

3 CROMOZOMUL LA ORGANISME PRO- SI EUCARIOTE

3.1 CROMOZOMUL BACTERIAN

3.1.1 Dimensiune

Imensa majoritate a bacteriilor dețin ca material genetic esențial o moleculă ADN dublu catenar circular covalent închis, suprarăsucit negativ și împachetat. Deși complexarea cu proteine și împachetarea nu este identică cu cea din cromozomii de tip eucariot, totuși această moleculă este denumită tot “cromozom” sau *nucleoid*. Cromozomul bacterian este atașat la membrana plasmatică a celulei bacteriene în aproximativ 20 de puncte, dar o funcție deosebită o are punctul de atașare de lângă regiunea de origine a replicării acestei molecule de ADN (*ori C*).

Cei mai mici cromozomi de tip procariot măsoară mai puțin de 1 Mpb și se întâlnesc mai ales la bacteriile fără perete celular (bacteriile din genurile *Mycoplasma*, *Ureaplasma* au cromozomi cu dimensiuni cuprinse între 600 și 800 kpb). La cealaltă extremă se află bacteriile din genurile *Mycococcus* și *Calothrix*, care au cromozomi foarte mari (aproximativ 12-13000 kpb). *Escherichia coli* are un cromozom de dimensiune intermediară: 4700 kpb.

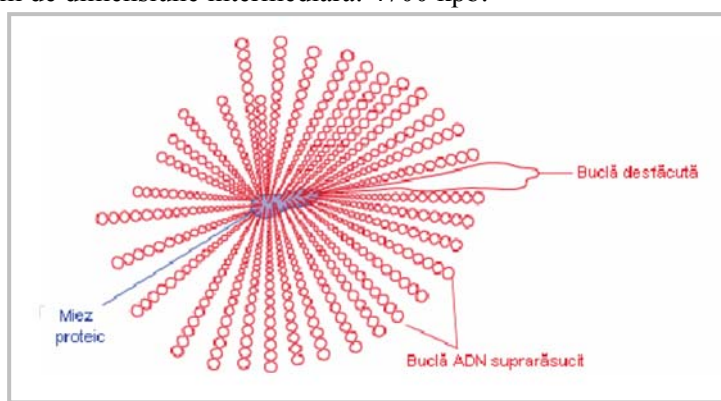


Figura 3.1 Modelul Pettijohn de structură a nucleoidului din bacteriile *Escherichia coli*. Între 40 și 50 de bucle suprarăsucite radiază dintr-un miez proteic (după Brown, 2002).

3.1.2 Modelul Pettijohn

În general, molecula ADN ce formează cromozomul bacterian este de aproximativ 1000 de ori mai lungă decât celula bacteriană. Această moleculă este complexată cu proteine, suprarăsucită și împachetată formând o structură conformă cu modelul elaborat de **Pettijohn** în 1974 (Figura 3.1). Conform acestui model, prin asocierea ADN cu proteine și cu ARN (de regulă, ARN nascent) se formează aproximativ 50 de domenii topologice, semi-independente (denumite și bucle), per genom de *E.coli*. Fiecare domeniu (bucă) este suprarăsucit separat de celelalte domenii și poate fi relaxat independent de celelalte prin introducerea unei rupturi monocatenare. Gradul de suprarăsucire negativă este dat de balanța dintre activitatea a doua enzime - ADN giraza și ADN topoizomeraza I – amândouă reglând densitatea helicală a moleculelor de ADN. ADN giraza crește numărul de spire per kilopereche de baze azotate (suprarăsucește molecula de ADN), iar ADN topoizomeraza I scade numărul de spire.

3.1.3 Configurații ADN în cromozomul bacterian

Principalul tip de structură secundară a ADN ce formează cromozomul bacterian este **forma B de dreapta**, dar această configurație nu se găsește uniform de-a lungul întregului cromosom bacterian. Alte configurații, prezente pe distanțe scurte, sunt ADN-Z (în dublu helix de stânga), structuri cruciforme (în regiunile cu secvențe invers repetate) și regiuni triplu-catenare (în zone extrem de bogate în purine sau pirimidine).

Este de reamintit faptul că, în general, expresia genică este afectată nu numai de secvența primară a ADN (succesiunea de nucleotide), ci și de structura secundară și terțiară a acestuia.

3.1.4 Compoziția în nucleotide

În majoritatea cazurilor, compoziția globală în nucleotide a unei molecule de ADN se exprimă prin procentul molar de guanină + citozină (%mol GC), restul până la 100% fiind, evident, reprezentat de adenină + timină (datorită criteriului de complementaritate între cele două catene ADN, respectiv între bazele purinice și cele pirimidinice). Compoziția în nucleotide a cromozomului bacterian variază în limite foarte largi: 25 %mol GC la *Mycoplasma capricolum* și 75 %mol GC la *Micrococcus luteus*.

Procentul molar de guanină – citozină în cromozomul bacterian

Procentul molar de guanină + citozină este corelat și cu compoziția în codoni a genomului. Aceasta variază în sensul preferinței pentru 1-2 codoni dintr-un grup de codoni sinonimi. Se constată astfel că la *Mycoplasma capricolum* este favorizată prezenta A/T în poziția 3-a a codonilor sinonimi.

S-a mai constatat că în cromosomul bacterian există și o serie de gene a căror activitate poate afecta compoziția totală în nucleotide a unei molecule de ADN. Astfel, la *E.coli* au fost descrise două gene (*mut T* și *mut Y*) care afectează frecvența transversiiilor A-T / C-G și, respectiv, C-G / A-T. Echilibrul între exprimarea acestor două gene afectează compoziția globală în nucleotide a moleculei de ADN ce reprezintă cromosomul de *E.coli*.

O altă problemă o reprezintă contextul de citire a unui anumit codon, în speță, configurația codonilor adiacenți. Numărul teoretic posibil de codoni (sens) adiacenți este foarte mare ($612 = 3721$), dar, examinând 237 de gene de la *E.coli* s-a constatat că perechile de codoni adiacenți nu sunt distribuite randomizat, ci anumite perechi de codoni sunt mai abundente decât altele. S-a dedus astfel că, compoziția în nucleotide a unui codon este corelată și cu compoziția codonilor adiacenți, corelat probabil cu procesele de atașare la situsurile A (*Aminoacil*) și P (*Peptidil*) ale ribozomilor. Se pare deci, că aparatul de traducere a informației genetice ar fi putut să determine evoluția unor anumite trăsături ale matricei genetice.

3.1.5 Proteine asociate cu cromozomul bacterian

În afară de ARN polimeraza, care datorită ratei înalte de transcriere la procariote, rămâne asociată cvasi-permanent cu moleculele de ADN, cromosomul bacterian este asociat cu o serie de proteine bazice, similare cu histonele de la organismele eucariote, denumite **proteine histone-like**.

La *E.coli* au fost descrise 9 specii moleculare de proteine bazice, cu greutate moleculare între 9 și 28 kd, care se atașează la ADN într-o manieră situs-nespecifică și care au funcții similare cu histonele de la organismele eucariote. Dintre cele 9 specii moleculare proteice, cea mai importantă este o proteină denumită **HU** (« Helix–Unwinding » = desfacerea dubului helix). Din complexarea moleculelor HU cu ADN cromozomal bacterian se formează structuri similare cu nucleosomii de la

eucariote, denumite **structuri nucleosom-like** (chiar și în condiții *in vitro*), în care sunt cuprinse circa 200 pb/nucleosom. Structurile *nucleosom-like* nu sunt statice, ci se află într-un echilibru dinamic.

Studii mai ample au demonstrat faptul că proteina HU nu este o simplă proteină structurală a arhitecturii cromozomului bacterian, ci are rol foarte important în diverse procese din celula bacteriană – în mod special în procese ce implică interacțiuni ADN-proteine, situație în care moleculele HU favorizează atașarea altor proteine la ADN.

Deși proteinele *histone-like* au fost studiate extensiv la *E.coli*, totuși au fost identificate proteine cu funcții omoloage și la alte genuri și specii bacteriene. Pe de altă parte, s-a constatat că o serie de microorganisme din grupul *Archaea* prezintă proteine *histone-like* cu structură intermediară între cele de la *Bacteria* și histonele de la *Eukarya*.

Figura 3.2 Organizarea genomului la câteva specii de procariote.

Denumirea speciei	Molecule ADN	Dimensiune (Mb)	Numărul de gene
<i>Escherichia coli</i> K-12	1 moleculă circulară	4.639	4397
<i>Vibrio cholerae</i> El	2 molecule circulare		
	1 cromozom	2.961	2770
	1 megaplasmid	1.073	1115
<i>Deinococcus radiodurans</i> R1	4 molecule circulare		
	Cromozom 1	2.649	2633
	Cromozom 2	0.412	369
	Megaplasmid	0.177	145
<i>Borrelia burgdorferi</i> B31	Plasmid	0.046	40
	7 – 8 molecule circulare		
	11 molecule lineare		
	Cromozom linear	0.911	853
	Plasmid circular cp9	0.009	12
	Plasmid circular cp26	0.026	29
	Plasmid circular cp32	0.032	necunoscut
	Plasmid linear lp17	0.017	25
	Plasmid linear lp25	0.024	32
	Plasmid linear lp28-1	0.027	32
	Plasmid linear lp28-2	0.030	34
	Plasmid linear lp28-3	0.029	41
	Plasmid linear lp28-4	0.027	43
	Plasmid linear lp36	0.037	54
Plasmid linear lp38	0.039	52	
Plasmid linear lp54	0.054	76	
Plasmid linear lp56	0.056	necunoscut	

Proteina HU

- este cea mai importantă proteină *histone-like* de la *E.coli*
- au fost descrise proteine tip HU la toate bacteriile studiate
- ca și histonele de la organismele eucariote, proteina HU se atașează la molecule de ADN d.c. indiferent de secvența de nucleotide, deci situs - nespecific
- are un caracter bazic și o greutate moleculară de 9,7 kd
- funcționează ca dimer
- un monomer este format din 2 polipeptide, denumite **HU- α** și **HU- β** și codificate de gene diferite: ***hup A*** și, respectiv, ***hup B***.
- are proprietăți fizice și compoziție totală în aminoacizi similară cu histonele de la eucariote, dar secvența de aminoacizi este diferită de acestea.
- ca și histonele de la eucariote, proteina HU are un grad înalt de conservare în lumea bacteriană
- există aproximativ 25000 de molecule HU per genom de *E.coli*, dar aceste molecule nu sunt dispuse uniform de-a lungul cromozomului bacterian, ci sunt mai dense la periferia nucleoidului (în zonele extrem de active transcripțional)

Proteina IHF

Alte proteine *histone-like* de la *E.coli* sunt proteina H (similară cu histona H2A de la eucariote), proteina H1, proteina H-NS, proteina IHF.

Proteina IHF (*Integration Host Factor*, g.m.=20 kd)

- ✓ a fost descoperită în studiile privind integrarea genomului fagului □ în cromozomul de *E. coli* ; ulterior, s-a constatat că această proteină intervine și într-o serie de alte procese din celula bacteriană
- ✓ este formată din 2 subunități codificate de genele *him A* și *him B*.
- ✓ spre deosebire de alte proteine *histone-like*, moleculele de IHF se atașează la ADN într-o manieră situs-specifică (adică doar la anumite secvențe de nucleotide)
- ✓ moleculele de IHF nu sunt necesare pentru recombinarea bacteriană omoloagă, dar par să intervină în fenomene de recombinare neomoloagă, specializată, cum sunt de altfel, procesele de transpoziție și de conjugare bacteriană

3.1.6 Secvențe ADN repetate

În cromosomul bacterian au fost identificate secvențe ADN repetate, dintre care unele nu codifică proteine/ARN (secvențe repetate necodificatoare), iar altele codifică o serie de proteine sau specii moleculare de ARN (ARNr, ARNt) – secvențe repetate codificatoare.

3.1.6.1 Secvențe repetate necodificatoare = ADN repetitiv

Acestea sunt reprezentate de secvențe scurte ce apar de regulă în afara ORF-urilor (*Open Reading Frames*). Marea majoritate a acestor secvențe servesc ca situsuri de interacțiune ADN-proteine (inclusiv în procese de recombinare, inversie, excizie, transpoziție) și ca situsuri modificate ce identifică catena matrită în timpul replicării semiconservative a cromozomului bacterian. Dintre cele mai cunoscute asemenea secvențe sunt secvențele *REP*, *Chi*, *Dam*.

Secvențe REP (*Repeated Extragenic Palindromes*)

Sunt formate din palindroame repetate, majoritatea aflate în regiuni extragenice. O secvență REP conține o secvență consensus de 38 bp. Grupuri (denumite “clusteri”) de 2-4 secvențe REP separate între ele prin 20 pb formează un **element REP** (**REPE** = *REP Element*). La *E.coli* și *Salmonella typhimurium* există 100-200 structuri REPE (reprezentând deci, aproximativ 0.5% din cromozomul bacterian), localizate însă diferit. La secvențele REP se leagă molecule de proteină HU și

de ADN girază, aceste secvențe având deci rol în împachetarea nucleoidului. Tot la aceste secvențe se leagă și ADN polimeraza intervenind în procese de replicare și reparare ADN.

Situsuri Chi Un situs *Chi* are o lungime de 8 pb:

5' GCTGGTGG 3'

Datorită faptului că reprezintă situsul de recunoaștere și tăiere a moleculelor ADN de către **endonucleaza RecBCD**, secvențele *Chi* stimulează recombinarea omoloagă mediată de **RecA** și **RecBCD**. S-a mai constatat că efectul recombinativ al situsului *Chi* este polar: stimulează recombinarea la capatul 5' al lui *Chi*, și nu la 3'. Situsuri *Chi* au fost identificate atât pe cromosom, cât și pe unele plasmide de la *E.coli*. Mai mult chiar, s-a constatat că secvențele *Chi* reprezintă situsuri de stimulare a recombinării genetice la toți membrii familiei *Enterobacteriaceae*. În contrast, la *Pseudomonadaceae*, complexul enzimatic RecBCD nu acționează la situsuri *Chi*.

În cromosomul de *E.coli* au fost identificate aproximativ 950 de situsuri *Chi*, ce par să fie distribuite relativ uniform (1 situs *Chi* la 5 kbp), cu excepția zonei adiacente regiunii *oriC*, unde există un cluster de 22 de situsuri *Chi*.

Situsuri Dam sunt formate din 4 bp (5' GATC 3') și reprezintă situsurile de metilare a adeninei (adenina este metilată la poziția N6). Cromosomul de *E.coli* conține foarte multe situsuri *Dam* (peste 18.000) care, dacă ar fi dispuse randomizat ar există 1 situs *Dam*/250 bp. Și în acest caz s-a constatat un număr foarte mare de situsuri *Dam* în, și în jurul, regiunii *oriC*: în cele 245 bp ce definesc *oriC* la *E.coli* există 8 situsuri *Dam*, iar în cele 350 bp ce flanchează *oriC* există încă 12 situsuri *Dam*.

3.1.6.2 Secvențe repetate codificatoare

Asemena secvențe au dimensiuni mult mai mari decât cele necodificatoare.

operoni *rrn* există la toate bacteriile și cuprind secvențe ce codifică pentru molecule de ARN ribozomal (ARNr). La cele mai multe dintre speciile bacteriene, *linkage*-ul (ordinea) în cadrul unui operon *rrn* este 16S – 23S – 5S. *E.coli* și *S.typhimurium* au câte 7 loci (operoni) *rrn*, localizați în poziții echivalente. Cei 7 operoni sunt notați *rrnA...rrnG*, fiecare cuprinzând secvențe 16S-23S-5S cotranscrise într-o moleculă ARN de 30S. Numarul operonilor *rrn* variază de la un gen/specie bacteriană la alta: *Bacillus subtilis* are 10 operoni *rrn*, *Mycobacterium* sp. are 1-2 operoni *rrn*. Chiar în cadrul aceleiași tulpini bacteriene, locii *rrn* pot diferi între ei în ceea ce privește prezența/absența genelor pentru ARNt în regiunile “*spacer*” dintre genele ARNr. Astfel, toți cei 7 operoni *rrn* de la *E.coli* au 1-2 gene pentru ARNt între secvențele pentru 16 și 23S, și 0-2 gene ARNt după 5S.

Distribuția locilor *rrn* pe cromosom variază la diverse specii bacteriene: la *E.coli*, majoritatea locilor *rrn* sunt localizați în jumătatea cromosomală ce are în centru regiunea *oriC*; la *B.subtilis*, locii *rrn* sunt grupați într-o zonă ce reprezintă 30% din cromosom.

genele pentru ARNt

Cromosomul de *E.coli* conține 41 de gene/operoni pentru ARNt distribuiți în tot cromosomul. Unii din acești operoni codifică pentru o moleculă de ARNt, alții pentru mai multe molecule de ARNt, iar alții au și secvențe ce codifică diverse proteine.

secvențe *rhs* (“*rearrangement hot-spots*”) sunt capabile să genereze duplicații genice prin procese de crossing-over inegal între secvențe repetate. La *E.coli* K-12 au fost identificate 4 secvențe *rhs*, notate *rhs A, B, C și D*. Fiecare din aceste 4 secvențe au dimensiuni între 8 și 9 kbp și sunt compuse dintr-o regiune “*core*” (*miez*) (cu secvență conservată de aproximativ 3700 bp) și din segmente flancatoare. Prin exprimarea regiunilor “*core*” ale secvențelor *rhs A, B și C* sunt sintetizate 2 proteine cu funcție deocamdată necunoscută. În contrast cu *E.coli* K-12, la *E.coli* B și *E.coli* C, precum și la *S.typhimurium*, nu au fost descrise secvențe *rhs*.

3.1.7 Numărul de copii genice

La bacterii, în afară de genele sau operonii aflați în copii multiple, numărul de copii ale unei gene cromosomale poate varia în funcție de următorii factori:

1. *poziția genei pe cromosom, de-a lungul unui gradient pornind de la oriC spre regiunea ter.* În timpul replicării cromosomului bacterian, în funcție de poziția lor pe cromosom, unele gene sunt deja replicate, în timp ce altele încă nu. Astfel, unele gene se găsesc într-un număr mai mare de copii decât altele.

2. *copii ale unei gene cromosomale pot exista și pe plasmide.*

În această situație se folosește termenul de *meroploidie parțială*, care se referă la o ploidie parțială ce afectează anumite gene, fără să existe multipli de cromosomi întregi (este cazul genelor prezente și pe cromosom și pe un plasmid).

3. *într-o celulă bacteriană pot exista mai multe copii ale întregului cromosom.*

La anumite specii bacteriene există mai multe copii cromosomale în aceeași celulă: *Desulfovibrio vulgaris* (4 copii), *D.gigas* (17 copii), *Azotobacter chroococcum* (20-25 de copii), *A.vinelandii* (40-80 de copii cromosomale). Este de subliniat faptul că la majoritatea speciilor bacteriene cu copii cromosomale multiple există procese de inactivare cromosomală ce asigură funcționarea unui singur cromosom și inactivarea celorlalți.

3.1.8 Densitatea informației în cromosomul bacterian

Densitatea informației biochimice, adică numărul de reacții biochimice catalizate per kb de material genetic, este un parametru extrem de variabil în genomul bacterian, în primul rând datorită faptului că nu întotdeauna genele au o relație 1-la-1 cu reacțiile biochimice. În mod uzual, este valabilă relația 1 genă => 1 polipeptidă => 1 reacție biochimică. Există însă și foarte multe excepții, de exemplu:

- *enzime formate din mai multe lanțuri polipeptidice:* de exemplu, enzima succinat dehidrogenaza este formată din 4 polipeptide codificate de 4 gene diferite – *sdh A, B, C și D*. În acest caz, relația este 4 gene => 1 reacție biochimică;

- *enzime polifuncționale, ce catalizează mai multe reacții biochimice:* de exemplu, complexul enzimatic FAD ce oxidează acizii grași este format din 2 polipeptide codificate de 2 gene diferite – *fadA și fadB*. Acest complex enzimatic catalizează 5 reacții biochimice (din care 4 sunt catalizate de **FadA**). În acest caz, relația este 1 genă => 4 reacții biochimice.

Gene suprapuse

În cadrul genomului bacterian variază și densitatea de citire a informației genetice prin transcriere. Astfel, deși marea majoritate a genelor bacteriene sunt contigue, totuși cromosomul bacterian cuprinde și gene cu diverse grade de suprapunere:

- *gene cu promotor plasat în regiunea terminală a genei "upstream":* **trpA-trpB, ilvA-ilvD**
- *gene cu promotor plasat în interiorul genei "upstream":* **mioA-mioD**
- *gene cu grad ridicat de suprapunere:*

1. **gene suprapuse total**, codificate pe cele 2 catene ale ADN; de exemplu, locusul **CysE** (codifică serin-acetil-transferaza), ce este implicat în biosinteza cisteinei, are 2 ORF (**cys X** și **cys E**) codificate pe cele 2 catene ale locusului **CysE**.

2. **secvențe codificatoare ce sunt folosite de mai multe ori**, pornind transcrierea din poziții diferite; de exemplu, locusul **McrB** (codifică proteine implicate în restricția ADN la 5-metil-citozină) conține 3 ORF-uri pe aceeași catenă ADN, pornind din 3 situsuri diferite de inițiere a transcrierii. Traducerea celor 3 transcripse duce la formarea a 3 proteine (de 51, 53 și, respectiv, 54 kdal) cu secvență carboxi-terminală identică, dar diferită în regiunea amino-terminală.

3. **gene suprapuse total**, cu molecule transcript identice, dar traduse diferit prin procese de "frameshift" translațional: de exemplu, gena **trpR** care codifică 2 proteine și gena **dnaX**, care codifică 2 subunități ale ADN polimerazei III.

Asemenea gene – cu grad ridicat de suprapunere – sunt foarte bogate în informație genetică (unele au chiar informație dublă). Multe din ele (de exemplu, genele **mcrB** și **dnaX**) nu sunt tipice pentru cromosomul de *E.coli*, în ceea ce privește procentul molar de guanină + citozină. Astfel, gena **mcrB** are 48% mol GC, iar gena **dnaX** are 58 % mol GC, în timp ce cromosomul de *E.coli* are, *in toto*, 51-51.5 % mol GC. S-a emis ipoteza că asemenea gene care prezintă un procent molar de guanină + citozină extrem de diferit față de media cromosomală, ar fi fost achiziționate de *E.coli* din alte genomuri, prin procese de transfer de material genetic pe orizontală (transformare, conjugare, transducție).

3.1.9 Redundanța genelor și a produselor genetice

Pentru foarte multe funcții celulare, există câte 2 gene, ca și cum programul genetic al acestor bacterii ar avea sisteme de “*backup*” (*copii de siguranță*). Această redundanță ar putea fi considerată ca fiind o densitate redusă a informației genetice. Cu toate acestea, cel puțin în unele cazuri, această informație este reglată diferit și folosită în scopuri metabolice diferite.

În cromosomul de *E.coli* există 58 de perechi (de câte 2 gene) sau clusteri (mai mult de 2 gene), ce codifică 125 de produse genice. Din acestea, în 56 de perechi/clusteri reacțiile catalizate sunt identice. În cazul anumitor grupuri genice, reglajul exprimării lor este diferit, de exemplu:

- genele *aroF*, *aroG* și *aroH* codifică, toate trei, o aceeași enzimă: 3-deoxi-D-arabino-heptulosonată-7 – fosfat-sintetaza (DAHPaza). Enzima codificată de *aroF* este sensibilă la concentrația de tirozină, cea codificată de *aroG* este sensibilă la concentrația de fenil-alanină, iar cea codificată de *aroH* este sensibilă la concentrația de triptofan ;

- genele *glpA* și *glpD* codifică amandouă glicerol-3-fosfat-dehidrogenaza. Enzima codificată de *glpA* este sintetizată și folosită în condiții de anaerobioză, iar cea codificată de *glpD* – în condiții de aerobioză.

Unele din aceste gene pereche au secvență similară una cu cealaltă (și proteinele corespunzătoare au secvență similară în aminoacizi) și este probabil că au apărut prin procese de duplicație genică, urmată de o oarecare divergență funcțională.

Altele însă, deși enzimele desfășoară aceeași funcție, totuși au secvență diferită (atât ca proteine, cât și genele corespunzătoare). Este posibil ca asemenea gene să fi apărut fie prin evoluție convergentă, fie prin transfer lateral de la alți repliconi bacterieni (această ultimă situație poate fi identificată prin determinarea procentului molar de guanină + citozină).

3.1.10 Situsuri de inserție fagică în cromosomul bacterian

Mulți bacteriofagi își inseră genomul în cromosomul bacterian, fie prin transpoziție în poziții randomizate în cromosom (bacteriofagul **Mu**), fie prin recombinare situs-specifică numai în anumite poziții (fagii **lambdoizi** ce constituie fagii înrudiți cu fagul λ). Situsurile cromosomale în care fagul **lambda** (și fagii *lambdoizi*) își inseră genomul, sunt denumite situsuri **attB** (“*attachment on bacteria*”) și corespund unor situsuri **attP** (“*attachment on phage*”) pe genomul fagic. Fagul lambda necesită un situs **attB** de minimum 21 bp, din care 15 sunt identice cu 15 bp din **attP** (234 bp). Lambda are un situs preferat **attB** în cromosomul de *E.coli* K-12 (dar și situsuri secundare) ce e situat intergenic. Alți fagi lambdoizi (**21**, **e14**, **P22**) se inserează în situsuri attB cu localizare intragenică. Pe cromosomul de *E.coli* K-12 au fost cartate situsurile pentru 15 fagi lambdoizi și s-a constatat că ele sunt grupate între pozițiile 6 min și 44 min.

Această organizare a situsurilor în care se inseră genomul fagilor lambdoizi reprezintă încă un argument în favoarea ipotezei conform căreia cromosomul de *E.coli* (și nu numai) ar avea o origine himerică, unul din segmente derivând dintr-o gazdă ancestrală a fagilor lambdoizi.

Pe de altă parte, s-a constatat că foarte multe elemente genetice cu caracter “mobil” se inseră în genele pentru ARN_t, fapt pentru care s-a sugerat că acestea ar fi reprezentat situsurile folosite de fagii ancestrali.

3.1.11 Localizarea cromosomală a genelor înrudite fiziologic

În cromosomul de *E.coli*, majoritatea genelor (operonilor) par să fie distribuiți randomizat (indiferent de înrudirea lor fiziologică), cu excepția celor implicați în catabolismul glucozei ce par să fie dispuși în 4 zone echidistante pe harta circulară a cromosomului. Această randomizare a dispunerii operonilor a fost verificată și prin rearanjamente cromosomale induse experimental – caz în care operonii au continuat să funcționeze normal.

Totuși, în treimea cromosomală ce are în centru regiunea **ter** (ca, de altfel, și în regiunea **oriC**), poziția operonilor pare să fie destul de fixă, orice rearanjament deranjând total funcționarea acestora.

Orice aranjament nerandomizat al genelor duce la ipoteza că poziția genelor ar putea afecta funcționarea lor, deducându-se astfel o organizare mai înaltă decât operonul sau regionul.

La cele mai bine studiate specii de *Pseudomonas* (*P.aeruginosa*, *P.putida*) s-a constatat că genele cu funcții “housekeeping” (gene implicate în metabolismul central al oricărei celule, precum și în metabolismul anabolic) sunt grupate în jumătate din harta circulară a cromosomului. În același timp, genele implicate în metabolismul catabolic par să fie dispuse randomizat în cealaltă jumătate a cromosomului bacterian.

Acest mod de dispunere a genelor ar putea reflecta istoria construcției cromosomului din părți separate: genele pentru metabolismul esențial ar fi existat probabil în acel genom ancestral mai mic, care a fost ulterior mărit la actuala dimensiune prin fuziune cu un alt element genetic sau prin transfer de material genetic (transformare, conjugare, transducție).

În general, s-a constatat că la foarte multe bacterii clusterii genici din jurul regiunilor *oriC* și *ter* sunt foarte similari între taxoni bacterieni foarte îndepărtați filogenetic, fapt ce susține ipoteza că inițierea și terminarea replicării cromosomului bacterian sunt procese extrem de vechi și sunt coordonate prin mecanisme similare la marea majoritate a bacteriilor.

3.2 PLASMIDE BACTERIENE

În prezent, plasmidele sunt definite ca fiind repliconi neesențiali, capabili să se replice fizic independent de cromozom și transmiși în formă extracromosomală, pe verticală (de la o generație de indivizi la alta) sau pe orizontală (de la un individ la altul în cadrul aceleiași generații). În general, plasmidele reprezintă un fenomen al lumii procariote. Cu toate acestea, au fost descrise structuri genetice plasmidiale și la unele grupuri de eucariote inferioare (drozdii, fungi, mușcagăiuri) și chiar și la eucariote superioare (plasmidele din mitocondrii și cloroplaste la unele celule vegetale).

Există cercetători care încadrează plasmidele în categoria organismelor acelulare. În contrast cu această teorie, alți cercetători consideră plasmidele ca nefiind organisme, ci alături de virusuri, transposoni și secvențe de inserție, ar aparține unor specii moleculare speciale, de elemente genetice mobile (ADN accesoriu) (P. Sykora 1992).

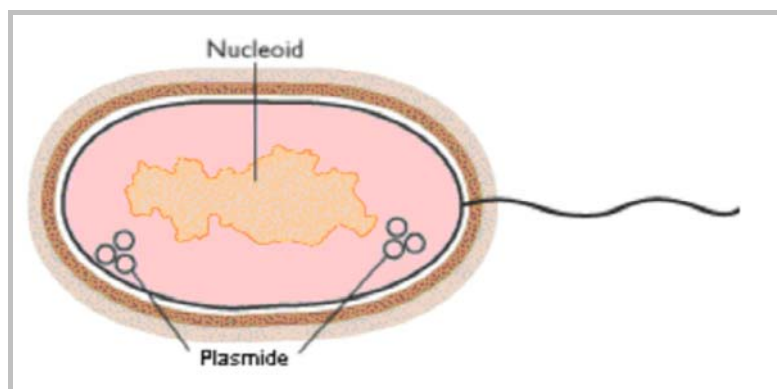


Figure 3.3. Reprezentare schematizată a nucleoidului și a plasmidelor bacteriene.

3.2.1 Principalele categorii de gene codificate de plasmide

Fiind repliconi fizic independenți, toate plasmidele poartă gene implicate în propria **replicare**, **stabilitate** și **partiție**. În afară de acestea, majoritatea genelor codificate de plasmide conferă caractere adaptative, avantajoase gazdelor bacteriene. Cele mai frecvente categorii de gene purtate de plasmide sunt :

- gene de rezistență la antibiotice

Plasmidele de rezistență au fost descoperite în Japonia în anii 50 la izolate clinice de *Shigella dysenteriae* și s-a constatat că aceste gene nu erau plasate pe cromosomul bacterian. Au fost denumite **factori R**. Folosirea abuzivă și/sau incorectă a antibioticelor au determinat,

pe de o parte, răspândirea în populațiile de bacterii patogene a determinantilor de antibioretistență, iar pe de altă parte, dezvoltarea unor mecanisme de antibioretistență extrem de eficiente.

În ceea ce privește originea genelor de antibioretistență, au fost emise 2 ipoteze majore:

- a-genele de antibioretistență existente în prezent în foarte multe bacterii patogene ar avea originea în gene similare prezente în bacteriile ce produc antibiotice (și care, de regulă, sunt bacterii din sol); în această situație se presupune că ar fi avut loc evenimente de transfer de gene pe orizontală între bacterii din sol și bacterii patogene;
- b-genele de antibioretistență ar fi evoluat din gene "house-keeping" (de exemplu: proteinele ce efluxează tetraciclina ar putea să fi evoluat din proteinele de transport al nutrienților).
 - gene de toleranță la metale grele (ioni mercurici, derivați organomercurici, nichel, cobalt, plumb, cadmiu, bismut, antimoniu, zinc, argint)
 - gene de rezistență la agenți intercalanți, radiații UV, radiații X, bacteriofagi, bacteriocine
 - gene ce codifică factori de patogenitate și de virulență
 - gene ce codifică bacteriocine
 - gene ce codifică antibiotice
 - gene implicate în metabolismul opinelor (plasmidele *Ti* de la *Agrobacterium*)
 - gene ce codifică capacități metabolice deosebite, de exemplu folosirea unor substanțe xenobiotice ca sursă unică de carbon și energie la specii de *Pseudomonas*
 - gene fixatoare de azot (genele *nif* de la *Rhizobium* și *Bradyrhizobium*)
 - gene ce codifică enzime proteolitice, amilolitice etc
 - gene ce codifică transferul plasmidelor pe orizontală prin proces de conjugare bacteriana (regiunea *tra*)

Deci, în afară de genele implicate în propria replicare, stabilitate și partiție, plasmidele mai pot purta și gene ce codifică caractere - unele esențiale, altele neesențiale pentru bacteria gazdă. **Caracterele esențiale** reprezintă căi metabolice centrale ale celulei bacteriene, plasmidele purtătoare fiind prezente la marea majoritate a izolatelor naturale ale unor tulpini bacteriene (de exemplu, unele plasmide mari de la *Agrobacterium tumefaciens*, *Rhizobium leguminosarum*).

Unele **caractere neesențiale** conferă avantaje selective gazdelor purtătoare, permițând supraviețuirea acestora în anumite condiții de mediu. În această categorie intră genele plasmidiale ce codifică rezistența la antibiotice și toleranța la metale grele (asemenea gene sunt prezente la multe tulpini de bacterii patogene). În absența antibioticelor sau a metalelor grele ca factori de presiune selectivă, procesele de replicare, menținere și partiție a plasmidelor reprezintă un efort metabolic suplimentar pentru celulele bacteriene. Ca urmare, în doar câteva pasaje, aceste plasmide sunt eliminate din populația bacteriană prin subreplicare. Ele se mențin doar într-o proporție scăzută a populației bacteriene, fapt ce ar putea fi considerat o "economie" metabolică la nivelul întregii populații bacteriene.

Există însă și așa-numite **plasmide "criptice"**, care nu par să ofere gazdei nici un caracter avantajos. Interesant este faptul că, deși pot fi încadrate în categoria "*selfish DNA*" ("*ADN egoist*"), totuși și aceste plasmide sunt menținute într-o anumită proporție din populațiile bacteriene.

În ansamblu, genomul bacterian se caracterizează printr-o **intensă circulație pe orizontală** a genelor, între diverși indivizi bacterieni (H.Tschape 1994). Această circulație se desfășoară prin trei procese centrale - **conjugare, transformare și transducție** - dintre care, plasmidele sunt implicate în primele două.

Acest flux continuu de gene pare să fie o caracteristică definitorie a lumii bacteriene și să fi contribuit esențial la extrema versatilitate metabolică a bacteriilor în aproape toate mediile naturale.

3.2.2 Clasificarea plasmidelor

Plasmidele pot fi clasificate funcție de diverse criterii :

- a) **criteriul autotransferabilității**. Autotransferabilitatea reprezintă capacitatea unui element genetic (în speță, un plasmid) de a se autotransfera de la un individ bacterian la altul. De

regulă, genele implicate în autotransferabilitate sunt grupate într-un operon (denumit operon *tra*), iar elementul genetic purtător este denumit **conjugon**. Conform acestui criteriu există 3 clase de plasmide :

- 1) **plasmide conjugative** – sunt plasmidele care prezintă întreg operonul *tra* și sunt autotransferabile, putând genera proces de conjugare bacteriană;
- 2) **plasmide non-conjugative** – sunt plasmidele ce nu prezintă operonul *tra*, nu sunt autotransferabile (nu pot genera conjugare bacteriană) și nici mobilizabile;
- 3) **plasmide mobilizabile** – sunt plasmidele ce nu prezintă operonul *tra* (și, ca atare, nu sunt autotransferabile), dar au gene **mob** care le permit mobilizarea în evenimente de conjugare inițiate de alte plasmide ce sunt conjugative.

b) **criteriul structurii moleculare**. Marea majoritate a plasmidelor sunt formate din ADN dublu catenar. Mai mult decât atât, din punct de vedere al structurii moleculare, există două clase de plasmide :

1) **plasmide circulare**. Plasmidele circulare sunt alcătuite din ADN d.c. circular, iar forma *in vivo* pare să fie CCC (“*circular covalently closed*”); foarte rar și temporar se întâlnește forma CO (“*circular open*”), în care una din cele două catene nu este continuă având o legătură fosfodierică lipsă. Tot *in vivo* se mai întâlnesc și concatemi (oligo/multimeri).

2) **plasmide lineare**. Asemenea plasmide sunt alcătuite din ADN d.c. linear (L). Au fost descrise două tipuri de plasmide lineare :

1. **Plasmide lineare cu telomere hairpin** (telomere în ac-de-păr). Acestea sunt întâlnite cu precădere la plasmidele lineare de la *Borrelia* (Figura 3.3). Spirochetele din genul *Borrelia* au structură genomică lineară: toate speciile de *Borrelia* examinate au un cromosom linear de 950-1000 kb și mai multe plasmide lineare cu dimensiuni cuprinse între 5 și 200 kb (unele au și plasmide CCC). Toate plasmidele lineare de la *Borrelia* descrise până în prezent au telomere de tip *hairpin*. Telomerele *hairpin* se caracterizează prin palindroame terminale bogate în A/T, la capete existând o continuitate în legătura de tip covalent între cele 2 catene. Asemenea structuri (ADN d.c. L cu telomere de tip *hairpin*) mai există și la ADNmt de la reprezentanții genului *Pichia*, la poxvirusuri și la plasmidele lineare din mitocondriile de fungi.

2. **Plasmide cu telomere invertroni** Asemenea telomere sunt întâlnite la plasmidele lineare de la genul *Streptomyces* – plasmide cu dimensiune cuprinsă între 9 și 600 kbp. Ca și cele *hairpin*, telomerele *invertroni* prezintă și ele palindroame terminale bogate în A/T. Diferența constă însă în faptul că nu există legătura covalentă între capetele celor 2 catene, și, ca urmare, capetele 5’ și 3’ sunt libere. La capul 5’ al fiecărei catene este atașată covalent o moleculă proteică denumită **proteina TP** (“*Telomeric Protein*”) (Figura 3.3). Asemenea structuri ADN, cu palindroame terminale bogate în A/T și cu proteina TP atașată covalent la capul 5’ au fost definite de K.Sakaguchi (1990) ca fiind **elemente genetice mobile cu replicare autonomă** și denumite INVERTRONI, similare funcțional cu transposonii.

Telomerele *invertroni* sunt caracteristice repliconilor lineari de la eucariote, în mod special la plasmidele lineare din mitocondriile de fungi și plante și din citoplasma de fungi.

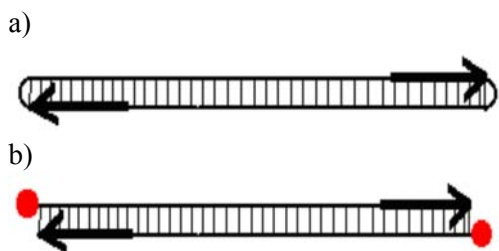


Figura 3.4 Structura telomerelor *hairpin* la *Borrelia* (a) și *invertroni* la *Streptomyces* (b) (după H.Hinnebusch, 1992).
⇒ palindroame terminale, bogate în A-T

● = la capul 5’ al fiecărei catene există o proteină atașată covalent

c) criteriul incompatibilității plasmidiale

De-a lungul anilor, **conceptul de incompatibilitate plasmidială** a ridicat foarte multe probleme. Actualmente, acest concept este definit ca fiind incapacitatea a două sau mai multe plasmide de a se menține și propaga stabil în aceeași linie celulară. Plasmidele din același grup de incompatibilitate ce conțin origini de replicare similare (deci, au același **replicon-type**), nu sunt moștenite stabil în descendență și segregă una de cealaltă. Cel mai probabil acest proces se datorează unei competiții moleculare între plasmide cu același **replicon-type** pentru aceleași proteine implicate în replicarea respectivelor plasmide – proteine codificate cromozomal. La *Enterobacteriaceae* au fost descrise peste 30 de grupuri de incompatibilitate plasmidială. Un tip de incompatibilitate s-a bazat pe competiția între secvențele repetate RS (la care se atașează proteina **RepA**), dintre diferite plasmide existente în celula bacteriană. Bazându-se pe tipul de replicon a fost imaginată o metodă nouă de identificare plasmidială folosind o colecție de sonde **rep** (Sykora P. 1992).

Procesul de "**gene-flow**" în lumea bacteriană ar putea avea un rol important în evoluția bacteriilor. Cu toate acestea, procesele de transfer de material genetic pe orizontală nu par să afecteze stabilitatea speciilor taxonomice bacteriene. Din acest punct de vedere, plasmidele se găsesc într-o situație total diferită: "**gene-flow**"-ul între plasmide neînrudite pare să fie atât de ridicat și de răspândit, încât nu se poate alcătui o identificare taxonomică a plasmidelor doar pe baza unei liste a caracterelor pe care le codifică. Se pare că biologia plasmidelor nu implică mecanisme de izolare taxonomică asemănătoare cu speciile de plante și animale. În contrast, în lumea plasmidelor se observă un fenomen total diferit: izolarea plasmidelor individuale sau foarte asemănătoare. Acest fenomen este cunoscut ca fenomenul de **incompatibilitate plasmidială**.

3.3 STRUCTURA CROMOZOMULUI LA EUCARIOTE

Materialul genetic este reprezentat dintr-un genom nuclear și un genom extranuclear (reprezentat din genomul mitocondrial și, în cazul celulelor vegetale, și din genomul cloroplastic).

Genomul nuclear cuprinde un set de molecule lineare de ADN, fiecare reprezentând un cromozom. Fără excepție, toate eucariotele au cel puțin 2 cromozomi care întotdeauna sunt lineari. Singura variație care se înregistrează este cea referitoare la numărul de cromozomi, care însă nu este corelată cu caracteristicile biologice sau cu poziția evolutivă a organismului respectiv. Astfel, la *Saccharomyces cerevisiae* (drojdia de bere - un eucariot unicelular, inferior) numărul haploid de cromozomi este 16, de 4 ori mai mare decât la muscă (*Musca domestica*). De asemenea, numărul de cromozomi nu este corelat nici cu dimensiunea genomului; de exemplu, la salamandră genomul este de 30 ori mai mare decât la om, dar are jumătate din numărul haploid de cromozomi de la om.

Cromozomii eucarioti sunt separați de citoplasmă prin membrana nucleară și, ca urmare, transcrierea și traducerea sunt separate în timp și spațiu una de cealaltă; în contrast, la procariote, cele două procese se desfășoară aproximativ simultan.

Nucleul este format din 4 componente: **nucleolemă**, **carioplasmă**, **cromonemata** (singular, cromonema) și **nucleol**. Cromonemata este o substanță cromatică despiralizată în interfază și condensată în mitoză. Din această substanță se diferențiază cromozomii.

Nucleolul este o substructură nucleară cu rol în biogeneza ribozomilor. În această zonă a nucleului are loc sinteza de ARN ribozomal prin transcrierea genelor corespunzătoare.

La eucariote ADN este complexat cu o clasă de proteine specializate, cu caracter bazic, denumite histone. Acest complex poartă numele de **cromatină** (denumirea are la bază tinctorialitatea față de coloranți bazici).

Cromatina se prezintă sub 2 stări:

- **eucromatina** – cromatină decondensată în interfază și condensată în timpul diviziunii celulare; se replică la începutul fazei S a interfazei
- **heterocromatina** – condensată permanent și se replică la sfârșitul fazei S.

Heterocromatina se clasifică în:

- **heterocromatină constitutivă** – este caracteristică zonelor centromerice și telomerice
- **heterocromatină facultativă** - este caracteristică unuia din cromozomii X la femelele de mamifer și constituie cromatina sexuală; apare datorită necesității compensației de doză a genelor esențiale prezente pe cromozomul X
- **heterocromatină de citodiferențiere** – apare prin inactivarea selectivă a unor gene în timpul proceselor de citodiferențiere

Eucromatina este caracteristică genelor active transcripțional, în timp ce heterocromatina are rol preponderent regulator și structural.

3.3.1 Clase de secvențe ADN la eucariote

1. **secvențe unice**, o copie per genom haploid, corespunzătoare genelor codificatoare pentru lanțuri polipeptidice, excepție făcând genele ce codifică histone, care se găsesc în câteva sute de copii per genom haploid (este necesară o sinteză rapidă într-un timp scurt corespunzătoare perioadei S de replicare a ADN).
2. **secvențe moderat repetitive**: sunt într-un număr de $10^2 - 10^3$ copii per genom haploid și codifică pentru ARN ribozomal, ARN de transfer și pentru histone.
3. **secvențe înalt repetitive**: ajung până la 10^6 copii per genom haploid

În general, eucariotele conțin mult mai mult ADN decât procariotele. O celulă de om conține de peste 1000 de ori mai mult ADN decât o celulă de *Escherichia coli*. Astfel, ADN-ul dintr-o celulă de om, aflat în stare decondensată atinge o lungime de 4 cm și nu poate încăpea într-un nucleu (de diametru 5 μ m) decât în formă condensată.

Forma cromozomilor dintr-o celulă eucariotă se modifică în timpul ciclului celular.

3.3.2 Proteinele cromozomale

Principalele proteine ce complexează ADN-ul au greutate moleculară mică, o cantitate mare de aminoacizi cu caracter bazic și se numesc histone. Au fost descrise 5 clase de histone: H1, H2A, H2B, H3 și H4.

Histona H1 este alcătuită din 3 domenii structurale distincte:

- un capăt amino-terminal cu 39 aminoacizi bazici
- o regiune globulară centrală cu un diametru de 2,8 nm, cu aminoacizi cu caracter acid; interacționează cu alte proteine amino-terminal ce conține 40% lizină (cu puternic caracter bazic) și are afinitate de interacțiune cu molecula de ADN.

Histonele H2A și H2B au cantitate mai redusă de lizină. H2A are leucină, iar H2B are serină și prolină. Au doar 2 domenii, dintre care unul interacționează cu ADN. Histonele H3 și H4 sunt bogate în arginină și, în mod similar cu H2A și H2B, au tot doar 2 domenii funcționale.

În general, histonele prezintă o structură înalt conservată în lumea vie. Cea mai heterogenă este histona H1, acest parametru scăzând de la H1 către H4. Genele pentru histone sunt organizate în unități repetitive dispuse în tandem. De exemplu, la *D.melanogaster* unitatea repetitivă este formată din genele pentru: H1 – H3 – H4 – H2A – H2B



Figura 3.5 Organizarea genelor pentru histone la *D.melanogaster*. regiuni reglatoare ale genelor (situsuri hipersensibile la nuclează)

3.3.3 Nivelele de organizare ale cromatinei

Elementul structural de bază al cromatinei este **nucleozomul** care este format dintr-un octamer histonic central și o histonă de legătură (*linker*). Nucleozomul are 145 perechi de baze corespunzătoare la aproximativ 2 ture de spirală suprarăsucite negativ în jurul miezului histonic. Miezul histonic prezintă domeniul hidrofob carboxi-terminal de interacțiune cu celelalte proteine spre partea centrală, iar domeniul amino-terminal se găsește la suprafața de interacțiune cu ADN.

Între 2 nucleozomi adiacenți se interpune ADN linker de 60 perechi de baze și este complexat cu histona H1.

Cromatina extinsă apare ca un șirag de mărgele cu un diametru de 11 nm. Prin intermediul histonei H1 implicată în superspiralizare interacționează 6 – 10 nucleozomi pentru a forma solenoidul de 30 nm. Acesta este unitate de bază a condensării materialului genetic în interfază.

Forma superioară de condensare este reoprezentată de cromozomii metafazici, perioada în care materialul genetic atinge un nivel maxim de condensare de 1400 nm. Procesul de condensare a cromatinei este realizat prin atașarea ei la o clasă de proteine denumite „proteine scaffold” (proteine de eșafodaj). Atașarea se realizează cu formarea unor bucle (domenii) care au fiecare 100 kpb, comparabile cu cele de la cromozomul bacterian.

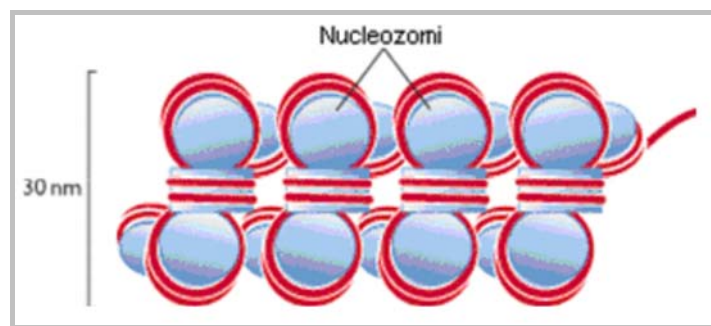


Figura 3.6 Reprezentarea schematică a structurii nucleosomilor (după Strachan, 1999).

În organizarea materialului genetic nuclear la organisme eucariote intervin și proteine non-histonice. Acestea se subîmpart în mai multe clase:

- enzime ale metabolismului ADN: ADN polimeraze, ADN ligaze, ADN topoisomerase, terminal-ADN-nucleotidil-transferaze
- enzime ale metabolismului ARN: ARN polimeraze, poliA polimeraze, enzime de *splicing*, RNaze
- enzime ce modifică histonele: protein-kinaze, metilaze, acetilaze, esteraze, proteaze
- proteine structurale: proteinele fusului de diviziune, proteinele matricei nucleare, proteinele „scaffold”, proteinele particulelor RNP („*RiboNucleoProtein*”)
- proteinele HMG (*High Mobility Group*) ce intervin în modificarea structurii cromatinei în vederea funcționării genelor în replicare și transcriere:
 - HMG 1 și HMG 2 – proteine omoloage cu g.m. 29 kd;
 - HMG 14 și HMG 17 - proteine omoloage cu g.m. 10-12 kd

Proteinele HMG înlocuiesc histona H1 determinând, alături de topoisomerase, relaxarea helixului ADN în inițierea transcrierii genelor.