

### 1.1. La infecció vírica a *Bacillus subtilis*

La infecció vírica és un fenomen universal en totes les cèl·lules, incloses les procariotes. En essència consisteix en l'entrada d'un material genètic estrany a la cèl·lula (Figura 1.1), que s'apodera de la maquinària molecular cel·lular per realitzar el cicle vital víric, fins que la cèl·lula mor i s'allibera al exterior el material genètic víric, després de ser amplificat i encapsulat en una coberta generalment proteica.



Figura 1.1: Imatge de microscòpia electrònica de l'injecció del material genètic per part d'un bacteriòfag de cua no contràctil a l'interior de la cèl·lula procariota, produint-se unes invaginacions de la membrana plasmàtica a l'interior del citoplasma.

El gènere *Bacillus* inclou diverses espècies de bacteris grampositius, aeròbics, formadors d'endospores, que normalment habiten el sòl i material vegetal en descomposició.

En aquest hàbitat s'han aïllat una gran varietat de fags que infecten al gènere *Bacillus*, presentant tots ells característiques comunes respecte al seu material genètic i a la morfologia de la partícula vírica.

El material genètic d'aquests fags és DNA de doble cadena i la morfologia que presenten consisteix en un cap icosaèdric prolat i la presència de cua.

Ens centrarem en els fags del gènere  $\phi 29$ -like, que inclou, a part de  $\phi 29$ , els fags PZA,  $\phi 15$ , BS32, B103, M2Y (M2), Nf i GA-1. Tots ells són fags lítics que pertanyen a la família *Podoviridae*, que es caracteritzen per presentar una cua curta. La majoria d'aquests fags infecten *Bacillus subtilis*, però sovint infecten també altres espècies relacionades, com *Bacillus pumilus*, *Bacillus amyloliquefaciens*, i *Bacillus licheniformis*.

Basant-nos en diferents propietats d'aquests fags, com són els mapes físics de DNA i de proteïnes i propietats serològiques, podem classificar-los en tres subgrups. El primer grup inclou els fags  $\phi 29$ , PZA,  $\phi 15$  i BS32; el segon grup inclou B103, Nf i M2Y; i el tercer grup conté GA-1.

Els fags  $\phi 29$ -like són els fags de *Bacillus* més petits aïllats fins a la data, i són els fags coneguts amb dsDNA més petits. Les dimensions de la partícula fàgica són de 415x315 Å el cap i de 325x60 Å la cua. Una anàlisi feta per microscòpia electrònica (Hagen i Anderson, 1975) mostra com les partícules víriques de  $\phi 29$ , B103 i GA-1 tenen una simetria radial d'ordre sis i una cua curta, no contràctil en forma de tub.

Analitzant el rang d'hoste, ens mostra que  $\phi 29$  és capaç d'infectar les soques de *B. subtilis* 168, 110NA i Marburg, les soques de *B. amyloliquefaciens* H i varies soques de *B. licheniformis* i *B. pumilus*. El rang d'hoste de B103 no està ben estudiat, però se sap que infecta *B. subtilis* 9/3. Finalment, GA-1 és conegut que infecta líticament les soques de *Bacillus* G1R.

## **1.2. El fag $\phi 29$**

El genoma del fag  $\phi 29$  consisteix en una molècula de DNA lineal de doble cadena (dsDNA) de 19.285 parells de bases. Cada extrem 5' del DNA té unit covalentment

una proteïna codificada pel fag, anomenada proteïna terminal (Salas, 1991). Aquesta proteïna està implicada en la replicació del DNA víric, interaccionant amb la DNA polimerasa i amb l'ajuda d'una proteïna viral de tipus histona anomenada p6, la qual forma un complex de nucleoproteïna probablement implicat en el desenrotllament de la doble cadena de DNA (Serrano *et al.*, 1994).

### ***1.2.1. Organització genètica i transcripcional***

En els genomes dels fags, generalment, els gens es troben agrupats en clústers segons la funció que realitzen. Aquests clústers, transcripcionalment, estan organitzats en operons (diversos gens sota el control d'un sol promotor).

Podem distingir en el genoma de  $\phi 29$ , dos tipus d'operons segons l'etapa del cicle infectiu en la que s'expressen. Anomenarem operons primerencs (*early operons*) els que s'expressen en el primer estadi de la infecció, i operons tardans (*late operons*) els que s'expressen en un segon estadi de la infecció. L'expressió dels gens tardans es realitza en el moment en que ocorre un canvi en la transcripció, reprimint-se els promotors primerencs més importants i activant-se l'únic promotor tardà que existeix a  $\phi 29$ .

En els operons primerencs la transcripció es realitza en sentit de dreta a esquerra, segons el mapa genètic de  $\phi 29$  (Mellado *et al.*, 1976), i en els tardans en sentit d'esquerra a dreta.

Els operons primerencs se situen en els dos extrems del genoma, mentre que l'operó tardà està localitzat en el centre del genoma (Sogo *et al.*, 1979)

#### ***1.2.1.1. Situació dels promotors en el mapa genètic***

A l'extrem esquerre del genoma de  $\phi 29$  (Figura 1.2) trobem dos promotors

primerencs dèbils, el A1 i el A1IV, que codifiquen pel pRNA, un RNA de 174 nucleòtids requerit en la translocació del DNA i en l'especificitat de la polaritat d'empaquetament (Guo *et al.*, 1987).

També trobem els promotors A2b i A2c, els promotors primerencs més importants. Es troben situats en una regió intergènica compartida entre 3 promotors: A2b, A2c i A3. Els promotors A2b i A2c són promotors molt forts i transcriuen els gens del 1 al 6, que codifiquen, la majoria, per proteïnes de replicació del DNA.

El promotor A3, que comparteix espai amb A2b i A2c, és l'únic promotor tardà del genoma de  $\phi 29$ . Transcriu els gens compresos entre el 7 i el 16, que codifiquen per proteïnes estructurals.

A l'extrem dret del genoma de  $\phi 29$  trobem els promotors C2 i C1, que codifiquen per les proteïnes p16.7 i p17, necessàries per a la replicació del DNA i el procés d'encapsidament del profag.

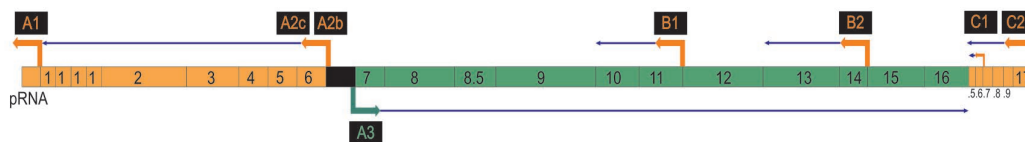


Figura 1.2: Esquema del mapa genètic del bacteriòfag  $\phi 29$ . En taronja els gens primerencs i en verd els tardans.

Fent una alineació dels promotors (Figura 1.3) observem que tots els promotors mostren una bona conservació de les caixes -10 i de la caixa -35, a excepció del promotor A3, que no la presenta.

A1	TTAATGT	<u>TTGACA</u>	ACTATTACAGAGTATGC	<u>TATAAT</u>	GGTAGTATCAAT
A2c	AAAAGTG	<u>TTGAAA</u>	ATTGTGGAACAGGGTGA	<u>TATAAT</u>	AAAAGAGTAGAA
A2b	AAAGTC	<u>TTGCAAAA</u>	AGTTATACAGGTGTGG	<u>TTAAAT</u>	AGAGAACGTAGA
A2a	GGTTTTA	<u>ATGGCA</u>	TATGTTTCACCTCTTT	<u>TATAAT</u>	CTATTAGTACTA
B2	TCCGA	<u>TACACA</u>	CAAAGCCGTATAAACCGT	<u>TATAAT</u>	AGGGGTAACCCG
C1	ATCAACG	<u>TTTACA</u>	AAGTGAACAGGAAGTGT	<u>TAAACT</u>	TATAGAGACAC
C2	AAAGGG	<u>TAGACA</u>	AACTATCGTTTAACATGT	<u>TATACT</u>	TATAATAGAAGTA
A3	CAAATCC	<b>TTATGT</b>	ATCAAGGGTTCACGTGG	<u>TATAAT</u>	TAAGTAGTACTA
Consens		TTGACA		TATAAT	

Figura 1.3: Seqüències dels promotors de  $\phi 29$  (Mellado, 1986)

### 1.2.2. Funcionament general de la transcripció a $\phi 29$

La transcripció del genoma  $\phi 29$  s'inicia en el precís moment d'entrada del seu material genètic en el citoplasma bacterià.

Els promotors primerencs més forts A2b, A2c i C2 són transcrits intensament per l'RNA polimerasa de l'hoste que conté la subunitat  $\sigma A$ , la majoritària a *Bacillus subtilis*, equivalent a  $\sigma 70$  de *Escherichia coli*.

La transcripció, des dels promotors A2b i A2c, dels gens 6, 5, 3, 2 i 1, promou la síntesi de les proteïnes corresponents p6, p5, p3, p2 i p1. Aquestes proteïnes són essencials per la replicació del DNA fàgic. És en aquest moment quan s'expressa el gen 4, el producte del qual, la proteïna p4, és el principal responsable de la commutació de la transcripció i l'objecte d'estudi en aquest treball.

La transcripció des del promotor C2 té com a resultat la síntesi de la proteïna p17, que s'acumula en grans quantitats dins la cèl·lula i és necessària per l'inici de la replicació del genoma de  $\phi 29$ .

Durant aquesta primera etapa transcripcional, s'expressen les proteïnes necessàries per replicar el DNA víric, acumulant-se una gran quantitat de material genètic viral a l'interior de la cèl·lula hoste.

La commutació transcripcional ocorre quan se sintetitzen, en suficient quantitat, les proteïnes primerenques p4 i p6. p4 s'uneix al DNA en la regió intergènica on es localitzen els promotors A2b, A2c i A3, mentre que p6 s'uneix a bona part del genoma, especialment als seus extrems (Prieto *et al.*, 1988).

La transcripció en els operons primerencs s'atura i s'inicia en el promotor tardà, transcrivint-se els gens del 7 al 16. Aquests gens codifiquen per proteïnes que constitueixen la càpside i la cua del fag, proteïnes involucrades en la morfogènesi de la càpside i la cua del fag i proteïnes requerides en la lisi de l'hoste.

La síntesi de proteïnes està separada, doncs, en dues etapes, obeint a les diferents necessitats del fag, de manera que, en la primera etapa, les proteïnes sintetitzades es dediquen a replicar el DNA fàgic, i quan aquest està en una quantitat suficient, la síntesi proteica es dedica a proteïnes estructurals per envoltar les molècules de DNA fàgic sintetitzades.

#### ***1.2.2.1. Regulació transcripcional***

La regulació transcripcional en el fag  $\phi 29$  consisteix, bàsicament, en un canvi en el patró d'expressió dels promotors situats en una regió intergènica de 219 parells de bases (Figura 1.4), és a dir, els promotors A2b, A2c i A3.

El canvi en el patró de transcripció ve determinat per la repressió dels 2 promotors majoritaris que s'expressaven en els primers estadis de la infecció, els promotors primerencs A2b i A2c, i l'activació d'un promotor silenciats fins aquest moment, el promotor A3.

En la regió intergènica, on es localitzen els 3 promotors, trobem, agrupats per parelles, 4 llocs d'unió de la proteïna p4, la principal responsable de la commutació en la transcripció, actuant de manera diferent en cadascun dels 3 promotors.

El lloc d'unió de la proteïna p4 al promotor tardà A3 se superposa parcialment al del promotor A2b, que està orientat en la direcció contrària a A3, i per tant, la proteïna p4 i l'RNA polimerasa no poden unir simultàniament el promotor A2b. Com a conseqüència, la proteïna p4 exclou l'RNA polimerasa del promotor A2b, reprimint-lo (Rojo *et al.*, 1991).

Finalment, el promotor primerenc A2c, que es localitza 96 parells de bases *downstream* del promotor A2b, també és reprimint per la proteïna p4. El mecanisme pel qual es reprimeix el promotor A2c implica que s'uneixin simultàniament al promotor la proteïna p4 i l'RNA polimerasa (Monsalve *et al.*, 1995; Monsalve *et al.*, 1996).

En l'activació del promotor A3 i la repressió del promotor A2c es requereix una interacció entre la proteïna p4 i l'RNA polimerasa. Les regions de p4 i de l'RNA polimerasa implicades en les interaccions són les mateixes en els dos casos.



Figura 1.4: Seqüència intergènica del DNA de  $\phi 29$  que inclou els promotors A2b, A2c i A3, inspirat en Camacho *et al.*, 2004. Es mostren els llocs d'unió de p4, i les caixes -10 i -35 dels diferents promotors.

### 1.2.2.2. Els llocs d'unió de p4

S'han descrit 4 llocs d'unió de p4 localitzats en dues regions. La primera regió comprèn els llocs 1 i 2, mentre que la regió 2 comprèn els llocs 3 i 4. Les dues regions comprenen 63 parells de bases, des de les posicions -25 a -84 relatives a l'inici de transcripció del promotor A2c i des de les posicions -27 a -87 relatives a l'inici de transcripció del promotor A2b (Camacho *et al.*, 2004).

En un alineament dels 4 llocs d'unió a p4 (Figura 1.5) observem com el lloc d'unió consisteix en una repetició invertida de 31 parells de bases de consens TTGAAAAA-15pb-TTTTACAA (Barthelemy *et al.*, 1989).

Lloc 1	GCTGTTAAAAGTTGTGAAAAGATATTTTTCAA
Lloc 2	GCTGTTTCAAATGTTCCAAAATTATTTTCCA
Lloc 3	ATTGAAAAACGTTCTGAAAAAATATTTTACAA
Lloc 4	CTGCAAAAAGCTGTTCTTGCGGGTGTTTAGGA
Consens	<b>TGAAAAA-----TTTTACAA</b>

Figura 1.5: Alineament dels 4 llocs d'unió a p4

Les dues regions tenen una diferent curvatura intrínseca (Barthelemy *et al.*, 1989; Rojo *et al.*, 1990). Mentre la regió 1 (llocs 1 i 2) no es troba corbada (Pérez-Lago *et al.*, 2005), la regió 2 (llocs 3 i 4) presenta una curvatura de 45° (Rojo *et al.*, 1990). En unir-se p4 la curvatura de les dues regions augmenta dràsticament fins als 85° (Rojo *et al.*, 1990; Pérez-Lago *et al.*, 2005).

La proteïna p4 presenta una afinitat molt diferent pels 4 llocs d'unió. El lloc pel qual presenta una major afinitat és el lloc 3. P4 uneix el lloc 1 amb la meitat d'afinitat que el 3 i el lloc 2 cinc cop menys que el 1, o deu menys que el 3 (Pérez-Lago *et al.*, 2005). El lloc 4 però, no és unit de manera independent, sinó que requereix que la proteïna p4 estigui unida anteriorment al lloc 3.



La proteïna p4 és un dímer en solució, i uneix el DNA com a dímer també. En la unió a DNA es contacta únicament una banda de la hèlix de DNA (Rojo *et al.*, 1991).

L'augment de la curvatura en la unió de p4 té com a conseqüència que l'hèlix de DNA es distorsioni i que certes posicions es tornin hipersensibles a l'atac amb DNasa I. Aquestes posicions cauen en la cara més externa de la curvatura, on no hi ha protecció per la proteïna p4 i on el solc menor del DNA s'eixampla com a conseqüència de la torsió. Les posicions hipersensibles es veuen cada 10-12 parells de bases, al llarg de tot el lloc d'unió de p4, cada volta de la doble hèlix.

### ***1.2.3. Control de la transcripció per la proteïna p4***

#### ***1.2.3.1. Repressió del promotor primerenc A2b***

Els promotors A2b i A2c són reprimits coordinadament per la proteïna p4 (Rojo *et al.*, 1990; Monsalve *et al.*, 1995), però per mecanismes totalment diferents.

La proteïna p4 uneix una regió del promotor tardà A3, on hi ha el lloc 3, el de més afinitat per p4, corbant el DNA considerablement (Barthelemy *et al.*, 1989; Rojo *et al.*, 1990). El lloc 3 es solapa parcialment amb el promotor primerenc A2b. L'RNA polimerasa és incapaç de desplaçar la proteïna p4 per unir-se al promotor, quedant el promotor A2b inhibit (Rojo *et al.*, 1991).

El promotor A2b conté una regió UP (Meijer i Salas, 2004) que és la zona solapada amb el lloc d'unió de la proteïna p4. La modificació d'aquesta regió UP o la seva inaccessibilitat tenen efectes severes en la transcripció del promotor A2b (Meijer i Salas, 2004).

Un increment progressiu en la curvatura intrínseca del lloc d'unió de p4 causa un descens en l'activitat del promotor A2b. La curvatura provocada per la unió de la

proteïna p4 pot ajudar a incrementar l'eficiència de repressió, modificant l'estructura del promotor per què no sigui reconegut per l'RNA polimerasa (Rojo *et al.*, 1991).

#### ***1.2.3.2. Repressió del promotor primerenc A2c***

Al contrari de com passa en el promotor primerenc A2b, s'ha vist que la proteïna p4 no desplaça l'RNA polimerasa, sinó al contrari; totes dues proteïnes s'uneixen a ell cooperativament (Monsalve *et al.*, 1996).

La proteïna p4 s'uneix en el lloc 1, centrat a la posició -39 relativa al lloc d'inici de la transcripció. En unir-se l'RNA polimerasa, la proteïna p4 es veu desplaçada al seu lloc 2 (Monsalve *et al.*, 1998), centrat a la posició -72 relativa al lloc d'inici de transcripció.

En els primers estadis de la infecció, en absència de la proteïna p4, l'RNA polimerasa s'uneix al promotor A2c, transcrivint-se els gens primerencs de l'1 al 6. En presència de la proteïna p4, l'RNA polimerasa es continua unint al promotor formant el complex tancat d'inici de la transcripció. Aquest complex pot isomeritzar cap al complex obert, permetent la incorporació de nucleòsids trifosfats (NTPs) per transcriure fragments curts d'RNA. En aquest moment l'enzim RNA polimerasa no pot escapar del promotor per iniciar la fase d'elongació dels trànscrips.

La cooperativitat en la unió de la proteïna p4 i de l'RNA polimerasa, suggereix que el mecanisme de repressió ha de estar relacionat amb el mecanisme de cooperativitat.

Diversos mutants en la regió carboxi-terminal de la proteïna p4 tenen una capacitat reduïda de reprimir el promotor A2c. Els mutants R120Q i R120A són incapaçs de reprimir el promotor A2c i de unir-se cooperativament amb l'RNA polimerasa al promotor (Monsalve *et al.*, 1996), essent aquest residu el més important en la interacció entre la proteïna p4 i l'RNA polimerasa.

### 1.2.3.3. Activació del promotor tardà A3

L'activació de la transcripció del promotor tardà A3 requereix la unió de la proteïna p4 al seu lloc d'unió, concretament el lloc 3. Aquest lloc està centrat a la posició -82 relativa al lloc d'inici de la transcripció (Barthelemy *et al.*, 1989).

Com ja hem vist en l'anàlisi dels promotors de  $\phi 29$ , el promotor A3 presenta una bona seqüència a la regió -10 per l'RNA polimerasa de *Bacillus subtilis* que conté la subunitat  $\sigma$  vegetativa, la  $\sigma A$ . En canvi, a aquest promotor, li manca la regió -35, que és reconeguda també per la subunitat  $\sigma A$ . La mancança de la regió -35 és la responsable de que el promotor A3 no es transcriu en absència de la proteïna p4. En presència de p4, però, la unió de l'RNA polimerasa al promotor s'estabilitza i la transcripció del promotor A3 esdevé possible. Els contactes que realitza la proteïna p4 permeten estabilitzar l'RNA polimerasa com a complex tancat de transcripció (Nuez *et al.*, 1992).

De manera equivalent al promotor A2c, la cooperativitat en la unió de les dues proteïnes (Nuez *et al.*, 1991), requereix de contactes entre la proteïna p4 i l'RNA polimerasa. El posicionament de les dues proteïnes al llarg de la cadena de DNA es realitza en la mateixa cara del DNA (Rojo i Salas, 1991), permetent el contacte entre elles. El lloc d'unió de p4 al promotor A3, el lloc 3, es troba centrat a la posició -82 de l'inici de transcripció, mentre que la subunitat  $\alpha$  de l'RNA polimerasa se situa en les posicions centrades a -49 i -40 (Camacho *et al.*, 2004), solapant amb el lloc 4 d'unió de p4.

Al' hora d'estudiar les superfícies implicades en les interaccions, trobem novament que els mutants de la proteïna p4 R120Q i R120A són incapaços d'estabilitzar l'RNA polimerasa al promotor.

#### 1.2.4. Regions comunes activen i reprimeixen promotors

La proteïna p4 interacciona amb l'RNA polimerasa en 3 promotors diferents, A2b, A2c i A3, i en cadascun d'ells té efectes diferents. En el promotor A2b, el contacte entre l'RNA polimerasa i la proteïna p4 es redueix a una incapacitat per part de l'RNA polimerasa de desplaçar la proteïna p4 unida al seu lloc d'unió. Per contra, en el promotor A2c, aquest desplaçament és aconseguit, però la interacció entre l'RNA polimerasa i la proteïna p4 impedeix la transcripció del promotor. Finalment, en el promotor A3, aquest contacte és el que permet la transcripció del promotor.

En els promotors A2c i A3, la superfície de la proteïna p4 implicada en les interaccions amb l'RNA polimerasa és clara. Es tracta de la regió carboxi-terminal, focalitzada en el residu Arg120, el més important per a aquesta interacció.

La regió implicada en el contacte per part de l'RNA polimerasa és la subunitat  $\alpha$ , concretament el domini carboxi-terminal ( $\alpha$ CTD). Diversos mutants de la subunitat  $\alpha$  que els mancaven els últims 15, 37 i 59 residus poden reconèixer i transcriure diferents promotors virals no activats per p4, però no poden reconèixer el promotor A3 en presència de p4 (Mencía *et al.*, 1996).

El  $\alpha$ CTD (Figura 1.6) conté diferents residus àcids que són candidats a ser dianes del residu Arg120 de la proteïna p4.

<b>MIEIEKPKIE</b>	<b>TVEISDDAKF</b>	<b>GKFFVVEPLER</b>	<b>GYGTTLGNSL</b>	<b>RRILLSSLPG</b>
<b>AAVTSIQIDG</b>	<b>VLHEFSTIEG</b>	<b>VVEDVTTIIL</b>	<b>HIKKLALKIY</b>	<b>SDEEKTLEID</b>
<b>VQEGGTVTAA</b>	<b>DITHSDSVEI</b>	<b>LNPDLHIATL</b>	<b>GENASFRVRL</b>	<b>TAQRGRGYTP</b>
<b>ADANKRDDQP</b>	<b>IGVIPIDSIY</b>	<b>TPVSRVSYQV</b>	<b>ENTRVGQVAN</b>	<b>YDKLTLDVWT</b>
<b>DGSTGPKEAI</b>	<b>ALGSKILTEH</b>	<b>LNIFVGLTDE</b>	<b>AQHAEMIVEK</b>	<b>EEDQKEKVL</b>
<b>MTIEELDLSV</b>	<b>RSYNCLKRAG</b>	<b>INTVQELANK</b>	<b>TEEDMMKVRN</b>	<b>LGRKSL</b> <b>EEVK</b>
<b>AKLEELGLGL</b>	<b>RKDD</b>			

Figura 1.6: Seqüència de la subunitat  $\alpha$  de l'RNA polimerasa, en vermell els residus àcids

Tenint en compte que les regions implicades en la proteïna p4 i el  $\alpha$ CTD de l'RNA polimerasa són presumiblement les mateixes en totes les interaccions entre ambdues proteïnes en els diferents promotors o llocs d'unió, la diferència l'hem de buscar en l'estructura dels promotors (Monsalve *et al.*, 1997). La presència o l'absència de la caixa -35 pot canviar el caràcter del promotor, essent reprimat o activat respectivament. En efecte, les interaccions de l'RNA polimerasa amb la proteïna p4 i amb la regió -35 del promotor causen una sobreestabilització de l'RNA polimerasa en el promotor que li impedeixen d'iniciar l'etapa d'elongació.

### 1.2.5. La proteïna p6

S'ha descrit que en la regulació dels promotors A2b, A2c i A3 hi participa una altra proteïna, la p6, a més de la p4.

La proteïna p6 és una proteïna associada a nucleòide de 103 aminoàcids, que participaria tant en la replicació del DNA viral *in vivo* com en el control de la transcripció. S'ha proposat, també, que juga un paper en l'organització i en la compactació del genoma. P6 forma dímers, però a concentracions properes a 1mM, s'associa formant oligòmers d'alt pes molecular. Aquests oligòmers es poden observar per microscòpia electrònica de transmissió i processament d'imatge (Abril *et al.*, 1999). Els oligòmers presenten una forma torçada, mostrant que poden interaccionar cap i cua per formar estructures circulars (Figura 1.7).

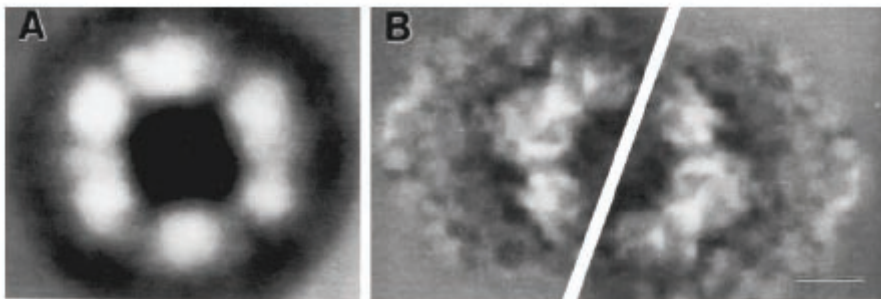


Figura 1.7: Imatge de microscòpia electrònica de la proteïna p6 formant superestructures circulars, segons Abril (Abril *et al.*, 1999).

P6 s'uneix a DNA de doble cadena, obtenint-se complexos multimèrics de nucleoproteïna. El DNA, en unir-se a un nucli de proteïna p6, forma una superhèlix dextrògira que compacta el DNA. La unió de la proteïna al DNA es fa pel solc menor, contactant les dues cadenes de DNA cada 12 parells de bases.

#### ***1.2.5.1. Complex p6-nucleoproteïna***

El complex de nucleoproteïna és un complex regular en el que la proteïna p6 s'uneix com a dímer a un segment de 24 parells de bases. Aquesta unió provoca un eixamplament del solc menor del DNA com ho suggereix la hipersensibilitat que mostra per la DNasa I.

La superhèlix dextrògira es proposa a partir dels contactes cada 12 parells de bases que realitza cada monòmer de p6. Aquesta repetició de l'hèlix cada 12 parells de bases és major que la repetició de la conformació B del DNA, que és de 10,5 parells de bases (Serrano *et al.*, 1993a) per volta.

El reconeixement del DNA per part de la proteïna p6 no és específic de seqüència, sinó que reconeix una característica estructural depenent de la seqüència. La majoria de regions reconegudes per p6 contenen seqüències amb una predicció de propietats deformables (Serrano *et al.*, 1989) capaç de ser corbades cada 12 parells de bases. La proteïna p6 reconeixeria la deformabilitat del DNA (Serrano *et al.*, 1993b; González-Huici *et al.*, 2004), amb especial predilecció pels extrems del genoma. Així, les regions per les quals p6 té més afinitat podrien actuar com a punts de nucleació en la formació del complex p6-DNA.

En unir-se p6, la molècula de DNA pateix una gran reducció en la seva llargada, fruit de l'empaquetament de la molècula de DNA al voltant de la proteïna p6 (Figura 1.8).

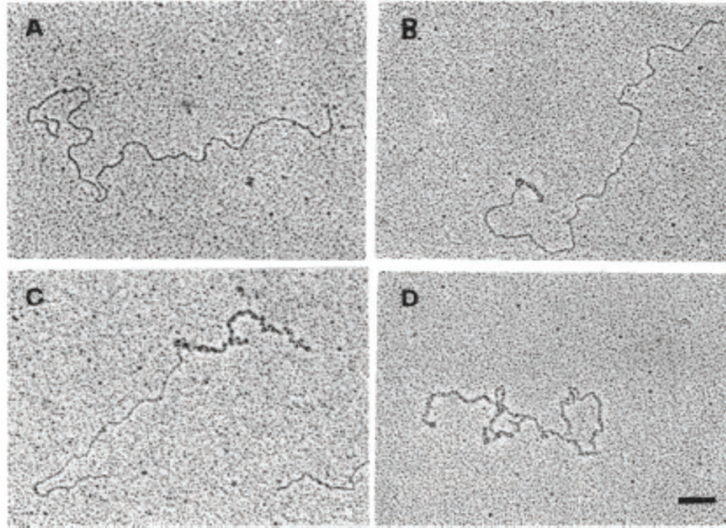


Figura 1.8: Mitjançant microscòpia electrònica es pot observar la reducció que pateix la molècula nua de DNA al unir-se la proteïna p6 (Serrano *et al.*, 1993b )

El model proposat de la superhèlix de DNA (Figura 1.9) és de 63 parells de bases per volta (Serrano *et al.*, 1993b), implicant 2,6 dímers de p6. L'amplada de la superhèlix seria de 5,1 nm i tindria un diàmetre de 6,6 nm.

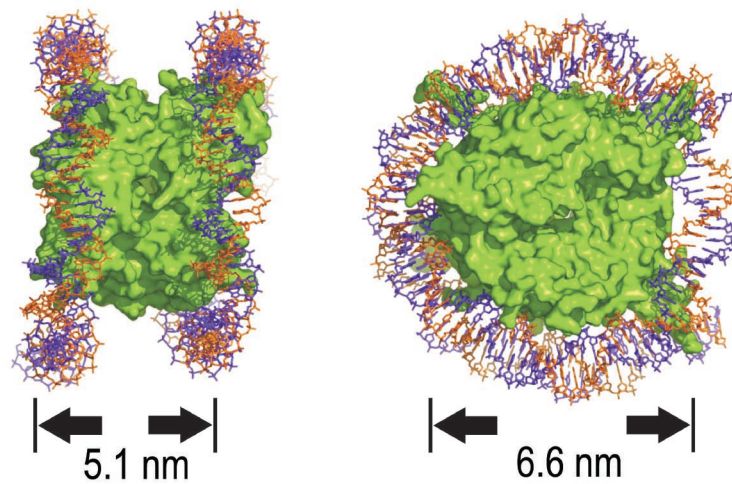


Figura 1.9: esquema de model d'empaquetament proposat, on el DNA s'enrotllaria al voltant d'un nucli format per proteïna p6.

El DNA es corbaria 66° cada 12 parells de bases i es trobaria relaxat, fent un pas de volta de 11,5 parells de bases (Serrano *et al.*, 1993b), en comparació amb els 10,5 del DNA en conformació B.

#### ***1.2.5.2. Implicacions funcionals del complex de nucleoproteïna***

La formació del complex nucleoproteic té funcions de suport en la replicació del DNA i en el control transcripcional. La proteïna p6 és absolutament imprescindible per a la replicació *in vivo* del DNA viral (Bravo *et al.*, 1994; Camacho *et al.*, 2000). *In vitro*, p6 estimula l'inici de la replicació (Blanco *et al.*, 1986; Pastrana *et al.*, 1985), així com la transició a l'elongació (Blanco *et al.*, 1988). Concretament, per afavorir la replicació es requereix la formació d'un complex de la proteïna p6 amb els fragments terminals del DNA de  $\phi$ 29 (Abril *et al.*, 2000; Freire *et al.*, 1994).

El mecanisme pel que la proteïna p6 activa la iniciació de la replicació del DNA no és clar. S'ha proposat que la gran alteració de l'estructura del DNA produïda per p6 pot facilitar el desenrotllament del DNA en llocs específics (Serrano *et al.*, 1993a).

La proteïna p6 també té un efecte en el control de la transcripció a través de la repressió del promotor primerenc C2. El complex de nucleoproteïna travessa el promotor C2 (Prieto *et al.*, 1988; Serrano *et al.*, 1989) i això explicaria la incapacitat de l'RNA polimerasa de formar el complex tancat (Camacho *et al.*, 2001a).

Per altra banda, la proteïna p6 s'uneix dèbilment a la regió que conté els promotors A2b, A2c i A3, sense modificar l'expressió d'aquests promotors (Elías-Arnanz *et al.*, 1999). En presència de p4, però, les dues proteïnes s'uneixen cooperativament, augmentant l'efecte de p4 en el control transcripcional.



### ***1.2.6. Control de la transcripció pel complex p4-p6***

P6 ajudaria p4 a reprimir A2b i A2c, i a activar A3 a través de la formació de complex de nucleoproteïna (Calles *et al.*, 2002; Camacho *et al.*, 2001b; Elías-Arnanz *et al.*, 1999).

La unió de la proteïna p6 a la regió del DNA on es troben els llocs d'unió de p4 és molt pobre. Aquesta unió es veu afavorida en presència de la proteïna p4, establint-se una cooperativitat en la unió (Elías-Arnanz *et al.*, 1999). Per a aquesta cooperativitat en la unió, en que es forma un complex de nucleoproteïna p4-p6, es requereix que la proteïna p4 tingui l'extrem carboxi-terminal íntegre (Calles *et al.*, 2002), tal i com ocorre en la repressió de A2c i l'activació de A3.

Els residus més importants dins de l'extrem carboxi-terminal de la proteïna p4 són R116, R118, R120, R121 i K122. La modificació d'aquests residus impedeix la formació del complex p4-p6, suggerint que el contacte entre les proteïnes implica dues superfícies carregades (Calles *et al.*, 2002).

#### ***1.2.6.1. Repressió del promotor A2b pel complex p4-p6***

El promotor A2b, en absència de la proteïna p6, és reprimat per un solapament entre el lloc d'unió de p4 i la regió -35 del promotor A2b. La incapacitat de l'RNA polimerasa de desplaçar la proteïna p4 impedeix la transcripció del promotor.

En el cas del complex p4-p6 la impossibilitat de desplaçar p4 es veu més accentuat. La unió de les proteïnes p4 i p6 al DNA es realitza de manera cooperativa. La repressió de A2b es veuria augmentada per dos motius. El primer seria la major afinitat de la proteïna p4 pel seu lloc d'unió i el segon seria el cobriment del promotor A2b, degut a la polimerització de la proteïna p6 en direcció cap al lloc d'inici de la transcripció de A2b.

### 1.2.6.2. Repressió del promotor A2c pel complex p4-p6

En absència de p6, la proteïna p4 s'uneix al lloc 1, solapant la regió -35 del promotor A2c, d'igual manera que en el promotor A2b. En unir l'RNA polimerasa, la proteïna p4 és moguda al lloc 2 (Figura 1.10). Aquest moviment permet que la proteïna p4 contacti amb la subunitat  $\alpha$  de l'RNA polimerasa i que aquesta es quedi en el promotor per una sobreestabilització del complex RNA polimerasa-promotor-proteïna p4.

En presència de la proteïna p6, la proteïna p4 és fixada en el lloc 1 d'unió. La unió de la proteïna p4 al lloc 1 actua com a punt de nucleació per la polimerització de p6 continuant *downstream* del promotor A2c.

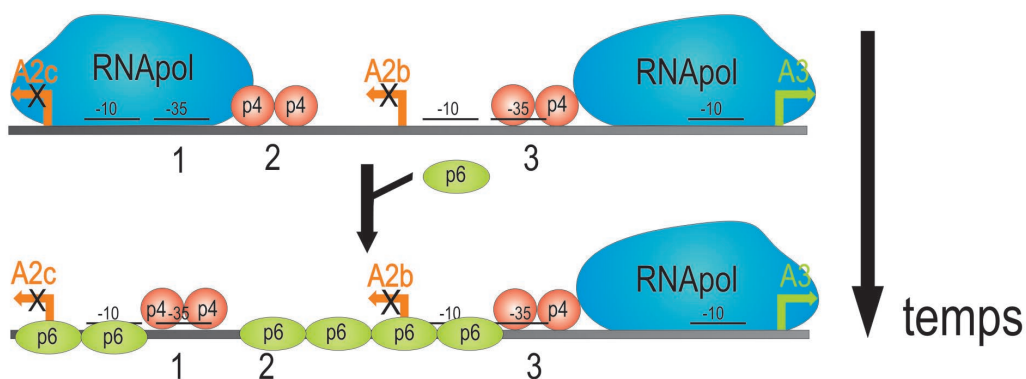


Figura 1.10: Esquema del canvi de posició de la proteïna p4 que impedeix la unió de l'RNA polimerasa en presència de la proteïna p6.

El promotor A2c és reprimat pel mateix mecanisme que el promotor A2b en presència de la proteïna p6 degut a la incapacitat de l'RNA polimerasa de desplaçar la proteïna p4, estabilitzada per p6, del lloc 1, que solapa amb la regió -35 del promotor A2c.

### ***1.2.6.3. Activació del promotor A3 pel complex p4-p6***

La presència del complex p4-p6 en el promotor A3 no modifica el mecanisme pel qual és activat per la proteïna p4 sola. La polimerització de la proteïna p6 a partir del lloc 3 d'unió de la proteïna p4 es realitza en direcció *upstream* del promotor A3, en direcció al promotor A2c (Calles *et al.*, 2002). En canvi, aquesta polimerització no ocorre en direcció *downstream*, quedant la zona d'unió de l'RNA polimerasa totalment accessible.

Així, tal i com ocorria en absència de la proteïna p6, l'RNA polimerasa contacta amb la proteïna p4 i s'estabilitza al promotor com a complex tancat (Nuez *et al.* 1992). L'única diferència entre els dos mecanismes d'activació del promotor A3 estaria en l'augment de l'afinitat de p4 pel seu lloc d'unió en presència de la proteïna p6 degut a la cooperativitat en la unió que presenten aquestes dues proteïnes.

### ***1.2.7. Coexistència de 2 mecanismes de regulació in vivo***

La demostració de l'existència de dos mecanismes de regulació de la transcripció *in vitro* ens condueixen a plantejar la solució a diverses incògnites, com saber si els dos mecanismes coexisteixen en el temps, si hi ha una separació temporal entre ells o només un d'ells és el que ocorre realment *in vivo*.

La proteïna p6 facilita un mecanisme alternatiu que és més efectiu en la regulació transcripcional, ja que estabilitza la proteïna p4 en els seus llocs d'unió, però no sembla que sigui l'únic que ocorre *in vivo*. L'explicació més versemblant és que els dos mecanismes ocorren *in vivo*, però estan separats temporalment depenent de les concentracions de les proteïnes p4 i p6.

En un primer moment, durant la fase primerenca d'infecció en què es comencen a transcriure els gens que codifiquen per les proteïnes p4 i p6, el mecanisme principal seria el que implicaria la proteïna p4 sola. La concentració de p6 seria baixa i les

molècules de proteïna p6 estarien localitzades en els extrems del genoma, dedicades a iniciar de la replicació vírica.

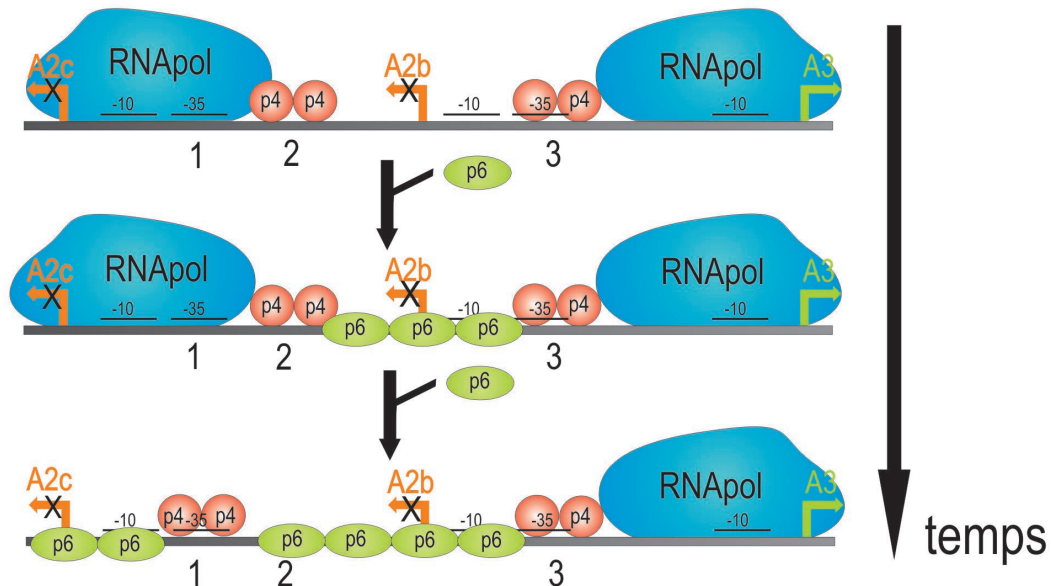


Figura 1.11: esquema de la transició entre els dos mecanismes de control de la transcripció.

La repressió dels promotors A2b i A2c i l'activació dels promotors A3, vindria intervinguda per l'acció de la proteïna p4 únicament. Aquesta regulació transcripcional no seria molt efectiva, existint transcripció des dels promotors primerencs; produint-se més quantitat de les proteïnes p4 i p6.

A mesura que la concentració de les proteïnes p4 i p6 augmentés (Figura 1.11), es faria el canvi de mecanisme, implicant les dues proteïnes i fent una regulació més forta i efectiva que reprimís completament els promotors A2b i A2c (Calles *et al.*, 2002).

### 1.3. L'enzim RNA polimerasa

Les molècules d'RNA de tots els organismes cel·lulars són sintetitzades per una maquinària molecular complexa, l'RNA polimerasa dependent de DNA. El seu

nucli catalític, constituït per les subunitats  $\alpha_2\beta\beta'\omega$  (Chamberlin, 1976) i amb un pes molecular de 450 kDa, està conservat en seqüència, estructura i funció des dels bacteris als humans (Archambault *et al.*, 1993). L'inici de la transcripció en promotors específics requereix una subunitat addicional, la subunitat  $\sigma$ , que s'uneix al nucli de l'RNA polimerasa per constituir l'holoenzim.

Al llarg de les diferents etapes que constitueixen la transcripció, l'RNA polimerasa canvia el seu comportament respecte al DNA. Durant l'inici de la transcripció reconeix el DNA de la regió del promotor enmig de tota la resta de DNA, de manera altament específica. En un pas posterior separa la doble cadena de la regió de l'element -10 per exposar la cadena motlle i poder iniciar la síntesi d'RNA utilitzant només oligonucleòtids. Abans de començar l'elongació hi ha un canvi en el seu comportament com a proteïna d'unió a DNA específica de seqüència per passar a comportar-se com a una proteïna d'unió a DNA no específica de seqüència que li permet moure's al llarg del DNA.

Diversos estudis estructurals (Zhang *et al.*, 1999; Cramer *et al.*, 2000; Murakami *et al.*, 2002; Vassylyev *et al.*, 2002) han aportat una nova visió al funcionament de l'enzim.

### **1.3.1. L'holoenzim**

L'estructura obtinguda per difracció de raigs X del nucli de l'RNA polimerasa de *Thermus aquaticus* mostra que l'enzim presenta un canal central de 27 Å de diàmetre. La molècula mesura 150 Å de llarg, 115 Å d'alt i 110 Å d'ample (Figura 1.12).

L'RNA polimerasa té una forma de pinça de cranc i el seu lloc actiu es troba a la base de les dues pinces que la formen. L'espai entre les dues pinces forma el canal central on es situa el DNA. Precisament és en aquest lloc on se separen les dues cadenes del DNA, quedant la cadena codificant molt a prop del centre actiu per poder ser utilitzada com a motlle per ser transcrita.

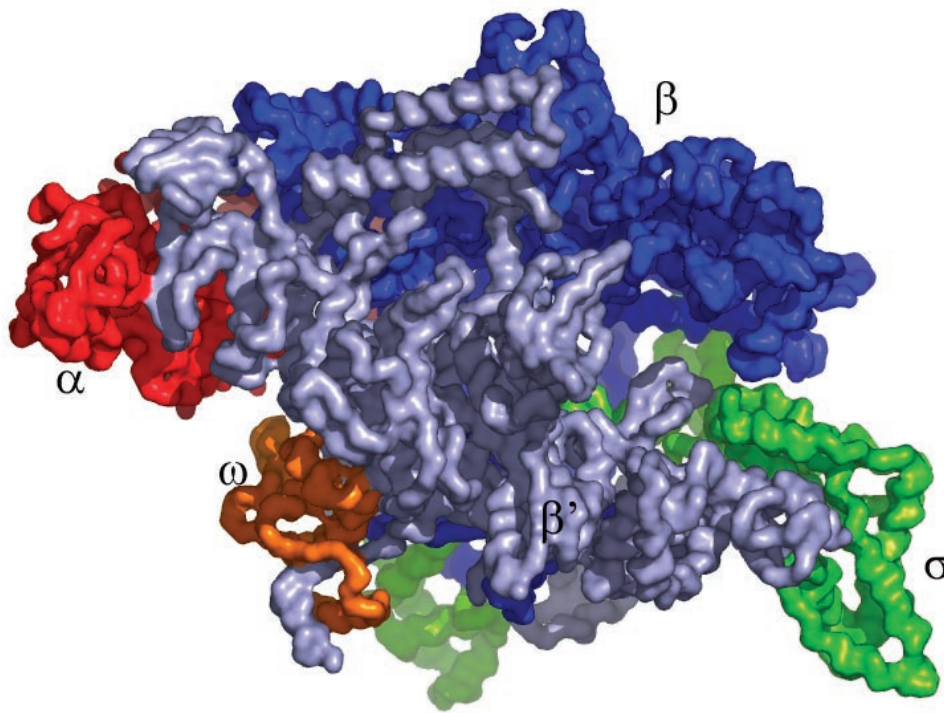


Figura 1.12: Representació en superfície de l'estructura de l'holoenzim de l'RNA polimerasa (Murakami *et al.*, 2002) Codi PDB: 1L9U.

### ***1.3.2. Les subunitats de l'RNA polimerasa***

#### ***1.3.2.1. La subunitat $\alpha$***

La subunitat  $\alpha$ , de 37kDa, és la responsable de reconèixer els elements UP (element suplementari del promotor localitzat *upstream* de la regió -35 en alguns promotors; Ross, 1993), i respon a gran varietat d'activadors, repressors, factors d'elongació i factors de terminació (Busby *et al.*, 1994; Ebright *et al.*, 1995). Estructuralment, consisteix en dos dominis plegats independentment (Blatter, 1994).

El domini amino-terminal o  $\alpha$ NTD està format pels residus 8-235 a *Escherichia coli* i pels residus 4-229 a *Bacillus subtilis*. Aquest domini conté les regions principals per la dimerització de  $\alpha$ . El  $\alpha$ NTD interacciona amb l'RNA polimerasa i pot interaccionar amb alguns activadors.

El domini carboxi-terminal o  $\alpha$ CTD està format pels residus 249-329 a *Escherichia coli* i pels residus 245-314 a *Bacillus subtilis*. El  $\alpha$ CTD conté una petita regió implicada en la dimerització de  $\alpha$  però la seva funció principal és interaccionar amb el DNA i amb diferents factors proteics que intervenen en la transcripció com activadors, repressors, factors d'elongació i factors de terminació.

Els dos dominis estan units per un connector de 13 aminoàcids de llargària, totalment flexible i desestructurat (Blatter *et al.*, 1994; Negishi *et al.*, 1995) que permet que el dominis es puguin separar fins a 44 Å. Aquest llarg i desestructurat connector permetria al  $\alpha$ CTD ocupar diferents posicions relatives al  $\alpha$ NTD i a la resta de l'RNA polimerasa, en diferents complexes de transcripció (Blatter *et al.*, 1994; Busby *et al.*, 1994).

### **1.3.2.2. Les subunitats $\beta$ i $\beta'$**

Les subunitats  $\beta$ , de 151 kDa, i  $\beta'$ , de 155 kDa, són les responsables de l'activitat catalítica de l'RNA polimerasa, però també respon a activadors, repressors, factors d'elongació i factors de terminació (Chamberlin *et al.*, 1976; Severinov *et al.*, 1994; Miller *et al.*, 1997; Necheav *et al.*, 1999).

Les subunitats  $\beta$  i  $\beta'$  formen les pinces i interaccionen amb diferents regions de  $\sigma$ , essent responsables del canvi d'obertura que s'observa en el canal central, com també intervenen en el pas cap a l'elongació per interacció amb  $\sigma_4$  i amb el *loop*  $\sigma_{3,2}$  (Murakami *et al.*, 2002).

### 1.3.2.3. La subunitat $\omega$

La subunitat  $\omega$ , de 12kDa, no té una funció clara. Per una part no és una subunitat essencial per la viabilitat del bacteri sota condicions de laboratori (Gentry *et al.*, 1991), encara que la seva mancança condueix a fenotips de creixement lent (Mukherjee *et al.*, 1997). Per altra part tots els organismes vius seqüenciats presenten proteïnes homòlogues a  $\omega$ , suggerint una important funció conservada al llarg de l'evolució.

La funció més clara associada a la subunitat  $\omega$ , segons el que s'observa en reconstitucions *in vitro* de l'RNA polimerasa (Mukherjee *et al.*, 1997), seria l'afavoriment del muntatge de les altres unitats per estabilització de la subunitat  $\beta'$  (Minakhin *et al.*, 2001).

### 1.3.2.4. La subunitat $\sigma$

La majoria dels factors  $\sigma$  pertanyen a una família homologa que és molt propera a  $\sigma^{70}$  de *Escherichia coli*, amb diferents regions de seqüència altament conservada (Gruber *et al.*, 1997; Lonetto *et al.*, 1992). Aquesta subunitat, de pes molecular variable, és la responsable del reconeixement dels elements -35 i -10. Hi ha diversos tipus de subunitats  $\sigma$ , 7 a *E.coli* i 8 a *B.subtilis*, essent  $\sigma^{70}$  i  $\sigma^A$  la majoritària respectivament.

El canvi de subunitat és el mecanisme principal per canviar el patró de transcripció, sobretot en resposta a factors ambientals (taula 1.1). La competició entre les diferents subunitats  $\sigma$  disponibles hauria de ser clau per saber quins gens són transcrits.

Les anàlisi estructurals del factor  $\sigma^{70}$  (Murakami *et al.*, 2002) mostren que els factors  $\sigma$  presenten quatre dominis connectats per un *linker* flexible. Aquests quatre dominis són  $\sigma_{1.1}$ ,  $\sigma_2$ ,  $\sigma_3$  i  $\sigma_4$ , que contenen les regions conservades 1.1, 1.2-2.4, 3.0-



3.1 i 4.1-4.2, respectivament (Campbell *et al.*, 2002).

Subunitat $\sigma$	Gen	Funció
$\sigma^{43}$ , $\sigma^A$	rpoD, sigA	factor $\sigma$ principal
$\sigma^B$	sigB	regulació de gens de la fase estacionària
$\sigma^D$	sigD	transcripció de gens flagel·lars
$\sigma^E$	sigE (spoIIGB)	expressió de gens en la cèl·lula mare, generadora d'espores
$\sigma^F$	sigF (spoIIAC)	expressió de gens en l'espore
$\sigma^G$	sigG (spoIIIG)	expressió de gens en l'espore
$\sigma^H$	sigH (spoOH)	transcripció dels gens primerencs d'esperulació
$\sigma^K$	sigK (spoIVCB-spoIIIC)	transcripció dels gens tardans en la cèl·lula mare, generadora d'espores

Taula 1.1: Les diferents subunitats  $\sigma$  són codificades per gens diferents i la seva funció està fortament lligada a factors ambientals.

En l'holoenzim, els dominis globulars de  $\sigma$  estan repartits per tota la part de la pinça de l'RNA polimerasa (Murakami *et al.*, 2002 ; Vassylyev *et al.*, 2002). Les regions del factor  $\sigma$  que s'uneixen al promotor,  $\sigma_2$  a l'element -10 i  $\sigma_4$  a l'element -35, estan exposades al solvent i separades entre si el corresponent a la distància que hi ha entre els dos elements. Els dominis  $\sigma_3$  i  $\sigma_4$  estan separats per 45 Å en l'holoenzim. Aquesta distància és recorreguda per un connector estès de 33 residus que compren la regió 3.2 de  $\sigma$  ( el *loop*  $\sigma_{3,2}$ ) que s'introdueix dins del canal del lloc actiu de l'RNA polimerasa, per després sortir pel canal de sortida del RNA.

La interacció entre  $\sigma$  i el nucli de l'RNA polimerasa és molt estable. Cada un dels dominis de  $\sigma$ , així com els connectors que els uneixen, fan interaccions amb el nucli de l'RNA polimerasa. No obstant, quan l'RNA polimerasa entra en la fase d'elongació de la transcripció, normalment el factor  $\sigma$  és alliberat. Aquestes aparents contradiccions s'expliquen per l'arquitectura del factor  $\sigma$ .

### **1.3.2.5. Funcions del factor $\sigma$**

#### **1.3.2.5.1. El domini $\sigma_{1,1}$**

Els factors  $\sigma$  del grup 1 tenen una única extensió amino-terminal d'uns 90 residus anomenada regió 1.1. Aquesta regió està pobrament conservada en seqüència, encara que la seva acidesa característica es manté. S'han descrit dues funcions per aquesta regió. La primera és que autoinhibeix el reconeixement del promotor per part d'altres factors  $\sigma$  lliures (Camarero *et al.*, 2002; Dombroski *et al.*, 1992). La segona funció consisteix en accelerar la formació del complex obert en molts dels promotors (Vuthoori *et al.*, 2001).

El domini  $\sigma_{1,1}$  es posiciona dins del canal del centre actiu de l'RNA polimerasa. Un cop es forma el complex obert en el promotor, el domini  $\sigma_{1,1}$  és desplaçat fora del canal (Mekler *et al.*, 2002), explicant com el domini  $\sigma_{1,1}$  afecta la cinètica de la formació del complex obert. S'ha proposat que el posicionament de  $\sigma_{1,1}$  en el canal de l'RNA polimerasa podria eixamplar el canal per facilitar l'entrada del DNA de doble cadena (Murakami *et al.*, 2002)

#### **1.3.2.5.2. El domini $\sigma_2$**

La regió 2.4 reconeix l'element -10 del promotor. La regió 2.3 és essencial per "fondre" el DNA a través d'un reconeixement específic de seqüència de simple cadena de la cadena no codificant de l'element -10 (Marr *et al.*, 1997), que estabilitza la bombolla de l'inici de la transcripció.

Els residus de la regió 2.4 de  $\sigma$  estan posicionats en la posició -12, que és la única porció de doble cadena de l'element -10. Diversos residus aromàtics altament conservats de la regió 2.3, que són importants per la funció de "fondre" el promotor, es posicionen adequadament per interaccionar amb les bases exposades de la cadena no codificant de l'element -10.

#### **1.3.2.5.3. El domini $\sigma_3$ i el loop $\sigma_{3.2}$**

El domini  $\sigma_3$  és un domini compacte de tres hèlix  $\alpha$ . Els aminoàcids de l'hèlix amino-terminal, que constitueixen la regió 3.0, estan implicats en el reconeixement dels elements -10 estesos (Barne *et al.*, 1997) estabilitzant així el complex obert (Voskuil *et al.*, 2002).

Variants trucades en l'extrem carboxi-terminal de  $\sigma$ , que els mancava el loop  $\sigma_{3.2}$  i el domini  $\sigma_4$ , presenten poca activitat transcripcional en promotors amb la regió -10 estesa. La proximitat del loop  $\sigma_{3.2}$  al lloc actiu suggereix que l'eliminació del loop  $\sigma_{3.2}$  és l'origen del dèficit en la unió del nucleòtid iniciador.

#### **1.3.2.5.4. El domini $\sigma_4$**

El domini  $\sigma_4$  compren les regions 4.1 i 4.2 i té forma de ferradura, amb una butxaca còncava envoltada per residus hidrofòbics de la regió 4.1. La funció que realitza és la d'unir l'element -35 induint una curvatura en el DNA que permet situar-lo més proper a l'RNA polimerasa. Així facilita les interaccions entre el domini carboxi-terminal de la subunitat  $\alpha$  i el DNA *upstream* (Ross *et al.*, 1993) i les interaccions amb activadors que s'uneixen *upstream* de l'element -35.

### **1.3.3. El camí de l'inici de la transcripció a l'elongació**

L'inici de la transcripció en els organismes procariotes és un procés que consta de diversos passos. En el primer pas l'RNA polimerasa uneix específicament el promotor, formant el que s'anomena complex tancat. Llavors, el complex tancat pot isomeritzar formant el complex obert, transcripcionalment actiu, en el que les cadenes del DNA del voltant del lloc d'inici de transcripció són separades. En aquest punt es poden incorporar nucleòtids i es comencen a sintetitzar petits trànscrips abortius. Normalment, aquests trànscrips no són majors de 11 nucleòtids i l'RNA polimerasa roman unida al promotor. Per poder realitzar trànscrips de gran llargària,

l'RNA polimerasa ha d'escapar del promotor i procedir amb la fase productiva d'elongació. En aquesta fase es transcriuen les molècules de mRNA senceres. La interacció de l'holoenzim de l'RNA polimerasa amb el promotor inicia una sèrie de transicions estructurals que comencen amb el complex tancat inicial i acaba amb el complex d'elongació ternari (Murakami i Darst, 2003).

#### ***1.3.3.1. El complex tancat***

Els dominis  $\sigma_4$  i  $\sigma_2$  contacten amb els elements -35 i -10 respectivament. L'holoenzim se situa cobrint el DNA desde la posició -54 fins a la -6. El *loop*  $\sigma_{3,2}$  es situa a l'interior del canal del lloc actiu de l'RNA polimerasa, a l'igual que el domini  $\sigma_{1,1}$ , que s'hi posiciona mitjançant interaccions electrostàtiques (Murakami *et al.*, 2002) (Figura 1.13).

#### ***1.3.3.2. Passos intermedis***

Els residus aromàtics de la regió 2.3 de  $\sigma$  es posicionen perfectament per poder interaccionar amb les bases de la cadena no codificant de l'element -10. Aquesta interacció es dona l'exposició d'aquestes bases provocada per la distorsió natural que presenta aquesta regió rica en AT (Spassky *et al.*, 1988). Aquestes interaccions entre la proteïna i el DNA de cadena simple estabilitzaria un petit segment de DNA que formaria l'inici de la bombolla de transcripció. La separació de les dues cadenes està associada a un desenrotllament del DNA que aporta flexibilitat al DNA de la bombolla i permet que el DNA es curvi cap a l'entrada del canal del lloc actiu (Figura 1.13). Aquest procés d'entrada del DNA en el canal de l'RNA polimerasa ve acompanyat per la sortida de  $\sigma_{1,1}$  (Murakami *et al.*, 2002).

#### ***1.3.3.3. El complex obert***

En el complex obert de l'RNA polimerasa i el promotor, la separació de les cadenes de DNA, que forma la bombolla de transcripció, s'estén més enllà del punt

d'inici de la transcripció. La cadena codificant s'introdueix en el lloc actiu dirigit per un canal carregat positivament (Murakami *et al.*, 2002). El DNA de doble cadena que compren les posicions +5 a +12 s'introdueix entre les subunitats  $\beta$  i  $\beta'$ .

Quan es forma el complex obert, el lloc actiu de l'RNA polimerasa rep el subministre de NTPs per un canal secundari i es comença a sintetitzar la cadena d'RNA (Figura 1.13). El petit trànscrip de pocs nucleòtids de llargada es troba en el seu camí de sortida el *loop*  $\sigma_{3,2}$  començant-se el procés de síntesi abortiva (Murakami *et al.*, 2002). La cadena elongada d'RNA ha de desplaçar el *loop*  $\sigma_{3,2}$  o dissociar-se del complex i alliberar-se. Un cop la cadena d'RNA aconsegueix desplaçar el *loop*  $\sigma_{3,2}$  s'acaba el procés de síntesi abortiva.

#### ***1.3.3.4. Escapament del promotor***

El desplaçament del *loop*  $\sigma_{3,2}$  provocaria una desestabilització de les interaccions entre  $\sigma_4$  i el nucli de l'enzim. L'alliberament de  $\sigma_4$  desestabilitzaria les interaccions entre  $\sigma_4$  i l'element -35, permetent a l'RNA polimerasa de sortir del promotor i començar a elongar l'RNA (Figura 1.13). Aquesta transició a l'elongació no requereix l'alliberament complet de  $\sigma$  (Bar-Nahum *et al.*, 2001), per que la unió de  $\sigma_2$  i  $\sigma_3$  a l'RNA polimerasa no és incompatible amb l'elongació de la cadena d'RNA.

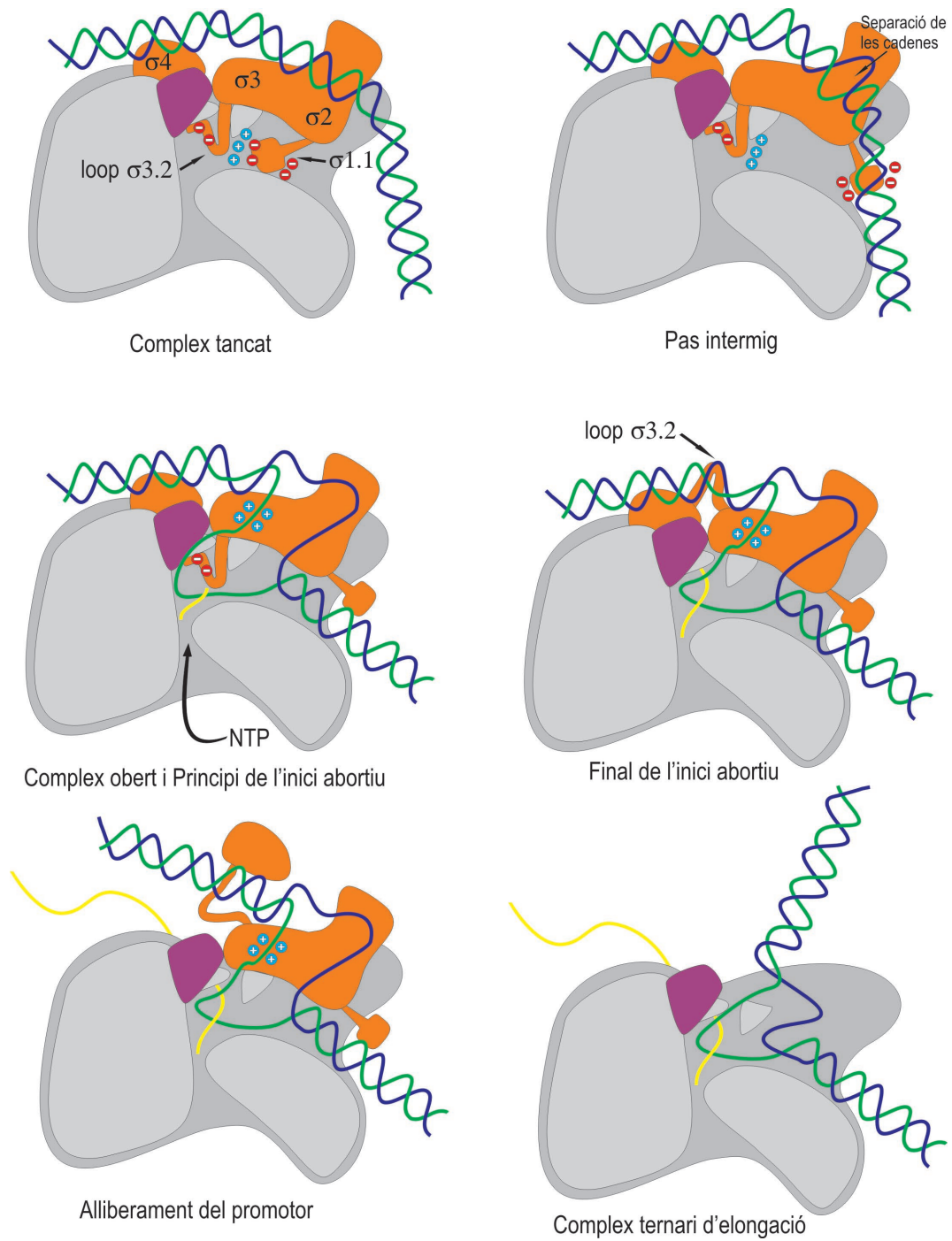


Figura 1.13: Esquema dels processos realitzats per l'holoenzima de l'RNA polimerasa durant l'inici de la transcripció. (Figura elaborada a partir de Murakami i Darst, 2003)

## ***1.4. Regulació de la transcripció***

### ***1.4.1. Elements que intervenen en la regulació de la transcripció***

En els promotors de gens no expressats constitutivament, sotmesos a regulació, trobem proteïnes accessòries a l'RNA polimerasa. Gran part d'aquestes proteïnes contacten directament o indirectament amb l'RNA polimerasa actuant com a activadors o com a repressors de la transcripció depenent de l'efecte que tinguin sobre ella. L'especificitat funcional de l'holoenzim de l'RNA polimerasa és modulada per interacció d'un, o en alguns casos, dos dels aproximadament 100-150 diferents factors de transcripció presents en la cèl·lula procariota.

Aquests factors es poden classificar basant-nos en quines subunitats de l'RNA polimerasa contacten. Originalment aquestes proteïnes reguladores es van classificar en activadors i en repressors depenent de la seva acció sobre la transcripció. Estudis posteriors han indicat que tant els activadors com els repressors poden tenir funcions duals, activant o reprimint la transcripció de diferents promotors depenent d'on uneixen el DNA (Aiba, 1983; Choy *et al.*, 1995; Ishizuka *et al.*, 1994), però tots dos sembla que realitzen contactes directes amb l'RNA polimerasa per funcionar.

Existeixen 4 classes depenent de la subunitat de l'RNA polimerasa contactada per la proteïna reguladora. La classe 1 correspon a proteïnes que contacten la subunitat  $\alpha$ , la classe 2 a les que contacten  $\sigma$ , la classe 3 a les que contacten  $\beta$  i la classe 4 a les que contacten  $\beta'$ .

### ***1.4.2. Reguladors de classe 1: el domini carboxi-terminal de la subunitat $\alpha$***

El  $\alpha$ CTD ha estat intensivament estudiat en el sistema de promotors activats per CAP, podent-se diferenciar 2 tipus de promotors on la seva transcripció depèn de CAP: el primer tipus comprèn promotors on la interacció entre CAP i  $\alpha$ CTD és suficient per activar la transcripció. El segon tipus comprèn promotors on la

interacció entre CAP i  $\alpha$ CTD és una de les múltiples necessàries per activar la transcripció.

Substitucions en cada aminoàcid particular van servir per identificar 3 regions crítiques en el  $\alpha$ CTD. Aquestes 3 regions reben els noms del residu més important, obtenint així la *regió 265*, la *regió 261* i la *regió 287*.

La *regió 265*, que comprèn els residus R265, N268, N294, G296, K298, S299 i E302, contacta amb el DNA (Gaal *et al.*, 1996) mitjançant contactes amb fosfats en la zona del solc menor del DNA.

La *regió 261*, que comprèn els residus V257, D258, D259, E261, no és essencial per la interacció  $\alpha$ CTD-DNA (Tang *et al.*, 1994) però és important per la transcripció dependent de CAP, contactant amb la subunitat sigma i amb elements UP.

La *regió 287*, que comprèn els residus T285, E286, V287, L289, G315, A317 i L318, no és necessària per la interacció amb el DNA i no és crucial per a la transcripció dependent d'element UP (Gaal *et al.*, 1996). És essencial per les interaccions proteïna-proteïna entre CAP i  $\alpha$ CTD al promotor. Mutacions en aquesta regió eliminen la cooperativitat de  $\alpha$ CTD i CAP.

Comparant les seqüències de les subunitats  $\alpha$  de l'RNA polimerasa de *E.coli* i de *B.subtilis* veiem com la regió d'unió a DNA, la *regió 265*, es troba totalment conservada, i les altres regions, *261* i *287*, no mantenen la identitat però sí una gran homologia (Taula 1.14).



<i>E. coli</i>	FDPIILLRPVDDLELTVRSANCLKAEATHYIGDLVQRTEVEL	289
<i>B. subtilis</i>	KEKVLEMTIEELDLSVRSYNCLKRAGINTVQELANKTEEDM	285
	: :* .: :*:*:*** ***** .*: : :*: :*:** ::	
<i>E. coli</i>	LKTPNLGKKSLEIKDVLASRGLSLGMRLLENWPPASIADE	329
<i>B. subtilis</i>	MKVRNLGRKSLEEVKAKLEELG--LGLRKDD-----	314
	:* . ***:*** *:* * . * **:* ::	

Figura 1.14: Alineament dels dominis carboxi-terminals de les subunitats  $\alpha$  de les RNA polimerases de *E. coli* i de *B. subtilis*. En vermell es mostra la regió 265, en groc la regió 261 i en blau la regió 287.

Els últims 15 residus de la subunitat  $\alpha$  de l'RNA polimerasa de *Bacillus subtilis*, essencials per la interacció amb la proteïna p4, corresponen a la regió 287, que a *Escherichia coli* està implicada amb la unió amb CAP.

#### 1.4.3. Estructura del complex DNA-CAP- $\alpha$ CTD

Existeix una estructura cristal·lina del complex format per la proteïna CAP i el domini carboxi-terminal de la subunitat  $\alpha$ , units a la seva regió del promotor (Benoff *et al.*, 2002).

En l'estructura es mostra que el  $\alpha$ CTD és un domini globular (Figura 1.15), que s'uneix al solc menor del DNA per la regió 265. Els contactes que realitza el  $\alpha$ CTD amb les bases són a través d'una molècula d'aigua, però no existeix cap interacció directa amb les bases. Diversos contactes amb els fosfats estableixen la unió del  $\alpha$ CTD al DNA.

El contacte entre la proteïna CAP i el  $\alpha$ CTD (Figura 1.16), tal i com s'esperava, es dona en la regió 287 de  $\alpha$ CTD, essencial per la interacció entre CAP i  $\alpha$ CTD (Gaal *et al.*, 1996). La superfície de contacte entre les dues proteïnes és molt petita, però intervien una gran quantitat de interaccions de tipus pont d'hidrogen.

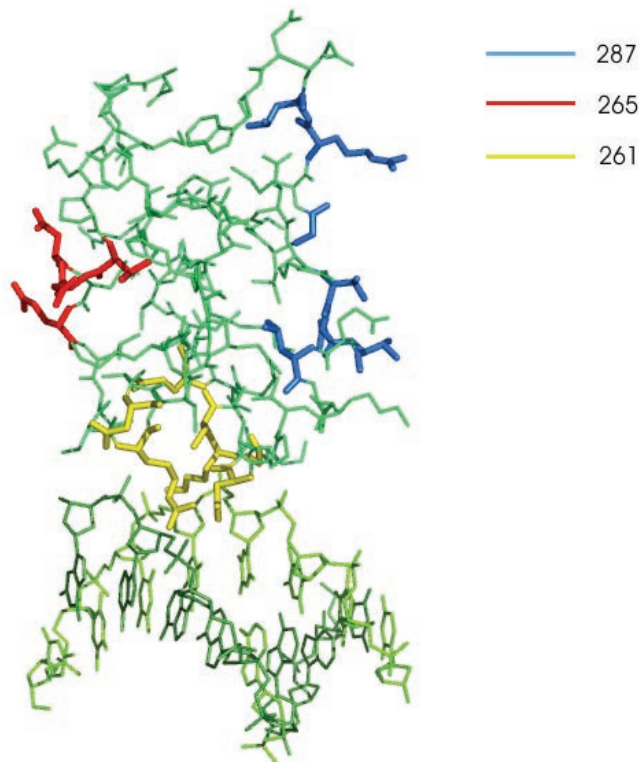


Figura 1.15: representació del model obtingut per difracció de raigs X del  $\alpha$ CTD unit a DNA. Se senyalen les diferents regions per colors (Codi PDB: 1LB2) (Benoff *et al.*, 2002).

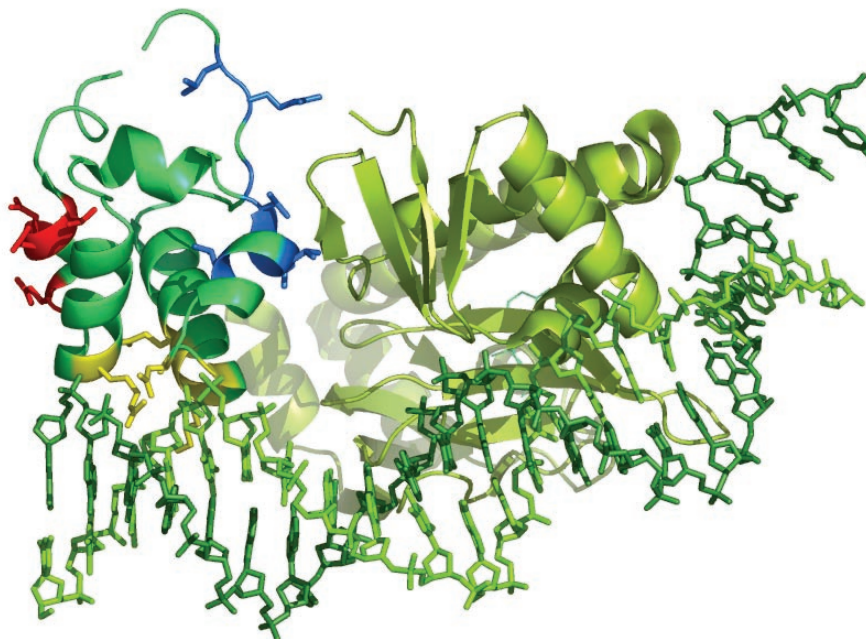


Figura 1.16: representació del model obtingut per difracció de raigs X del complex CAP- $\alpha$ CTD. Es marquen les regions de  $\alpha$ CTD amb diferents colors (Codi PDB: 1LB2) (Benoff *et al.*, 2002).