003)

ADN helicazele degradează punţile de hidrogen dintre cele două catene de ADN, asigurând desfacerea moleculei. Ruperea punţilor de hidrogen prin acţiunea helicazelor are loc atât în procesul de replicare, cât şi în cel de transcripţie, facilitând astfel intrarea enzimelor specifice în acţiune, la nivelul monocatenelor de ADN.

ADN polimerazele sunt enzime care catalizează biosinteza noilor catene de ADN adăugând nucleotidă după nucleotidă în direcţia 5'-3', pe bază de complementaritate cu catena matriţă, întotdeauna în continuarea unei secvenţe primer. Viteza cu care polimerazele fac sinteza ADN poate fi destul de ridicată, aceasta fiind estimată la aproximativ 1000 nucleotide/secundă. Numărul de nucleotide pe care îl poate adăuga polimeraza în timpul procesului de replicare se numeşte **capacitate de procesare**. Capacitatea este specifică fiecărei polimeraze şi poate varia foarte mult, de la câteva nucleotide până la 50.000 (Watson şi colab., 2003). În timpul sintezei noilor catene de ADN polimeraza poate împerechea greşit nucleotide în lanţul în creştere. Aceste greşeli sunt corectate de enzime care prezintă activitate de excizare şi înlocuire de nucleotide,denumită exonucleazică.

Generic, **nucleazele** sunt o categorie de enzime care degradează moleculele de ADN prin clivarea legăturilor fosfodiesterice dintre două nucleotide învecinate. Degradarea legăturilor poate avea loc în interiorul duplexului ADN (endonucleaze) sau la capetele acestuia (exonucleaze).

În replicarea ADN exonucleazele au rol corector al moleculei aflată în sinteză, iar acest rol este îndeplinit chiar de către una dintre subunităţile ADN polimerazei, denumită **exonucleză de corectare**.

La exonucleaze una din activităţi este în direcţia 3'-5' prin care sunt corectate greşelile de împerechere, iar cealaltă este în direcţia 5'-3' prin care se îndepărtează nucleotidele din primer, în aceeaşi direcţie în care are loc şi sinteza ADN. Activitatea exonucleazică 5'-3' este necesară la replicarea ADN pentru îndepărtarea primerilor de natură ribonucleotidică. Înlocuirea cu secvenţe dezoxiribonucleotidice, specifice ADN, se realizează prin sinteză cu ajutorul ADN polimerazei. Schematizat, cele trei activităţi posibile ale ADN polimerazei sunt prezentate în figura următoare (Fig.3.6).

Sinteza primerilor, secvenţe scurte de circa 5-10 ribonucleotide la procariote şi 8-12 ribonucleotide, plus alte 10-20 dezoxiribonucleotide la eucariote, este asigurată de o ARN polimerază modificată, denumită ADN **primaza**. Procesul de sinteză a primerilor este demarat înaintea replicării ADN, deoarece ADN polimeraza nu îşi poate exercita funcţia polimerazică fără o structură anterioară oligonucleotidică cu gruparea 3'-OH liberă, după

care să lege precursori nucleotidici. Excepţie de la această regulă face enzima telomeraza specializată în sinteza capetelor telomerelor, care conţine secvenţa primer chiar în structura proprie şi sintetizează prin reverstranscripţie capătul mai scurt al unei catene de ADN.

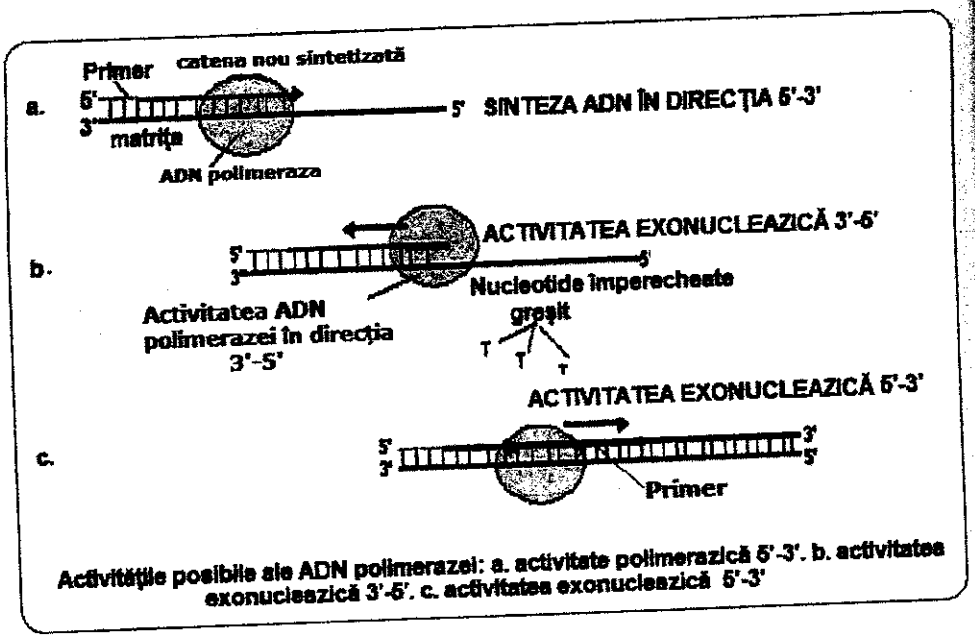
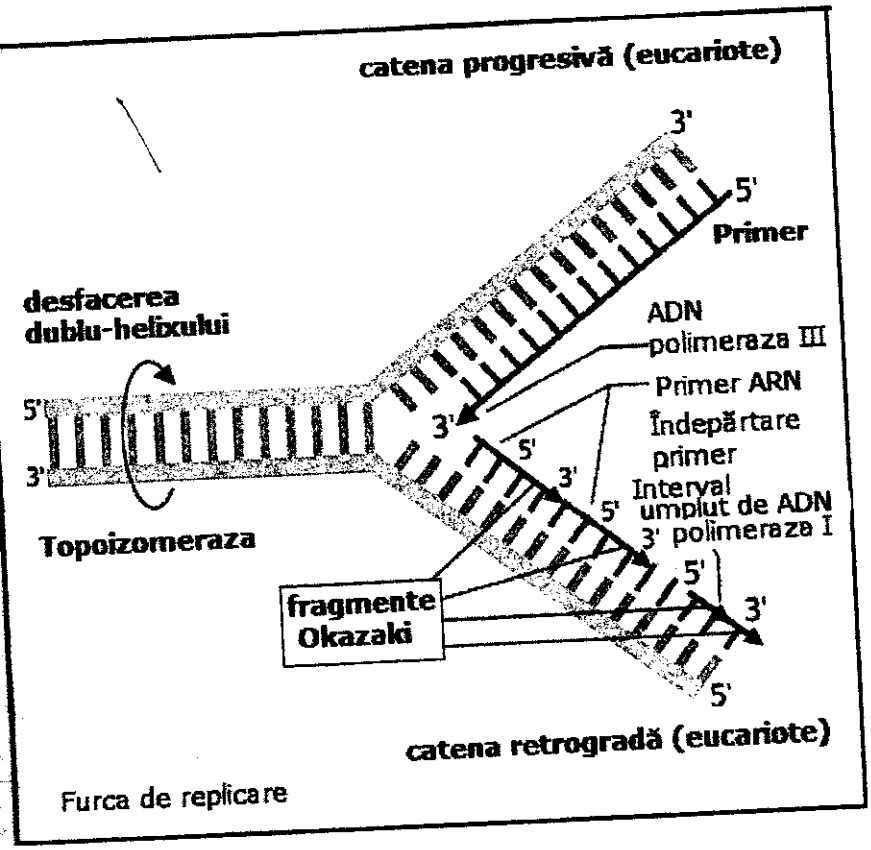


Fig.3.6 Activităţile ADN polimerazei: a) polimerazică 5'-3', b) exonucleazică 3'-5', c) exonucleazică 5'-3'. (Coşier, 2008)

Sinteza celor două catene la nivelul furcii de replicare se face diferit, una dintre catene se va sintetiza continuu, iar cealaltă discontinuu, pe fragmente scurte Okazaki. Din acest motiv modelul de replicare este denumit **modelul replicării semidiscontinue**. Cele două catene au denumiri diferite la procariote şi la eucariote. Catena 3'-5' care se sintetizează continuu la procariote se numeşte **conducătoare,** iar la eucariote **progresivă**, iar catena 5'-3' care se sintetizează discontinuu, pe fragmente Okazaki, este denumită convenţional **decalată** la procariote şi **retrogradă** la eucariote. (Fig.3.7). După terminarea sintezei are loc înlocuirea primerului ribonucleotidic cu o secvenţă specifică ADN, iar fragmentele vor fi unite covalent de ADN ligază pentru obţinerea unei secvenţe neîntrerupte.

ADN-ligazele sunt enzime care refac legăturile fosfodiesterice dintre două nucleotide învecinate, iar în cazul replicării ADN ligazele unesc fragmentele Okazaki de pe catena sintetizată discontinuu. Ele intervin la sfârşitul sintezei când primerii sunt deja înlocuiţi cu secvenţe dezoxiribonucleotidice, refăcând legăturile fosfodiesterice

dintre fragmente şi dând naştere astfel unei secvenţe ADN neîntrerupte.



3.6 Modelul sintezei semidiscontinue a ADN la nivelul furcii de replicare. ena 3'-5' este sintetizată continuu iar catena 5'-3' discontinuu, pe fragmente te Okazaki (sursa: Passarge, 2001)

3.2 Replicarea ADN la bacterii

3.2.1 Enzimele implicate în replicarea ADN la bacterii

Procariotele sunt entităţi unicelulare la care aparţin reprezentanţii upului Bacteria, alge unicelulare şi reprezentanţii Grupului Archea. upul Bacteria cuprinde unicelularele cele mai răspândite în toate mediile viaţă de pe glob.

Genomul nucleoidal al majorităţii bacteriilor este reprezentat în rea majoritate a cazurilor de un singur cromozom circular ce conţine o gură moleculă de ADN dublu-catenar circular. În citoplasma unor

3.3 Replicarea ADN la eucariote

3.3.1 Mecanismul molecular al replicării ADN la eucariote

Mecanismul de replicare a ADN la eucariote se remarcă prin gradul ridicat de similitudine cu cel al replicării cromozomului bacterian, fiind tot un proces **semidiscontinuu** şi **semiconservativ**.

Genomul nuclear eucariot este format din cromozomi cu structură liniară, fiecare conţinând o singură moleculă de ADN de lungime enormă. Fiecare moleculă de ADN este complexată cu proteine histonice şi compactată prin spiralizare, pliere şi superpliere.

Replicarea ADN cromozomial la eucariotelor are loc în interfaza mitozei, în perioada sintetică S, deşi unele procese încep chiar din perioada presintetică G1. Conform modelului semiconservativ de replicare, dintr-o moleculă parentală bicatenară vor rezulta două molecule fiice bicatenare, fiecare cu o catenă veche şi una nou sintetizată şi cu aceeaşi informaţie ereditară ca şi molecula parentală. După replicarea ADN, cromozomii monocromatidici devin bicromatidici.

Deoarece dimensiunea ADN de la eucariote este mult mai mare decât la procariote, replicarea este iniţiată în *mai multe origini ale replicării* simultan, care vor da naştere la mai multe *ochiuri de sinteză*. De exemplu omul are între 10.000 şi 100.000 origini ale replicării, cu alte cuvinte câte una la fiecare 300 kb (kilobaze), iar drojdia *Saccharomyces cerevisiae* are circa 300 origini, câte una la fiecare 40 kb (Kornberg, A şi Baker, T.A, 1992; Brown, T.A, 2002).

În linii mari procesul de replicare este similar din punct de vedere biochimic şi molecular cu cel bacterian, prin iniţierea replicării la nivelul originilor de replicare; prin intervenţia enzimelor şi proteinelor corespondente; prin sinteza primerilor care se deosebesc de cei de la procariote fiind secvenţe ARN-ADN; prin replicarea bidirecţională şi semidiscontinuă a ADN; înlăturarea secvenţelor primer şi umplerea intervalelor cu dezoxiribonucleotide, iar la final, prin legarea covalentă a fragmentelor Okazaki cu ADN ligaza.

Deoarece cromozomii de la eucariote sunt liniari, replicarea capetelor la nivelul ADN telomeric este particulară, prin intervenţia unei enzime specializate, telomeraza.

Fiecare unitate în care este iniţiată sinteza pe cromozomii de la eucariote se numeşte **replicon** sau **unitate de replicare.** La nivelul fiecărui replicon iau naştere furcile de replicare care avansează în ambele direcţii, pe

distanţe mult mai scurte decât la procariote la care există o singură unitate de replicare.

La procariote viteza de sinteză a noilor catene de ADN este de circa 1000 pb/s, iar întregul cromozom (de 4,6 x 10^6 pb) este replicat în circa 40 de minute. La eucariote viteza de sinteză este mult mai mică, de circa 100pb/s. Astfel la om, întregul genom nuclear (care conţine 3,2 x 10^9pb) este replicat în 8 ore, iar la alte specii replicarea poate dura o lună de zile (Kornberg A şi Baker, T.A, 1992; Russell, P., 2010).

Desfăşurarea procesului replicativ la nivelul repliconilor parcurge tot trei etape ca la procariote, acestea fiind: iniţierea replicării, elongaţia şi terminarea replicării.

Iniţierea replicării la nivelul cromozomilor eucarioţi are loc simultan în mai multe *origini ale replicării*. **Proteinele iniţiator** la eucariote aparţin complexului de recunoaştere a originii - **ORC** (*Origin Recognition Complex* - ORC). Complexul ORC este o structură formată din nouă subunităţi:

- trei dintre subunităţi (orcp1, orcp4 şi orcp5) îndeplinesc rolul proteinei iniţiator ADN-A de la procariote, de recunoaştere şi legare la regiunea originea replicării;

- o subunitate de legare la origini care se ataşează în maniera situs - specific şi împreună cu altă proteină (cdc6p) procesează originea replicării şi recrutează helicaza eucariotă MCM (Watson şi colab., 2003).

Ciclul celular al unei celule eucariote cuprinde două stadii:

Interfaza cu faza **presintetică G1** în care celula se pregăteşte pentru replicarea ADN; **S-sintetică**, în care are loc replicarea propriu-zisă şi - **postsintetică G2**, în care celula se pregăteşte pentru diviziunea mitotică şi cel de-al doilea stadiu al ciclului celular este diviziunea propriu-zisă a celulei sau **mitoza**.

Pentru duplicarea corectă a materialului genetic în metafaza mitozei, replicarea din faza S trebuie să decurgă precis, astfel că fiecare origine a replicării este activată o singură dată în ciclul celular.

Primul pas în iniţiere este recunoaşterea simultană a tuturor originilor de replicare, proces care se desfăşoară în faza G1 din interfază prin intermediul complexului proteic ORC. Imediat după legarea complexului ORC la originile replicării sunt recrutate alte proteine care formează **complexele pre-replicative (pre-RC)**.

În contrast cu procesul de replicare bacterian, denaturarea nu are loc imediat, ci este amânată până în momentul în care celula trece în faza S când acest complex pre-replicativ este activat (Russell, P., 2010). Amânarea inițierii replicării până în faza S este controlată de anumite proteine (***licensing factors***). Sinteza lor are loc tot în faza G1, când proteinele trec din citoplasmă unde au fost sintetizate, în nucleu și se leagă la complexul pre-replicativ (pre-RC). În faza sintetică S, complexului pre-replicativ i se alătură alte proteine (helicaza MCM, clema glisantă), care vor iniția denaturarea ADN și vor da startul replicării. În momentul inițierii replicării proteinele *licensing factors* sunt inactivate astfel încât inițierea replicării are loc simultan, la nivelul tuturor repliconilor.

Elongația este procesul de sinteză a noilor catene de ADN care se realizează în manieră asemănătoare cu cel de la procariote. Și la eucariote sinteza poate fi inițiată numai după o secvență primer care de această dată are structură mixtă de tip ARN-ADN. De asemenea, cele două catene se vor sintetiza diferit și cu enzime diferite: cea progresivă cu direcția 3'-5' se va sintetiza continuu, iar cea retrogradă cu direcția 5'-3', discontinuu pe fragmente Okazaki.

3.3.2 Enzimele implicate în replicarea ADN la eucariote

Numărul de enzime implicate în replicarea ADN la eucariote este mult mai mare decât la procariote, iar mecanismele complete de replicare sunt mai puțin cunoscute.

Până în prezent, se cunosc mai mult de 15 ADN polimeraze implicate în replicarea ADN la eucariote, ale căror funcții însă sunt incomplet înțelese. În mod curent în replicarea ADN nuclear intervin trei ADN polimeraze:

- **ADN polimeraza α/primaza** (Pol α/primaza)
- **ADN polimeraza δ** (Pol δ/delta)
- **ADN polimeraza ε** (Pol ε/epsilon)

La eucariote, sinteza celor două catene (progresivă și retrogradă) este asigurată de enzime diferite spre deosebire de procariote la care dimerii miez ai ADN polimerazei III catalizează replicarea întregului cromozom bacterian.

ADN polimeraza ε realizează sinteza continuă a catenei progresive cu direcția 3'-5', imediat după o secvență primer sintetizat de enzima Pol α/primaza.

ADN polimeraza δ realizează sinteza discontinuă a catenei retrograde cu direcţia 5'-3', pe fragmente Okazaki. Pe lângă enzimele Pol α/primaza şi Pol δ, în replicarea catenei retrograde la eucariote intervin mai multe enzime şi proteine. Clema glisantă a ADN sau complexul proteic PCNA (*Proliferating Cellular Nuclear Antigen*), factorul de replicare C (RFC), complexul proteic de replicare A (*Replicating Protein A*) din categoria complexelor proteice de stabilizare a monocatenelor SSB, nucleazele FEN 1 şi RN-aza H precum şi ADN ligaza.

Enzima **ADN polimeraza α/primaza** are rolul de a sintetiza primerii ARN-ADN cu direcţia 5'-3', formaţi din 8-12 ribonucleotide după care enzima adăuga la capătul 3' şi o secvenţă scurtă dezoxiribonucleotidică formată din 10-20 nucleotide. Secvenţa primer ARN-ADN este recunoscută de **Factorul de replicare C** (RFC) care se ataşează complexului şi eliberează Pol α/primaza. După disocierea enzimei este recrutat complexul proteic PCNA (*Proliferating Cellular Nuclear Antigen*) sau clema glisantă, proces care are ca rezultat legarea enzimei polimeraza δ după secvenţa primer, unde va fi iniţiată sinteza catenei decalate (retrograde) şi fixarea enzimelor în timpul înaintării pe molecula de ADN.

Înlăturarea primerului şi înlocuirea lui cu o secvenţă dezoxiribonucleotidică se realizează diferit la eucariote faţă de procariote. La procariote, acest proces este realizat integral de ADN polimeraza I prin exercitarea funcţiilor exonucleazice 5'-3 şi polimerazice 5'-3.

ADN polimeraza δ realizează sinteza discontinuă pe fragmente Okazaki a catenei decalate până întâlneşte următorul fragment Okazaki. În acest moment, polimeraza iniţiază dislocarea capătului 5' al fragmentului Okazaki producând o regiune monocatenară sub forma unei clapete. Digerarea primerului ribonucleotidic este asigurată de două nucleaze FEN1 şi RN-aza H (Russell, P., 2010).

La eucariote, există mai multe mecanisme de îndepărtare şi reparare a ADN. Într-unul dintre acestea se pare că **Pol δ**, lucrează în cooperare cu **endonucleaza Fen1** pentru îndepărtarea şi înlocuirea primerului. Endonucleaza îndepărtează câte o ribonucleotidă prin mecanismul *nick translation* (de tăiere), iar polimeraza adaugă dNTP- ul corespunzător. Imediat ADN ligaza reface legătura fosfodiester (Watson şi colab., 2003).

Terminarea replicării. Cromozomii eucariotelor au structură liniară iar la capete sunt prezente secvenţele ADN telomeric. Structura telomerelor este formată din unităţi SSR hexanucleotidice (3-TTAGGG-5), repetate în tandem de 100-1000 de ori. ADN-ul telomeric nu este dublu-catenar pe toată lungimea. La sfârşitul cromozomului ADN-ul

telomeric este monocatenar şi formează o buclă, bucla-T care pătrunde în secvenţa dublu-catenară formând o buclă de înlocuire – bucla D. La capătul cromozomului înlocuirea ultimului primer ribonucleotidic de pe catena nou sintetizată nu are loc, deoarece înaintea lui lipseşte o secvenţă ADN. Cum acest primer este la capătul cromozomilor, degradarea lui înseamnă o scurtare a cromozomului în fiecare ciclu. Din aceste motive procesul de replicare a telomerelor se desfăşoară diferit.

Telomeraza este o enzimă cu funcţie de transcriptază inversă care este activă doar în anumite perioade ale vieţii celulei (celule germinale) şi în unele celule proliferative (celule stem, canceroase, etc.). Telomeraza este formată din două subunităţi: una cu rol structural de primer şi cealaltă cu rol enzimatic de transcriptază inversă. Telomeraza se ataşează la capătul mai lung al unei catene din ADN prin secvenţa primer din structura proprie şi copiază propria secvenţă prin reverstranscripţie cu cealaltă subunitate din structură, cu funcţie enzimatică de reverstranscriptază. În continuare, ADN polimeraza completează catena mai scurtă cu secvenţe dezoxiribonucleotidice în direcţia 5'-3'. (Fig.3.16).

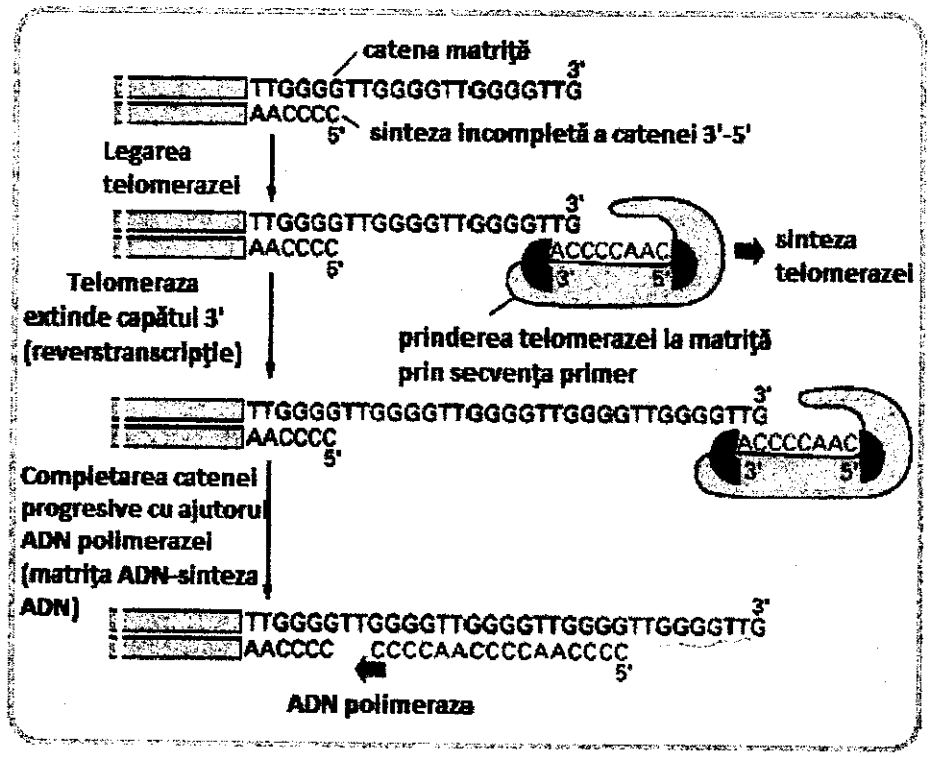


Fig.3.16 Replicarea capetelor cromozomilor cu ajutorul telomerazei şi a ADN polimerazei

Terminatorii *Rho independenţi* sau intrinseci nu interacţionează cu ARN polimeraza decât după transcrierea secvenţelor din structura proprie. În urma transcrierii informaţiei din cadrul acestor structuri, cele două secvenţe cu structură palindromică formează o buclă *ac de păr* care întrerupe comunicarea cu complexul transcripţional. Eliberarea definitivă a polimerazei de la matriţă are loc doar atunci când după bucla *ac de păr* se sintetizează secvenţa poli-U, care rămâne foarte slab ataşată de secvenţa formată din şase adenine din matriţă (Fig.4.12).

4.5 Structura genelor la eucariote şi controlul expresiei genice

Genomul eucariot este reprezentat în totalitatea cromozomilor cu structură liniară din nucleu. Genomul vegetal şi animal citoplasmatic este reprezentat de molecule circulare de ADN, prezente în mitocondrii (ADNmt) şi în plastide (ADNcl) în cazul celulelor vegetale.

Dimensiunea cu mult superioară a genomului eucariot de la mamifere şi de la plante şi asocierea ADN cu proteinele histonice, determină un tip particular de organizare, mult mai complex comparativ cu cel procariot. Cu toate acestea doar o mică parte din genom este acoperită de gene structurale care codifică proteinele, aşadar complexitatea genomului nu este corelată cu dimensiunea. Aşa de pildă la om, la care genomul este complet descifrat, mărimea este de $3,2 \times 10^9$ (în genom haploid), din care secvenţele codificatoare reprezentate de exonii din cele 25.000 de gene acoperă aproximativ 2% din genom. Acest procent reprezintă exomul celulei care este transcris în molecule de ARN.

Totuşi, practic proporţia ocupată de genele structurale în genom este mai mare, deoarece în interior genele eucariote conţin intronii, secvenţe necodante din genă care sunt excizate din transcript în procesul de prelucrare a ARNm precursor.

Intercalat printre secvenţele necodante se găsesc exonii, secvenţele codificatoare din genă care se transcriu şi se traduc. Din acest motiv, spunem despre genele eucariote că au structura discontinuă sau mozaicată. Secvenţe necodificatoare (introni) ocupă circa 26% din genomul uman, cu alte cuvinte cea mai mare parte din genele de la eucariote este necodantă însă această informaţie genetică este eliminată. (Fig.4.13).

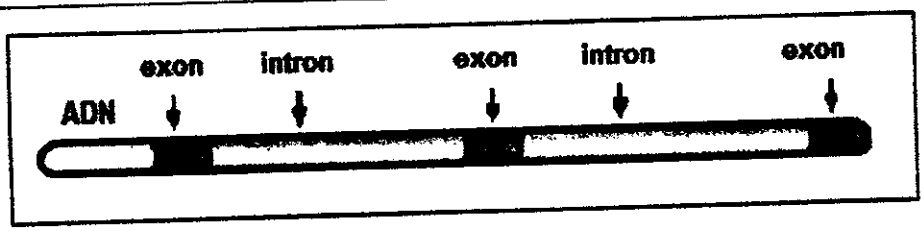


Fig.4.13 Reprezentarea schematizată a structurii discontinue caracteristică genelor eucariote

Expresia genelor eucariote este un proces mult mai complex decât cel întâlnit la procariote datorită complexării ADN cu proteinele histonice, a organizării cromozomilor în structuri liniare de dimensiuni mari, localizate în nucleul individualizat, astfel încât procesele care au loc pentru decodificarea informației genetice necesită intervenția mai multor molecule în evenimente desfășurate separat în timp și spațiu. Cu alte cuvinte, decodificarea informației genetice din genele structurale eucariote începe în nucleu unde are loc transcripția și prelucrarea ARNm și se definitivează în citoplasmă, la ribozomi, unde are loc procesul de translație și sinteza proteinelor. La procariote, lipsa membranei nucleare nu delimitează procesele, ci dimpotrivă acestea au loc nu în același compartiment și apoximativ consecutiv, înaintea finalizării transcripției începe și translația informației genetice.

4.5.1 Structura genelor eucariote

Gena structurală de tip eucariot este un segment de ADN care specifică un produs polipeptidic (proteină, enzimă sau hormon). În afara genelor structurale, genomul eucariot conține și alte secvențe genice care nu se transcriu și/sau nu se traduc întotdeauna. Genele pentru sinteza speciilor de ARN din celulă sunt astfel de exemple care deși se transcriu nu sunt decodificate în produse polipeptidice. Alte categorii de gene cum sunt și pseudogenele, fie nici nu se transcriu sau dacă se transcriu, nu sunt decodificate într-un produs funcțional. Ele sunt copii genice truncate inserate în genom, care au devenit nefuncționale din cauza mutațiilor sau în unele cazuri, au rămas parțial funcționale.

Spre deosebire de procariote, la care sunt decodificate simultan mai multe gene adiacente organizate sub forma unui operon, decodificarea informației genetice într-un eveniment transcripțional la eucariote are loc de pe o singură genă. Prin urmare ARNm eucariot este monocistronic iar cel procariot policistronic.

Transcripţia genelor structurale eucariote se realizează cu enzima ARN polimeraza II, în timp ce transcripţiile genelor pentru ARNr, ARNt şi a speciilor mici de ARN, sunt realizate de celelalte două polimeraze, I şi III. La procariote dimpotrivă, există o singură ARN polimerază care transcrie toate genele din celulă.

Expresia unei gene la eucariote este un proces similar ca parcurs cu cel de la procariote, adică se desfăşoară în două etape succesive: transcripţia şi translaţia informaţiei genetice. Procesele moleculare în mare măsură sunt asemănătoare însă enzimele care intervin sunt mai numeroase la eucariote, fiecare îndeplinind rol specific într-o anumită etapă din proces (vezi scap. Enzimologia transcripţiei).

ARNm procariot este molecula ce intră în procesul de translaţie pe măsură ce este sintetizat, fără modificări suplimentare. ARNm eucariot dimpotrivă, este o moleculă care este mai întâi prelucrată în nucleu, sub formă de ARNm matur, care apoi trece în citoplasmă pentru decodificare la ribozomi. Procesul de prelucrare are loc în etape, unele încep aproape concomitent cu transcripţia iar altele după finalizarea ei. În prima etapă sunt modificate capetele moleculei de ARNm precursor (primar) cu formarea unui ARNm intermediar, din care vor fi excizaţi intronii pentru a se forma molecula de ARNm matur (Fig.4.14). Aceasta din urmă este molecula care intră efectiv în procesul de translaţie, părăsind mai întâi nucleul, pentru a ajunge în citoplasmă la ribozomii, locul sintezei proteice.

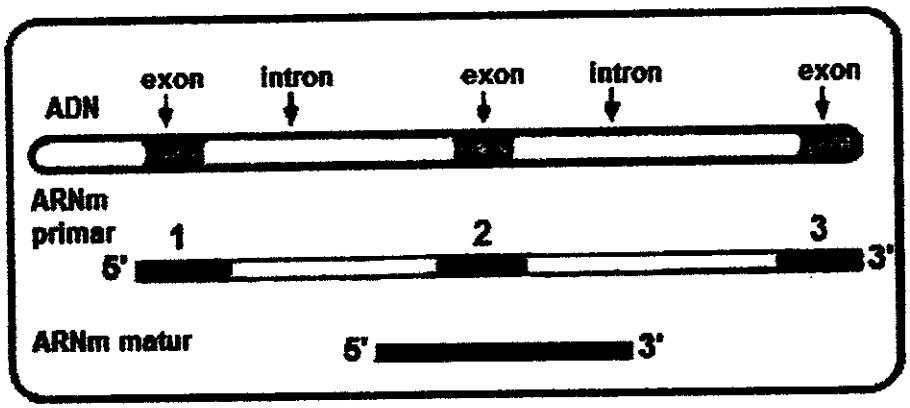


Fig.4.14 Formarea ARNm matur în urma excizării intronilor din transcript (adaptată după Watson şi colab., 2003)

Structural, gena de tip eucariot este alcătuită din trei regiuni importante: ***regiunile cu rol reglator din amonte, segmentul structural şi regiunea cu rol reglator din aval de genă.***

Regiunile de reglaj din vecinătatea genei au rolul de a semnaliza enzima ARN polimeraza II cu privire la locul de legare şi la cel de eliberare de la matriţa ADN. O parte din aceste secvenţe se transcriu în ARNm dar nu se traduc în secvenţe polipeptidice. Din aceste motive au primit denumirile de regiunile 5' UTR, respectiv 3' UTR (UnTranslated Region).

Unele secvenţe din amonte de genă sunt situate în proximitatea startului transcripţiei şi în poziţie *cis*, în timp ce alte secvenţe care participă în proces ca reglatori ai transcripţiei sunt situate în poziţie trans şi uneori la distanţe mari faţă de genă.

Regiunea reglatoare din amonte de genă care se transcrie în ARNm este lipsită de secvenţele de legare la ribozom aşa cum se întâmplă în cazul genelor procariote. Această funcţie va fi asigurată prin modificarea capătului 5' al ARNm. Imediat după această regiune urmează segmentul structural care va fi transcris şi translatat în produsul polipeptidic, urmat de regiunea de reglaj din aval de genă cu rol de terminare a transcripţiei prin secvenţele poli-A şi nu de un terminator aşa cum întâlnim la procariote.

4.5.1.1 Regiunea cu rol de reglaj din amonte de genă

Atât la procariote, cât şi la eucariote, elementele de reglaj din amonte de genă care semnalizează enzimei ARN polimeraza iniţierea transcripţiei, sunt situate în secvenţa promotor.

Elementele din promotorii de la eucariote au fost analizaţi prin procedee mutaţionale pentru a se vedea care regiuni îndeplinesc rol efectiv în proces. Mutaţiile care afectează efectiv transcripţia aparţin regiunii din genă situată cu aproximativ 200 pb în amonte de startul transcripţiei. În această zonă se găsesc două regiuni importante: (1) **miezul promotorului** şi (2) **elementele proximale promotorului**.

Miezul promotor cuprinde un set de elemente ce acţionează în poziţia *cis*, necesare pentru ca ARN polimeraza II să ia startul transcrierii informaţiei genetice din poziţie corectă. Elementele *cis* reglatoare afectează doar procesul transcripţional al genei învecinate. Miezul promotor ocupă regiunea situată cu circa 40 nucleotide în amonte de punctul start +1 şi conţine mai multe secvenţe *consensus* care au rolul de a semnaliza ARN polimerazei II locul de legare la ADN, prin intermediul unor factori proteici denumiţi generic **factori generali de transcripţie GTF** (***General Transcription Factors***). Aceştia vor forma împreună cu ARN polimeraza II **complexul de preiniţiere**, similar cu cel format de miezul ARN polimerazei bacteriene şi factorul σ.

degradării enzimatice, prin urmare cozile poli-A protejează transcriptul de acţiunea endonucleazelor, pe lângă multe alte roluri pe care le îndeplinesc în transcripţie dar şi în translaţie. Toate acestea vor fi discutate în secţiunea *Expresia genelor eucariote.*

Prin urmare, transcriptele stabile ale genelor structurale sunt doar acele molecule de ARNm care prezintă la capătul 3' secvenţele poli-A şi care au capătul 5' metilat.

4.6 Transcripţia la eucariote

Transcripţia unei gene eucariote este un proces complex datorită asocierii dintre ADN cu proteinele histonice în nucleozomi. Spre deosebire de procariote, complexarea ADN cu proteinele histonice determină intervenţia mai multor factori de remodelare a cromatinei şi a altora care vor facilita accesul enzimelor la matriţa ADN. În timp ce replicarea este un proces care se realizează simultan pentru tot genomul nuclear, transcripţia este un proces care are loc numai în anumite perioade, nu implică toate genele simultan şi are loc diferit în funcţie de necesităţile celulei.

Procesul de transcripţie are loc în trei etape, similar cu cel întâlnit la procariote, acestea fiind iniţierea transcripţiei, elongarea şi terminarea transcripţiei.

Pentru ca enzima ARN polimeraza II, ce realizează transcripţia tuturor genelor structurale eucariote, să aibă acces la matriţa ADN, în proces este ajutată de mai mulţi factori proteici (GTF) cu care interacţionează formând *complexul de preiniţiere*. Complexul la rândul lui interacţionează cu alţi factori proteici denumiţi *mediatori.*

Semnalele pentru legarea factorilor proteici în vederea iniţierii transcripţiei sunt declanşate de anumite regiuni cu rol de reglaj din genă, situate în amonte şi în aval de promotor, mai precis ***miezul promotor*** şi ***elementele din proximitatea promotorului.***

În cadrul procesului de iniţiere, ARN polimeraza se asociază cu factorii proteici de transcripţie care recunosc secvenţele cu rol de reglaj din genă, la care se leagă şi semnalizează enzima să se asocieze într-un complex la nivelul promotorului, locul iniţierii procesului.

În prima fază se formează **complexul de preiniţiere închis** în care enzima este asociată factorilor TFB, legaţi stabil de secvenţa promotor. Ceilalţi factori sunt adăugaţi secvenţial şi atunci când în complex intră şi factorul TFHII cu funcţie de helicază, are loc denaturarea ADN în regiunea

de inițiere a transcripției. **Complexul de inițiere** devine astfel **deschis** și ARN polimeraza II începe sinteza unor secvențe scurte din transcript care mai întâi sunt descărcate de la matriță. Abia după ce enzima a sintetizat secvențe mai lungi de 10 nucleotide procesul de transcripție devine continuu și are loc trecerea la faza de elongare în care GTF sunt descărcați din complex, iar la ARN polimeraza II se atașează la factorii de elongare.

Copierea informației de pe o catenă matriță de ADN, se realizează întotdeauna în direcția 5'-3', direcția de activitate a ARN polimerazei. În funcție de orientarea promotorului poate fi copiată oricare din cele două catene cu direcția 3'-5'.

Practic, transcriptul eucariot primar începe cu regiunea 5'- UTR și se termină cu regiunea 3' – UTR. În cadrul procesului este transcrisă toată informația din secvența structurală, începând cu codonul START până la codonul STOP, inclusiv secvențele necodante din introni. Terminarea transcripției are loc când ARN polimeraza întâlnește secvențele *consensus* transcrise –AAUAAA- și poli-U(U/GU) de după codonul STOP, care sunt semnale pentru clivajul situs-specific în vecinătatea regiunii 3'- UTR și secvențele pentru sinteza poliadenilatului (4.17b). Aceste modificări fac parte din procesul de prelucrare a ARNm precursor și nu din procesul de transcripție. Ele se desfășoară în paralel cu procesul de transcripție și sunt semnale pentru ARN polimeraza II să se elibereze de la matrița ADN și să termine transcripția (Pierce, 2008).

Așa cum am arătat, molecula de ARNm precursor este supusă prelucrării la capete, proces prin care rezultă un ARNm intermediar. După sinteza primelor 20-40 nucleotide din transcript, este inițiată deja prelucrarea capătului 5' a ARNm precursor prin formarea unei structuri „cap" de 7- metilguanozină. Capătul 3' al ARNm intră și el în procesul de prelucrare, imediat după terminarea copierii informației din genă, ce constă în sinteza unor cozi poliadenilate. Transcriptul care prezintă la capătul 5'- 7 metilguanozină și cozile poli-A la capătul 3' este un transcript intermediar care va continua procesul de prelucrare. În lipsa poliadenilatului la molecule va fi inițiat un proces de degradare enzimatică a ARNm, când are loc încheierea prematură a expresiei genei.

Ultima etapă în prelucrarea transcriptului constă în excizarea intronilor și asamblarea exonilor sub forma unui ARNm matur. Procesul are loc cu participarea moleculelor de ARNsn (*small nuclear*) cu funcție de enzimă. După prelucrarea finală, ARNm matur părăsește nucleul prin porii membranei nucleare, ajunge în citoplasmă unde va fi recrutat la ribozomi pentru decodificarea informației genetice (translația).

Iniţierea transcripţiei are loc la nivelul miezului promotor, cu participarea enzimei ARN polimeraza II. *Miezul promotor* intervine în iniţierea transcripţiei la eucariote prin elementele constitutive (BRE, TATA, Inr şi DPE), care sunt semnalele pentru legarea factorilor generali de transcripţie GTF. Aceste elemente din ADN, cerute de enzima ARN polimeraza II pentru poziţionarea corectă a startului transcripţiei şi pentru legarea efectivă a enzimei la ADN, sunt iniţial recunoscute de factorii generali de transcripţie GTF. Aceştia semnalizează ARN polimeraza II. În marea majoritate a cazurilor genelor eucariote sunt necesare cel puţin două, trei sau chiar patru elemente pentru legarea efectivă şi activarea ARN polimerazei. Acestea sunt elementele BRE, cutia TATA, iniţiatorul (Inr) şi elementele din aval de promotor (DPE). De aceste secvenţe *consensus* se leagă mai întâi factorii generali de transcripţie (GTF) în ordinea TBP (TATA binding proteins), TFIIB şi TFIID şi apoi ARN polimeraza II, cu care formează **complexul de preiniţiere**.

Deci formarea complexului de preiniţiere are loc la nivelul miezului **promotor,** situat cu circa 40 pb înaintea startului transcripţiei, prin recunoaşterea factorilor de legare la cutia TATA – TBF, care de fapt sunt subunităţi ale complexului TFIID. Imediat după legarea TBF la cutia TATA promotorul suferă modificări conformaţionale care vor iniţia formarea unei platforme de recrutare şi a altori factori generali necesari în transcripţie, precum şi a polimerazei însăşi care astfel se leagă efectiv la promotor.

Ceilalţi factori generali de transcripţie recrutaţi în complex, sunt în ordine: TFIIA, TFIIB, TFIIF (de care se vor lega în continuare cei din complexul de mediere) şi enzima ARN polimeraza. În această stare, **complexul de preiniţiere este închis,** iar pentru deschidere şi iniţierea transcripţiei, complexul adiţionează ceilalţi doi factori proteici TFIIE şi TFIIH, situaţi în amonte de ARN polimeraza II. Spre deosebire de procariote denaturarea ADN eucariot în regiunea promotor se realizează cu consum energetic, prin participarea precursorilor ATP şi a factorului TFIIH, care îndeplineşte funcţia de helicază. În această structură se formează complexul de iniţiere care cuprinde toţi factorii generali de transcripţie GTF, de care se leagă cei ai complexului de mediere şi ARN polimeraza II. În momentul intrării în acţiune a factorilor TFIIH din complex care denaturează local regiunea, se trece de la forma închisă la **forma deschisă a complexului** şi la iniţierea transcripţiei (Fig.4.18).

ARN polimeraza II formează împreună cu factorii generali de transcripţie complexul de preiniţiere, similar factorului σ de la procariote. Deşi ARN polimeraza nu necesită o secvenţă nucleotidică precursor aşa

cum cere ADN polimeraza în replicare, totuşi iniţierea transcripţiei are loc numai cu holoenzima care are ataşaţi factorii generali de transcripţie. Denaturarea locală a regiunii pentru formarea complexului deschis corespunde iniţierii efective a transcripţiei. Procesul de transcripţie constă în sinteza enzimatică a moleculei de ARNm, pe bază de complementaritate cu matriţa ADN.

Aşa cum se întâmplă şi la procariote, mai întâi are loc o fază de iniţiere abortivă în care sunt sintetizate de mai multe ori secvenţe scurte de circa 10 ribonucleotide până când ARN polimeraza „scapă" de regiunea promotor şi trece în faza de elongare efectivă. Factorii de transcripţie TBF sunt cei care ajută ARN polimeraza II să treacă de promotor şi să iniţieze elongarea transcriptului. Elongarea se realizează numai cu miezul enzimei, ceea ce presupune detaşarea factorilor generali de transcripţie din complex şi ataşarea altor factori proteici. Detaşarea miezului ARN polimerazei II de GTF are loc prin fosforilarea ARN polimerazei, proces diferit faţă de procariote. Coada polimerazei cu secvenţele heptapeptidice repetitive este cea care va suferi procesul de fosforilare, în fapt fiecare unitate repetitivă va intra în acest proces. Aceste reziduuri fosforice vor ajuta ARN polimeraza II să se elibereze de factorii generali de iniţiere a transcripţiei (GTF) şi să treacă singură la faza de elongare.

Pentru decondensarea cromatinei şi eliberarea histonelor cu care este complexat ADN-ul eucariot, ca ARN polimeraza să aibă acces la genă este necesar să intervină şi alţi factori în proces. În acest scop sunt recrutaţi ***factorii de remodelare a cromatinei şi activatorii,*** molecule de reglaj ale transcripţiei care facilitează procesul.

Prin urmare, în iniţierea transcripţiei intră în acţiune atât complexul de iniţiere a transcripţiei (GTF), cât şi *complexul de mediere,* cu care interacţionează şi elementele din proximitatea promotorului, cum sunt activatorii. Proteinele reglatoare ale iniţierii transcripţiei, denumite activatori, recrutează polimeraza la promotor stabilizând legarea la această secvenţă. Recrutarea se realizează prin intermediul proteinelor reglatoare (mediatorii), legate atât la ADN cât şi la factorii generali de transcripţie (Fig. 4.19). Deoarece la complexul de iniţiere este legat complexul de mediere, cel mai frecvent interacţiunea activatorilor cu ARN polimeraza se realizează prin intermediul complexului de mediere.

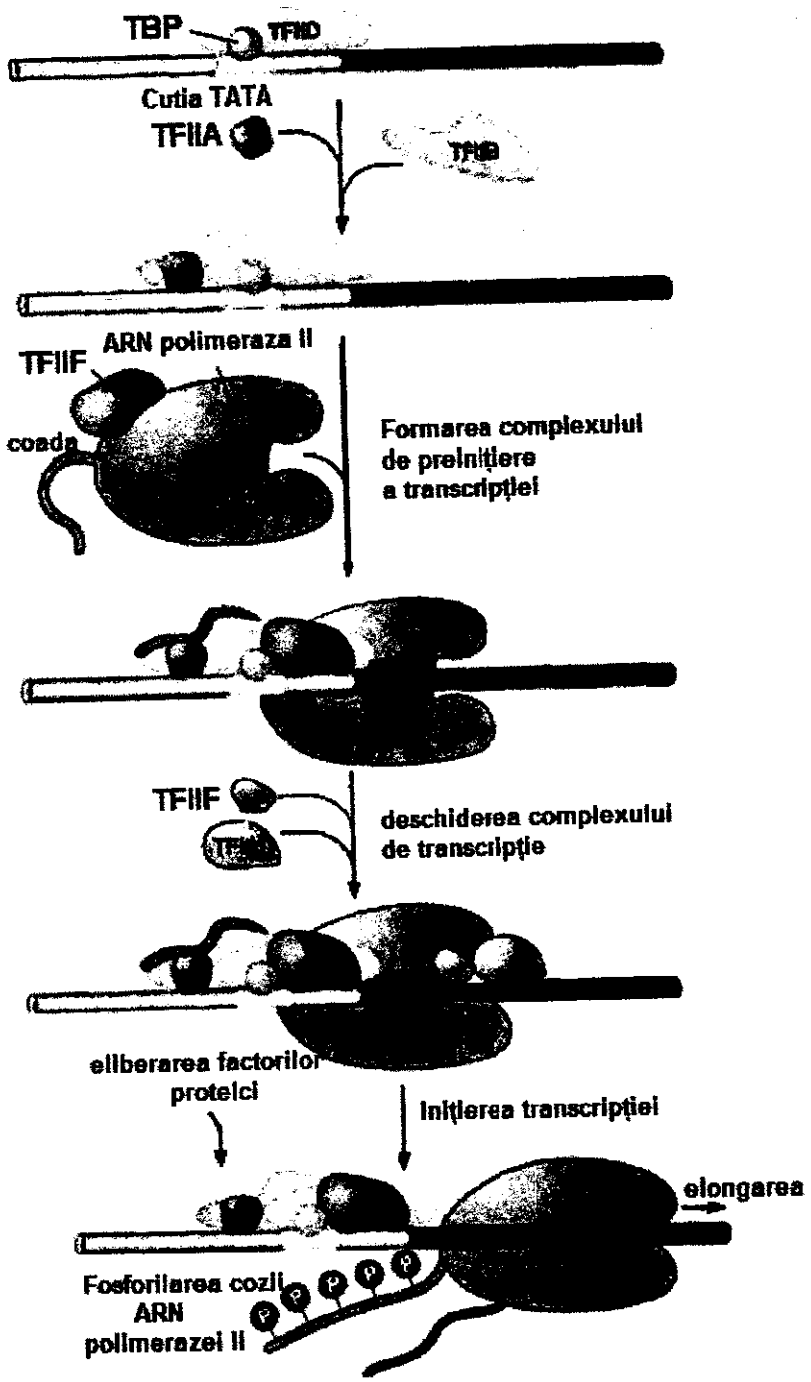


Fig.4.18 Inițierea transcripției la eucariote (adaptată după Watson și colab., 2003)

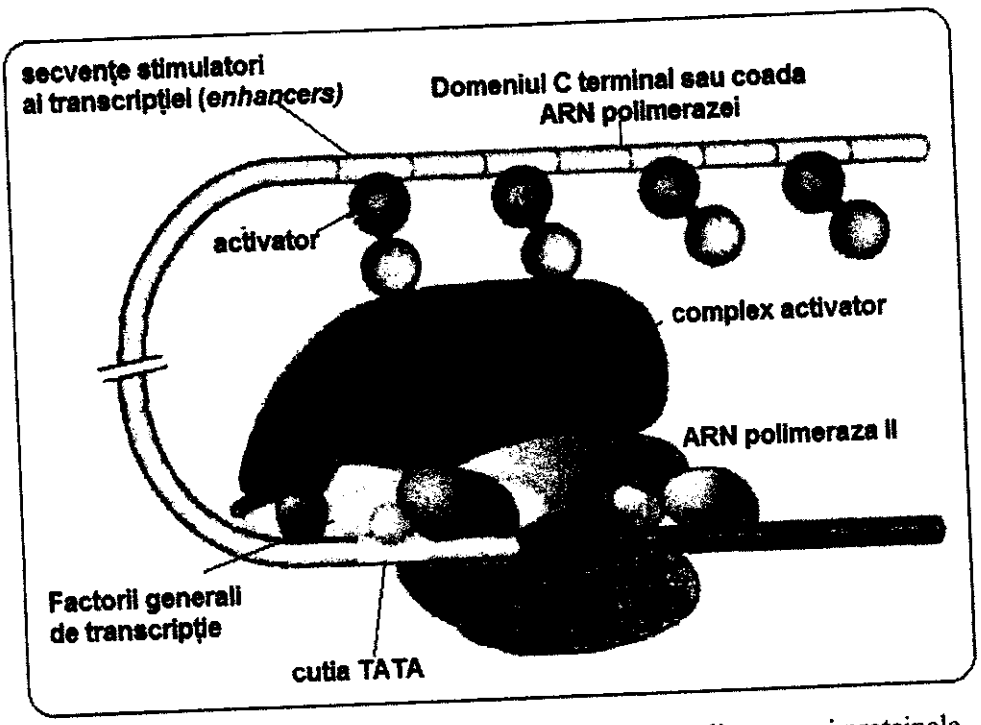


Fig.4.19 Interacţiunea complexului mediator cu ARN polimeraza şi proteinele reglatoare (activatorii) (adaptată după: Watson şi colab., 2003)

Elongarea

După ce se produce deschiderea complexului de iniţiere prin acţiunea helicazei (TFIIH), are loc iniţierea transcripţiei însă procesul efectiv de elongare începe abia după câteva încercări repetitive ale enzimei de sinteză, proces asemănător unei amorsări a regiunii de iniţiere. În prima fază sunt sintetizate transcripte scurte de circa 10 pb care sunt descărcate rapid de la matriţă, iar procesul este reluat de mai multe ori. În momentul în care enzima scapă de la secvenţa promotor şi trece la sinteza unor transcripte mai lungi, are loc şi trecerea la faza de elongare.Aceasta corespunde detaşării factorilor generali de transcripţie, respectiv a mediatorilor, de ARN polimeraza miez, proces facilitat de TBF.

În schimb, în etapă sunt adiţionaţi alţi factori proteici la enzimă, care participă la procesul de elongare (TFIIS şi hSPT5). Recrutarea factorilor de elongare are loc prin intermediul domeniul C terminal al polimerazei (CTD) care suferă fosforilări la anumite reziduuri din domeniului C terminal. De exemplu, serina din poziţia 2 a monomerului heptapeptidic este fosforilată de factorii proteici de stimulare a elongării P-TEFb, care la rândul lor vor determina adiţionarea altor factori de elongare şi activarea proteinei -

hSPT5. Factorii TFIIS intervin atât în intensificarea procesului de elongare, cât şi în activarea funcţiei de corectare a ARN polimerazei II, pentru îndepărtarea nucleotidelor greşit încorporate şi adăugarea celor corecte (funcţia exonucleazică).

Procesul de elongare constă în transcrierea corectă a întregii informaţii din ADN, începând cu regiunea 5' UTR şi terminând cu regiunea 3' UTR, iar rezultatul este o moleculă de ARNm precursor.

Procesul de elongare este parţial suprapus cu cel de prelucrare a capetelor moleculei de ARNm precursor, procese ce constau în metilarea capătului 5', după adiţia unei guanine şi transformarea în 7-metilguanozină sub forma unei conformaţii denumită „cap" şi prin sinteza poliadenilatului la capătul 3' (poli-A). Când toată informaţia din gena structurală, de la codonul START la codonul STOP, dar şi aceea din regiunile 5' UTR şi 3' UTR este transcrisă, procesul de elongare este terminat.

Terminarea transcripţiei

Ultima etapă în procesul de transcripţie este terminarea elongării sau a transcripţiei. Secvenţele din aval de genă cu rol de reglaj care contribuie la finalizarea procesului, sunt situate după codonul STOP. Aşa cum am arătat, ARN polimeraza II continuă transcripţia mult după ce întâlneşte codonul STOP şi recunoaşte anumite secvenţe ca fiind terminatori ai transcripţiei. Acestea sunt secvenţa consensus AAUAAA, situată cu 11-30 nucleotide înainte de situsul de clivaj şi secvenţa poli-U (U/GU), situată în aval de acesta. Ajutată de anumiţi factori proteici, enzima ARN polimeraza II va încheia transcripţia după clivaj şi după sinteza cozilor poli-A, cu enzima poliadenil-polimeraza (PAP). Reţinem că sinteza cozilor poli-A la transcript nu este determinată prin transcripţia informaţiei conţinută în ADN, ci este un proces de prelucrare postranscripţional în care cozile poli-A sunt adiţionate *de novo*, cu ajutorul unei polimeraze speciale. Urmare a poliadenilării transcriptului, se produce eliberarea ARN polimerazei II de la matriţa ADN şi încheierea efectivă a transcripţiei.

4.6.1 Prelucrarea transcriptului primar (ARNm precursor)

4.6.1.1 Prelucrarea capetelor ARNm precursor

Prelucrarea capătului 5' al ARNm precursor

Capătul 5' al ARNm precursor conţine secvenţa 5' – UTR care nu va intra în procesul de translaţie efectivă, însă înainte va suferi modificări care au rolul de a iniţia translaţia informaţiei genetice la ribozomi.

CAPITOLUL V

EXPRESIA GENELOR –TRANSLAŢIA

5.1 Codul genetic şi translaţia informaţiei genetice, la procariote şi la eucariote

Translaţia este procesul prin care informaţia genetică conţinută în ARNm, este decodificată la nivelul ribozomilor în proteine. Proteinele, enzimele şi hormonii sunt biomolecule sintetizate pe baza informaţiei genelor structurale. Procesul de decodificare a informaţiei genetice sau expresia genelor are loc în două etape, în care informaţia din ADN este transcrisă în ARNm şi apoi translatată la ribozomi în secvenţe polipeptidice.

Intuirea relaţiei dintre genotip şi fenotip aparţine încă din anul 1908 lui Archibald Garrot. Neexplicată însă la acea vreme, relaţia a fost demonstrată experimental abia în anii 1940 de către Beadle şi Tatum. Cei doi cercetători au reuşit să demonstreze în experimentele efectuate asupra fungului *Neurospora* că mutantele auxotrofe nu pot creşte pe medii de cultură minimale, decât dacă în mediu sunt adăugaţi compuşi pentru care tulpinile sunt defective. Inducând mutaţii fungilor, aceştia au devenit defectivi în anumite sinteze necesare creşterii (mutante auxotrofe). Adăugând mediilor minimale diferiţi compuşi, cum este de exemplu aminoacidul arginină, cei doi cercetători au reuşit să izoleze diferite tulpini auxotrofe deficiente în arginină, care cresc doar pe mediile în care aminoacidul era adăugat.

Concluziile studiului au fost deosebit de valoroase la acea vreme pentru că demonstrau că mutaţia indusă tulpinilor a inactivat anumite gene, responsabile de sinteza proteinelor sau enzimelor pentru care fungii prezentau deficienţe metabolice. Prin urmare, gena care a suferit mutaţia putea fi intuită prin intermediul compusului esenţial adăugat mediului minimal. Aceste rezultate dovedeau pentru prima dată că între genotip şi fenotip există o relaţie de cauzalitate. În prezenţa compusului adăugat în mediu, fungii defectivi îşi pot asigura necesarul de compuşi esenţiali pentru creştere şi diviziune, pe aceea pe care nu îi pot sintetiza din cauza mutaţiei. Repetând experimentele şi cu alţi compuşi, cei doi cercetători identifică genele unor mutante auxotrofe care controlează câteva căi de biosinteză la *Neurospora*. Rezultatele au fost concretizate în ipoteza cunoscută sub denumirea **„o genă – o enzimă"**. Ipoteza încearcă să explice faptul că enzimă este codificată de o singură genă care dacă este alterată prin mutaţi

mutaţia va face imposibilă biosinteza enzimei. Câţiva ani mai târziu conceptul a fost modificat deoarece s-a demonstrat că anumite enzime şi proteine sunt formate din mai multe lanţuri polipeptidice, fiecare codificate de gene proprii. Astfel ipoteza a devenit **„o genă – o polipeptidă"** (Pierce, 2008).

Calea prin care informaţia genetică din gene este decodificată la ribozomi în secvenţe polipeptidice poartă numele de translaţie. Trecerea de la informaţia conţinută în ADN, la produsul polipeptidic obţinut în urma translaţiei, se produce prin procese moleculare complexe, în care are loc succesiv un transfer informaţional între moleculele ADN → ARN → polipeptide. În mod firesc, aceste transferuri între molecule necesită corespondenţe specifice între limbajele de codificare. Descifrarea acestor relaţii explică de fapt corespondenţa la nivel molecular dintre genotip şi fenotip.

Primul transfer al informaţiei are loc între două molecule de acizi nucleici: ADN şi ARNm. Corespondenţa între cele două molecule este *complementaritatea* dintre bazele azotate. Informaţia genetică conţinută de ADN este determinată de succesiunea nucleotidelor, respectiv de succesiunea celor 4 baze azotate (A, G, C, T) care îi dau specificul structural. Prin transcripţie informaţia genetică din ADN este transferată (transcrisă) moleculei informaţionale de ARNm, cu succesiunea complementară de baze azotate. Cantitatea de informaţie genetică transcrisă într-o anumită moleculă de ARNm poartă denumirea de **mesaj genetic** (Vlaic A., 2011).

Al doilea transfer se produce între o moleculă de acid nucleic şi o polipeptidiă. Molecula de acid nucleic (ARNm) duce acest mesaj genetic la ribozomi, unde va fi decodificat într-o moleculă proteică prin procesul de translaţie. Aceasta presupune iarăşi existenţa unei corespondenţe precise între limbajul alfabetului conţinut în ARNm (succesiunea nucleotidelor din acizi nucleici) şi limbajul proteinei (succesiunea aminoacizilor).

Proteinele sunt formate din aminoacizi legaţi între ei prin legături peptidice. În lumea vie numărul total al aminoacizilor care compun proteinele este de 20, denumiţi aminoacizi esenţiali. (Fig.5.1) Toţi aminoacizii esenţiali se aseamănă în structură prin trei grupări identice, excepţie făcând radicalului R sau catena laterală a aminoacidului, care îi diferenţiază structural (Fig.5.2). La fel cu acizii nucleici şi proteinele prezintă polaritate, un capăt conţine gruparea liberă amino- protonată (NH_3^+) iar la celălalt gruparea carboxil deprotonată (COO^-).

Structura moleculară a unei proteine prezintă mai multe nivele de organizare, similar cu acizii nucleici.

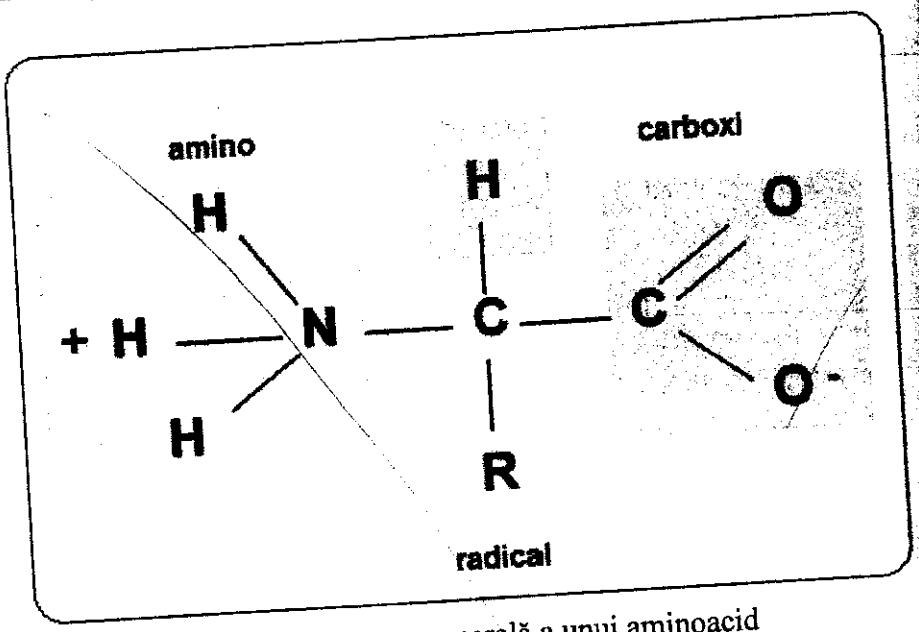


Fig.5.1 Structura generală a unui aminoacid

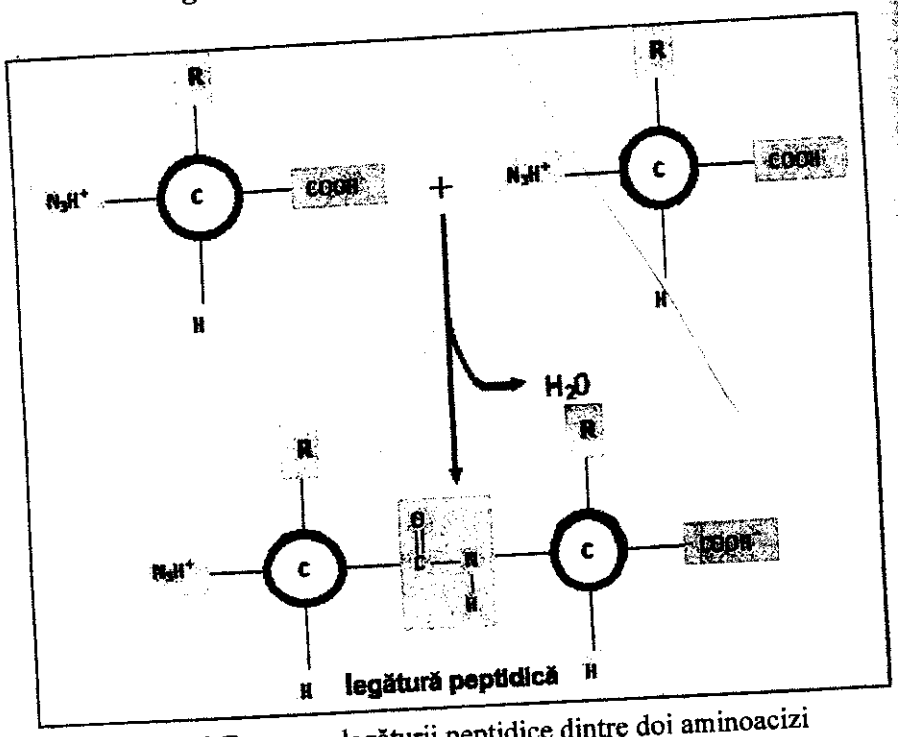


Fig.5.2 Formarea legăturii peptidice dintre doi aminoacizi

Structura primară a proteinei se referă la secvenţa sa în aminoacizi. Prin interacţiunile dintre aminoacizi învecinaţi, un lanţ polipeptidic se pliază formând falduri şi coturi care determină *structura secundară* a proteinelor. Structurile secundare comune proteinelor sunt conformaţia pliată β şi conformaţia α helix. Aceste structuri pot în continuare să interacţioneze pentru a forma o *structură terţiară,* care în ansamblu reprezintă structura tridimensională a proteinei. Ambele structuri, secundară şi terţiară, sunt în cele din urmă determinate de structura primară aminoacidică. Unele proteine sunt formate din mai multe lanţuri polipeptidice ce se asociază în forma unei structuri *cuaternare* (Fig.5.3) (Pierce şi colab.,2008).

Regula după care se încatenează aminoacizii în catena polipeptidică a unei proteine, corespunzător naturii mesajului genetic din ARNm, poartă numele de **cod genetic.**

Primul cercetător care a încercat să explice corespondenţa între limbajul mesajului genetic conţinut de ARNm şi succesiunea aminoacizilor din proteine, a fost Francis Crick. Ipoteza de la care a pornit Crick pentru stabilirea corespondenţei a fost aceea că trebuia să existe o moleculă care să facă legătura între cele două limbaje, deoarece aminoacizii din proteine şi acizii nucleici nu prezintă afinitate. Această moleculă, a cărei natură nu era încă,cunoscută a fost denumită iniţial moleculă adaptor, pentru ca mai târziu să-i fie determinată natura ca fiind o molecula de ARNt (ARN de transport).

Cunoscând structura nucleotidică a ARNm format din succesiunea celor 4 baze azotate, Crick emite ipoteza că regula după care se încatenează aminoacizii în catena polipeptidică trebuie să fie dictată de succesiunea bazelor azotate din ARNm. Prima întrebare pusă când s-a încercat descifrarea codului genetic a fost: de câte nucleotide este nevoie ca să codifice un aminoacid? Prin cercetările ulterioare s-a stabilit că unitatea de bază a codului genetic este codonul, reprezentat prin numărul nucleotidelor care specifică un aminoacid.

Cercetările efectuate anterior emiterii ipotezei au arătat ca plauzibilă combinaţia a câte trei nucleotide pentru un codon. Probabilitatea codonului cu o singură nucleotidă era imposibilă deoarece există doar patru baze azotate diferite care compun nucleotidele, iar numărul aminoacizilor esenţiali este de 20. De asemenea, combinaţiile a câte două nucleotide determină un număr maxim de codoni diferiţi egal cu $4^2= 16$, insuficient pentru a codifica cei 20 aminoacizi. În fine, cazul codonului cu trei nucleotide părea cel mai plauzibil deoarece din combinarea celor patru nucleotide luate câte trei, se pot obţine $4^3 = 64$ combinaţii diferite. Prin urmare, tripletul părea să fie combinaţia corectă pentru un codon, deşi numărul lor era mult mai mare decât numărul aminoacizilor esenţiali.

În anul 1961 Francis Crick demonstrează experimental prin mutaţii induse la bacteriofagul Λ, alcătuirea codului genetic din tripleţi, adică faptul că fiecare trei nucleotide codifică un aminoacid. Înaintea demarării cercetărilor s-a pornit de la cele două evidenţe şi anume că numărul aminoacizilor esenţiali este de 20 (Fig.5.4) iar numărul codonilor este de 64.

5.1.1 Descifrarea codului genetic

Descifrarea codului genetic sau determinarea corespondenţei dintre codonii din ARNm şi aminoacizii din proteine a pornit mai întâi de la ideea identificării succesiunii aminoacizilor dintr-o proteină, produsă pe baza unei molecule de ARNm cu secvenţă cunoscută (Russell, P., 2010).

Marshall Nirenberg şi Johann Heinrich Matthaei au obţinut în laborator primele molecule sintetice homopolimere de ARNm, cu ajutorul enzimei polinucleotid fosforilaza. Prima moleculă produsă a fost de tip poliuracil (poli-U). Deoarece proteina care a fost sintetizată pe baza acestei molecule de ARNm era formată doar din aminoacidul fenilalanină, s-a tras concluzia că aminoacidul fenilalanină este codificat de codonul UUU. Următoarea moleculă sintetică produsă a fost de tip poli-C, pe baza căreia s-a sintetizat un lanţ polipeptidic format doar din prolină, prin urmare codonul corespondent aminoacidului prolină s-a dedus că este tripletul CCC. Continuând în această manieră sinteza de molecule poliribonucleotidice homopolimere, cei doi cercetători au reuşit să descifreze corespondenţa pentru câţiva dintre codoni.

În continuare s-a trecut la adăugarea în tubul de reacţie a câte două nucleotide diferite, însă în urma procesul de combinare de nucleotide, din reacţie s-au obţinut lanţuri polipeptidice cu aminoacizi diferiţi în secvenţă. De exemplu, prin adăugarea nucleotidelor A şi C în tubul de reacţie, polinucleotid fosforilaza leagă randomizat nucleotidele, din a căror combinaţie pot rezulta mai mulţi codoni: AAA, AAC, ACC, ACA, CAA, CCA, CAC şi CCC (Pierce şi colab., 2008). Încatenarea lor în molecula sintetică de ARNm este deci aleatorie, rezultând astfel proteine cu aminoacizi diferiţi în structură, a căror corespondenţă nu putea fi stabilită. Din aceste considerente s-a încercat optimizarea procesului prin adăugarea nucleotidelor în proporţii variabile, astfel încât pe baza probabilităţilor aşteptate să poată fi construite molecule sintetice într-o singură combinaţie de câte două baze azotate. Chiar şi aşa, metoda nu a dat rezultatele scontate deoarece enzima a adăugat nucleotidele randomizat şi uneori a încorporat alte nucleotide decât cele aşteptate teoretic, ceea ce a schimbat mesajul genetic şi implicit aminoacizii din proteină.

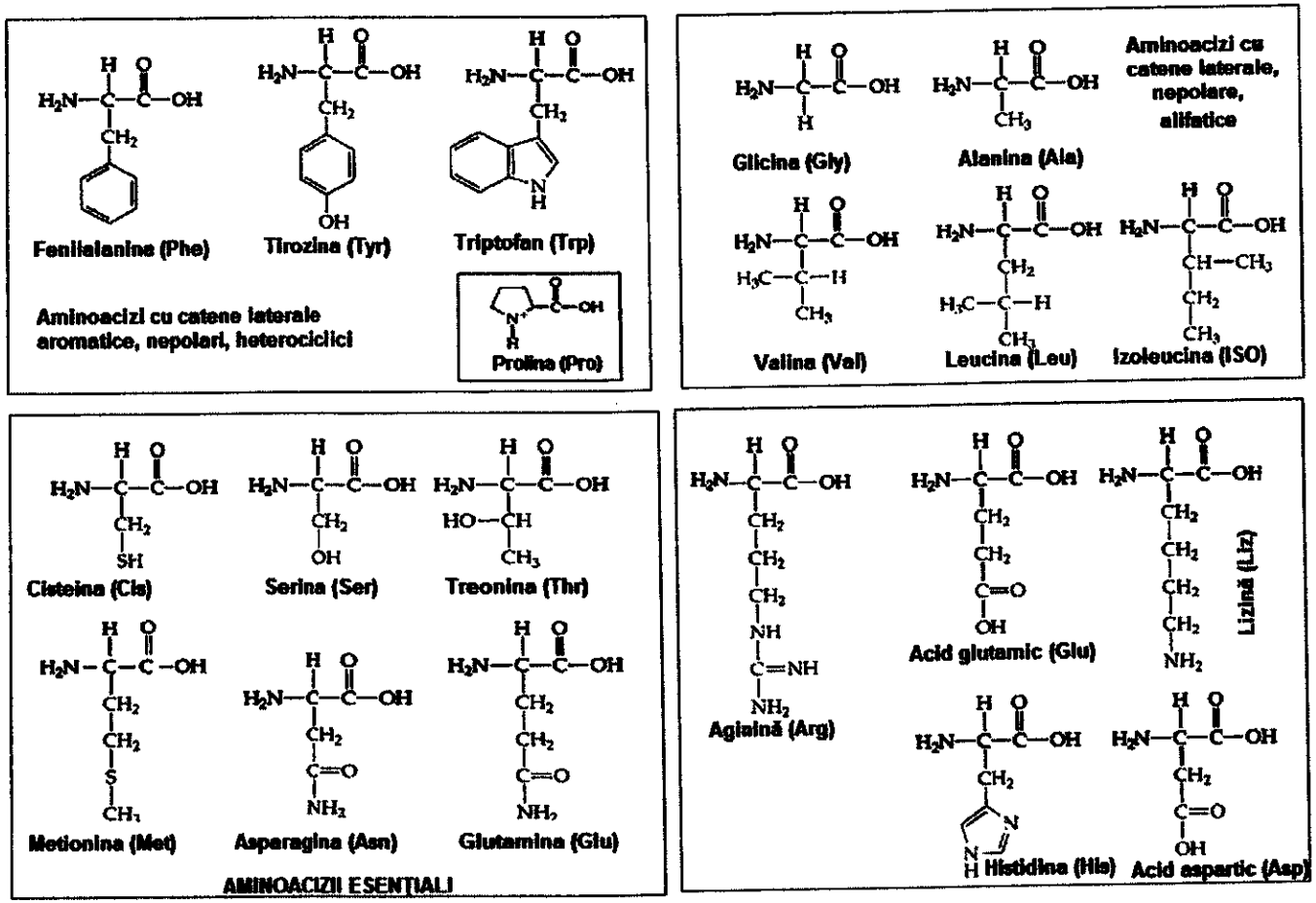


Fig.5.4 Aminoacizii esențiali

Câţiva ani mai târziu, Nirenberg şi Leader aplică o metodă nouă pentru descifrarea codului genetic în care se folosesc moleculele de ARNt. Aceştia produc circa 50 molecule sintetice scurte de ARNm (formate din trei nucleotide) în a căror secvenţă sunt două nucleotide diferite. Încatenarea celor trei nucleotide se face însă controlat, astfel încât fiecare triplet să fie cu secvenţă cunoscută. Fiecare oligoribonucleotidă este apoi adăugată separat unui amestec format din toţi aminoacizi şi o moleculă de ARNt. Prin bucla anticodon a moleculei, de ARNt are loc recunoaşterea pe bază de complementaritate a tripletul potrivit din mesajul genetic, la care se ataşează încărcat cu un singur aminoacid din amestec.

Deoarece existau dovezi suficiente privind codificarea aceluiaşi aminoacid de mai mulţi codoni, experimentele s-au bazat pe stabilirea corespondenţei, pornind de la codonii din ARNm pentru identificarea aminoacizilor încărcaţi la moleculele de ARNt. Cei doi cercetători au descifrat astfel ce aminoacizi au fost încorporaţi în polipeptida sintetizată şi ce codoni corespund unui anumit aminoacid. De exemplu, ribonucleotida cu codonul GUU a fost recunoscută de un ARNt încărcat cu aminoacidul valină, iar tripletul UUU, de ARNt care transportă fenilalanina (Fig.5.5). Bazându-se pe acest principiu, cei doi cercetători stabilesc astfel corespondenţa aminoacizilor pentru 50 de codoni.

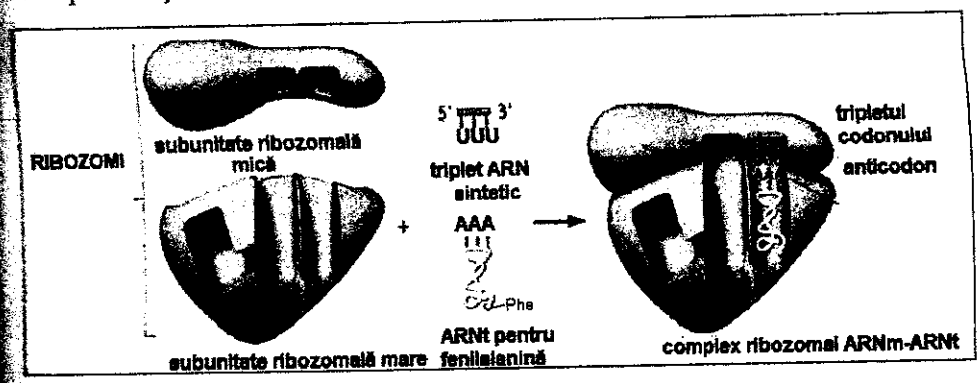


Fig.5.5 Descifrarea corespondenţei dintre codoni şi aminoacizi cu ajutorul ribonucleotidelor sintetice şi a moleculelor de ARNt (***, 2006)

A treia metodă de descifrare a codului genetic a fost pusă la punct de Gobind Khorana şi colegii săi, care au sintetizat chimic ribonucleotide copolimere, prin alternarea a două baze azotate. De exemplu din combinaţia uracilului cu guanina se obţine o ribonucleotidă copolimer de tip UGUGUGUGUGUG. Rezultatul obţinut de cercetători a fost surprinzător deoarece cei doi au obţinut o polipeptidă formată prin alternarea a doi

aminoacizi, cisteină şi valină. Descifrând înşiruirea succesiunii de câte trei nucleotide din lanţ, aceştia au observat că în ribonucleotidă se repetă constant doi tripleţi diferiţi UGU GUG UGU GUG, fiecăruia corespunzându-i câte un aminoacid din cei doi care se repetă alternativ în proteină, mai precis codonul UGU căruia îi corespunde cisteina, cu codonul GUG care specifică valina.

Din aceste cercetări reiese şi una dintre caracteristicile codului genetic, caracteristici formulate imediat după descifrare şi anume că acesta este ***nesuprapus***, adică doi codoni vecini nu îşi împrumută unul altuia nucleotide în mesajul genetic.

În anul 1968 întregul cod genetic a fost descifrat folosind tehnici combinate pentru stabilirea corespondenţei dintre codoni şi aminoacizi.

În întreaga lume vie, începând cu bacteriile, până la cele mai evoluate eucariote, inclusiv la om, un anumit codon din codul genetic codifică acelaşi aminoacid (Fig.5.6), iar această caracteristică a codului genetic se numeşte ***universalitate.***

CODUL GENETIC

a doua literă

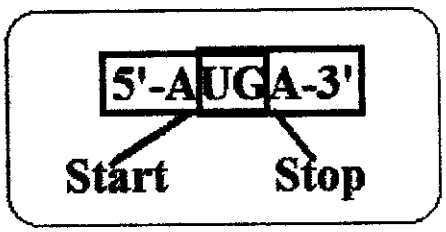
prima literă	U		C		A		G		a treia literă
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC		UCC		UAC		UGC		C
	UUA	Leu	UCA		UAA	Stop	UGA	Stop	A
	UUG		UCG		UAG	Stop	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Gln	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU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

Abrevierile standard ale aminoacizilor: Phe – fenilalanină; Leu – leucină; Ile – izoleucină; Val – valină, Ser – serină; Pro – prolină; Met – metionină; Thr – treonină; Ala – alanină; Tyr – tirozină; His – histidină; Gln – glutamină; Asn – asparagină; Lys – lizină; Asp – acid aspartic; Glu – acid glutamic; Cys – cisteină; Trp – triptofan; Arg - arginină; Gly – glicină;

Fig. 5.6 Codul genetic

Din cei 64 codoni, unul îndeplineşte funcţia de codon START care marchează începutul mesajului genetic. La toate genele de la eucariote, cu excepţia celor mitocondriale, dar şi la genele grupului Archea şi la majoritatea grupului Bacteria, acesta este codonul AUG care specifică aminoacidul metionină (Met), sau metionina modificată în cazul bacteriilor (formil metionina/fMet) (Touriol şi colab., 2003). Din cei 64 codoni, doar doi aminoacizi sunt specificaţi de câte doi codoni; este vorba de aminoacizii metionină şi triptofan. Restul aminoacizilor sunt specificaţi fiecare de trei sau mai mulţi codoni diferiţi, ceea ce semnifică faptul că codul genetic este ***degenerat*** (Watson şi colab., 2003).

În afara codonul START AUG, 3 codoni marchează sfârşitul mesajului genetic, fiind denumiţi codoni STOP sau nonsens deoarece nu specifică nici un aminoacid. Aceştia sunt codonii UAA, UAG şi UGA. În final, din cei 64 codoni rămân 61 codoni sens care codifică cei 20 aminoacizi esenţiali. În anumite proteine, în funcţie de anumite secvenţe semnal asociate la ARNm, doi codoni stop codifică doi aminoacizi nestandard, care nu sunt incluşi printre cei 20 aminoacizi esenţiali. De exemplu, codonul UGA codifică aminoacidul selenocisteină şi UAG codifică pirolizina. În prezent, selenocisteina este considerat cel de - al 21 aminoacid esenţial. (Elzanowski şi Ostell,, 2010). Codonii înşiraţi de-a lungul catenei de ARNm se citesc continuu, deoarece între codoni nu există virgule sau alte semne de punctuaţie. Citirea unui mesaj genetic ***lipsit de virgule,*** format din mai mulţi codoni înşiraţi unul după altul, pentru a fi întotdeauna exactă, necesită acelaşi punct de plecare. O citire care ar începe de exemplu de la a doua nucleotidă a primului codon, ar duce la schimbarea mesajului genetic şi formarea unei proteine modificate în structură şi funcţie. Acest proces se întâmplă în mod normal în cazul genelor parţial suprapuse, care au fost descoperite la virusuri şi mai apoi şi la bacterii. La acestea, codonul start se suprapune parţial peste codonul stop al genei următoare. Cel mai frecvent, această suprapunere apare prin secvenţa –AUGA-.



5.1.2 Excepţiile codului genetic de la regula generală de corespondenţă

Cele câteva excepţii de la regula generală de corespondenţă codon-aminoacid pot fi întâlnite la bacterii la care şi codonii GUG şi UUG pot fi codoni START pentru unele gene. De exemplu, în cazul genelor *lac*A (ce

doar individul, în timp ce mutaţiile germinale se transmit prin intermediul gameţilor la descendenţi.

- *După natura materialului genetic afectat de mutaţie,* acestea se clasifică în **genomice, cromozomiale** şi **genice**.
- *După acţiunea asupra genei mutante,* mutaţiile pot fi **dominante** sau **recesive.**
- *După efectul fenotipic sau gradul de viabilitate,* mutaţiile pot fi: **dăunătoare, letale** şi **viabile**.

Măsurarea efectelor mutaţiilor poate fi definită prin doi termeni: **rata** cu care se produc mutaţiile, care defineşte probabilitatea de apariţie a mutaţiei, funcţie de timp (de exemplu: număr de mutaţii/genă/generaţie sau număr de mutaţii/pb/generaţie) şi **frecvenţa** mutaţiei, care defineşte numărul de apariţii ale unui anumit tip de eveniment mutaţional, exprimat ca proporţie faţă de numărul de celule sau indivizi dintr-o populaţie (număr de mutaţii/100.000 indivizi sau număr de mutaţii/1 milion de gameţi).(Watson şi colab., 2003; Araten colab., 2005)

6.1 Mutaţiile genice

Gena poate fi definită ca un fragment discret şi continuu de ADN care specifică un anumit produs polipeptidic. Gena cuprinde secvenţele cu rol de reglaj din amonte şi din aval de segmentul structural format din exoni şi introni. Mutaţia genică constă în orice modificare produsă în structura ADN, fie în segmentul structural, fie în regiunile de reglaj. Mutaţiile care afectează gena pot să se producă la nivelul unei nucleotide sau perechi de nucleotide, caz în care vorbim de **mutaţii punctiforme** (Russell, P., 2010).

6.1.1. Mutaţiile genice punctiforme

Mutaţiile punctiforme constau în *substituţia* unei singure nucleotide cu altă nucleotidă prin mecanismul de tranziţie sau transversie, sau prin *inserţia* sau *deleţia* unei nucleotide din genă.

6.1.1.1 Clasificarea mutaţiilor punctiforme după mecanismul de producere

Mutaţia punctiformă prin tranziţie este o mutaţie de înlocuire a unei perechi purină-pirimidină cu o altă pereche purină-pirimidină, un exemplu fiind înlocuirea perechii A-T cu G-C. Concret, aceasta înseamnă că adenina de pe o catenă a lanţului ADN este înlocuită cu cealaltă purină – guanina, în timp ce pe cealaltă catenă a ADN pirimidina T este înlocuită cu pirimidina C (Fig.6.2. a).

Mutaţia punctiformă prin transversie constă în înlocuirea u perechi purină-pirimidină cu o pereche pirimidină – purină, cum ar fi exemplu modificarea G-C în C-G, sau A-T în C-G (Fig. 6.2.b). Proce constă în înlocuirea unei baze azotate pirimidinice de pe o catenă a ADN o bază azotată purinică (de exemplu A din exemplul 2 cu C), în timp ce cealaltă catenă a ADN are loc înlocuirea unei baze azotate pirimidinice una purinică (de exemplu T cu G).

Substituţia bazelor azotate în cadrul lanţului ADN poate avea efec diferite, în funcţie de modificarea produsă la nivelul codonilor din mesa genetic. În urma substituţiei bazelor azotate mutaţiile pot să nu determi efecte în proteina sintetizată, în alte cazuri pot produce modific nesemnificative, iar în alte cazuri modificările pot fi foarte semnificativ (Brown, 2002; Russell, P., 2010).

6.1.1.2 Clasificarea mutaţiilor punctiforme după efectul asupra proteinei sintetizate

O mutaţie nesinonimă este o mutaţie la nivel genic în c schimbarea unei perechi de baze azotate din codon va determi specificarea altui aminoacid. Fenotipul proteinei produse în urma u mutaţii nesinonime poate sau nu să fie alterat, în funcţie de aminoacid implicat. În figura 6.2c, tranziţia A-T în G-C modifică codonul

5'- AAA-3' în codonul 5'- GAA -3',

3'- TTT -5' 3'- CTT- 5', astfel încât aminaocidul lizi specificat de codonul iniţial al genei (AAA), este înlocuit cu aminoacid acidului glutamic specificat de noul codon (GAA).

Mutaţia nonsens este o mutaţie în care o pereche de baze azotate schimbă la nivel ADN şi în urma căreia în locul unui codon sens d mesajul genetic apare unul nonsens sau STOP. În figura 6.2d es exemplificat acest tip de mutaţie prin transversia A-T în T-A, în urma căre în locul codonul 5'- AAA-3' din ADN apare codonul 5'- TAA -3'. P transcripţia mesajului genetic, din acest codon modificat în ADN apare nivelul ARNm tripletul nonsens 5'-UAA -3' care va înlocui în protei codonul lizină specificat de codonul 5'- AAA-3'. Mutaţiile nonse determină terminarea prematură a sintezei lanţului polipeptidic al protein care în cele mai multe cazuri devine nefuncţională.

Mutaţia neutră este acel tip de mutaţie punctiformă din genă c nu afectează funcţia proteinei sintetizate de pe molecula de ARN modificată. În urma înlocuirii unei baze azotate în ADN codonul modific

prin mutaţia neutră va specifica un aminoacid echivalent în proteina sintetizată. Mutaţia neutră este de fapt o variantă a mutaţiei nesinonime, în care un codon este modificat în alt codon, însă în acest caz aminoacizii specificaţi de cei doi codoni sunt echivalenţi (fig.6.2e). De exemplu, tranziţia A-T în G-C, care se produce la nucleotida 2 în codonul 5'-AAA-3'(rezultând 5'-AGA-3') va determina schimbarea aminoacidului lizină în arginină. Deoarece lizina şi arginina au proprietăţi similare, ambii fiind aminaocizi bazici, această modificare nu schimbă semnificativ funcţia proteinei (Russell, P., 2010).

Mutaţia sinonimă, cunoscută şi cu denumirea de mutaţie tăcută, este acea mutaţie care schimbă o pereche de baze azotate în genă, însă aminoacidul specificat de cei doi codoni (nemodificat şi modificat) este acelaşi în proteina sintetizată. De exemplu, dacă în acelaşi codon 5'-AAA-3' apare mutaţia prin tranziţia A-T în G-C la nivelul celei de-a treia nucleotide, va rezulta codonul 5'-AAG -3'. Ambii codoni, **AAA şi AAG,** specifică acelaşi aminoacid – lizina, prin urmare proteina sintetizată în urma unei mutaţii tăcute este aceeaşi în structură şi funcţie cu proteina originală (Fig.6.2f).

Mutaţia prin decalarea cadrului de citire a mesajului genetic apare ca urmare a inserţiei sau deleţiei unei perechi de nucleotide din genă, prin urmare citirea mesajului genetic se schimbă cu o nucleotidă în plus sau în minus. În general acest tip de mutaţie duce la sinteza unei proteine nefuncţionale, mai scurte, deoarece de foarte multe ori se produc codoni stop care întrerup mesajul genetic (Fig.6.2g) (Russell, P., 2010).

6.1.1.3.Clasificarea mutaţiilor punctiforme după direcţia de producere a mutaţiei (Mutaţia înainte-înapoi)

Direcţia mutaţiei face referire la transformarea fenotipul „sălbatic" determinat de o genă normală, într-unul modificat ca urmare a producerii mutaţie. După acest criteriu mutaţiile punctiforme se clasifică în două categorii, în funcţie de fenotipul afectat de mutaţie:

a) *mutaţia înainte* este acea mutaţie prin care gena normală se transformă în genă mutantă şi b) *mutaţia înapoi* sau inversă (retromutaţie) este mutaţia prin care gena mutantă revine total sau parţial la fenotipul iniţial determinat de gena normală.

În cazul mutaţiei nonsens, codonul stop apărut ca urmare a unei mutaţii punctiforme se transformă într-un codon sens, iar gena redevine funcţională. În cazul în care mutaţia înapoi determină ca gena să revină la structura iniţială, vom spune despre această mutaţie că este o **mutaţie inversă reală**. În alte cazuri însă, mutaţia înapoi schimbă codonul stop cu alt codon sens decât cel din secvenţa originală a genei, caz în care vom spune despre mutaţia înapoi că este una **parţială**.

Efectul mutaţiei poate fi abolit sau suprimat de o **mutaţie supresoare,** care se produce în alt situs, diferit de cel al mutaţiei originale. În acest caz, efectul mutaţiei supresoare maschează sau compensează efectul mutaţiei înainte produsă în alt situs, însă nu anulează mutaţia în sine. După situsul în care se produce mutaţia supresoare acestea sunt de două tipuri: intragenice, când ambele mutaţii se produc la nivelul aceleiaşi gene sau intergenice, când mutaţia supresoare apare în afara genei, mai precis în regiuni intergenice. În cazul mutaţiei intragenice supresoare, aceasta poate apărea în acelaşi codon cu cel al mutaţiei înainte sau în alt codon. Un exemplu de acest tip este mutaţia prin inserţia unei nucleotide care se produce lângă o deleţie anterioară, care schimbă astfel cadrul de citire al genei (Fig.6.3) (Russell,P., 2010).

6.1.1.4.Clasificarea mutaţiilor punctiforme după modul de producere

Mutageneza, ca proces de modificare a structurii ADN, poate apărea spontan sau poate fi indusă. *Mutaţiile spontane* se produc natural iar cele *induse* apar ca urmare a expunerii accidentale sau voite la diferiţi agenţi fizici sau chimici.

Mutaţiile spontane

Mutaţiile punctiforme apar de obicei spontan în timpul replicării ADN, sau în alte faze ale creşterii şi diviziunii celulare. Mutaţiile care interesează segmente mai lungi de ADN sunt determinate de cele mai multe ori de elemente transpozabile care îşi schimbă poziţia în genom.

Mutaţiile spontane apar în genomul uman cu rate cuprinse între 10^{-4} - $4x10^{-6}$/genă/generaţie, în timp ce la eucariote această valoare medie se situează între 10^{-4} şi 10^{-6}/genă/generaţie, iar la bacterii şi la fagi valorile medii sunt cuprinse între 10^{-5}-10^{-7}/genă/generaţie (Russell, P, 2010). Ratele şi frecvenţele exemplificate se referă la mutaţiile fixate în genom, care se moştenesc şi nu la totalitatea mutaţiilor produse deoarece o mare parte din mutaţiile spontane sunt înlăturate prin diferite mecanisme de reparare care funcţionează la nivel ADN sau la nivel celular.

Cauzele mutaţiilor spontane sunt: erorile de replicare a ADN şi modificarea structurii chimice a ADN prin depurinarea şi dezaminarea bazelor azotate.

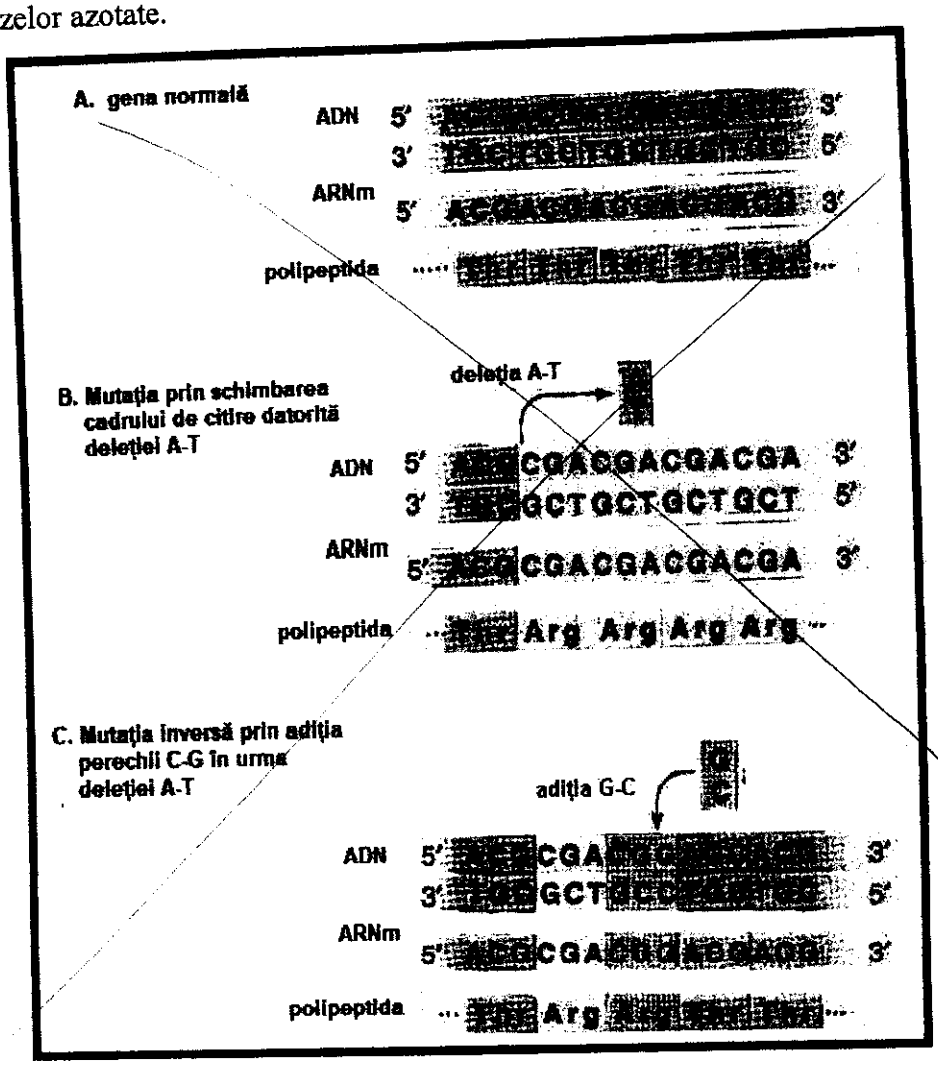


Fig.6.3 Mutaţia supresoare (sursa: Russell, P. 2010)

Mutaţiile spontane determinate de erorile de replicare a ADN

În timpul procesului de replicare a ADN, polimeraza introduce greşit nucleotide care nu se pot împerechea complementar. Mecanismele de reparare a erorilor de replicare sunt foarte variate, cel mai important fiind cel exonucleazic 3'-5' al ADN polimerazei care se întoarce în timpul catalizei enzimatice 5' - 3', recunoaşte greşelile introduse, le excizează şi le înlocuieşte cu nucleotide corecte (Brown, 2002).

În alte cazuri însă, apar spontan forme tautomere ale bazelor azotate prin migrarea unui proton în altă poziţie decât cea originală, care modifică proprietăţile de legare complementară dintre baze. Acestea nu sunt recunoscute însă de ADN polimerază şi prin urmare nu sunt corectate (Fig.6.4).

În afara mutaţiilor determinate de împerecherile necomplementare dintre baze azotate normale şi formele tautomere, uneori pot apărea spontan mici deleţii sau adiţii în timpul replicării ADN. Acestea apar datorită deplasării sau buclării bazelor, fie pe matriţa ADN, fie pe lanţul nou sintetizat, de regulă în regiuni în care apar baze azotate repetitive. În cazul în care buclarea se produce la nivelul matriţei, ADN polimeraza ocoleşte baza sau bazele azotate ieşite în afara catenei matriţă, producând deleţii. În cazul în care ADN polimeraza sintetizează baze azotate în plus, necomplementare cu matriţa, acestea se buclează în afară, producând adiţii. Deleţiile şi adiţiile de baze azotate produse prin aceste mecanisme în timpul replicării, determină decalarea cadrului de citire a mesajului genetic (mutaţii prin modificarea cadrului de citire a mesajului genetic).

Mutaţiile produse de ADN polimeraza pot fi reparate prin mecanisme celulare variate, care vor fi discutate în acest capitol.

Mutaţii spontane prin modificarea structurii chimice a ADN

Depurinarea şi dezaminarea bazelor azotate sunt două procese chimice care pot produce mutaţii spontane la nivel ADN.

Depurinarea constă în pierderea purinei de la nivelul nucleotidei, proces care se produce prin hidroliza legăturii chimice dintre bază azotată şi dezoxiriboză. Procesul are loc spontan deoarece legătura covalentă a purinei cu dezoxiriboza este mult mai slabă decât cea a pirimidinei cu aceeaşi pentoză, rezultatul fiind o nucleotidă apurinică. Aceste modificări se produc cu miile în genomul eucariot şi determină leziuni la nivel ADN, care dacă nu sunt reparate prin mecanisme interne, produc disocierea ADN polimerazei de la matriţă în timpul replicării ADN (Russell, P., 2010).

Dezaminarea constă în îndepărtarea grupării amino de la nivelul bazei azotate. De exemplu dezaminarea citozinei, prin pierderea grupării amino duce la transformarea ei în uracil, bază azotată din structura ARN. Dacă uracilul nu este înlocuit cu citozina prin mecanismele reparatorii interne, în ciclul următor de replicare în ADN va fi încorporată adenina, ceea ce poate duce la înlocuirea perechii C-G cu perechea T-A (mutaţie prin tranziţie) (Vlaic, 2011).

Altă mutaţie prin dezaminare poate să apară prin modificarea 5'-metilcitozinei, bază azotată modificată prezentă în cantităţi mici în ADN şi transformarea ei în timină. 5-metilcitozina (5^mC) apare în lanţul ADN în locul citozinei (G-5^mC) iar dezaminarea ei şi transformarea în timină va duce în final la înlocuirea perechii G-5^mC cu G-T. În ciclul următor de replicare, guanina de pe o catenă se va împerechea normal cu citozina în timp ce timina de pe cealaltă catenă se va împerechea cu adenina. Mutaţia produsă constă în tranziţia G-C în A-T. Atât depurinarea cât şi dezaminarea pot fi reparate prin mecanismul de excizie a bazelor azotate.

Mutaţiile induse

Mutaţiile induse sunt acele mutaţii care apar ca urmare a expunerii organismului la diferiţi agenţi mutageni de natură fizică şi chimică. Mutaţiile induse experimental sunt o resursă foarte valoroasă pentru studiul funcţiei genelor. Metoda, cunoscută sub denumirea de mutageneza *in vitro* presupune inducerea de mutaţii întâmplătoare sau dirijate în anumite situsuri din gene. Acestea sunt apoi înapoiate în celule bacteriene, sau în organismul donor, pentru studierea efectelor produse în urma procesului mutaţional (genetica inversă) (Vlaic, 1997).

Agenţii mutageni chimici

Agenţii mutageni chimici pot acţiona asupra moleculei de ADN în timpul replicării ADN, sau în oricare fază a ciclului celular. Din prima categorie fac parte analogii bazelor azotate şi agenţii intercalari, iar în a doua categorie sunt incluşi agenţii chimici de modificare a bazelor azotate (Brown, 2002).

Analogii bazelor azotate sunt similari bazelor azotate care intră în compoziţia ADN în mod natural, aici fiind incluşi 5 – bromuracilul, 5 – bromdeoxiuridina, fluoruracilul şi 2 – aminopurina. La fel cu bazele azotate din compoziţia ADN şi analogii lor pot apărea sub două forme tautomere sau izomere. Similitudinea foarte mare în structură, cu bazele normale, face ca şi încorporarea lor în ADN să se poată face în acelaşi mod. 5 – bromuracilul este un analog al timinei care prezintă în structură un reziduu de brom în locul grupării metil. În lanţul de ADN, acesta se leagă complementar de adenină sau, în forma sa izomeră rară se leagă de guanină. În cazul încorporării ca formă tautomeră, 5 – bromuracilul se va lega de guanină formând perechea 5BU-G, iar în ciclul următor de replicare va rezulta perechea C-G în locul perechii T – A. Aceasta este o mutaţie prin tranziţie de tip T-A în C-G, care poate să se producă şi în sens invers, C-G în T-A, dacă în prima etapă de replicare 5BU este în stare normală şi apoi trece în starea tautomeră rară, în cursul procesului de replicare. Mutaţiile

După amplificare, produsul PCR se vizualizează în urma electroforezei în gel de agaroză, obţinându-se profilele corespunzătoare secvenţei amplificate. În cazul în care primerii nu îşi găsesc regiuni complementare şi nu se ataşează la matriţă, amplificarea nu are loc iar în gel nu se vor distinge profile reprezentate prin benzi.

10.2 Secvenţierea acizilor nucleici

Secvenţierea acizilor nucleici presupune descifrarea naturii şi ordinii nucleotidelor dintr-un fragment de ADN.

Fragmentul de ADN supus secvenţierii este o monocatenă de ADN, însă în două reacţii separate pot fi descifrate ambele catene ale moleculei.

Tehnicile de secvenţiere se împart în trei categorii, diferenţiabile după principiul de evidenţiere a bazelor azotate în reacţie.

Prima metodă de secvenţiere, cunoscută cu denumirea de metoda terminatorilor de lanţ sau dideoxi (ddNTP) a fost dezvoltată de Fred Sanger în anul 1977. Principiul metodei se bazează pe întreruperea controlată a sintezei ADN cu ajutorul nucleotidelor terminatori de lanţ (ddNTP).

Cea de-a doua metodă, apărută aproape concomitent cu metoda terminatorilor de lanţ, are la bază principiul clivării chimice a nucleotidelor din catena de ADN aflată în reacţia de secvenţiere. Metoda a fost implementată de cercetătorii Allan Maxam şi Walter Gilbert în anul 1977 (1972,1973).

Pentru contribuţiile aduse la dezvoltarea acestor metode de descifrare a secvenţei acizilor nucleici, Fred Sanger şi Walter Gilbert au fost distinşi cu premiul Nobel în anul 1980.

Câţiva ani mai târziu aveau să apară şi primele echipamente de secvenţiere, complet automatizate. Atât primele modele de echipamente, cât şi cele automate s-au bazat pe principiul metodei terminatorilor de lanţ dezvoltat de Sanger.

În ultimii ani au apărut însă echipamentele de secvenţiere de mare capacitate, care pot descifra un genom într-o singură zi. Categoria metodelor dezvoltate pentru aceste echipamente a fost grupată sub denumirea de *Next Generation Sequencing* (NGS), echipamente care exploatează diferite principii de secvenţiere, inclusiv cel dezvoltat de Sanger.

Tehnologia NGS extinde principiul electroforezei capilare prin care bazele unui fragment de ADN sunt identificate secvenţial prin semnalele emise de un colorant fluorescent în timpul sintezei, la mii de reacţii de secvenţiere desfăşurate în parallel care produc în timp relativ scurt o secvenţă genomică pe fragmente. Dintre acestea vom prezenta în cele ce urmează pirosecvenţierea, metodă dezvoltată pentru echipamentul 454 al firmei Roche.

În continuare vor fi descrise principiile celor trei metode de secvenţiere dezvoltate până în prezent.

Metoda de secvenţiere a acizilor nucleici prin clivare chimică dezvoltată de MAXAM ŞI GILBERT deşi nu se mai utilizează în prezent, o descriem cu titlu informativ datorită valorii sale ştiinţifice.

Metoda de secvenţiere Maxam&Gilbert

Metoda de secvenţiere Maxam&Gilbert se bazează pe degradarea chimică a catenelor de ADN intrate în reacţia de secvenţiere. Cu ajutorul acestei metode a fost descifrată pentru prima dată secvenţa operonului *lac* de la *E. coli* şi mai apoi secvenţele genomului virusului SV 40 şi ale plasmidei pBR322.(Maxam şi Gilbert, 1977; 1980)

Ca etape de lucru, metoda porneşte de la izolarea materialului genetic reprezentat de un fragment de ADN dintr-un genom. La startul reacţiei de secvenţiere se porneşte de la populaţia unei catene care este tratată cu fosfataza alcalină pentru eliminarea radicalului 5' – fosfat. Defosfatarea are rolul de a permite marcarea radioactivă a capătului fragmentului de acid nucleic, cu unul din izotopii ^{32}P, ^{33}P. Populaţia monocatenei este împărţită în patru tuburi de reacţie, în fiecare tub existând câte un reactiv care îndepărtează (clivează) selectiv şi parţial o anumită bază azotată. În fiecare tub reactivii digeră catena de ADN la una sau două baze azotate specifice, rezultând fragmente de dimensiuni variabile la care ultima bază îndepărtată a fost, în exemplificarea din figura 9.5 guanina în tubul 1, adenina sau guanina în tubul 2, timina sau citozina în tubul 3 şi citozina în tubul 4. (Fig.9.5).

Cei patru reactivi, utilizaţi separat în cele patru tuburi de reacţie, produc degradarea selectivă şi parţială a uneia sau a două nucleotide diferite. Reactivul din tubul 1 produce o hidroliză mai blândă a guaninei, reactivul din tubul 2 produce metilarea guaninei în poziţia 7 şi a adeninei în poziţia 3 cu dimetilsulfat, iar piperidina distruge selectiv guanina. Reactivii

După reacţie are loc migrarea populaţiilor fragmentelor într-un gel de poliacrilamidă pentru separare pe baza greutăţii molecular iar la final se realizează autoradiografia gelului (Coşier, 2008).

Citirea şi interpretarea rezultatelor – Citirea secvenţei fragmentului supus secvenţierii se face de sus în jos, în partea superioară a autoradiografiei se regăsesc fragmentele de dimensiuni mai mari sau mai grele, iar în partea inferioară fragmentele mai uşoare. Citirea gelului se face simultan pe două câte două câmpuri de migrare (1 şi 2, respectiv 3 şi 4) pentru admiterea sau excluderea unei baze din şir.

Deoarece în fiecare din cele două tuburi există reactivi care distrug o bază azotată comună, de exemplu guanina în tubul 1 dar şi în tubul 2, respectiv citozina în tuburile 3 şi 4, prezenţa acestor două baze în secvenţa ADN este confirmată prin 2 benzi, câte una în fiecare câmp de migrare. În schimb prezenţa celorlalte două nucleotide, adenina şi timina este confirmată printr-o singură bandă în câmpul 2, respectiv 3 de migrare (Fig.10.2).

Metoda terminatorilor de lanţ (dideoxinucleotidele)

Spre deosebire de metoda Maxam şi Gilbert, despre care anterior menţionam că nu se utilizează în aplicaţiile echipamentelor de secvenţiere dezvoltate pe parcursul celor 40 de ani de când au fost puse la punct strategiile, metoda terminatorilor de lanţ sau a dideoxinucleotidelor dezvoltată de Sanger, a fost preluată şi implementată de toate companiile producătoare de echipamente de secvenţiere.

Primul genom la care s-a determinat succesiune în nucleotide prin această metodă, este cel al bacteriofagului φ X174. Acesta este format din ADN monocatenar circular, organizat în 10 gene şi 5386 nucleotide (Sanger, F. şi colab., 1977).

Principiul metodei Sanger se bazează pe întreruperea controlată a replicării enzimatice a ADN, prin utilizarea nucleotidelor terminatori de lanţ. Acestea sunt analogi ai nucleotidelor (dNTP) la care în poziţia 3' gruparea OH a dezoxiribozei este înlocuită cu H , prin urmare poziţia nu mai poate fi utilizată pentru ataşarea altei nucleotide (Fig.10.3).

În timpul sintezei ADN, după încorporarea unei ddNTP enzima blochează replicarea în acel punct, deoarece nu mai poate adăuga o nouă nucleotidă. La startul reacţiei de secvenţiere sunt generate prin PCR

populaţiile unei singure catene care intră în reacţie, sau secvenţierea se poate produce direct de pe o moleculă vectoare (fagul M13) în care este inclus fragmentul de descifrat.

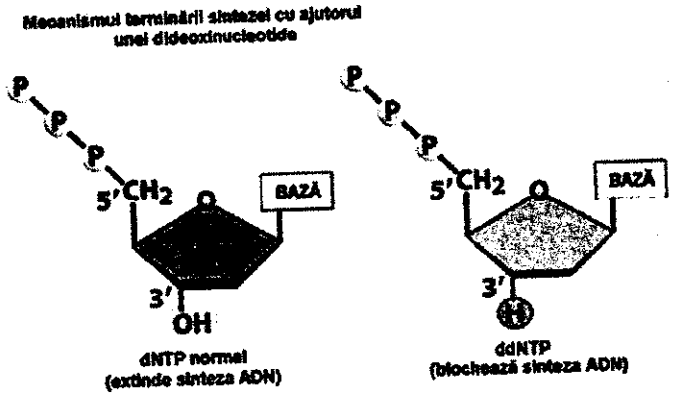


Fig.10.3 Structura unei dideoxinucleotide (ddNTP) comparativ cu o dezoxinucleotidă (dNTP)

Aşadar reacţia de secvenţiere are la bază replicarea ADN monocatenar produs prin reacţia PCR sau prin clonare în vectorul fagic M13, un primer universal, patru deoxinucleotidtrifosfaţi: dNTP (dATP, dGTP, dCTP şi dTTP) şi ¼ din analogii dNTP-urilor – dideoxinucleotidtrifosfaţi sau ddNTP (ddATP, ddGTP, ddCTP şi ddTTP) (Sanger, F., 1980, citat de Coşier, 2008).

Producerea populaţiei monocatenare de ADN care intră în reacţia de secvenţiere poate fi realizată prin mai multe strategii. Prima şi cea mai utilizată strategie exploatează una din variantele tehnicii PCR, denumită PCR Asymetric, în care se multiplică populaţia unei singure catene. Cealaltă strategie are la bază o reacţie PCR clasică, în care produsul PCR dublu catenar este supus unei reacţii de digestie a uneia din catene cu exonucleaza III, începând cu capătul 3' (tehnica pusă la punct de A.J. H Smith).

Altă variantă de obţinere a ADN monocatenar este clonarea fragmentelor de ADN în vectori fagici şi secvenţierea direct de pe molecula vectorului. Până în prezent au fost construiţi mai mulţi vectori care permit multiplicarea unei singure catene de ADN, pornind de la promotori fagici

(M13, SP6). Aşa de exemplu, pentru clonarea fragmentelor de ADN mitocondrial uman în vederea secvenţierii, Sanger a utilizat vectori derivaţi din bacteriofagul M13, construiţi de către Gronenborg şi Messing (1978), care conţin gena β-galactozidazei ca marker de selecţie şi un situs de restricţie pentru enzima *EcoR*I. Selecţia coloniilor *E.coli* transformate cu aceşti vectori se face după culoarea alb-albastră, în prezenţa indicatorului X-Gal. Din coloniile recombinate se extrage ADN–ul plasmidial şi molecula este supusă secvenţierii integral (Coşier, 2008).

Unul din avantajele secvenţierii direct de pe molecula vectoare este exploatat la descifrarea clonele librăriilor genomice, când secvenţierea se poate face direct din librărie fără să fie nevoie de recuperarea fragmentului din vector. Alt avantaj al metodei este faptul că sinteza decurge cu un primer universal generat din situsul de restricţie al vectorului (Heidecker şi colab., 1980, citat de Coşier, 2008). Pentru secvenţiere direct de pe molecula vectorului acesta se deschide la nivelul situsului unic al enzimei *Eco*RI unde a fost clonat fragmentul de ADN, de mărime circa 200 pb. Prima etapă constă în fracţionarea întregului ADN genomic cu enzime de restricţie, care sunt apoi migrate în gel de agaroză pentru separare pe baza greutăţii molecular, iar fragmentele de dimensiuni mici sunt recuperate şi clonate în vectorul M13 de unde sunt recuperate şi secvenţiate. (Sanger, F. şi colab., 1980; Barell, B.G. şi colab., 1980). În cazul în care se doreşte secvenţierea unui fragment particular de ADN, acesta este amplificat prin PCR şi clonat în vectorul M13, după care strategia decurge în mod similar.

Principiul secvenţierii dideoxi

În reacţia Sanger ADN polimeraza determină iniţierea sintezei ADN imediat după primer, adăugând pe baza de complementaritate un dNTP şi uneori câte un dezoxinucleotidtrifosfat (ddNTP), deoarece concentraţia acestora în tubul de reacţie este mai mică decât a nucleotidelor normale. În fiecare tub se includ toţi cei patru dNTP şi un singur ddNTP. În timpul sintezei ADN, la un moment dat polimeraza introduce analogul unui dNTP (de exemplu, ddATP) care va fi încorporat în locul nucleotide normale dATP şi va bloca sinteza *de novo* a catenei în acel punct. Dideoxinucleotidele (ddNTP) încorporate aleator vor stopa sinteza datorită faptului că fiecare ddNTP are în poziţia 3' gruparea H şi nu 3'-OH (Fig.10.3).

Ca rezultat al amplificării se va produce un fragment mai scurt de ADN, care se termină întotdeauna cu un ddNTP. În mod similar, reacţiile de sinteză şi de întrerupere au loc în toate cele patru tuburi de reacţie. Diferenţa între fragmentele din cele patru tuburi este aceea că fragmentele se termină

cu alt ddNTP. Migrând produşii de reacţie din cele patru tuburi în gel de poliacrilamidă, în fiecare câmp de migrare se va obţine numărul de benzi corespunzător numărului de nucleotide de acelaşi tip din catena supusă secvenţierii. Aşa de exemplu, de la un fragment de ADN care conţine să spunem 10 adenine în poziţii diferite ale matriţei se vor obţine 10 fragmente de lungimi diferite care toate se termină în ddTTP (Fig.10.4).

Metoda convenţională dezvoltată de Sanger utiliza marcarea radioactivă a capătului 5'. Deşi a fost exploatată mai bine de 20 de ani, la scurt timp după punerea ei la punct a fost îmbunătăţită considerabil prin utilizarea marcării fluorescente a dideoxinucleotidelor, electroforeaza capilară a fragmentelor de ADN şi automatizarea întregii proceduri de secvenţiere (Fig.10.5 a).

Instrumentele de secvenţiere automată a ADN decodifică secvenţa nucleotidică sub forma unei electroforegrame, după o electroforeză specială realizată în capilare fine în care este încărcat polimerul (Fig.9.8). Reacţia de decodificare are loc într-un singur tub de reacţie în care sunt prezenţi toţi deoxinucleotidtrifosfaţii şi toţi cei patru dideoxinuclotidtrifosfaţi (fiecare marcat fluorescent cu alt fluorofor care produce coloraţii diferite).

Enzima utilizată în reacţia de scevenţiere este o polimerază specială, cu capacitate ridicată de procesare a ADN în ceea ce priveşte încorporarea atât a dNTP-urilor cât şi addNTP-urilor şi cu fidelitate ridicată în sinteză.

Prin marcarea fluorescentă a dideoxinucleotidelor, fiecărei tip de bază azotată îi va corespunde altă coloraţie în electroforegramă, astfel adenina apare colorată în verde, timina în roşu, guanina în negru şi citozina în albastru) (Smith, 1986) (Fig.10.5B).

Platforma de secvenţiere *Next Generation Sequencing* -Pirosecvenţierea

Eforturile pentru descifrarea genomului uman au fost imense, privite nu doar prin prisma resurselor financiare utilizate dar mai ales a resurselor umane implicate în proiect. Deşi proiectul s-a desfăşurat în mai multe laboratoare simultan, timpul necesar pentru descifrarea completă a fost de 13 ani. Totul a depins de tehnologia disponibilă la acea vreme, dar mai ales de metoda de secvenţiere. Aşa cum am mai arătat descifrarea genomului uman s-a desfăşurat în paralel în două proiecte, unul finanţat din surse guvernamentale, condus de consorţiul HUGO care a utilizat metoda de secvenţiere ierarhică sus-jos (între anii 1990-2003) şi cel de-al doilea iniţiat de compania privată Celera genomics care a abordat studiul de secvenţiere a întregului genom prin metoda *shotgun* şi a finalizat proiectul după patru ani