# UNIVERSITA' DEGLI STUDI DI PARMA

**Dipartimento di Biochimica e Biologia Molecolare**

Dottorato di Ricerca in Biotecnologie, XVIII Ciclo


# MODULAZIONE DELLE FASI DI INIZIO E TERMINAZIONE DELLA TRASCRIZIONE DA PARTE DELLA RNA POLIMERASI III EUCARIOTICA


Coordinatore del corso:
Prof. Nelson Marmiroli

Tutori:
Prof. Simone Ottonello
Prof. Giorgio Dieci

Dottoranda:
**Priscilla Braglia**

# INDICE

*1. INTRODUZIONE* ____ *4*

**1.1 La trascrizione eucariotica** ____ **4**

**1.2 Il sistema trascrizionale della RNA polimerasi III** ____ **5**
1.2.1 Promotori dei geni di classe III ____ 5
1.2.2 Assemblaggio del complesso di trascizione Pol III ____ 8
1.2.3 RNA polimerasi III ____ 10
1.2.4 TFIIIB ____ 12

**1.3 Il ciclo trascrizionale della RNA polimerasi III** ____ **14**
1.3.1 Inizio ____ 14
1.3.2 Allungamento ____ 14
1.3.3 Terminazione ____ 16
1.3.4 Re-inizio ____ 18

**1.4 La trascrizione e l'organizzazione della cromatina** ____ **20**
1.4.1 Cromatina e trascrizione ____ 20
1.4.2 Effetti cromatinici su geni di classe III ____ 26
1.4.3 I geni di classe III come barriere cromatiniche ____ 27

**1.5 La regolazione della trascrizione Pol III-dipendente** ____ **32**
1.5.1 Esempi di regolazione del sistema Pol III ____ 32
1.5.2 Nhp6 e RNA polimerasi III ____ 33

*2. SCOPO DELLA RICERCA* ____ *38*

*3. MATERIALI E METODI* ____ *41*

**3.1 Terminazione della trascrizione: effetti del contesto sulla funzionalità del terminatore** ____ **41**
3.1.1 Analisi statistica delle regioni fiancheggianti i tDNA in 3' ____ 41
3.1.2 DNA templato ____ 42
3.1.3 Saggi di trascrizione in vitro in condizioni standard ____ 43
3.1.4 Saggi di trascrizione in vitro in assenza di TFIIIC ____ 44
3.1.5 Analisi in vivo ____ 45

**3.2 Inizio della trascrizione: sequenza a monte dei geni per tRNA ed effetti della proteina Nhp6** ____ **46**
3.2.1 Analisi statistica delle regioni fiancheggianti i tDNA in 5' ____ 46
3.2.2 DNA templato ____ 47
3.2.3 Saggi di trascrizione in vitro ____ 48
3.2.4 Analisi in vivo ____ 48
3.2.5 Chromatin immunoprecipitation ____ 50
3.2.6 Purificazione di proteina ____ 51
3.2.7 Saggi di mating ____ 52

*4. RISULTATI E DISCUSSIONE* ____ *54*

**4.1 Terminazione della trascrizione: effetti del contesto sulla funzionalità del terminatore** ____ **54**

**4.2 Inizio della trascrizione: sequenza a monte dei geni per tRNA ed effetti della proteina Nhp6 _____ 76**
4.2.1 Sequenza a monte dei geni per tRNA _____ 76
4.2.2 Effetti di Nhp6 sui geni di classe III _____ 85

***5. CONCLUSIONE _____ 100***

**5.1 Terminazione della trascrizione: effetti del contesto sulla funzionalità del terminatore _____ 100**

**5.2 Inizio della trascrizione: sequenza a monte dei geni per tRNA ed effetti della proteina Nhp6 _____ 101**

***6. BIBLIOGRAFIA _____ 105***

***7. APPENDICE _____ 124***

# 1. INTRODUZIONE

## 1.1 LA TRASCRIZIONE EUCARIOTICA

Negli organismi eucariotici il compito di riconoscere i promotori e di trascrivere i geni nucleari è suddiviso fra tre diversi enzimi, altamente correlati tra loro, le RNA polimerasi (Pol) I, II e III. Ciascuna di queste RNA polimerasi è dedicata alla trascrizione di specifici *set* di geni, e ciascuna dipende, per il riconoscimento dei propri promotori, da specifici fattori accessori, i cosiddetti fattori di trascrizione (TF).

La RNA polimerasi I ha la caratteristica, unica tra le polimerasi nucleari, di trascrivere un singolo *set* di geni, il *cluster* degli RNA ribosomali (rRNA); Pol I, pertanto, deve essere in grado di riconoscere soltanto una singola struttura di promotore.

La RNA polimerasi II trascrive i geni codificanti proteine (geni per mRNA) ed alcuni geni per *small nuclear* RNA (snRNA). I promotori per la RNA polimerasi II possono essere generalmente suddivisi in una regione *core*, definita come regione minima capace di dirigere la trascrizione *in vitro*, ed in una regione regolativa. Le regioni regolative dei promotori Pol II presentano una struttura molto varia, riflesso del grande numero di geni che codificano proteine e della necessità di una specifica e complessa regolazione dell'espressione di tali geni. Gli stessi promotori *core* si suddividono in diversi tipi e possono contenere combinazioni di diversi elementi come TATA-*box*, iniziatori, promotori a valle.

La RNA polimerasi III è dedicata alla trascrizione di un insieme di geni la cui principale caratteristica comune è quella di codificare RNA, strutturali o catalitici, generalmente più corti di 400 paia-basi (bp) e non tradotti in proteine. Tali geni codificano per molecole di RNA implicate in processi metabolici fondamentali, come la biosintesi delle proteine e lo *splicing* nucleare, oltre che per alcuni RNA di funzione ignota. I promotori per la RNA polimerasi III sono stati suddivisi in tre tipi principali, due dei quali sono intragenici e generalmente privi di una TATA-*box*, ed uno dei quali è extragenico e caratterizzato dalla presenza di una TATA-*box*. E' stato raggiunto un buon livello di conoscenza del meccanismo attraverso il quale la RNA polimerasi III viene reclutata su ciascuno di questi

tipi di promotori. Pol III pertanto costituisce un buon modello che consente di comprendere alcune regole generali attraverso le quali uno stesso enzima viene reclutato su diverse strutture di promotore, attraverso diverse combinazioni di interazioni proteina-proteina e proteina-DNA (vedi (Schramm and Hernandez, 2002) per *review*).

Un'importante questione che riguarda la trascrizione eucariotica è la regolazione coordinata dei tre sistemi polimerasici nucleari; tale regolazione ha un'importanza fondamentale nell'economia della cellula in quanto un mancato bilanciamento risulterebbe in un enorme ed inutile dispendio di energia. In particolare, i due sistemi della RNA polimerasi I e III sono strettamente accoppiati fra loro ed influenzati dalla crescita cellulare in quanto implicati nella biogenesi dei ribosomi.

Si noti che a livello quantitativo le RNA polimerasi I e III sono responsabili della produzione di più dell'80% degli RNA cellulari, a fronte del piccolo numero di geni diversi che esse trascrivono. La trascrizione da parte di questi due enzimi, pertanto, viene condotta con alta efficienza; si può stimare che, mentre in media un gene trascritto da Pol II produce 1-2 trascritti in un'ora, nello stesso tempo un gene di classe III produce in media ~3000 trascritti.

## 1.2 IL SISTEMA TRASCRIZIONALE DELLA RNA POLIMERASI III

### *1.2.1 Promotori dei geni di classe III*

Esistono tre diversi tipi di promotore per la RNA polimerasi III, chiamati, rispettivamente, di tipo1, di tipo 2 e di tipo 3; la loro particolare struttura è schematizzata in *Figura 1.1*.

Il promotore dei geni codificanti il 5S rRNA è il solo esempio di promotore di tipo 1, mentre quelli dei geni che codificano gli RNA *transfer* (tRNA) sono tipici promotori di tipo 2. Questi due tipi di promotori hanno una caratteristica localizzazione intragenica.

Il promotore del gene *5S* di *X. laevis* è costituito da una A-*box*, da un elemento intermedio (IE) e da una C-*box*, conservata in diverse specie; nel loro insieme questi diversi elementi costituiscono le *Internal Control Regions* (ICR) (Bogenhagen, 1985) (Pieler et al., 1985a;

Pieler et al., 1985b; Pieler et al., 1985c). Nel gene *5S* di *Saccharomyces cerevisiae* soltanto la C-*box* è necessaria per dare inizio alla trascrizione (Challice and Segall, 1989).

I promotori della maggior parte dei geni per tRNA sono costituiti da una A-*box* e da una B-*box* (Galli et al., 1981) (Hofstetter et al., 1981) (Sharp et al., 1981) (Allison et al., 1983). Questi elementi di sequenza sono conservati nei tDNA di diverse specie eucariotiche, probabilmente anche per il fatto che codificano il D- ed il T-*loop* dei tRNA, elementi strutturali necessari per la funzione dei tRNA. Varia molto, tuttavia, la distanza che intercorre fra A- e B-*box*, anche perché in alcuni casi fra esse si trovano degli introni.

Le A-*box* dei promotori di tipo 1 e 2 sono elementi strutturalmente correlati ed intercambiabili in *X. laevis* (Ciliberto et al., 1983); questo fatto, tuttavia, riflette più probabilmente una similarità di sequenza piuttosto che una funzione conservata, dal momento che le A-*box* del 5S rRNA e dei tRNA si legano in modo diverso ai fattori di trascrizione (Braun et al., 1992a).

I promotori *core* di tipo 3 originariamente sono stati identificati nel gene *U6* dei mammiferi, gene che codifica lo snRNA U6 facente parte dello spliceosoma (Krol et al., 1987) (Das et al., 1988) (Kunkel and Pederson, 1988), e nel gene umano *7SK* (Murphy et al., 1986), il cui RNA prodotto è implicato nella regolazione del complesso CDK9/ciclina T (Nguyen et al., 2001) (Yang et al., 2001). Promotori di tipo 3 si trovano inoltre nel gene *H1*, che codifica la componente RNA della RNasi P umana (Baer et al., 1990), nel gene che codifica l'RNA della RNasi MPR umana (Topper and Clayton, 1990) ed in alcuni geni di funzione ignota.

I promotori di tipo 3 sono extragenici; essi si trovano nella regione in 5' e sono costituiti da un *Proximal Sequence Element* (PSE) e da una TATA-*box* che si trova ad una distanza fissa a valle del PSE (Hernandez and Lucito, 1988) (Mattaj et al., 1988) (Kunkel and Pederson, 1989) (Lobo and Hernandez, 1989). E'interessante il fatto che, nei vertebrati, la specificità di polimerasi può essere cambiata da Pol III a Pol II, e viceversa, eliminando oppure inserendo una TATA-*box* nei promotori dei geni per snRNA (Lobo and Hernandez, 1989).

Nei promotori di tipo 3, a monte del PSE, si trova un ulteriore elemento denominato *Distal Sequence Element* (DSE), in grado di attivare la trascrizione dipendente dal promotore *core*.

Anche se la presenza di una TATA-*box* è caratteristica dei promotori di tipo 3, si possono trovare TATA-*box* anche nelle regioni in 5' di alcuni geni con promotori intragenici. Questo tipo di promotore ibrido si trova, ad esempio, nel gene *SNR6* di *Saccharomyces cerevisiae*. In esso si trovano una A-*box*, una B-*box* localizzata in posizione anomala, 120 bp a valle della regione codificante lo snRNA, e di una TATA-*box* che si trova a monte del sito di inizio della trascrizione. Tutti e tre questi elementi promotori sono risultati necessari per ottenere un'efficiente trascrizione *in vivo* (Brow and Guthrie, 1990; Eschenlauer et al., 1993).

Altri esempi di promotori ibridi comprendono alcuni geni per tRNA in piante (Yukawa et al., 2000), in lievito (Dieci et al., 2000) ed in baco da seta (Ouyang et al., 2000). In *Schizosaccharomyces pombe* il gene *5S* e quasi tutti i geni per tRNA presentano una TATA-*box*, a monte del sito di inizio, che è richiesta per la trascrizione (Hamada et al., 2001).

*In vitro*, promotori artificiali costituiti soltanto da una TATA-*box* sono in grado di dirigere la trascrizione da parte della RNA polimerasi III (Mitchell et al., 1992) (Roberts et al., 1995) (Wang and Stumph, 1995) (Whitehall et al., 1995) (Huang et al., 1996).

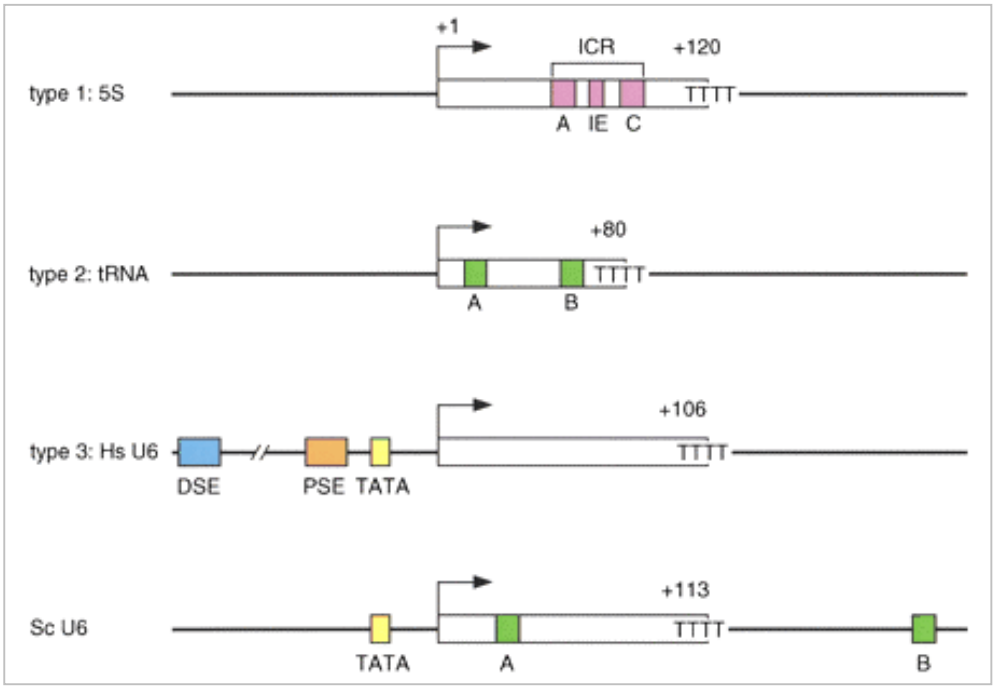



**Fig. 1.1: Diversi tipi di promotore per la RNA polimerasi III.**
Il promotore di tipo 1 del gene *5S* rRNA di *Xenopus laevis* è costituito da una *internal control region* (ICR), che può essere suddivisa in A-*box* (da +50 a +60), elemento intermedio (IE, da +67 a +72) e

C-*box* (da +80 a +90). Il promotore di tipo 2 del gene $tRNA^{Leu}$ di *X. laevis* è costituito da una A-*box* (da +8 a +19) ed una B-*box* (da +52 a +62). Il promotore di tipo 3 del gene *U6* snRNA di *Homo sapiens* è costituito da un *distal sequence element* (DSE, da -215 a -240), che potenzia la trascrizione, e da un promotore *core* composto da un *proximal sequence element* (PSE, da -65 a -48) ed una TATA-*box* (da -32 a -25). Il promotore per lo stesso gene di *Saccharomyces cerevisiae* è un promotore ibrido costituito da una TATA-*box* (da -30 a -23), una A-*box* (da +21 a +31) ed una B-*box* posizionata a valle della regione codificante (da +234 a +244 rispetto al sito di inizio della trascrizione).
*(Da Schramm and Hernandez, 2002).*

### *1.2.2 Assemblaggio del complesso di trascizione Pol III*

La caratterizzazione dei fattori di trascrizione per la RNA polimerasi III ha avuto inizio dal frazionamento di un estratto di cellule HeLa in tre frazioni, denominate A, B, e C, e dall'osservazione che la trascrizione a partire dai promotori di tipo 2 richiedeva le frazioni B e C, mentre la trascrizione dai promotori di tipo 1 richiedeva tutte e tre le frazioni (Segall et al., 1980). Per la trascrizione sui promotori di tipo 3 sono richieste le frazioni B e C oppure la frazione B più una diversa frazione D (Lobo et al., 1991). La maggior parte delle attività presenti in queste frazioni sono state purificate e caratterizzate, sia a partire da cellule di lievito che da cellule umane.

Nei promotori di tipo 2 (vedi *Fig. 1.2, A*), le A- e B-*box* sono riconosciute da un complesso, costituito da diverse subunità, che si trova nella frazione C, denominato TFIIIC (Lassar et al., 1983).

TFIIIC in *Saccharomyces cerevisiae* è costituito da 6 subunità. La struttura di questo fattore di trascrizione è adattata a svolgere il compito di legarsi alle A- e B-*box* che sono variamente distanziate nei promotori di tipo 2; TFIIIC, infatti, è costituito di due domini separati da un *linker* flessibile. Il dominio $\tau_B$ si lega strettamente alla B-*box* mentre il dominio $\tau_A$ si lega in modo più debole con la A-*box* (Marzouki et al., 1986).

Questa interazione DNA-proteina iniziale consente il reclutamento di una seconda attività, presente nella frazione B, chiamata TFIIIB (Bieker et al., 1985) (Setzer and Brown, 1985).

TFIIIB è costituito di tre polipeptidi, uno dei quali è la *TATA Binding Protein*, TBP. Una descrizione più dettagliata del fattore TFIIIB si trova più avanti in *sezione 1.2.4*.

Il legame di TFIIIB, dipendente da TFIIIC, a sua volta consente il reclutamento della RNA polimerasi III sul gene che deve essere trascritto. La polimerasi III viene reclutata principalmente attraverso interazioni proteina-proteina con TFIIIB, ma non si può escludere che anche contatti con TFIIIC possano dare un contributo.

Nei promotori di tipo 1 (vedi *Fig. 1.2, B*), la ICR è riconosciuta dal fattore TFIIIA (Engelke et al., 1980) che consente il successivo legame di TFIIIC (Lassar et al., 1983). TFIIIA, dunque, può essere considerato come un fattore di specificità in grado di alterare le proprietà di riconoscimento del promotore di TFIIIC e di dirigerlo verso il promotore 5S. Dopo il legame di TFIIIC, il *pathway* di reclutamento della polimerasi è simile a quello visto per i promotori di tipo 2.

Nei promotori di tipo 3 (vedi *Fig. 1.2*, *C*), il PSE è riconosciuto da un complesso multiproteico chiamato in vario modo *PSE-Binding Protein* (PBP) o *snRNA Activating Protein complex* ($SNAP_{c)}$, mentre la TATA-*box* viene riconosciuta dalla componente TBP di un'attività specializzata TFIIIB-*like* (Murphy et al., 1992; Sadowski et al., 1993; Schramm et al., 2000; Teichmann et al., 2000; Waldschmidt et al., 1991; Yoon et al., 1995). Queste interazioni DNA-proteina sono rafforzate da interazioni proteina-proteina tra $SNAP_c$ e TBP (Ma and Hernandez, 2001; Mittal and Hernandez, 1997). Al legame di $SNAP_c$ e dell'attività TFIIIB-*like* segue il reclutamento della RNA polimerasi III (Chong et al., 2001).

Esiste un ulteriore *pathway* di reclutamento della polimerasi (vedi *Fig. 1.2*, *D*) in cui TFIIIB si lega direttamente sulla TATA-*box*, senza l'aiuto di contatti proteina-proteina con TFIIIC o con $SNAP_c$. Questo *pathway* può essere osservato *in vitro* con TFIIIB di *Saccharomyces cerevisiae* e, anche se non si osserva *in vivo*, rivela un aspetto importante della trascrizione Pol III-dipendente, cioè che il solo fattore TFIIIB è sufficiente per il reclutamento della RNA polimerasi III. La TATA-*box* può dirigere cicli multipli di trascrizione quando vengono forniti soltanto TFIIIB ricombinante e Pol III purificata (Kassavetis et al., 1995) (Ruth et al., 1996).

Nel loro insieme, i diversi promotori dei geni di classe III costituiscono un sistema unico grazie al quale si è giunti a conoscere come diversi fattori proteici possono combinarsi per reclutare, in ultima analisi, un'attività TFIIIB e la RNA polimerasi III.

La RNA polimerasi III e TFIIIB sono descritti più in dettaglio nelle sezioni *1.2.3* e *1.2.4*.

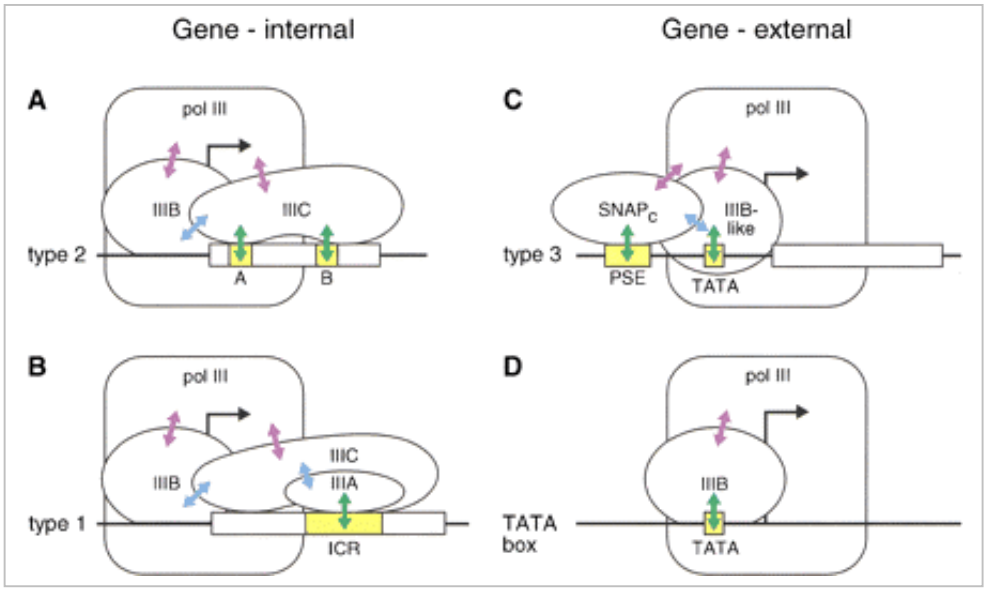



**Fig. 1.2: Diversi sistemi per il reclutamento di TFIIIB ed RNA polimersasi III.**
Sono mostrati complessi di inizio formati su promotori di tipo 1, di tipo 2, di tipo 3 e su un promotore artificiale costituito soltanto da una TATA-*box*. Le frecce verdi rappresentano le interazioni proteina-elementi promotori, le frecce blu i contatti proteina-proteina fra i diversi fattori di trascrizione e le frecce viola i contatti proteina-proteina fra la RNA polimerasi III ed i fattori di trascrizione.
*(Da Schramm and Hernandez, 2002).*

### *1.2.3 RNA polimerasi III*

La RNA polimerasi III è la più grande delle polimerasi nucleari; in *Saccharomyces cerevisiae* essa è costituita da 17 subunità ed ha una massa molecolare complessiva di 712 kDa (vedi *Tab. 1.1*).

Tutti i geni corrispondenti alle diverse subunità, eccetto *RPC37*, sono risultati essere geni essenziali in lievito (vedi (Chedin et al., 1998a) per review).

| S. cerevisiae RNA Pol III subunits | MW (kD) | Accession number | Guide for nomenclature | Corresponding RNA Pol II subunits |
|---|---|---|---|---|
| C160 (β'-like) | 162.1 | P04051 | ScRPC1 | RPB1 |
| C128 (β-like) | 129.3 | AAB59324 | ScRPC2 | RPB2 |
| **C82** | 73.6 | CAA45072 | ScRPC3 | |
| **C53** | 46.6 | P25441 | ScRPC4 | |
| **C37** | 32.1 | NP 012950 | ScRPC5 | |
| **C34** | 36.1 | P32910 | ScRPC6 | |
| **C31** | 27.7 | P17890 | ScRPC7 | |
| C25 | 24.3 | P35718 | ScRPC8 | RPB7 |
| C17 | 18.6 | P47076 | ScRPC9 | RPB4 |
| C11 | 12.5 | AAD12060 | ScRPC10 | RPB9 |
| AC40 (α-like) | 37.6 | P07703 | ScRPAC1 | RPB3 |
| AC19 (α-like) | 16.1 | P28000 | ScRPAC2 | RPB11 |
| ABC27 | 25.1 | P20434 | ScRPABC1 | RPB5 |
| ABC23 (ω-like) | 17.9 | AAA34989 | ScRPABC2 | RPB6 |
| ABC14.5 | 16.5 | CAA37383 | ScRPABC3 | RPB8 |
| ABC10α | 7.7 | AAA64417 | ScRPABC4 | RPB12 |
| ABC10β | 8.2 | P22139 | ScRPABC5 | RPB10 |

**Tab. 1.1: Subunità della RNA polimerasi III di** *Saccharomyces cerevisiae***.**
Le subunità indicate in grassetto non hanno paraloghi nelle RNA polimerasi I e II; quelle in grassetto sottolineato formano un complesso separabile dal resto dell'enzima. Sono indicate anche le subunità che corrispondomo alle subunità β', β, α e ω di *E. coli.*
*(Adattata da Schramm and Hernandez, 2002).*

Delle 17 subunità di Pol III, 5 sono comuni a tutte e tre le polimerasi eucariotiche, 2 sono comuni a Pol I e Pol III e le restanti 10 sono subunità caratteristiche della RNA polimerasi III; a queste ultime, probabilmente, si devono le proprietà specifiche di questo enzima, come la capacità di interagire con il proprio apparato trascrizionale e con gli specifici enzimi deputati al processamento dei trascritti o la capacità di dare termine alla trascrizione in modo indipendente da fattori accessori.

In *Figura 1.3* è rappresentata un'ipotetica struttura della RNA polimerasi III di *S. cerevisiae*, basata su dati relativi alla RNA polimerasi di *T. acquaticus* ed alla RNA polimerasi II di *S. cerevisiae* (Geiduschek and Kassavetis, 2001).

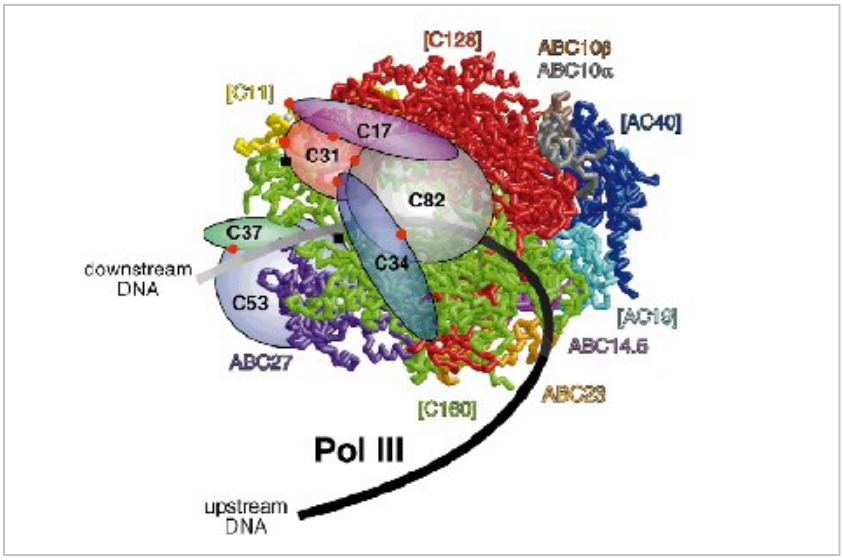



**Fig 1.3: Ipotetica struttura della RNA polimerasi III di** *Saccharomyces cerevisiae***, basata su dati strutturali relativi a Pol II ed alla Taq polimerasi.**
Le subunità C27(Rpb5), C23(Rpb6), C14.5(Rpb8), C10α(Rpb12) e C10β(Rpb10) sono identiche a quelle di Pol II; le subunità indicate tra parentesi (C160(Rpb1), C128(Rpb2), C40(Rpb3), C19(Rpb11) e C11(Rpb9)) sono omologhe a quelle di Pol II. Nell'immagine il DNA più a monte si proietta verso il lettore, il DNA più a valle si allontana dal lettore; fra le subunità specifiche di Pol III, il complesso C82-C34-C31 e la subunità C17 si trovano al di sopra della polimerasi, le subunità C53 and C37 si trovano dietro a C160 and C27.
*(Da Geiduschek and Kassavetis, 2001).*

## *1.2.4 TFIIIB*

Il fattore TFIIIB di lievito è costituito di due attività separabili cromatograficamente, denominate B' e B", che contengono polipeptidi, di 70 e 90 kDa rispettivamente, che possono essere *cross-linkati* al DNA (Bartholomew et al., 1991) (Kassavetis et al., 1991).

Il polipeptide di 70 kDa è denominato Brf1 (*TFIIB-related factor 1*). Esso si trova nella frazione B' associato alla proteina TBP (Kassavetis et al., 1992b) che, attraverso metodi biochimici, è stata vista fare parte delle attività TFIIIB sia di lievito che di mammifero (Margottin et al., 1991) (Huet and Sentenac, 1992) (Kassavetis et al., 1992b) (Lobo et al., 1992) (Taggart et al., 1992) (White et al., 1992) (Chiang et al., 1993) (Meyers and Sharp, 1993).

La componente da 90 kDa presente nella frazione B" è denominata Bdp1 (*B double prime 1*) (Willis, 2002).

Nei promotori costituiti soltanto da una TATA-*box*, TFIIIB si lega al DNA attraverso il riconoscimento della TATA-*box* da parte della sua subunità TBP. Una mutazione nella TATA-*box*, che diminuisce il livello di trascrizione da parte della RNA polimerasi III, può essere infatti compensata da un'opportuna mutazione in TBP tale da alterarne la specificità di legame al DNA (Strubin and Struhl, 1992) (Whitehall et al., 1995).

Il compito di riconoscere il complesso TATA-TBP è portato avanti dalla proteina Brf1; in essa possono essere distinte due porzioni, quella N-terminale, TFIIB-relata, e quella C-terminale, specifica di Brf1, che gioca il ruolo principale nel riconoscimento del complesso TATA-TBP (Kassavetis et al., 1997) (Kassavetis et al., 1998b) (Colbert et al., 1998).

In *S. cerevisiae*, il successivo legame di Bdp1 al complesso TATA-TBP-Brf1 induce un ripiegamento del DNA tra la TATA-*box* ed il sito di inizio della trascrizione, in fase con il ripiegamento imposto dal legame di TBP (Leveillard et al., 1991) (Braun et al., 1992b) (Grove et al., 1999). E' stato teorizzato che questa conformazione ripiegata possa contribuire, impedendo lo scivolamento del DNA al di fuori del complesso, alla stabilizzazione, Bdp1-dipendente, dei complessi TFIIIB-DNA (Grove et al., 1999).

Nella trascrizione da parte della RNA polimerasi III, l'apertura della bolla di trascrizione avviene in modo ATP-indipendente dopo il reclutamento di Pol III da parte di TFIIIB (Kassavetis et al., 1990) (Kassavetis et al., 1992b). Bdp1 gioca un ruolo chiave in questa fase dell'inizio della trascrizione. E' stato mostrato che molecole di Bdp1 che mancano della regione conservata a monte del dominio SANT possono in qualche modo dirigere livelli ridotti di trascrizione su templati superavvolti; esse però risultano del tutto inattive su templati lineari e non generano sensibilità al permanganato attorno al sito di inizio della trascrizione (Kassavetis et al., 1998a). Bdp1, inoltre, è risultata essere non necessaria per la trascrizione in condizioni in cui non è richiesta l'apertura del promotore (Kassavetis et al., 1999); ciò significa che Bdp1 gioca un ruolo essenziale proprio in questa fase.

## 1.3 IL CICLO TRASCRIZIONALE DELLA RNA POLIMERASI III

### *1.3.1 Inizio*

Tramite studi *in vitro* condotti su un gene per tRNA di lievito, è stato calcolato che, una volta formatosi il complesso di pre-inizio, sono necessari altri 5' a 22°C perché abbia inizio la sintesi dell'RNA trascritto (Dieci and Sentenac, 1996). Durante questo periodo avvengono tre passaggi successivi: il reclutamento della polimerasi, il *melting* dell'elica di DNA e l'inizio della sintesi dell'RNA. Il reclutamento ed il *melting* sono due passaggi lenti, mentre la successiva sintesi di un trascritto di 17 bp viene completata entro 10 secondi.

Quando si assembla su un complesso di pre-inizio, la RNA polimerasi di lievito genera un *footprint* che si estende dalla posizione -3 alla posizione +21 sul gene *SUP4* $tRNA^{Tyr}$ e dalla posizione –10 alla posizione +13 sul gene per il 5S rRNA (Kassavetis et al., 1990).

E' stato mostrato che esiste una debole omologia di sequenza tra i vari geni di classe III a livello del sito di inizio della trascrizione e che mutazioni in questa regione possono, in qualche caso, causare una forte riduzione della trascrizione (Raymond and Johnson, 1983) (Johnson and Raymond, 1984; Raymond et al., 1985) (Kassavetis et al., 1992a) (Fruscoloni et al., 1995) (Zecherle et al., 1996).

### *1.3.2 Allungamento*

I *footprints* generati dai complessi di trascrizione sui geni per tRNA e 5S rRNA cambiano dopo che ha avuto inizio la trascrizione e Pol III si sposta verso valle insieme alla bolla di trascrizione (Kassavetis et al., 1990). Se l'inizio della trascrizione avviene in assenza di un *set* completo di NTPs, tuttavia, in condizioni tali da non consentire la formazione di un trascritto più lungo di 5 nucleotidi, la bolla di trascrizione non risulta spostata (Kassavetis et al., 1992a). Quando l'RNA viene allungato fino a 17 nucleotidi, l'estensione della bolla di DNA disavvolto diminuisce, passando da 22 a 13-17 basi (Kassavetis et al., 1992a).

La RNA polimerasi III, che viene facilmente *strippata* dai complessi di pre-inizio per addizione di eparina, diventa assai più resistente a tale trattamento dopo che ha avuto inizio la trascrizione (Kassavetis et al., 1990; Zecherle et al., 1996). Tale stabilizzazione, tuttavia, non viene raggiunta se non dopo che sono stati sintetizzati trascritti di almeno 5 nucleotidi (Zecherle et al., 1996).

E' stato misurato che l'estensione dell'ibrido DNA-RNA in una bolla di trascrizione arrestata ha un'estensione di sole 7 bp, un'estensione molto inferiore rispetto a quella osservata in *E. coli* (Kassavetis et al., 1990). Il fatto che durante la trascrizione Pol III-dipendente si mantenga un ibrido DNA-RNA così breve potrebbe essere collegato con il processo relativamente semplice con il quale avviene la terminazione della trascrizione in questo sistema.

Il tasso medio di allungamento per la RNA polimerasi III in un sistema *in vitro*, a 20°C, è stato stimato attorno ai 20-22 nt/sec (Matsuzaki et al., 1994; Thuillier et al., 1996), una velocità simile a quella di Pol II *in vivo* (Matsuzaki et al., 1994). L'allungamento, tuttavia, non procede a velocità uniforme (Matsuzaki et al., 1994). E' stato proposto che Pol III oscilli tra due diverse conformazioni, una ad allungamento veloce ed una ad allungamento lento, e che le subunità C37 e C53 della RNA polimerasi giochino un ruolo critico nell'interconversione tra questi due stati, favorendo quello ad allungamento lento (Landrieux et al., 2005).

Mentre l'aggiunta di CMP o GMP a trascritti che terminano con U procede piuttosto rapidamente, l'aggiunta di UMP o AMP tende ad essere un passaggio lento, per cui anche serie di sole 3 T sul filamento codificante del DNA costituiscono un sito di pausa per la polimerasi (Matsuzaki et al., 1994).

Si ritiene che la RNA polimerasi III, in corrispondenza dei siti di pausa, entri in una modalità di "*inchworming*" e che la distanza fra il sito attivo ed il margine più avanzato dell'enzima si riduca (Thuillier et al., 1996); al superamento del sito di pausa il margine in 3' della polimerasi farebbe un balzo in avanti di diverse basi (Thuillier et al., 1996).

Ci si potrebbe aspettare che il voluminoso complesso di trascrizione Pol III assemblato all'interno della regione trascritta dei geni di classe III costituisca un ostacolo alla progressione della polimerasi in fase di allungamento, o che tale complesso venga spiazzato dal passaggio dell'enzima. E' stato verificato, tuttavia, che i complessi di

trascrizione assemblati non vengono rimossi dai geni di classe III anche a seguito di passaggi multipli della RNA polimerasi (Bogenhagen et al., 1982) (Jahn et al., 1987).

### 1.3.3 Terminazione

La RNA polimerasi III possiede la caratteristica unica di riconoscere una semplice serie di T sul filamento codificante del DNA come segnale di terminazione, e di riconoscere il terminatore in assenza di fattori proteici accessori. Per quest'ultima caratteristica la terminazione Pol III-dipendente si avvicina maggiormente alla terminazione rho-indipendente dei batteri piuttosto che a quella di Pol I o di Pol II. La pausa dell'enzima, tuttavia, costituisce il pre-requisito essenziale per la terminazione da parte di tutte le RNA polimerasi DNA-dipendenti.

E' stato proposto che l'antigene La, che si lega alla coda poli-(U) alla fine dei trascritti Pol III, sia implicato nella terminazione dalla trascrizione da parte della RNA polimerasi III (Gottlieb and Steitz, 1989b) (Gottlieb and Steitz, 1989a). Tuttavia, se La è risultata influenzare la trascrizione in sistemi *in vitro*, non è chiaro se essa gioca un ruolo simile anche entro la cellula.

La terminazione di Pol III entro il *cluster* di T che costituisce il terminatore è generalmente eterogenea e progressiva (Matsuzaki et al., 1994) e l'efficienza di terminazione cresce con la lunghezza della serie di T (Allison and Hall, 1985).

La lunghezza minima del terminatore varia nei diversi eucarioti: sono sufficienti per la terminazione segnali $T_4$ in *Xenopus* ed in uomo (Bogenhagen and Brown, 1981; Cozzarelli et al., 1983), segnali $T_4/T_5$ in *S. pombe* (Hamada et al., 2000) e segnali $T_5/T_6$ in *S. cerevisiae* (Allison and Hall, 1985).

Come già accennato, a livello di meccanismo la terminazione da parte di Pol III coinvolge un'ampia pausa, dettata dal *cluster* di T (Campbell and Setzer, 1992; Matsuzaki et al., 1994). La pausa di Pol III sul *cluster* di T determina lo stabilirsi di cicli multipli di idrolisi e ri-sintesi dell'RNA trascritto alla sua estremità 3' (Bobkova et al., 1999).

Sia la pausa che il taglio idrolitico dell'RNA vengono influenzati da mutazioni nelle subunità C160, C128, e C11 della RNA polimerasi III (Bobkova et al., 1999; Chedin et al., 1998b; Thuillier et al., 1996). Nel caso di C128 e C11, le stesse mutazioni sono state

mostrate influenzare le proprietà di terminazione di Pol III (Chedin et al., 1998b; Shaaban et al., 1996). L'attività di *cleavage* propria della subunità C11, tuttavia, non è implicata nel riconoscimento del segnale di terminazione (Landrieux et al., 2005). Al contrario, studi di *add-back* ad una forma incompleta di RNA polimerasi III hanno rivelato il ruolo chiave di C37 e C53 in questo riconoscimento (Landrieux et al., 2005).

Il riconoscimento del segnale di terminazione può essere sperimentalmente disaccoppiato dal rilascio della RNA polimerasi. Il rilascio dell'RNA trascritto e della polimerasi, infatti, sono dipendenti dallo spiazzamento del filamento di RNA durante la fase di allungamento, mentre il riconoscimento del segnale di terminazione non lo è (Campbell and Setzer, 1992). Ciò significa che la terminazione può essere considerata come la somma di queste due fasi distinte.

L'estrema semplicità del segnale di terminazione, costituito anche da sole 4 T, si accompagna con il fatto che i geni di classe III sono generalmente molto corti, cosa che riduce la probabilità di incontrare potenziali segnali di terminazione entro le regioni trascritte. Non mancano, tuttavia, casi di questo tipo.

Un esempio è costituito da un $tDNA^{Lys}$ di *Xenopus laevis*; in questo gene si trovano due tratti $T_4$, uno nella canonica posizione a valle, l'altro all'interno della *coding sequence*; questi due tratti $T_4$ determinano entrambi terminazione della trascrizione, ma a diverso livello, con il $T_4$ intragenico che si comporta come un terminatore molto più debole rispetto a quello extragenico (Mazabraud et al., 1987); qualche diverso elemento, di sequenza o posizionale, deve evidentemente essere evocato per spiegare un simile risultato.

L'idea che il contesto di sequenza possa influenzare l'efficienza di serie oligo-(dT) come terminatori è emerso per la prima volta da studi pionieri sulla terminazione di Pol III (Bogenhagen and Brown, 1981) ed i pochi casi documentati in letteratura possono difficilmente essere inseriti entro una regola generale; gli stessi nucleotidi fiancheggianti, per esempio, sono risultati produrre effetti differenti sui terminatori di diversi geni di classe III (Bogenhagen and Brown, 1981; Gunnery et al., 1999; Mazabraud et al., 1987).

E' stata infine rivelata l'esistenza di alcuni segnali di terminazione che non si conformano alla regola generale della serie di T (Emerson and Roeder, 1984; Gunnery et al., 1999; Hess et al., 1985).

### 1.3.4 Re-inizio

Il processo noto come re-inizio della trascrizione, osservato nel sistema trascrizionale Pol III, prevede che uno o più passaggi di associazione DNA-proteina vengano *by-passati* a seguito del primo ciclo di trascrizione (vedi *Fig. 1.4*). Il processo di re-inizio facilitato contribuisce in modo significativo all'efficienza di trascrizione caratteristica di Pol III.

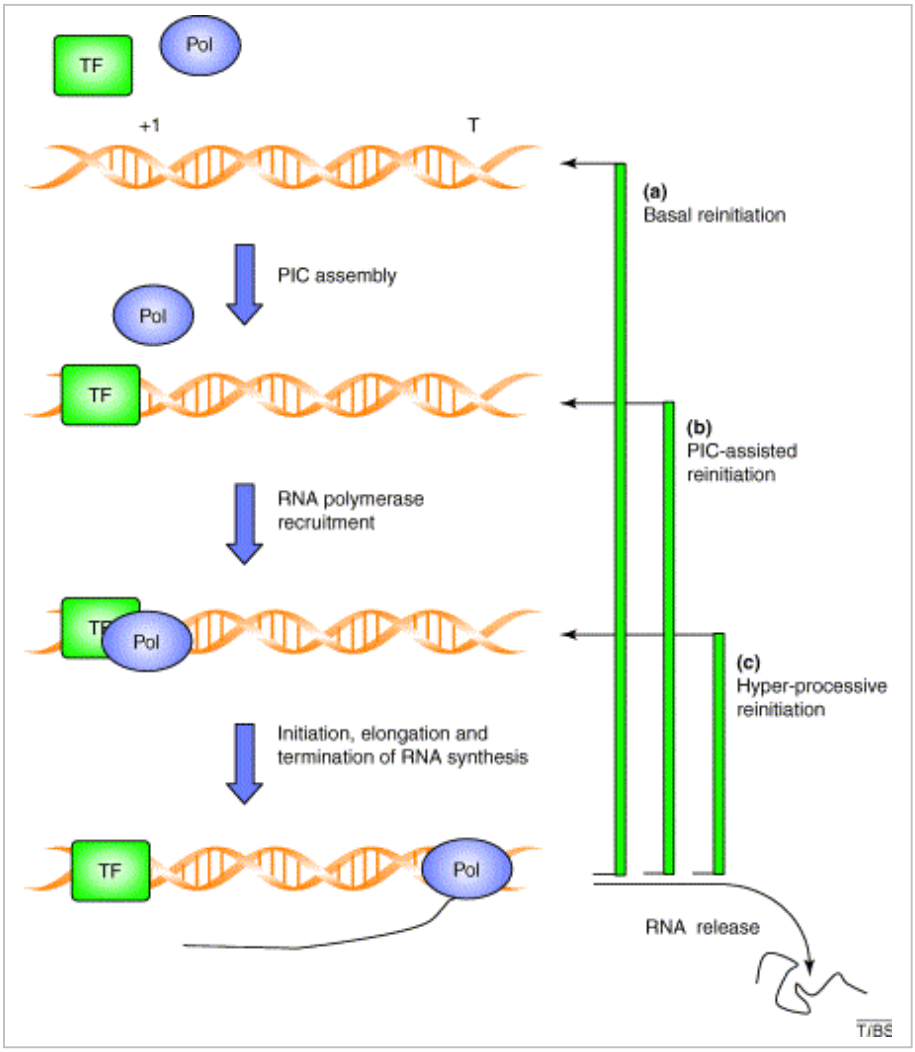



**Fig. 1.4: Meccanismi di re-inizio basati sul templato.**
Il ciclo di trascrizione ha inizio con un passaggio di associazione fattore di trascrizione (*TF*)-DNA. Questo *step*, generalmente non necessario nei batteri, genera un complesso di pre-inizio (*PIC*) necessario per il reclutamento della RNA polimerasi (*Pol*) negli eucarioti. Nel *pathway* di re-inizio basale *(a)*, tutti i passaggi necessari per il primo ciclo di trascrizione devono essere ripetuti ad ogni ciclo successivo. Nei *pathways* di re-inizio assistito dal complesso di pre-inizio *(b)* e iper-processivo *(c)*, uno o più di questi passaggi viene saltato. Nel re-inizio PIC-assistito alcuni complessi TF-DNA vengono mantenuti stabilmente sul promotore durante cicli multipli di trascrizione. Complesso di pre-

inizio, in questo caso, indica qualsiasi complesso intermedio proteina-DNA precedente e necessario al reclutamento della RNA polimerasi. Nel re-inizio iper-processivo, la RNA polimerasi può condurre diversi cicli di trascrizione sullo stesso templato senza essere rilasciata dallo stesso. Il re-inizio PIC-assistito è un meccanismo centrale della trascrizione da parte di Pol I, Pol II e Pol III. Il re-inizio iper-processivo è stato documentato solo per Pol III ma è stato ipotizzato anche per la trascrizione, Pol I-dipendente, delle ripetizioni di rDNA. I fattori di allungamento e di terminazione non sono stati rappresentati per chiarezza.
*(Da Dieci and Sentenac, 2003).*

Questo *pathway* di riciclo della polimerasi richiede che la terminazione della trascrizione abbia luogo al naturale segnale di terminazione; la terminazione *run-off* su geni troncati, infatti, non consente di ottenere un efficiente riciclo dell'enzima (Dieci and Sentenac, 1996). L'accoppiamento tra terminazione e re-inizio è supportato inoltre dall'osservazione, ottenuta in sistemi di insetto e di uomo, che i fattori di trascrizione di Pol III prendono contatto anche con regioni a valle dei geni di classe III, compreso il terminatore (Wang and Roeder, 1996; Young et al., 1991).

E' stato mostrato che *in vitro* Pol III viene ripetutamente ri-catturata dal primo gene trascritto anche in presenza di un promotore competitore; ciò suggerisce che il riciclo facilitato della polimerasi non sia dovuto ad un mero processo di ri-associazione favorito dalla piccola taglia dei geni trascritti da Pol III. E' stato inoltre mostrato che TFIIIC è richiesto per il re-inizio trascrizionale sui geni di classe III più lunghi (>300 bp) (Ferrari et al., 2004).

Uno studio recente, infine, ha rivelato che la subunità C11 della RNA polimerasi III svolge un ruolo nel re-inizio della trascrizione; è stato proposto che tale subunità sia necessaria per indurre un importante mutamento conformazionale nella RNA polimerasi pausata a livello del terminatore (Landrieux et al., 2005).

## 1.4 LA TRASCRIZIONE E L'ORGANIZZAZIONE DELLA CROMATINA

### *1.4.1 Cromatina e trascrizione*

Entro il nucleo eucariotico, il materiale genetico si trova strutturalmente e funzionalmente organizzato in strutture nucleoproteiche che nel loro insieme prendono il nome di cromatina.

L'unità base che compone la cromatina è il nucleosoma. Ciascun nucleosoma, attorno al quale 146 bp di DNA si avvolgono quasi due volte, è composto da due molecole di ciascun istone H2A, H2B, H3 e H4; gli istoni si organizzano come due dimeri H2A/H2B che si associano con un tetrametro $(H3/H4)_2$ centrale (Richmond et al., 1984) (Burlingame et al., 1985) (Pruss et al., 1995) (Luger et al., 1997). Le code N-terminali degli istoni *core* protrudono al di fuori del nucleosoma (Luger et al., 1997) (Wolffe and Pruss, 1996).

In corrispondenza della regione in cui il DNA entra ed esce dai nucleosomi *core* si trova una molecola di istone *linker* H1; la risultante catena di nucleosomi è ulteriormente compattata nella fibra da 30 nm e nelle strutture cromatiniche di ordine superiore.

La struttura cromatinica svolge la duplice funzione di impaccare strettamente il DNA e di mantenere un corretto funzionamento di quest'ultimo. La compatta organizzazione nucleoproteica, tuttavia, pone seri problemi a processi come la replicazione, la ricombinazione, il riparo e la trascrizione del DNA, il legame di proteine regolatrici. Si sono pertanto evoluti meccanismi che consentono la modulazione della struttura cromatinica, modulazione che gioca un ruolo centrale nella regolazione di quei fondamentali processi cellulari (Kornberg and Lorch, 1999).

I cambiamenti nella struttura cromatinica che accompagnano (e che spesso regolano) l'attivazione o la repressione della trascrizione vengono nel loro insieme chiamati rimodellamento della cromatina.

Responsabili di questo processo sono specifici complessi proteici ATP-dipendenti.

I geni trascritti da Pol III, relativamente corti, sono sostanzialmente coperti dai fattori di trascrizione; ciò potrebbe impedire la formazione di nucleosomi sui geni attivi (Gerlach et al., 1995). Almeno uno dei geni di lievito trascritti da Pol III, tuttavia, *RPR1*, che codifica

l'RNA1 dell'RNasi P, è sufficientemente lungo (369 bp) da poter contenere almeno un nucleosoma (Lee et al., 1991). Esistono altri esempi di nucleosomi posizionati entro unità trascrizionali Pol III; casi specifici di effetti cromatinici sui geni di classe III sono riportati con più dettaglio nella sezione *1.4.2*.

I nucleosomi risultano distrutti sui geni di classe I attivamente trascritti (Dammann et al., 1993), anche se il meccanismo di rimodellamento della cromatina dipendente da Pol I non è ancora noto.

Al contrario, sui geni trascritti da Pol II, sono presenti strutture nucleosomiche (Hartzog et al., 2002). I nucleosomi trascritti da Pol II, però, sono transitoriamente depleti di circa il 50% degli istoni H2A e H2B (Nacheva et al., 1989), ed il posizionamento dei nucleosomi viene spesso modificato (Hartzog et al., 2002). E' stato infatti osservato un estensivo scambio trascrizione-dipendente di istoni H2A e H2B (Hartzog et al., 2002), e forse anche di istoni H3 e H4 (Ahmad and Henikoff, 2002). Nei nucleosomi trascritti da Pol II, inoltre, i gruppi -SH dell'istone H3 sono transitoriamente accessibili a varie sonde (Hartzog et al., 2002); è stato suggerito che l'alta accessibilità di questi gruppi -SH rifletta il momentaneo spiazzamento degli istoni H2A e H2B durante la trascrizione (Bazett-Jones et al., 1996).

La struttura di più alto ordine della cromatina viene distrutta durante la trascrizione Pol II-dipendente (Hartzog et al., 2002); questo processo, così come la parziale distruzione della struttura del nucleosoma, è trascrizione-dipendente e risulta in una de-condensazione dello specifico dominio cromatinico che è attivamente trascritto (Mahy et al., 2002). Pol II, quindi, si comporta come un potente rimodellatore della cromatina, capace di modificare sia la struttura dei nucleosomi, sia la struttura di più alto ordine della cromatina. E' possibile che i cambiamenti nella struttura cromatinica possano, in certe circostanze, risultare in cambiamenti della struttura del cromosoma a più lungo termine. E' stato suggerito, per esempio, che la trascrizione attraverso un dominio cromatinico da parte di una polimerasi pioniera possa servire come evento regolativo che rende accessibile quel particolare dominio per la seguente attivazione della trascrizione (Gribnau et al., 2000).

Esistono sono almeno due meccanismi di rimodellamento della cromatina dipendenti dalla trascrizione: la mobilitazione dei nucleosomi (cioè il movimento degli ottameri istonici lungo il DNA) e la deplezione di dimeri H2A-H2B.

Due possibili meccanismi sono stati proposti per dare spiegazione del primo di questi due processi, la *nucleosome mobilization*: il "*bulging model*" ed il "*twisting model*" (Flaus and Owen-Hughes, 2001) (vedi *Fig. 1.5*).

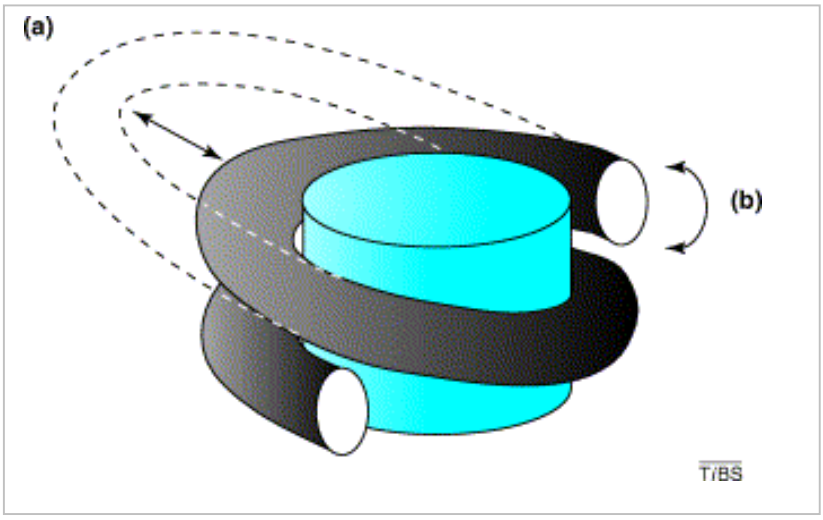



**Fig. 1.5: Possibili meccanismi che mediano la mobilitazione dei nucleosomi.**
I nucleosomi possono risultare spostati lungo il DNA a seguito di uno dei seguenti processi. *(a) "Bulging model"*. Il DNA viene parzialmente spiazzato dalla superficie dell'ottamero istonico e poi ri-legato formando un *loop* intra-nucleosomico. Il *loop* ruota attorno all'ottamero rompendo le interazioni DNA-istone su un lato del *loop* e ri-formandole sull'altro lato *(b) "Twisting model"*. La rotazione del DNA lungo il suo asse longitudinale, iniziato ad un lato del nucleosoma, ne cambia il *twist* e viene propagata lungo tutto il DNA nucleosomico. Una rotazione del DNA di 36° risulta in una traslocazione di 1 bp.
*(Da Studitsky et al, 2004).*

Secondo il *bulging model* i nucleosomi si muovano sul DNA attraverso la formazione e rotazione di un *bulge* sulla superficie dell'ottamero istonico (Studitsky et al., 1994). Secondo il *twisting model*, invece, una specifica attività di rimodellamento applica una torsione al DNA tale da alterare il suo *twist* sulla superficie del nucleosoma (van Holde, 1985); la propagazione direzionale del *twist* lungo il DNA nucleosomale risulta nella mobilitazione del nucleosoma.

E' stato osservato recentemente che i meccanismi di mobilitazione condotti dalle polimerasi o dai complessi di rimodellamento ATP-dipendenti sono molto simili.

La *nucleosome mobilization* si applica alla trascrizione da parte della RNA polimerasi del batteriofago SP6 e della RNA polimerasi III; nucleosomi interi, infatti, resistono alla

trascrizione *in vitro* da parte di questi due enzimi (Studitsky et al., 1997). In accordo con questi studi *in vitro*, esperimenti *in vivo* condotti in lievito ed in *Drosophila* hanno suggerito che i nucleosomi non vengono distrutti durante la trascrizione da parte di queste due polimerasi (Sathyanarayana et al., 1999; Saunders et al., 2003).

Sperimentalmente, gli studi *in vitro* sono stati condotti con corti templati lineari che portavano un singolo ottamero istonico posizionato (Studitsky et al., 1997) (O'Donohue et al., 1994). Dopo la trascrizione da parte della polimerasi del batteriofago, l'ottamero risultava trasferito all'indietro. L'analisi del destino di singoli nucleosomi posizionati diversamente su DNA templato identici, inoltre, rivelava che ogni isomero posizionale, sottoposto a trascrizione, generava un *set* distinto di posizioni dell'ottamero (Studitsky et al., 1997). E' difficile immaginare come una simile memoria della posizione di partenza possa essere mantenuta se l'ottamero è rilasciato in soluzione. Inoltre, nelle normali condizioni di trascrizione, gli ottameri liberi dal DNA si dis-assemblano (in 1 sec o meno (Feng et al., 1993)) e non possono essere ri-assemblati *in vitro*. L'abilità dell'ottamero di resistere al passaggio della polimerasi, quindi, suggerisce che esso sia trasferito direttamente.

Nel caso della RNA polimerasi II, il rimodellamento della cromatina trascrizione-dipendente prevede un parziale dis-assemblaggio dei nucleosomi: è stato mostrato su templati mononucleosomici che un dimero H2A-H2B viene spiazzato dall'ottamero istonico durante la trascrizione; come risultato i nucleosomi sono convertiti in esasomi (Kireeva et al., 2000). A seguito di questo processo, gli istoni non cambiano la loro posizione lungo il DNA (Kireeva et al., 2000).

E' stato dimostrato che gli stessi nucleosomi costituiscono una barriera più grande per la RNA polimerasi II piuttosto che per le RNA polimerasi III o SP6 (Kireeva et al., 2002). Un singolo nucleosoma, infatti, può costituire una barriera assoluta per la trascrizione da parte di Pol II (Izban and Luse, 1991; Kireeva et al., 2002) ed è altamente probabile che le barriere nucleosomiche siano utilizzate per la regolazione dell'espressione genica, come elementi in grado di cambiare il tasso di allungamento del trascritto.

L'incapacità di superare una barriera nucleosomica non è una caratteristica esclusiva di Pol II; la RNA polimerasi di *E. coli* è notevolmente simile a Pol II per questo aspetto (Walter et al., 2003). L'altezza della barriera costituita da un nucleosoma, pertanto, dipende dalle

proprietà generali del complesso di allungamento della trascrizione ed è quindi polimerasi-specifica.

Il superamento delle barriere nucleosomiche richiede il coinvolgimento di meccanismi che facilitano la trascrizione attraverso la cromatina (vedi *Fig.1.6*).

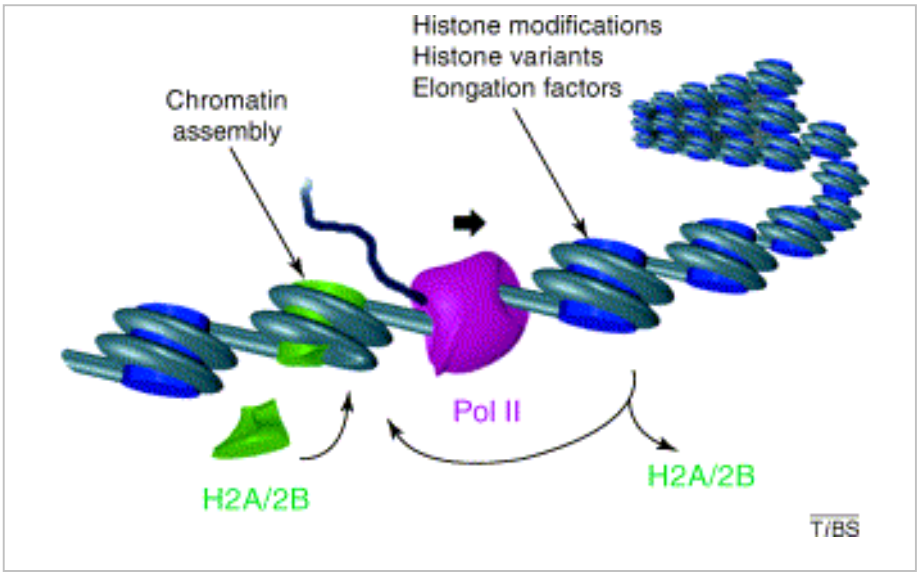



**Fig. 1.6: Ipotetico meccanismo per la trascrizione Pol II-dipendente attraverso la cromatina.** Mentre si muove lungo un gene eucariotico, Pol II (*viola*) converte gli ottameri istonici (*blu*) che incontra sul proprio percorso in esameri (*verdi*). Questo processo è facilitato da fattori di allungamento come FACT (*Facilitates Chromatin Transcription*) e da modificazioni degli istoni e/o varianti istoniche. Il momentaneo distacco dei dimeri H2A-H2B dai nucleosomi consente l'accesso di fattori di rimodellamento della cromatina e/o di proteine che si legano al DNA. A meno che questo stato non venga stabilizzato, alla fine si formano nuovamente nucleosomi completi. La trascrizione può anche risultare in una redistribuziome dei nucleosomi ed in una parziale deplezione di nucleosomi.
*(Da Studitsky et al, 2004).*

Tali meccanismi coinvolgono fattori che si legano alle proteine istoniche, fattori di allungamento per la RNA polimerasi II, complessi che inducono il rimodellamento cromatinico, non covalente e ATP-dipendente, e fattori responsabili di modificazioni covalenti degli istoni.

Il solo fattore noto in grado di stimolare la trascrizione attraverso la cromatina in un sistema altamente purificato *in vitro* è il complesso eterodimerico FACT (*Facilitates Chromatin Transcription*) (Orphanides et al., 1998). Una significativa evidenza genetica,

inoltre, connette FACT con l'allungamento della trascrizione *in vivo* (Orphanides et al., 1999).

Diversamente da quella indotta dai complessi di rimodellamento della cromatina, la riorganizzazione dei nucleosomi indotta da FACT non richiede l'idrolisi di ATP.

Il complesso FACT facilita lo spiazzamento di dimeri H2A-H2B, indotto da Pol II, probabilmente attraverso un'interazione diretta con quegli istoni (Belotserkovskaya et al., 2003). *In vitro*, FACT interagisce specificamente con tutti gli istoni *core* e mostra di possedere un'intrinseca attività di *chaperone* (Belotserkovskaya et al., 2003; Orphanides et al., 1999). Questa attività di *chaperone* potrebbe dare spiegazione delle osservazioni, ottenute *in vivo*, che suggeriscono che FACT non solo partecipa al parziale dis-assemblaggio del nucleosoma, ma contribuisce anche al ri-assemblaggio del nucleosoma alle spalle dell'enzima in trascrizione (Formosa et al., 2002) (Kaplan et al., 2003).

Il complesso FACT di lievito (yFACT), che fa parte di una famiglia altamente conservata negli eucarioti, è costituito dalle proteine Spt16, Pob3 e Nhp6. Spt16 (nota anche come Cdc68) e Pob3 costituisco un eterodimero stabile, denominato SP o CP (Brewster et al., 1998; Formosa et al., 2002), associato solo debolmente a Nhp6 (Brewster et al., 2001; Formosa et al., 2001).

Nhp6 fa parte della famiglia di proteine HMG che contengo un dominio di legame al DNA. Nhp6 è in grado di legarsi ai nucleosomi e di indurre un cambiamento nella loro struttura, ma per questo cambiamento è richiesta l'azione simultanea di ~10 monomeri di Nhp6 ogni nucleosoma (Ruone et al., 2003).

Non è ancora chiaro come il motivo HMG funzioni nella riorganizzazione nucleosomica portata avanti da FACT; è stato proposto che Nhp6 porti avanti il primo passaggio di un meccanismo a due *step* (Rhoades et al., 2004). Il legame di diverse molecole di Nhp6, simultaneo o in un breve spazio di tempo, indurrebbe un cambiamento nel nucleosoma che lo convertirebbe in substrato per Spt16-Pob3 o anche per altri fattori di rimodellamento della cromatina.

### 1.4.2 Effetti cromatinici su geni di classe III

La trascrizione dei geni per tRNA pare essere relativamente resistente alla repressione da parte degli istoni. La manipolazione del livello di istoni H1 in embrioni di rana, infatti, genera un piccolo effetto sulla sintesi dei tRNA (Bouvet et al., 1994). La rimozione dell'istone H1 da cromatina di topo, inoltre, genera una differenza piccola o nulla nel numero di geni per tRNA che risultano accessibili ai fattori di trascrizione (Carey and Singh, 1988).

In lievito, una stessa deficienza di nucleosomi che risulta attivare diversi geni di classe II, non produce alcun effetto sull'espressione dei tRNA (Han and Grunstein, 1988). Il gene SUP4 $tRNA^{Tyr}$ può rimanere attivo addirittura quando fuso con segnali di posizionamento nucleosomico in grado di sopprimere sia la trascrizione Pol II-dipendente, sia l'inizio della replicazione del DNA (Morse et al., 1992). Un gene $tRNA^{Glu}$ rimane attivo anche dopo incubazione prolungata con un estratto di ovocita di *Xenopus* che assembla i nucleosomi con una distanza regolare e fisiologica (Burnol et al., 1993).

Anche se i geni per tRNA risultano capaci di superare certi segnali di posizionamento dei nucleosomi, essi possono, in certe situazioni, essere incorporati entro la cromatina inattiva. Ad esempio, un gene per tRNA di lievito è risultato represso a seguito della sua inserzione entro il *locus* silente *HMR* (Schnell and Rine, 1986).

Un caso particolare è costituito dal gene *SNR6* di lievito. In questo gene i due elementi promotori A- e B-*box* sono separati da una distanza superiore al normale e che genera condizioni non ottimali per il legame di TFIIIC. Ci si potrebbe aspettare pertanto che, in questo gene, la delezione di parte della regione che si trova tra A- e B-*box* possa promuovere il legame di TFIIIC. Tale delezione è in realtà risultata dannosa ai fini della trascrizione; questo ha suggerito che la cromatina nella regione posta fra le due *box* sia organizzata in modo tale da facilitare il reclutamento di TFIIIC (Eschenlauer et al., 1993; Gerlach et al., 1995) (Shivaswamy et al., 2004).

A sostegno di questa idea, esperimenti di *footprint* hanno mostrato che ~100 bp poste fra le due *box* risultano protette dall'azione della nucleasi micrococcica (Gerlach et al., 1995). Questa protezione permane anche quando la B-*box* viene inattivata; è pertanto improbabile che sia dovuta a TFIIIC (Gerlach et al., 1995). La taglia della regione protetta non è

sufficiente per dare spazio nemmeno ad un nucleosoma posizionato, ma è possibile che proteine non istoniche siano implicate nell'organizzazione spaziale di questa estesa unità trascrizionale.

Il gene *U6* di lievito, infine, è fiancheggiato ad entrambi i lati da una serie di nucleosomi posizionati (Marsolier et al., 1995). L'inattivazione della B-*box*, tramite delezione di 2 bp, è risultata nella perdita di questa struttura (Marsolier et al., 1995). Questi risultati hanno fornito la prima evidenza del fatto che un gene di classe III può organizzare la cromatina attorno ad esso. Resta da stabilire se TFIIIB, TFIIIC od il processo di trascrizione sono responsabili di questo effetto di posizionamento. Si noti infine che l'abilità del gene *U6* di organizzare la cromatina che lo circonda è stata persa quando esso è stato trasferito dalla sua posizione cromosomica su un vettore plasmidico (Marsolier et al., 1995).

### *1.4.3 I geni di classe III come barriere cromatiniche*

Nel nucleo in interfase, la cromatina è fisicamente e funzionalmente organizzata in domini attivi e domini inattivi (Lamond and Earnshaw, 1998). L'eucromatina è trascrizionalmente attiva ed è mantenuta in questo stato aperto dall'azione di specifiche sequenze promotrici, *enhancer* e LCR (*Locus Control Region*) (Bulger and Groudine, 1999). L'eterocromatina, originariamente definita come "*loci* cromosomici che rimangono condensati durante l'interfase" (Eissenberg, 1995; Weiler and Wakimoto, 1995), è generalmente repressa a livello trascrizionale, anche se può contenere qualche gene attivo (Sun et al., 2000; Wakimoto and Hearn, 1990). La struttura e la funzione dell'eterocromatina sono mediate da sequenze che nucleano la formazione di complessi proteici di repressione che a loro volta silenziano la regione.

E' stato proposto che elementi detti barriere cromatiniche ed elementi isolatori funzionino nel delimitare i domini di funzione dei *silencer*s e degli *enhancers*, rispettivamente (Bell et al., 1999; Gerasimova and Corces, 1996; Geyer, 1997; Sun and Elgin, 1999).

Gli isolatori sono definiti come sequenze di DNA che, bloccando l'attivazione di un promotore da parte di un *enhancer* quando posizionate tra i due elementi, impediscono un'inappropriata attivazione genica. Le barriere sono invece sequenze che bloccano il

propagarsi dell'eterocromatina da un *locus* silenziato e che pertanto separano i domini di cromatina attiva ed inattiva (Sun and Elgin, 1999).

Il lievito *Saccharomyces cerevisiae* contiene *loci* cromosomici che condividono diverse caratteristiche molecolari con l'eterocromatina dei metazoi. Un lievito aploide *wild-type* codifica copie non espresse dei geni *mating-type MAT*a e *Mat*α ai *loci HMR* e *HML*, mentre una copia attiva si trova variamente trasposta al *locus MAT* (Haber, 1998; Loo and Rine, 1995; Lustig, 1998; Stone and Pillus, 1998). Anche se identici in sequenza ai geni espressi al *locus MAT*, i geni *HMR* e *HML* sono silenziati a causa della formazione di una struttura cromatinica repressiva.

Per il silenziamento di quei *loci* sono richieste due regioni fiancheggianti designate *HMR*-E e *HMR*-I. Questi elementi silenziatori sono composti di ARS (*Autonomously Replicating Sequences*), che legano il complesso di riconoscimento dell'origine, e di siti che legano le proteine Rap1 e Abf1. Queste proteine nucleano l'assemblaggio di uno specifico complesso sul DNA silenziato che contiene le proteine Sir (*Silent information regulators*) ed altri fattori coinvolti nel mantenimento della repressione (Loo and Rine, 1995).

Il silenziamento non gene-specifico in lievito si osserva anche vicino alle regioni telomeriche dei cromosomi (Gottschling et al., 1990) ed entro le ripetizioni degli rDNA (Smith and Boeke, 1997).

Attraverso un'analisi dettagliata, un gene codificante per un tRNA (*HMR* $tRNA^{Thr}$) è stato identificato come il confine più prossimo al telomero del dominio *HMR* (Donze and Kamakaka, 2001). Tale gene per tRNA, insieme con le sue sequenze fiancheggianti, è necessario e sufficiente ad impedire il diffondersi del silenziamento dal locus *HMR* e funziona pertanto come elemento barriera.

I geni per tRNA sono trascritti dalla RNA polimerasi III; mutazioni nei fattori TFIIIC e TFIIIB o nelle sequenze promotrici A- e B-*box* dell'*HMR* $tRNA^{Thr}$, risultano in un impedimento dell'attività di barriera di questo gene per tRNA; ciò suggerisce che sia richiesta l'integrità del complesso di trascrizione per un'efficiente attività di barriera. Il fatto che anche le sequenze fiancheggianti il tDNA siano richieste per la sua attività di barriera non è in realtà sorprendente, dal momento che numerosi studi hanno mostrato che le regioni fiancheggianti i geni per tRNA contribuiscono al loro potenziale di trascrizione, anche se gli elementi promotori si trovano all'interno dei geni (Dingermann et al., 1982;

Giuliodori et al., 2003; Ong et al., 1997; Raymond and Johnson, 1983; Shaw and Olson, 1984; Sprague et al., 1980; Young et al., 1991).

Come le sequenze barriera impediscono il propagarsi dell'eterocromatina?

Un possibile meccanismo prevede che un complesso di pre-inizio per la RNA polimerasi III, legato stabilmente al gene per tRNA, agisca come un blocco fisico al propagarsi del *silencing* (Donze and Kamakaka, 2001). Quello della RNA polimerasi III è un complesso ingombrante, che ha una massa complessiva di ~1,5 MDa e che genera *footprints* di ~150 bp su un gene per tRNA (Chedin et al., 1998a; Kassavetis et al., 1998b); un così grande complesso costituirebbe un impedimento al legame delle proteine del complesso SIR. Ai domini silenziati *HMR* e *HML*, infatti, si trova un distintivo *array* SIR-dipendente di di-nucleosomi posizionati, che potrebbe svolgere un ruolo importante per la propagazione e la stabilità dello stato eterocromatinico (Ravindra et al., 1999; Weiss and Simpson, 1998). Creando un largo *gap* in questo *array*, il complesso Pol III, stabilmente legato, distruggerebbe in modo passivo l'organizzazione dei nucleosomi mediata dalle proteine SIR (vedi *Fig. 1.7*).

Un meccanismo alternativo prevede che le regioni fiancheggianti del gene *HMR* $tRNA^{Thr}$ interagiscano con altri fattori implicati nell'attività di barriera (Donze and Kamakaka, 2001). L'acetilazione degli istoni e di altri substrati cromatinici gioca un ruolo fondamentale nella regolazione delle strutture cromatiniche di più alto ordine (Sterner and Berger, 2000). L'iper-acetilazione degli istoni è associata con una maggiore sensibilità al trattamento con DNasi I e con un maggiore potenziale di trascrizione dei domini cromosomici (Hebbes et al., 1994); l'ipoacetilazione, al contrario, è una caratteristica dell'eterocromatina inattiva (Lin et al., 1989). L'osservazione di un'apparente perdita dell'attività barriera in ceppi con mutazioni nei geni HAT (*Histone AcetylTransferase*) *GCN5* e *SAS2* (Donze and Kamakaka, 2001) ha suggerito il modello della competizione tra attività di modificazione della cromatina (vedi *Fig. 1.7*). Secondo questo modello, la deacetilazione mediata da Sir2 si propaga dal *silencer* fino al raggiungimento di un elemento barriera; questo recluta stabilmente una o più acetiltransferasi e questi enzimi competono attivamente con gli effetti della proteina Sir2. Anche se non è stato ancora dimostrato che alcuna acetiltransferasi sia associata con il complesso Pol III in lievito, è stata identificata un'attività HAT associata con il fattore TFIIIC di uomo (Kundu et al., 1999).

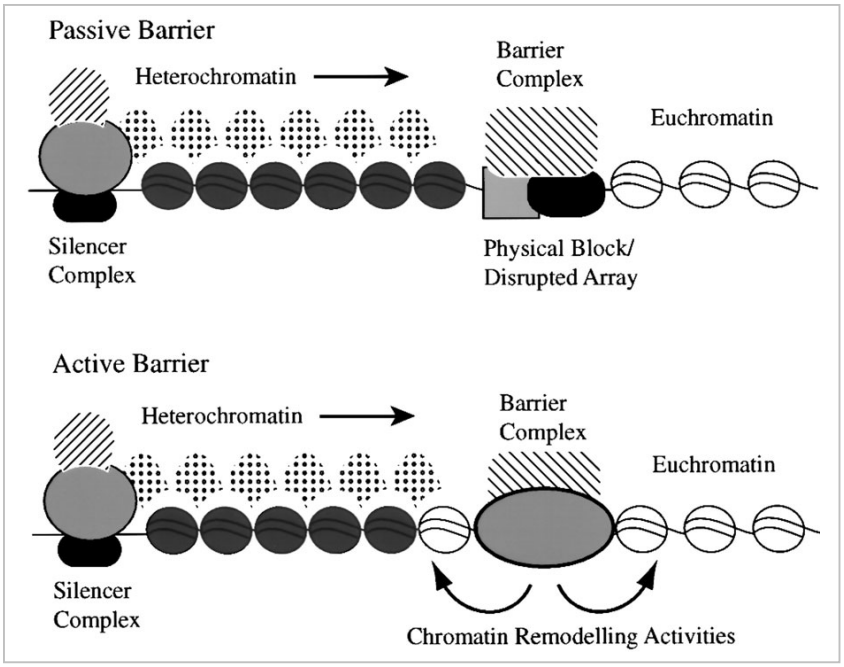



**Fig. 1.7: Modelli per l'attività di barriera al diffondersi dell'eterocromatina.**
I complessi proteici legati al *silencer* nucleano il propagarsi della de-acetilazione degli istoni ed il legame delle proteine associate all'eterocromatina lungo la regione silenziata. Questa struttura si propaga fino a quando non si raggiunge una barriera. Secondo il modello passivo (*sopra*), un complesso multiproteico ingombrante e legato stabilmente al DNA è in grado di interferire fisicamente con la propagazione della struttura eterocromatinica. Secondo il modello attivo (*sotto*), alcune attività di modificazione della cromatina (es. acetiltransferasi o complessi di rimodellamento dei nucleosomi) vengono reclutate dalla barriera cromatinica per modificare gli istoni e forse altri fattori associati alla cromatina. I nucleosomi così modificati sono substrati meno efficienti per l'incorporazione nell'eterocromatina.
*(Da Donze and Kamakaka, 2001).*

Sta emergendo recentemente il tema generale dei geni di classe III come confini dei domini cromatinici.

Geni per tRNA, ad esempio, possono causare un certo numero di effetti su processi cromatinici, come l'annullamento della fase dei nucleosomi, normalmente indotto da un elemento di posizionamento (Morse et al., 1992), o la pausa alla forca di replicazione del DNA in *S. cerevisiae* (Deshpande and Newlon, 1996).

Gli elementi Alu sono sequenze ripetute nei genomi dei primati che nella maggior parte dei casi contengono promotori funzionali per la RNA polimerasi III (Deininger and Batzer, 1999). E' stato dimostrato (Willoughby et al., 2000) che un elemento Alu umano che fiancheggia il gene cheratina-18 conferisce espressione posizione-indipendente e *copy*

*number*-dipendente del transgene in topo; la mutazione della B-*box* di questo elemento Alu abolisce questa protezione dagli effetti di posizione.

In *Schizosaccharomyces pombe*, la mappatura delle proteine centromeriche Swi6 e Mis6 ha definito due domini discreti di cromatina centromerica; due geni per tRNA costituiscono il confine di questi domini (Partridge et al., 2000).

Visto il numero non irrilevante di casi in cui geni di classe III demarcano distinti domini cromatinici, è possibile che i complessi Pol III, oltre alla loro taglia, posseggano qualche caratteristica non ancora identificata che li rende capaci di impedire il propagarsi di una data struttura cromatinica.

Al fine di identificare, oltre al $tDNA^{Thr}$, altri geni che, mutati, consentono un inappropriato propagarsi del *silencing* al *locus HMR*, è stato condotto uno *screening* di tipo genetico. Oltre a diversi fattori che erano già noti regolare il diffondersi dell'eterocromatina, è stato identificato un nuovo gene che contiene un bromodominio, *YTA7* (Jambunathan et al., 2005). In un indipendente studio di tipo proteomico, la stessa proteina Yta7 è stata identificata come proteina di barriera e ne è stata mostrata la localizzazione in siti vicini ai *loci HMR* e *HML*, alle estremità dei cromosomi ed in prossimità dei geni *FLO*, che sono regolati in modo epigenetico (Tackett et al., 2005).

Anche il complesso proteico RSC (*Remodels the Structure of Chromatin*) è implicato nella funzione barriera. RSC è un complesso di rimodellamento della cromatina costituito da diverse subunità, molto abbondante, essenziale e purificato per la prima volta da cellule di *S. cerevisiae* (Cairns et al., 1996). La delezione del gene *RSC2* e, in misura minore, del gene *RSC1*, porta al diffondersi del *silencing* attraverso il tRNA barriera (Jambunathan et al., 2005). Un' analisi di tipo genomico ha dimostrato che RSC è posizionato su molti geni per tRNA *in vivo* (Ng et al., 2002); questo suggerisce una possibile connessione diretta tra la funzione dell'*HMR*-tDNA ed il complesso RSC. Mentre *YTA7* ed *HMR*-tRNA funzionano in modo indipendente nell'impedire il diffondersi del *silencing*, è probabile che *RSC2* funzioni tramite il complesso RNA polimerasi III assemblato sul $tRNA^{Thr}$ (Jambunathan et al., 2005).

## 1.5 LA REGOLAZIONE DELLA TRASCRIZIONE POL III-DIPENDENTE

### *1.5.1 Esempi di regolazione del sistema Pol III*

Come processo fondamentale del metabolismo cellulare, la trascrizione da parte della RNA polimerasi III è soggetta ad una stretta regolazione; essa è controllata, per esempio, dalle condizioni di crescita, dal ciclo cellulare e da un certo numero di virus. Questi controlli coinvolgono i fattori di trascrizione specifici di questo enzima, soprattutto TFIIIB e TFIIIC, così come forse la stessa polimerasi III.

La regolazione del sistema Pol III in risposta alle condizioni di crescita è stata studiata in particolare nel sistema di lievito.

In condizioni di crescita bilanciate, il tasso di trascrizione della RNA polimerasi III risulta proporzionale al tasso di crescita; non tutti i trattamenti che rallentano o arrestano la crescita, tuttavia, de-regolano anche l'attività Pol III (Clarke et al., 1996).

La *starvation*, per esempio, può arrestare la crescita cellulare inibendo la trascrizione Pol III-dipendente. I lieviti cresciuti fino ad un'alta densità vanno soggetti ad uno *shift* diauxico durante il quale le cellule passano da un metabolismo di tipo fermentativo ad un metabolismo di tipo respiratorio. Durante questa transizione, la trascrizione dei geni per tRNA e 5S rRNA è fortemente *down*-regolata, con la trascrizione del gene per il 5S RNA che decresce più rapidamente rispetto a quella dei tRNA (Clarke et al., 1996; Sethy et al., 1995). Quando si aggiunge terreno fresco a queste cellule in fase stazionaria, tuttavia, si riprende la trascrizione da parte della RNA polimerasi III, mentre la crescita cellulare non fa altrettanto (Clarke et al., 1996).

Tramite esperimenti di aggiunta *in trans*, in accordo con osservazioni ottenute *in vivo* (Lopez-De-Leon et al., 1992), è stato dimostrato che Brf1, che fa parte del fattore TFIIIB, è il componente limitante dell'apparato di trascrizione Pol III in estratti preparati da cellule cresciute sia in fase logaritmica che in fase stazionaria (Sethy et al., 1995).

L'attività TFIIIB è regolata da un gran numero di fattori; un esempio ampiamente studiato è quello di *P53*.

Il gene *P53* è altamente conservato tra le specie eucariotiche e codifica per un importante soppressore tumorale (Chernova et al., 1995) (Cox and Lane, 1995) (Haffner and Oren, 1995) (Ko and Prives, 1996); *P53* risulta perso o mutato in più del 50% di tutti i tumori nell'uomo (Hollstein et al., 1991). La proteina p53 possiede una varietà di attività biochimiche, tra le quali è compresa l'abilità di regolare la trascrizione (Chernova et al., 1995) (Cox and Lane, 1995) (Haffner and Oren, 1995) (Ko and Prives, 1996). Legandosi al DNA in modo sequenza-specifico, essa è in grado di stimolare la trascrizione dei vicini geni di classe II (Bargonetti et al., 1991) (Kern et al., 1991) (Farmer et al., 1992) (Zambetti et al., 1992). Sulla trascrizione da parte della RNA polimerasi III, invece, p53 ha mostrato di possedere proprietà inibitorie. Over-espressa in cellule transfettate COS o 293, infatti, p53 reprime i geni *SNR6* ed Alu (Chesnokov et al., 1996). Gli stessi geni sono anche inibiti *in vitro* dall'aggiunta di p53 ricombinante (Chesnokov et al., 1996) (White, 1998). La repressione trascrizionale può essere specificamente recuperata dall'aggiunta di un eccesso del fattore TFIIIB, un effetto, questo, che è stato visto essere dipendente da TBP (White, 1998). Questi risultati suggeriscono che p53 si leghi direttamente a TFIIIB e che in questo modo ne diminuisca l'attività. Un meccanismo come questo potrebbe avere un effetto generale sul livello di trascrizione Pol III-dipendente, anche se *SNR6* e Alu sono i geni più sensibili; questo deriva probabilmente dal fatto che tali geni hanno promotori per la trascrizione relativamente deboli.

Sempre sul gene *SNR6*, è ampiamente documentato che la HMG-*protein* Nhp6 svolge un importante ruolo di attivatore trascrizionale. Maggiori informazioni su questa proteina e sulle sue interazioni con il sistema trascrizionale Pol III sono riportate in *sezione 1.5.2*.

### *1.5.2 Nhp6 e RNA polimerasi III*

Le proteine HMG1/2 sono proteine di architettura abbondanti e la cui principale funzione si rivela nella trascrizione da parte della RNA polimerasi II (Bustin, 1999); queste proteine si legano al solco minore del DNA con scarsa o nulla specificità di sequenza e piegano il DNA verso il solco maggiore di ~90° (vedi (Bustin, 1999) per review).

Le proteine Nhp6A e Nhp6B sono proteine HMG1-*like* di *Saccharomyces cerevisiae* identiche fra loro per l'80%. Nhp6A e Nhp6B hanno una massa molecolare di ~11 kDa e contengono un'unica HMG-*box*.

La struttura della proteina Nhp6A è stata determinata (Allain et al., 1999); essa assume una configurazione a forma di L costituita da tre α-eliche e contiene una coda non strutturata, N-terminale rispetto al dominio HMG, fortemente basica (vedi *Fig. 1.8*, *A*). Il segmento N-terminale è critico per il legame al DNA e per le proprietà funzionali di Nhp6A (Yen et al., 1998); proteine mutanti in questa regione, infatti, sono risultate incapaci di formare complessi stabili con il DNA, di co-attivare la trascrizione e di complementare il difetto di crescita di cellule di lievito *nhp6ΔΔ*.

E' stato costruito anche un modello, basato su NMR, del complesso di Nhp6A con il DNA (Allain et al., 1999). Il DNA in questo complesso risulta notevolmente piegato ed in due fessure adiacenti sono inserite le catene laterali di un residuo Met e di un residuo Phe importanti per il ripiegamento (vedi *Fig. 1.8*, *B*).

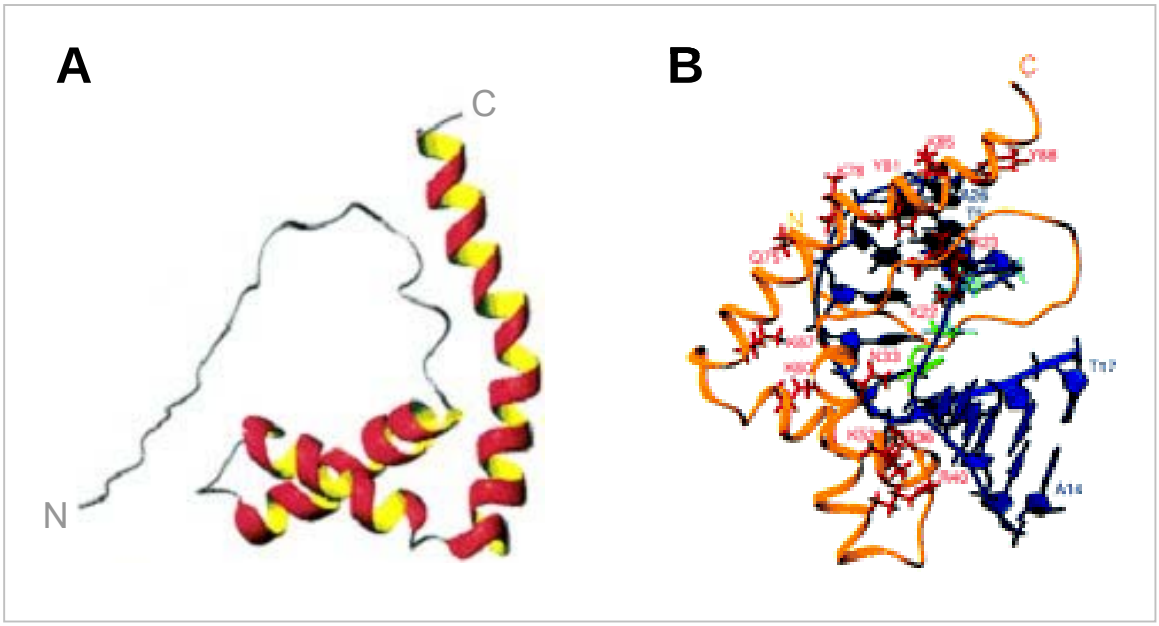



**Fig. 1.8: Struttura tridimensionale di Nhp6A da sola ed in complesso con il DNA.**
(A): Visione tridimensionale dello scheletro della struttura a più bassa energia di Nhp6A. Le regioni ad α-elica sono in rosso e giallo. Si noti che l'elica 3 è piegata al centro di 30°. Le estremità N- e C-terminali sono indicate. (B): Visione del solco minore del DNA che ne illustra i principali contatti con Nhp6A. Lo scheletro della proteina è rappresentato in giallo e le catene laterali degli amminoacidi responsabili del legame al DNA sono in rosso. Il DNA è rappresentato in blu.
*(Da Allain et al, 1999).*

Quantitativamente, Nhp6A è presente a livelli tre volte superiori rispetto a Nhp6B in cellule *wild-type* ed entrambi i geni sono soggetti a meccanismi di compensazione del dosaggio genico (Kolodrubetz et al., 2001).

L'importanza biologica di Nhp6A e Nhp6B è stata esaminata attraverso esperimenti di delezione genica. Entrambi i geni *NHP6A* e *NHP6B* devono essere deleti perché si osservi un fenotipo. I ceppi di lievito che mancano di Nhp6 crescono due volte più lentamente rispetto ai ceppi *wild-type* a 30°C, mentre non possono crescere a 37° in assenza di un supporto osmotico (1M sorbitolo) (Costigan et al., 1994). Questo fenotipo rimediabile osmoticamente è spesso associato a difetti nella biosintesi della parete cellulare; le cellule doppio mutanti (*nhp6ΔΔ*) mostrano anche una varietà di difetti morfologici a 37°C, compresa una distribuzione anomala di chitina e actina (Costigan et al., 1994). Questi fenotipi si accordano con il coinvolgimento di Nhp6 nella trascrizione di un certo numero di geni.

Nhp6 è coinvolta nella attivazione della trascrizione sia da parte della RNA polimerasi II che della RNA polimerasi III. In lievito è stato individuato un certo numero di mutanti in *TBP* che sono vitali in cellule *wild-type* ma che diventano letali in assenza di Nhp6 (Eriksson et al., 2004); molte di queste mutazioni sono localizzate in prossimità dei residui di TBP che sono noti contattare TFIIA o Brf1. Questo suggerisce che Nhp6 diventi essenziale quando una mutazione in TBP compromette l'abilità di quest'ultima di interagire con TFIIA o Brf1.

*NHP6A* e *NHP6B* sono stati identificati come soppressori multicopia di una mutazione nel promotore del gene *SNR6*, trascritto da Pol III (Lopez et al., 2001); cercando soppressori multicopia del difetto di crescita di mutanti *nhp6ΔΔ*, inoltre, è stato identificato il gene *BRF1*, che codifica per il componente limitante del fattore di trascrizione TFIIIB, specifico del sistema Pol III (Kruppa et al., 2001). Queste scoperte hanno fornito una forte evidenza *in vivo* del fatto che Nhp6 gioca un ruolo importante nella trascrizione di *SNR6*.

I livelli di U6 snRNA in stato stazionario diminuiscono di 10 o più volte in un ceppo *nhp6ΔΔ* dopo *shift* alla temperatura non permissiva (Kruppa et al., 2001), e studi di mutazione del promotore di *SNR6* confermano che diminuzioni di questo ordine di grandezza sono letali in lievito (Kaiser and Brow, 1995). Il basso livello di U6 snRNA, pertanto, è da solo sufficiente per dare spiegazione del fenotipo termosensibile dei mutanti *nhp6ΔΔ*.

Ulteriore supporto alla conclusione secondo la quale Nhp6 è coinvolta nella trascrizione di *SNR6* viene da esperimenti di *chromatin immunoprecipitation*, che mostrano che il legame di TBP a *SNR6* è ridotto in un mutante *nhp6ΔΔ* (Eriksson et al., 2004), e da trascrizioni *in vitro*, condotte sia con estratti di lievito che con un sistema ricostituito, che mostrano che l'aggiunta di rNhp6 risulta in uno specifico incremento della trascrizione di U6 (Kruppa et al., 2001) (Lopez et al., 2001).

Una caratteristica particolare di *SNR6* è la distanza sub-ottimale alla quale si trovano i due elementi promotori A- e B-*box*: la B-*box* si trova 202 bp a valle della A-*box*, in 3' rispetto al sito di terminazione della trascrizione (Brow and Guthrie, 1990; Eschenlauer et al., 1993). Studi biochimici hanno indicato che questa distanza anomala risulta in una riduzione della forza del promotore, derivante dal fatto che TFIIIC non è in grado di legarsi in modo efficiente ad entrambe le A- e B-*box* (Cannon et al., 1986; Dingermann et al., 1983; Eschenlauer et al., 1993; Fabrizio et al., 1987).

Come detto, l'over-espressione della subunità Brf1 di TFIIIB, così come l'espressione di una mutazione dominante (*PCF1-1*) nella subunità di TFIIIC che interagisce con Brf1, aumenta l'espressione dello snRNA U6 nel ceppo *nhp6ΔΔ*. Queste stesse manipolazioni aumentano anche la trascrizione Pol III-dipendente di un gene per tRNA difettivo nel suo promotore (Lopez-De-Leon et al., 1992; Rameau et al., 1994), facilitando uno *step* limitante nell'assemblaggio del complesso di trascrizione, cioè l'interazione tra TFIIIC, legato al promotore, e Brf1 (Moir et al., 1997; Sethy-Coraci et al., 1998). L'insieme di questi dati suggerisce che la stimolazione trascrizionale dipendente da Nhp6 derivi dalla sua capacità di facilitare il reclutamento di Brf1 oppure una fase dell'assemblaggio del complesso trascrizionale che precede questo passaggio.

Il *footprinting in vivo* del promotore di *SNR6*, insieme con le proprietà di mutanti per delezione che riducono la distanza fra A- e B-*box*, ha suggerito l'idea che il DNA fra questi elementi sia assemblato in un complesso nucleoproteico (Eschenlauer et al., 1993; Gerlach et al., 1995; Marsolier et al., 1995; Shivaswamy et al., 2004) ed è stato ipotizzato che questo complesso nucleoproteico possa essere un nucleosoma posizionato; il segmento di 90-100 bp protetto dalla digestione, tuttavia, è considerevolmente più piccolo di quello atteso per un nucleosoma (~146 bp). Il legame di una molecola di Nhp6, pertanto, potrebbe produrre la protezione osservata tra A- e B-*box*. La localizzazione della regione protetta, insieme con le proprietà strutturali e biochimiche di Nhp6 (Allain et al., 1999; Yen et al.,

1998), sono consistenti con l'idea che Nhp6 recuperi la distanza stericamente sfavorevole fra i due elementi promotori in *SNR6* (Kruppa et al., 2001).

In effetti, saggi di *bandshift* hanno mostrato che Nhp6B interagisce con complessi TFIIIC-*SNR6* DNA e TFIIIC-TFIIIB-DNA (Kruppa et al., 2001); la specificità di tale interazione, tuttavia, non è certa dal momento che Nhp6 *per se* causa la formazione di una *ladder* di complessi proteina-DNA sul gene *SNR6*.

Un'ulteriore caratteristica del gene *U6*, è la presenza di un tratto $(dT\text{-}dA)_7$ tra la TATA-*box* ed il sito di inizio della trascrizione. In assenza di Nhp6, mutazioni in questa regione risultano letali; Nhp6 ed il tratto dT-dA, pertanto, sembrano cooperare nel dirigere l'assemblaggio di un complesso di trascrizione produttivo su *SNR6* (Martin et al., 2001).

*SNR6* è ad oggi l'unico gene di classe III per quale è stato dimostrato un inequivocabile incremento della trascrizione Nhp6-dipendente. Qualche gene per tRNA è stato testato a questo riguardo, senza risultare chiaramente stimolato; va osservato, tuttavia, che queste osservazioni sono state condotte su intere famiglie di tDNA iso-codificanti e non su singoli geni. Anche *in vivo*, in un ceppo *nhp6ΔΔ*, il grosso della sintesi di tRNA e 5S rRNA non varia in modo rilevabile dopo shift alla temperatura non permissiva (Kruppa et al., 2001) (Lopez et al., 2001).

E' stato recentemente mostrato, infine, che Nhp6 funziona come fattore in grado di migliorare la selezione del sito di inizio della trascrizione su geni per tRNA (Kassavetis and Steiner, 2006); la stessa proteina, inoltre, riduce la produzione di trascritti aspecifici *in vitro* (Kassavetis et al., 2005). Questo risultato potrebbe derivare da una sorta di competizione fra Nhp6 e TBP per il legame di regioni A/T-*rich* entro i geni per tRNA: Nhp6 e TBP interagiscono entrambe con il solco minore legandosi preferenzialmente con le regioni di DNA più facilmente deformabili (Allain et al., 1999; Juo et al., 1996); Nhp6 funzionerebbe come una sorta di *filler* impedendo il legame di TBP (e quindi di TFIIIB) a TATA-*box* imperfette.

# 2. SCOPO DELLA RICERCA

Con il presente lavoro di ricerca ci siamo posti l'obiettivo di indagare più in profondità su alcuni aspetti del processo di trascrizione da parte della RNA polimerasi III di *Saccharomyces cerevisiae*.

Tale enzima ha il compito, fondamentale nell'economia della cellula eucariotica, di condurre la sintesi di alcuni RNA, generalmente non tradotti in proteine, tra i quali l'rRNA 5S, i tRNA ed alcuni snRNA. Da questo enzima, dunque, dipende la produzione di alcune componenti fondamentali del complesso apparato di biosintesi delle proteine.

Quelli di classe III sono generalmente geni piuttosto brevi, trascritti con alta efficienza e, probabilmente proprio per questa ragione, liberi da strutture cromatiniche. Per questo, e per il fatto che soltanto tre sono i principali fattori di trascrizione per la RNA polimerasi III (TFIIIA, TFIIIB e TFIIIC), la trascrizione Pol III-dipendente può apparire ad un'analisi superficiale un processo relativamente semplice.

Con la nostra analisi abbiamo indagato più in dettaglio due delle fasi del ciclo trascrizionale di Pol III, la terminazione e l'inizio della trascrizione, rivelando per esse nuovi elementi di regolazione che ne rivelano la reale complessità.

E' noto che il segnale di terminazione per la RNA polimerasi III è costituito da una semplice serie di T consecutive sul filamento codificante del DNA e che il segnale di terminazione è riconosciuto direttamente dalla polimerasi, senza l'intervento di fattori proteici accessori.

Sole 5 o 6 T consecutive costituiscono il segnale di terminazione in *S. cerevisiae* (Allison and Hall, 1985). Dati sparsi e che difficilmente possono essere tradotti in una regola generale, tuttavia, indicano che il segnale di terminazione per Pol III varia tra le diverse specie eucariotiche (Allison and Hall, 1985; Bogenhagen and Brown, 1981; Cozzarelli et al., 1983; Hamada et al., 2000). Il fatto che il segnale di terminazione sia costituito solamente da una breve serie di T pone il problema che simili serie possono, con una certa frequenza, ricorrere casualmente all'interno del genoma. Data la semplicità di questo segnale, un T-*run* che per effetto del caso si trovasse all'interno di un gene di classe III

potrebbe costituire un problema ai fini della produzione di un trascritto funzionale. Bisogna pertanto ammettere o che serie di T consecutive all'interno dei geni Pol III sono state contro-selezionate dall'evoluzione, o che il segnale di terminazione è qualcosa di più complesso di un semplice poli-(dT) e che il funzionamento di tale elemento come terminatore può essere variamente condizionato.

Con il nostro lavoro abbiamo condotto un'analisi dei segnali di terminazione per Pol III in diverse specie eucariotiche e, focalizzando l'attenzione su *S. cerevisiae*, abbiamo individuato alcuni elementi di sequenza e posizionali in grado di influenzare il comportamento della polimerasi III in fase di terminazione.

Riguardo alla fase di inizio della trascrizione, abbiamo focalizzato la nostra attenzione sui geni codificanti i tRNA di lievito. L'assemblaggio del complesso di pre-inizio su questi geni prevede l'intervento sequenziale di due fattori di trascrizione: TFIIIC, che riconosce i promotori intragenici (A- e B-*box*) e si lega ad essi, e TFIIIB che, in modo TFIIIC-dipendente, si lega sui tDNA in una regione a monte del sito di inizio della trascrizione; a sua volta, TFIIIB recluta Pol III per dare inizio alla trascrizione (Kassavetis et al., 1990; Kassavetis et al., 1989).

Il contatto TFIIIB-DNA nella regione a monte dei geni per tRNA è generalmente considerato sequenza-indipendente (Geiduschek and Kassavetis, 2001). Sono state mostrate, tuttavia, evidenze secondo le quali alcuni tDNA, appartenenti a diverse specie eucariotiche, posseggono particolari caratteristiche nella loro regione fiancheggiante in 5' in grado di contribuire alla funzionalità del promotore (Dieci et al., 2000; Hamada et al., 2001; Raymond et al., 1985). Abbiamo dunque condotto un'analisi sulle sequenze in 5' dei tDNA di *S. cerevisiae*, ipotizzando che la presenza di elementi di sequenza conservati in tale regione non fosse mai stata individuata a causa della sua natura degenerata e ritenendo improbabile che non ci sia stato alcun tipo di pressione selettiva verso l'ottimizzazione del legame di TFIIIB sui tDNA, che sono geni trascritti con alta efficienza.

Con la nostra analisi abbiamo difatti individuato particolari caratteristiche della regione in 5' dei tDNA che si sono rivelate influenzare l'espressione dei singoli geni per tRNA.

Abbiamo quindi verificato se la HMG *protein* Nhp6, che è nota aumentare specificamente l'espressione del gene di classe III *SNR6* (Kruppa et al., 2001; Lopez et al., 2001), ha un

analogo effetto sulla trascrizione dei geni per tRNA. E' infatti possibile che Nhp6 intervenga nella fase di assemblaggio del complesso di trascrizione sui geni per tRNA, come avviene sul gene *SNR6*; in particolare Nhp6, proteina che legandosi induce un notevole ripiegamento nel DNA, potrebbe svolgere un ruolo di tipo strutturale ed influenzare il legame del fattore TFIIIB sulla stessa regione a monte dei tDNA sulla quale abbiamo condotto la nostra analisi.

Abbiamo quindi indagato gli effetti della proteina Nhp6 sulla funzione non-codificante del $tRNA^{Thr}$ che si trova al locus *HMR*-E di *Saccharomyces cerevisiae*, quella di barriera al propagarsi dell'eterocromatina. Come proteina organizzatrice della cromatina, Nhp6 potrebbe modificare la funzione barriera del $tRNA^{Thr}$ sia influenzando l'assemblaggio del complesso di trascrizione Pol III su quel gene, sia generando effetti cromatinici su più larga scala.

# 3. MATERIALI E METODI

Per la nomenclatura dei geni per tRNA si fa riferimento al MIPS (*Munich Information center for Protein Sequences*, http://mips.gsf.de).

Tutti i geni clonati ed i relativi vettori plasmidici, insieme al nome dei *primers* utilizzati per il loro clonaggio o mutazione, sono elencati nelle tabelle *7.1* e *7.2* in appendice; le sequenze di tutti gli oligonucleotidi utilizzati come *primers* o come sonde sono elencate nella tabella *7.3* in appendice.

## 3.1 TERMINAZIONE DELLA TRASCRIZIONE: EFFETTI DEL CONTESTO SULLA FUNZIONALITÀ DEL TERMINATORE

### *3.1.1 Analisi statistica delle regioni fiancheggianti i tDNA in 3'*

Le coordinate per le regioni a valle dei tDNA sono state determinate direttamente dalle sequenze genomiche di *S. cerevisiae*, *S. pombe*, *M. musculus* e *H. sapiens* usando *Pol3scan* e *tRNAscan-SE* (Percudani et al., 1997) (Lowe and Eddy, 1997). Le sequenze individuate come tDNA sono state considerate per l'analisi se i confini del tRNA ed il tipo di anticodone sono stati riconosciuti identici dai due metodi; tale sistema di selezione, anche se non garantisce l'identificazione di tutti i tDNA genomici, consente di evitare falsi positivi.

Le regioni corrispondenti ai 100 nucleotidi immediatamente a valle della sequenza codificante i tRNA maturi sono state estratte come *files* FASTA ed usate per l'analisi dei segnali di terminazione. E' stato disegnato un elaborato Perl per l'analisi statistica di *T-run* di lunghezza arbitraria, facendo uso della *library BioPerl* (Stajich et al., 2002).

### 3.1.2 DNA templato

I tDNA P(AGG)CR, N(GTT)CR, N(GTT)KL, N(GTT)OL e L(CAA)LR2 sono stati amplificati tramite PCR da DNA genomico di lievito (ceppo S288C) usando *Deep Vent DNA polymerase* (New England Biolabs) e specifiche coppie di *primers* oligonucleotidici. I tDNA S(AGA)EL e M(CAT)E sono stati amplificati usando la *Pfu Turbo DNA Polymerase* (Stratagene). I prodotti di amplificazione, contenenti la regione codificante i tRNA + ~80 bp in 5' + ~50 bp in 3', sono stati inseriti in vettori plasmidici pUC-derivati e verificati tramite sequenziamento.

Il tRNA V(AAC)MR2 si trova nel plasmide pY7 precedentemente descritto (Baker et al., 1982).

Le forme mutate nella regione in 3' dei tDNA N(GTT)CR e L(CAA)LR2 sono state ottenute tramite PCR usando come templato i tDNA clonati ed appropriati *primers* mutagenici *reverse*.

Tutti i prodotti di amplificazione sono stati inseriti in vettori plasmidici pUC-derivati e verificati tramite sequenziamento.

In modo analogo i costrutti SCR1Δ_T7 e SCR1Δ_3'L sono stati ottenuti con opportuni *primers reverse* usando come templato il gene *SCR1* precedentemente clonato. In SCR1Δ_T7 sono comprese 32 bp in 5', in SCR1Δ_3'L sono comprese 246 bp in 5' rispetto alla sequenza codificante.

Per gli studi *in vivo* i costrutti derivati da *SCR1* sono stati clonati entro il vettore *shuttle* YEp352.

I costrutti 5'U6-L(CAA)LR2 e 5'U6-L_3'N sono stati creati per PCR utilizzando i corrispondenti costrutti clonati come templato ed il *primer* 5'U6_Sac_Leu186, che porta la sequenza che fiancheggia in 5' il gene *SNR6* di *Saccharomyces cerevisiae* seguita da un sito SacI, come *primer forward*. I frammenti amplificati così ottenuti sono stati inseriti in un plasmide pNEB193 modificato, in cui il sito SacI facente parte del *polylinker* è stato distrutto. Il costrutto 5'U6-SCR1Δ_3'L è stato ottenuto usando il *primer* U6_SCR1, che porta la sequenza a monte di *SNR6*, come *primer forward*; in un distinto costrutto il sito di restrizione per SacI è stato posto 5 bp a monte di SCR1Δ_3'L; entrambe le varianti di SCR1Δ_3'L sono state inserite nel vettore pNEB193 modificato.

L'unità trascrizionale *SNR52* è stata amplificata da DNA genomico usando la *Pfu Turbo DNA polymerase*.

Il mutante SNR52_mutT è stato costruito mediante PCR ricombinante (Higuchi, 1990). Tramite opportuni *primers* sono stati prodotti due amplificati primari di PCR che si sovrappongono in parte. Dopo gel purificazione, i prodotti primari di amplificazione sono stati mescolati ed usati in una seguente reazione di PCR, usando i due *primers* esterni; in questo modo è stato ottenuto l'amplificato secondario *full-length*.

Analogamente ai costrutti descritti sopra, è stato prodotto il costrutto 5'U6_SNR52, utilizzando il *primer* 5'U6_SNR52 ed il gene *SNR52* clonato come templato. Tutte le varianti di *SNR52* sono state inserite nel vettore pNEB193.

### *3.1.3 Saggi di trascrizione* in vitro *in condizioni* standard

Le condizioni di reazione per i saggi di trascrizione *in vitro* in *multiple* e *single round* e le procedure per la purificazione dell'RNA trascritto e l'analisi sono state condotte essenzialmente come già descritto (Ferrari et al., 2004) (Dieci et al., 2000).

E'stato utilizzato un sistema Pol III di *S. cerevisiae* ricostituito, comprendente le proteine TBP (40 ng) e Brf1 (80 ng) ricombinanti, purificate dopo over-espressione in *E. coli* (Huet et al., 1996), le frazioni B" (purificata fino allo *step* Bio-Rex 70 a partire da *pellets* cromatinici (Kassavetis et al., 1992b), 500 ng) e TFIIIC (purificata fino allo *step* DEAE-Sephadex A-25 (Huet et al., 1996), 150 ng), ed RNA polimerasi III altamente purificata (10 ng).

Nelle reazioni di trascrizione *standard* sono stati utilizzati 100 ng di DNA plasmidico templato. Salvo dove diversamente indicato, la concentrazione di ATP, CTP e GTP è 500 μM e quella di UTP 250 μM. I trascritti sono stati marcati radioattivamente con [α-$^{32}$P]UTP.

I DNA templato sono stati prima incubati con TFIIIC e TFIIIB per 20' a 20°C, quindi sono stati aggiunti Pol III ed NTPs e si è fatta procedere la trascrizione per 20' a 20°C.

Nell'esperimento in *Figura 4.4*, una stessa quantità di un RNA marcato radioattivamente di taglia diversa dai trascritti è stato aggiunto ad ogni singola reazione dopo che è stata fatta

avvenire la trascrizione ma prima delle procedure di purificazione e precipitazione dell'RNA. Tale RNA marcato funziona da *recovery marker* ed è stato utilizzato per normalizzare i dati relativi alla trascrizione.

I saggi di trascrizione *single round* sono stati ottenuti pre-incubando il DNA templato con solo tre dei quattro NTPs, scelti specificamente per ogni templato, tali da ottenere complessi ternari in fase di stallo comprendenti un RNA trascritto lungo almeno 5 nucleotidi; la reazione è stata quindi fatta procedere per aggiunta del quarto nucleotide in presenza di eparina ad una concentrazione di 100 μg/ml; l'eparina in fase di terminazione sequestra la RNA polimerasi III ed in tal modo impedisce il re-inizio trascrizionale da parte di quest'ultima.

### *3.1.4 Saggi di trascrizione* in vitro *in assenza di TFIIIC*

In *Figura 4.7*, i complessi di trascrizione sono stati assemblati alternativamente in presenza o in assenza di TFIIIC usando quantità incrementate dei diversi componenti di TFIIIB: 400 ng di TBP, 240 ng di Brf1 e 700 ng di frazione B" (*Fig. 4.7, B*) oppure 20 ng di Bdp1 di lievito ricombinante (Ferrari et al., 2004) (*Fig. 4.7, A*).

Nel esperimento di trascrizione in assenza di fattori in *Figura 4.7, B*, i frammenti di restrizione aventi un'estremità 3' protrudente sono stati ottenuti per digestione dei templati, descritti precedentemente, con SacI, che taglia 5 bp a monte del naturale sito di inizio della trascrizione, e con BamHI (per L(CAA)LR2) oppure NdeI (perL_3'N e SCR1Δ_3'L), che tagliano a valle dell'inserto clonato, nel *polylinker* del plasmide. I frammenti di restrizione sono stati gel purificati, trattati con fenolo-cloroformio e precipitati con etanolo prima di essere utilizzati nel saggio di trascrizione.

Le condizioni generali di trascrizione sono le stesse utilizzate nella trascrizione diretta da fattori. I frammenti 3'-protrudenti (30ng) sono stati prima incubati con 20 ng di Pol III purificata e con il di-nucleotide *primer* CpU ad una concentrazione 400 μM nel buffer di trascrizione contenente 160 μg/ml di BSA per 15' a 20°C, in un volume totale di reazione di 25 μl. Sono stati quindi aggiunti gli NTPs e l'[α-$^{32}$P]UTP (10 μCi) e si è fatta procedere la trascrizione per 30' a 20°C.

In tutti gli esperimenti di trascrizione, la quantificazione dell'RNA trascritto radioattivo è stata ottenuta tramite *Phosphorimager* utilizzando il *software* di scansione *QuantityOne Personal Imager FX* (BioRad) ed il *software* di analisi *Multi-Analyst* (BioRad). Ai fini della quantificazione, i segnali corrispondenti ad RNA di diversa lunghezza sono stati normalizzati per il numero di residui U incorporati.

### *3.1.5 Analisi* in vivo

Per l'analisi *in vivo* degli RNA trascritti, cellule di lievito del ceppo YPH500 sono state trasformate con i diversi costrutti "SCR1Δ" clonati entro il vettore *shuttle* YPH500, vettore ad alto numero di copie che fornisce la resistenza all'ampicillina ($Amp^r$) e che consente la selezione su terreno privo di uracile (*URA3*).

E' stata utilizzata la procedura di trasformazione con acetato di litio. Per ogni trasformazione, cellule provenienti da una singola colonia sono state fatte crescere fino ad un $O.D._{600}$ di ~0,8; le cellule (~8 O.D. totali) sono quindi state raccolte per centrifugazione, lavate prima con $H_2O$ poi con acetato di litio 0.1 M / Tris-HCl 10 mM / EDTA 1 mM, quindi risospese in 100 µl della stessa soluzione ed incubate a 30°C per 30'. Sono stati quindi aggiunti 15 µg di *salmon sperm carrier DNA* e 2 µg del DNA plasmidico d'interesse. Dopo incubazione a 30°C per 30', sono stati aggiunti 700 µl della soluzione sopra descritta + 40% $PEG_{4000}$; dopo ulteriore incubazione di 2 h a 30°C, le cellule sono state sottoposte a *shock* termico (15' a 42°C) e piastrate su terreno selettivo privo di uracile.

Gli RNA totali sono stati estratti con la procedura che fa uso di fenolo acido caldo. Cellule cresciute in fase logaritmica (10 O.D. totali) sono state raccolte e lavate con $H_2O$ *RNasi-free*, quindi risospese in 400 µl Tris-HCl 50 mM, pH 5.8 + 40 µl SDS 10% + 400 µl fenolo acido pre-riscaldato a 65°C; dopo energica mescolata ed incubazione a 65°C per 60', la fase acquosa, contenente la frazione nucleica del lisato cellulare, è stata trattata con cloroformio e gli RNA sono stati precipitati con acetato di ammonio ed etanolo, quindi quantificati tramite spettrofotometria; la quantità e la qualità degli RNA ottenuti è stata controllata su gel di poliacrilammide 6% - 7M urea colorato con etidio bromuro.

I campioni di RNA (10 µg ciascuno) sono stati sottoposti ad elettroforesi su gel di acrilammide 6% - 7M urea e trasferiti su membrana *GeneScreen Plus* (PerkinElmer Life Science); il filtro è stato quindi ibridato con la specifica sonda SCR1_prext_2, marcata in 5' e purificata su colonne *Micro Bio Spin* (BioRad), complementare al trascritto *SCR1* tra le posizioni +43 e +73; l'ibridazione è stata condotta *over-night* a 28°C in 5x SSC / 5X *Denhardt's solution* / 0,1 mg/ml *salmon sperm DNA* / 0,5% SDS; sono stati quindi condotti un breve lavaggio in 2X SSC / 0,1% SDS e due lavaggi in 1X SSC / 0,1% SDS.

I prodotti di ibridazione sono stati visualizzati e quantificati tramite *phosphorimager*.

## 3.2 INIZIO DELLA TRASCRIZIONE: SEQUENZA A MONTE DEI GENI PER TRNA ED EFFETTI DELLA PROTEINA NHP6

### *3.2.1 Analisi statistica delle regioni fiancheggianti i tDNA in 5'*

Le sequenze in 5' dei 274 tDNA di *S. cerevisiae*, dalla posizione –1 alla posizione –90 rispetto al primo nucleotide del tRNA maturo, sono state ricavate dalla sequenza del genoma di lievito prelevata dal sito *Web* del MIPS, usando le coordinate ricavate usando l'algoritmo *Pol3scan* (Percudani et al., 1997) (Hani and Feldmann, 1998). I geni per tRNA associati con retroelementi Ty, identificati sulla base dell'annotazione MIPS e sull'analisi del *cluster* basata sulla similarità (Parsons, 1995), sono stati eliminati dal *set* di tDNA utilizzati per l'analisi delle regioni in 5'.

#### 3.2.1.1 Ricerca di segnali conservati nelle regioni fiancheggianti in 5'

Le regioni fiancheggianti in 5' dei 221 tDNAs analizzati sono state analizzate con una procedura bioinformatica sviluppata specificamente per l'identificazione di brevi motivi conservati in sequenze non correlate, proposta originariamente da Waterman e Jones (Waterman and Jones, 1990). Dato un *set* di sequenze non allineate, una lunghezza di parola fissa *k*, un numero di possibili *mismatch d* ed una larghezza di finestra *W*, l'algoritmo identifica, per ogni posizione, la parola di lunghezza *k* con punteggio migliore.

I segnali che si distaccavano maggiormente dal *background* sono stati ottenuti con *k*=6, *d*=1 e *W*=9.

Per valutare la significatività dei risultati ottenuti nei picchi di frequenza sono stati generati 1000 *sets* di 221 sequenze ottenute per *shuffling* delle regioni in 5' dei tDNA; per ogni *set* è stata eseguita la scansione per la ricerca di motivi conservati e sono stati calcolati il valore medio di frequenza e i valori massimo e minimo.

#### 3.2.1.2 Analisi di *bendability* delle regioni fiancheggianti in 5'

Il profilo di *bendability* delle regioni in 5' dei tDNA di *S. cerevisiae* è stato ottenuto usando parametri già riportati (Brukner et al., 1995). La *bendability* media di ogni tri-nucleotide consecutivo è stata calcolata muovendosi dalla posizione –90 alla posizione –1. I valori di b*endability* sono stati analizzati ulteriormente usando un intervallo di 4 nucleotidi. I valori attesi e la significatività sono state determinate tramite simulazioni condotte su 1000 *datasets* di sequenze casuali.

### *3.2.2 DNA templato*

Tutti i geni per tRNA analizzati in questo studio sono stati amplificati per PCR da DNA genomico con *Deep Vent Polymerase* (New England Biolabs) ed opportuni *primers* che consentono di amplificare, oltre alla regione codificante il tRNA, ~80 bp in 5' e ~50 bp in 3'. I prodotti amplificati sono stati inseriti in plasmidi pUC-derivati e sottoposti a sequenziamento.

Il gene $tRNA^{Syn2}$ (variante artificiale del $tRNA^{Glu}$(TTC) che contiene uno pseudo-introne (Krieg et al., 1991)) è stato sintetizzato per amplificazione con PCR usando i due oligonucleotidi Glu(TTC)syn2_fw e Glu(TTC)syn2_rev, oligonucleotidi parzialmente complementari che generano un sito per BamHI all'estremità 5' dell'amplicone. Il risultante tDNA sintetico è stato fuso alle regioni 5' di tN(GTT)CR o tN(GTT)NR. A tal fine le due regioni in 5' (fino alla posizione –58) sono state amplificate individualmente con i *primers* Asn19-58 e Asn19_flank_rev (per [5'CR]) e Asn232_short e Asn232_flank_rev (per [5'NR]), che generano un sito BamHI all'estremità 3' degli

ampliconi. Ognuno di questi due ampliconi è stato quindi tagliato con BamHI e ligato al frammento $tDNA^{Syn2}$ precedentemente sottoposto a restrizione con BamHI. I risultanti prodotti di ligazione ([5'CR]Syn2 e [5'NR]Syn2) sono stati inseriti nel vettore YEp352 e successivamente sub-clonati anche in pFL38 e in pFL46S. Anche l'amplicone $tDNA^{Syn2}$, senza l'aggiunta di alcuna sequenza in 5', è stato inserito nel vettore Yep352.

Il gene *BRF1* (*coding sequence* + 194 bp in 5' + 619 bp in 3') è stato inserito entro il vettore *shuttle* pFL45S tramite sub-clonaggio dal vettore pRSM3 (Andrau et al., 1999).

### *3.2.3 Saggi di trascrizione* in vitro

I saggi di trascrizione mostrati nelle figure *4.11*, *4.12, A,* e *4.17*, sono stati condotti come descritto in *sezione 3.1.3*, in presenza di una concentrazione di UTP 25 µM.

Nell'esperimento in *Figura 4.17* in luogo della frazione B" è stata usata la proteina Bdp1 ricombinante; ~1000 cpm di un DNA di 283 bp marcato radioattivamente, ottenuto per PCR, sono stati utilizzati come *recovery marker*. Le quantità crescenti di proteina ricombinante Nhp6A (da 0 a 200 ng) sono state aggiunte alle singole reazioni nella fase di pre-incubazione con DNA templato e fattori di trascrizione.

### *3.2.4 Analisi* in vivo

Per l'analisi *in vivo* mostrata in *Figura 4.12, B*, pannello di sinistra, i costrutti [5'CR]Syn2 e [5'NR]Syn2, clonati entro il vettore plasmidico YEp352 (*$Amp^r$, URA3*) sono stati trasformati entro il ceppo di lievito W303-1b secondo la procedura già descritta nella sezione *3.1.5*. L'esperimento nel pannello di destra deriva dalla trasformazione degli stessi costrutti [5'CR]Syn2 e [5'NR]Syn2, clonati entro il plasmide centromerico pFL38 (*CEN6, $Amp^r$, URA3*) nel ceppo di lievito YPH500. 5 µg di RNA totale, estratto secondo la procedura già descritta, sono stati trasferiti sulla membrana *Hybond-N* (Amersham Biosciences) ed ibridati con la specifica sonda Glu_ttc_synt_minus, che si ibrida sia con il $tRNA^{Glu}$(TTC) naturale che con la sua versione modificata Syn2.

L'analisi *in vivo* mostrata in *Figura 4.12, C*, condotta nel contesto in cui *brf1* è mutato, è stata condotta utilizzando il ceppo di lievito SHy76 (*MATa leu2 ura3-52 brf1::HIS3*) (Colbert and Hahn, 1992), in cui la copia genomica del gene *BRF1* è deleta, ed i plasmidi pFL38 (*CEN6, Amp$^r$, URA3*), entro il quale si trovano i due costrutti [5'CR]Syn2 e [5'NR]Syn2, e pRSM3*II.6 (*Amp$^r$, LEU2*), entro il quale si trova la copia mutata (*D464A D466A*) del gene *brf1* (Andrau et al., 1999). L'RNA totale è stato estratto prima o dopo lo *shift* di 6h alla temperatura non permissiva di 37°C e sottoposto ad analisi *Northern* in modo analogo a quanto appena descritto.

Nell'esperimento in *Figura 4.13*, cellule alternativamente del ceppo mutante *nhp6ΔΔ* Y869 (*MATα ura3-52 trp1-289 his3-Δ1 leu2-3 112 gal2 gal10 nhp6A-Δ3::URA3 nhp6B-Δ3::HIS3*) o del suo corrispondente ceppo isogenico, *wild-type* per Nhp6, Y865 (Costigan et al., 1994), sono state trasformate con i costrutti [5'CR]Syn2 e [5'NR]Syn2 clonati entro il vettore ad alto numero di copie pFL46S (*Amp$^r$, LEU2*); i trasformanti sono stati selezionati su terreno privo di leucina; le cellule cresciute in terreno selettivo sono state fatte crescere a 30°C e raccolte prima o dopo lo shift di 2h alla temperatura di 37°C; le successive fasi di estrazione degli RNA ed ibridazione su filtro sono state condotte come in precedenza usando la membrana *GeneScreen Plus* (PerkinElmer Life Science) e la sonda Glu_ttc_synt_minus.

Per l'analisi in *Figura 4.14*, è stata eseguita una doppia trasformazione, alternativamente dei due ceppi Y865 o Y869 (wt e *nhp6ΔΔ*) utilizzando 1 µg di ciascuno dei due plasmidi pFL46S (*Amp$^r$, LEU2*), contenente alternativamente [5'CR]Syn2 o [5'NR]Syn2, e pFL39 (*CEN6, Amp$^r$, TRP1*) alternativamente vuoto o contenente il gene *NHP6A* (*coding sequence* + 381 bp in 5' + 253 bp in 3'). Le cellule trasformate sono state cresciute a 30°C in terreno selettivo privo di leucina e triptofano e campioni di 10 µg ciascuno di RNA totale sono stati sottoposti ad analisi *Northern*, in modo analogo a quanto appena descritto.

In *Figura 4.15* è stata eseguita una doppia trasformazione dei due ceppi Y865 e Y869 (wt e *nhp6ΔΔ*) con il plasmide pFL46S (*Amp$^r$, LEU2*) contenente alternativamente i due costrutti [5'CR]Syn2 o [5'NR]Syn2, e con il plasmide ad alto numero di copie pFL45S (*Amp$^r$, TRP1*), alternativamente vuoto o contenente il gene *BRF1* (*coding sequence* + 194 bp in 5' + 619 bp in 3').

Tale doppia trasformazione è stata ottenuta per elettroporazione. Cellule di lievito dei due ceppi desiderati sono state cresciute *over-night* in terreno massimo, quindi raccolte (a

partire da 1,5 ml di coltura per ogni trasformazione) e lavate prima con $H_2O$ poi con sorbitolo 1M; alle cellule raccolte dopo i lavaggi sono stati aggiunti 500 ng di ciascuno dei due plasmidi di interesse e l'elettroporazione è stata effettuata con una scarica elettrica a 1800 V. Le cellule trasformate sono state piastrate su terreno minimo privo di leucina e triptofano e contenete sorbitolo 1M.

Le fasi successive di estrazione degli RNA, *blotting* ed ibridazione con la sonda Glu_ttc_synt_minus sono state eseguite come in precedenza.

### *3.2.5 Chromatin immunoprecipitation*

Per gli esperimenti di immunoprecipitazione cromatinica (ChIP) sono stati utilizzati tre ceppi di lievito in cui sono marcate, per mezzo dell'inserzione di una cassetta di fusione TAP (*Tandem Affinity Purification*)-*tag* a livello cromosomico, le proteine Brf1, Nhp6A e Nhp6B. Come controllo negativo è stato utilizzato il ceppo BY4741 in cui nessuna proteina è marcata tramite TAP-*tag*. L'inserzione del TAP-*tag* consente la produzione di proteine funzionali e prodotte quantitativamente ai livelli naturali (Rigaut et al., 1999). Tali ceppi sono stati prelevati dalla libreria commerciale *Yeast TAP Fusion Collection* (Open Biosystem).

A 200 ml di coltura di lieviti cresciuti fino all'O.D. di ~0,5 è stata aggiunta formaldeide ad una concentrazione finale dell'1%. In tal modo è stato possibile ottenere il *cross-linking in vivo* delle proteine alle regioni di DNA con le quali esse interagiscono. Dopo 20' di incubazione a 20°C è stata aggiunta glicina ad una concentrazione finale 240 mM; dopo 5' le cellule sono state sottoposte a due lavaggi con tampone TBS (20mM Tris-HCl, 150 mM NaCl) e ad un lavaggio con *Fa-lysis buffer* (50 mM HEPES-KOH, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0,1% Na-deossicolato, 0,1% SDS, 1 mM PMSF). Le cellule, raccolte per centrifugazione, sono state quindi risospese in 500 µl di *Fa-lysis buffer* + 0,5% SDS; la lisi cellulare è stata ottenuta con biglie di vetro e, dopo diluizione con *FA-lysis buffer*, la cromatina "cross-linkata" è stata "pellettata" centrifugando per 20' a 20000 g. La cromatina è stata quindi lavata con *Fa-lysis buffer* e frammentata per sonicazione fino all'ottenimento di frammenti di taglia compresa mediamente fra 200 e 700 bp. Dopo ulteriore centrifuga di 20' a 20000 g, la cromatina è stata aliquotata e di ogni campione è

stata conservata un'aliquota "*input*"; alla restante cromatina sono stati aggiunte *beads IgG Sepharose 6 Fast Flow* (Amersham Bioscience), in cui sulla matrice di supporto sono covalentemente legate IgG umane che riconoscono la proteina A parte del *TAP-tag* (Rigaut et al., 1999); dopo il legame (*over-night* a 4°C), sono stati condotti due lavaggi con *Fa-lysis buffer* e quantità crescenti di NaCl (275 mM e 500 mM), un lavaggio con *wash buffer* (10 mM Tris-HCl, 0,25M LiCl, 1 mM EDTA, 0,5% NP-40, 0,5% Na-deossicolato) ed un lavaggio con TE (50mM Tris-HCl, 10 mM EDTA). L'eluizione dalle *beads* è stata ottenuta scaldando per 10' a 65°C in presenza di 1% SDS; il surnatante, insieme alla cromatina dell'aliquota "*input*" che non era stata soggetta ad immunoprecipitazione, è stato quindi sottoposto a trattamento con *Pronase* (Calbiochem) 0,8 mg/ml; questo *mix* di endo- ed eso-peptidasi reverte il *cross-linking* distruggendo le proteine; il DNA è stato quindi purificato con fenolo-cloroformio-alcool isoamilico, precipitato con Na-acetato ed etanolo e risospeso in Tris-HCl 10mM.

Regioni specifiche del DNA così ottenuto (immunoprecipitato e *input*) sono state sottoposte a PCR con coppie di oligonucleotidi che hanno consentito di amplificare diversi geni per tRNA ed il gene *SNR6*.

Per l'amplificazione, sono stati utilizzati per tutti i campioni 0,5 µl di DNA templato immunoprecipitato e 0,25 µl di DNA templato *input*; sono strati condotti 28 cicli di amplificazione con una temperatura di a*nnealing* per i *primer* di 54°C. I prodotti di amplificazione, marcati per incorporazone con [α-$^{32}$P]dCTP, sono stati fatti correre su gel di acrilammide al 6% e la visualizzazione è stata ottenuta tramite *phosphorimager*.

Il risultato relativo alle proteine Nhp6A e Nhp6B sugli specifici geni di classe III è stato valutato dal raffronto con il controllo negativo, BY4741, e con il controllo positivo, Brf1 *TAP-tagged*; i dati sono stati normalizzati rispetto ai relativi segnali "*input*".

### *3.2.6 Purificazione di proteina*

La proteina Nhp6A è stata prodotta in forma ricombinante e purificata da cellule di *E. coli*.

A tal fine la sequenza codificante la proteina Nhp6A, con un secondo codone di stop aggiunto dopo quello naturale, è stata clonata in modo direzionato entro il vettore di

espressione pET28-pme (*Kan*$^r$), in modo da ottenere la proteina fusa ad un *His-tag* N-terminale.

Il costrutto ottenuto è stato trasformato in cellule Rosetta, ceppo BL21-derivato in cui i tRNA per i codoni rari in *E. coli* sono portati dal plasmide pRARE (Chl$^r$) (Novy et al., 2001). Le cellule, cresciute in terreno LB addizionato di cloramfenicolo e canamicina, sono state cresciute fino all'O.D.$_{600}$ di ~0,6, quindi indotte con 1 mM IPTG ed incubate per 2h a 37°C.

La prova di induzione e di solubilità su piccola scala ha rivelato che Nhp6A in questo modo viene over-espressa e che la proteina è parzialmente solubile (il 45% circa della proteina si trova nel surnatante). Si è notata inoltre un'anomalia elettroforetica di Nhp6A che, sottoposta a *SDS-page*, ha mostrato la mobilità elettroforetica di una proteina di peso molecolare superiore a 14000 quando, compreso il *tag* N-terminale, ha un peso molecolare di ~13800.

Per la purificazione, la lisi cellulare è stata ottenuta per aggiunta di lisozima, ad una concentrazione finale di 0,1 mg/ml, e sonicazione. Dopo centrifugazione, la frazione surnatante è stata incubata in *batch* con la resina *Ni-NTA Agarose* (Qiagen) per 1h a 4°C; dopo opportuni lavaggi con tampone di lisi + 20 mM imidazolo, l'eluizione è stata ottenuta con 3 volumi di colonna di *buffer* di lisi + 250 mM imidazolo + 10% glicerolo.

La quantità di proteina presente nelle diverse frazioni è stata misurata con il saggio *Bradford* e l'assenza di proteine contaminanti nelle frazioni risultate contenere Nhp6A è stata verificata con *SDS-page*.

### *3.2.7 Saggi di* mating

I saggi di *mating* sono stati condotti essenzialmente come già descritto (Donze et al., 1999).

La capacità di *mating* delle cellule è stata monitorata tramite replica delle colonie su piastre selettive di YMD ricoperte dal *mating lawn* JRY19 e mantenendo la selezione per i plasmidi a seguito del *mating*.

I ceppi utilizzati per i saggi basati su plasmide sono i seguenti:

DDY171 (*MATα ADE2 his3 leu2 LYS2 trp1 ura3 hmrΔ*)

DDY591 (*MATα ade2 his3 leu2 LYS2 trp1 ura3 nhp6a:KanMX nhp6b:ADE2*)

I ceppi utilizzati per i saggi in contesto cromosomico sono i seguenti:

DDY689 (*MATα ADE2 his3 leu2 lys2Δ trp1 ura3 HMR-E-tDNA-a1*)

DDY714 (*MATα ADE2 his3 leu2 lys2Δ trp1 ura3 HMR-E-tDNA-a1 nhp6a::KanMX*)

DDY669 (*MATα ade2 his3 leu2 lys2Δ trp1 ura3 HMR-E-tDNA-a1 nhp6b::ADE2*)

DDY671 (*MATα ade2 his3 leu2 lys2Δ trp1 ura3 HMR-E-tDNA-a1 nhp6a::KanMX nhp6b::ADE2*)

DDY705 (*MATα ade2 his3 leu2 LYS2 trp1 ura3 HMR-E-tDNA-a1 nhp6a::KanMX nhp6b::ADE2 sir2Δ::TRP1*).

# 4. RISULTATI E DISCUSSIONE

## 4.1 TERMINAZIONE DELLA TRASCRIZIONE: EFFETTI DEL CONTESTO SULLA FUNZIONALITÀ DEL TERMINATORE

Parte dei risultati riportati in questa sezione sono stati pubblicati nell'articolo riportato in appendice (Braglia et al., 2005).

Le regioni fiancheggianti i geni per tRNA nei diversi eucarioti presentano *pattern* di sequenza conservati che divergono considerevolmente andando dai lieviti ai mammiferi (Giuliodori et al., 2003). Un ulteriore elemento di divergenza fra lieviti e mammiferi è costituito dal diverso modo di rispondere a segnali oligo-(dT) da parte della RNA polimerasi III in fase di terminazione (Hamada et al., 2000).

Al fine di approfondire le caratteristiche dei segnali di terminazione per Pol III su scala genomica, abbiamo scelto di analizzare e confrontare i terminatori trascrizionali nei genomi di due lieviti (*S. cerevisiae* e *S. pombe*) e di due mammiferi (*M. musculus* e *H. sapiens*). Sull'insieme di tutti i geni per tRNA di questi organismi, abbiamo come primo passo cercato serie di 4 o più T consecutive entro le prime 100 bp in 3' rispetto alla regione codificante i tRNA maturi; $T_4$, infatti, è la più breve serie oligo-(dT) che è nota causare terminazione da parte della RNA polimerasi III (Geiduschek and Tocchini-Valentini, 1988).

In *S. cerevisiae* e *S. pombe* non abbiamo mai trovato una serie di sole 4 T come unico segnale di terminazione: entro le prime 40 bp a valle della sequenza codificante è sempre presente una serie di 5 o più T consecutive; l'unica eccezione è costituita dai tDNA dimerici, nei quali l'unità trascrizionale tDNA posta più a monte è priva di terminatore (Mao et al., 1980) (Schmidt et al., 1980). Questo risultato indica che, in lievito, segnali più brevi di $T_5$ non costituiscono un segnale di terminazione per la RNA polimerasi III; si noti inoltre serie di 5 o più T consecutive non si trovano mai all'interno della regione codificante i tRNA di lievito.

Contrariamente a quanto appena descritto, molti dei tDNA di mammifero hanno una serie di sole 4 T come unico segnale di terminazione; questo risultato si accorda con studi precedenti in cui è stato riportato che la RNA polimerasi III di uomo è in grado di terminare la trascrizione a segnali $T_4$ e che, per un'efficiente terminazione, la polimerasi di lievito richiede serie di T più lunghe rispetto a quella dei vertebrati (Hamada et al., 2000) (Geiduschek and Tocchini-Valentini, 1988). Anche nei tDNA di mammifero la maggior parte degli oligo-(dT) si trova entro 40 bp in 3' rispetto alla sequenza codificante. Abbiamo quindi concluso che $T_5$ e $T_4$ costituiscono il minimo segnale di terminazione per Pol III nei lieviti e nei mammiferi, rispettivamente.

Limitando le ulteriori analisi alle prime 40 bp della sequenza fiancheggiante in 3', abbiamo selezionato, da ogni *set* di tDNA, quelli che possedevano una singolo T-*run* in quella regione; tali tDNA "a singolo terminatore" rappresentano il 56, 47, 69 e 67% di tutti i tDNA in *S. cerevisiae*, *S. pombe*, *H. sapiens* e *M. musculus*, rispettivamente. Abbiamo quindi calcolato, per ciascun organismo, la distribuzione di frequenza della lunghezza dei T-*run*, al fine di stabilire se la lunghezza degli oligo-(dT) nei diversi *set* di tDNA correla con differenze evolutivamente significative nel riconoscimento del terminatore da parte della RNA polimerasi III. I tDNA con più di un segnale entro le 40 bp in 3' sono stati esclusi in quanto la presenza di un potenziale terminatore "*back up*" potrebbe avere ridotto la pressione selettiva sui terminatori posti più a monte. Il risultato della nostra analisi è riportato in *Figura 4.1*.

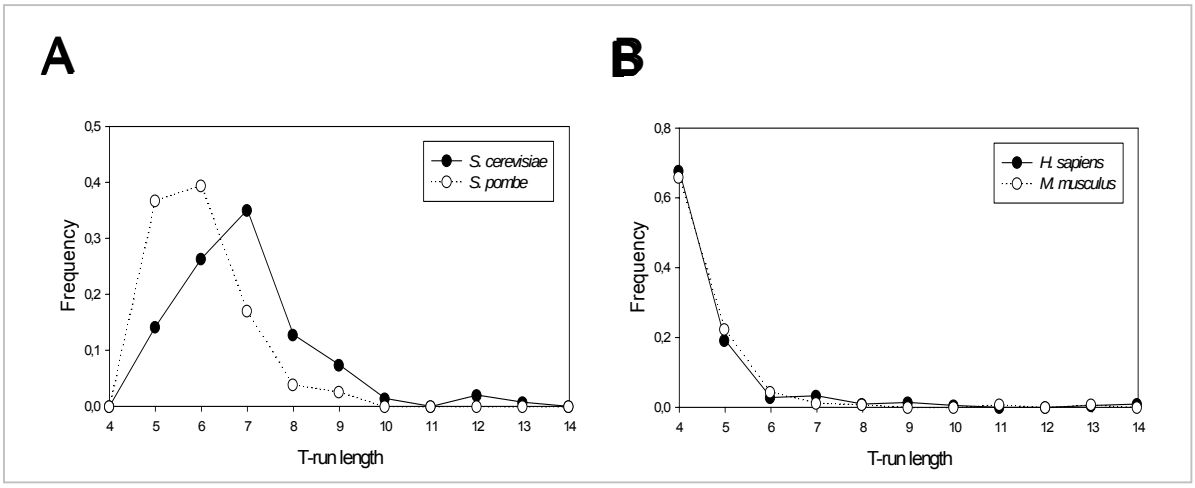



**Fig. 4.1: Distribuzione della lunghezza dei terminatori nei tDNA di diversi eucarioti.**
L'analisi della lunghezza dei terminatori poli-(dT) è stata condotta sul sottoinsieme di tDNA che presentano un singolo segnale di terminazione per Pol III entro le 40 bp che si trovano a valle della

regione codificante il tRNA maturo. (A): Distribuzione delle lunghezze dei T-*run* nei lieviti *S. cerevisiae* e *S. pombe*. (B): Distribuzione delle lunghezze dei T-*run* in *H. sapiens* e *M. musculus*.

In *S. cerevisiae* e *S. pombe*, le lunghezze più frequenti dei T-*run* sono 6 e 7, rispettivamente, e, come già detto, non si trovano terminatori più corti di $T_5$. I terminatori in *S. pombe* tendono ad essere di ~1 nucleotide (nt) più corti che in *S. cerevisiae*, in accordo con un'analisi precedente (Hamada et al., 2000). La distribuzione dei T-*run* nei mammiferi risulta completamente differente: $T_4$ è il terminatore più frequente ed i segnali più lunghi di $T_5$ sono rari.

Mentre la distribuzione della lunghezza dei terminatori nei lieviti è rappresentata da una curva a campana, la distribuzione osservata nei mammiferi è completamente diversa: se $T_4$ è il segnale più frequente, la presenza di terminatori più lunghi è da imputarsi semplicemente ad un evento casuale; per ogni lunghezza della serie di T (n), infatti, la probabilità di ottenere un terminatore di lunghezza (n+1), in assenza di una pressione selettiva, è del 25%, dal momento che quattro sono i nucleotidi che compongono il DNA; la distribuzione ottenuta per i mammiferi si avvicina molto a questa regola.

In *Saccharomyces cerevisiae*, 22 tDNA sono caratterizzati dalla presenza di un'unica serie $T_5$ come potenziale segnale di terminazione. Abbiamo pensato quindi che $T_5$ dovesse costituire un segnale efficiente per la terminazione da parte della RNA polimerasi III. Alcune osservazioni precedenti, tuttavia, avevano rivelato che 5 T consecutive nella regione in 3' dei tDNA funzionano soltanto in parte nell'indurre terminazione (Allison and Hall, 1985; Hamada et al., 2000).

Al fine di definire meglio le proprietà di terminazione della sequenza $T_5$, quindi, abbiamo analizzato 8 geni per tRNA di *S. cerevisiae* che portano $T_5$ come primo segnale di terminazione. Questi geni, insieme con le caratteristiche della loro sequenza in 3', sono riportati in *Figura 4.2, A*.

Due di questi geni (n$^o$ 1 e 2) sono classificati come tDNA "a singolo terminatore". Negli altri tDNA, invece, $T_5$ è seguito da un secondo terminatore *back up* localizzato da 3 a 20 bp più a valle. La presenza del T-*run* più a valle consente di monitorare facilmente eventuali polimerasi che oltrepassano il terminatore $T_5$, in quanto esse producono trascritti

più lunghi, terminati al *cluster* più a valle. Questo è vero anche per il tDNA 1, nel quale si trova un secondo T-*run* più lontano, in posizione +66 rispetto alla fine della sequenza codificante. Nel caso del tDNA 2, invece, il plasmide contenente il tDNA templato è stato linearizzato 121 bp a valle di $T_5$, in modo tale da rilevare l'eventuale *readthrough* come trascritti *run-off*. In modo simile, con la costruzione di templati *run-off*, abbiamo anche verificato che nessuna polimerasi in trascrizione oltrepassa il secondo T-*run*, ove presente; quelli visualizzati, pertanto, rappresentano la totalità dei trascritti prodotti da Pol III.

Questi geni templato sono stati utilizzati per programmare saggi di trascrizione *in vitro* in un sistema ricostituito di *S. cerevisiae* (vedi *Materiali e Metodi*). La trascrizione è stata condotta in presenza di una bassa concentrazione di UTP (25 μM) oppure di una concentrazione 10 volte più alta (250 μM), più vicina alla concentrazione intracellulare di UTP, che si stima essere nell'ordine del mM (Mathews, 1972; Reynolds et al., 1992).

Il risultato in *Figura 4.2, B* mostra che la forza di $T_5$ come terminatore varia in modo considerevole nei diversi contesti analizzati e che il superamento dell'elemento $T_5$ da parte della polimerasi è in alcuni casi fortemente favorito dalle più alte concentrazioni di UTP (vedi corsie 3 e 4, pannello di sinistra e di destra), in condizioni, cioè, che aumentano il tasso di polimerizzazione di Pol III e che pertanto diminuiscono la pausa dell'enzima sul $T_5$.

In sei degli otto geni testati, l'elemento $T_5$ ha mostrato di funzionare come un terminatore forte (corsie 1, 2 e 5-8). Al contrario, in due di quei tDNA (corsie 3 e 4), una frazione consistente delle polimerasi ha mostrato di oltrepassare l'elemento $T_5$ a favore del terminatore *back up* posto più a valle.

Osservando l'allineamento sul terminatore $T_5$ dei tDNA testati, abbiamo notato che entrambi gli elementi $T_5$ che si comportano come terminatori deboli sono seguiti dalla sequenza CT e da un secondo T-*run* più lungo, a poche bp di distanza. I sei elementi $T_5$ che funzionano come terminatori forti, invece, sono tutti seguiti immediatamente da A o G, con l'eccezione del tDNA 5. In quest'ultimo gene $T_5$ è seguito da C ed ha un'efficienza di terminazione un poco inferiore. Si noti l'elemento $T_5$ è fiancheggiato da sequenze a monte quasi del tutto identiche nei tre $tDNA^{Asn}$ in corsia 3, 6, e 7. La notevole differenza nella forza del terminatore tra il tDNA N(GTT)CR (corsia 3) e gli altri due geni per $tRNA^{Asn}$ (corsie 6 e 7) è quindi da attribuirsi al particolare contesto di sequenza che si trova a valle del terminatore.

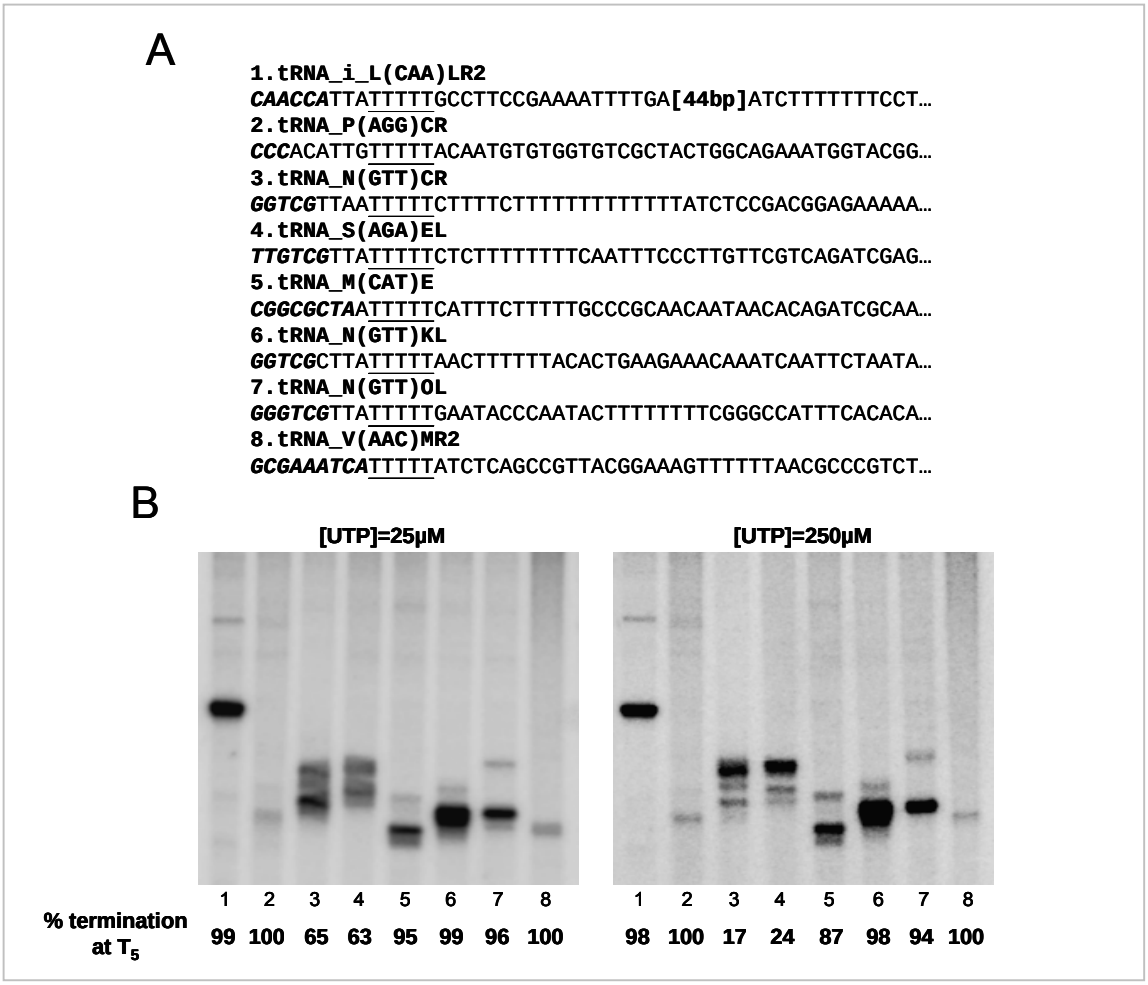



**Fig. 4.2: Trascrizione *in vitro* di tDNA di *S. cerevisiae* che presentano $T_5$ come primo segnale di terminazione.**
(A): Sequenza nucleotidica della regione del terminatore di otto geni per tRNA che presentano $T_5$ come primo segnale di terminazione. Le sequenze sono allineate sul terminatore $T_5$ (*sottolineato*); la sequenza che corrisponde al tRNA maturo è in grassetto e corsivo. Nel tDNA 1, ove indicato, è stata omessa una sequenza di 44 bp. (B): Trascrizione *in vitro* dei templati in *A*. Il numero di ciascuna corsia corrisponde a quello del relativo templato in *A*. Tutti i tDNA, eccetto #2, hanno un secondo terminatore più a valle che consente di monitorare le polimerasi che oltrepassano il $T_5$; per visualizzare eventuali trascritti *readthrough*, il templato #2 è stato linearizzato 121 bp a valle di $T_5$. Le reazioni di trascrizione sono state condotte in presenza di UTP 25µM (*sinistra*) oppure 250 µM (*destra*). Al di sotto di ogni corsia è riportata la percentuale di trascritti terminati al $T_5$.

Volendo investigare più in dettaglio l'effetto della sequenza a valle sulla forza del terminatore, abbiamo scelto di focalizzare l'attenzione su due dei tDNA che, in terminazione, avevano mostrato comportamenti ai due estremi: L(CAA)LR2 (corsia 1 in *Fig. 4.2*), il cui $T_5$ è un terminatore forte, e N(GTT)CR (corsia 3 in *Fig. 4.2*), il cui $T_5$ si

comporta come un terminatore debole ed è fiancheggiato da un secondo terminatore ad una distanza di 7 bp.

Abbiamo costruito e testato in trascrizione due costrutti ibridi (vedi *Fig. 4.3, A*): il templato L_3'N, nel quale la regione a valle di $T_5$ di L(CAA)LR2 è stata sostituita dalla sequenza $CT_4CT_{13}$, che si trova naturalmente a valle di $T_5$ in N(GTT)CR, ed il templato N_3'L in cui, simmetricamente, la sequenza a valle di $T_5$ di N(GTT)CR è stata sostituita dalle 20 bp che si trovano naturalmente a valle del terminatore in L(CAA)LR2 seguite da un terminatore *back up* $T_{10}$.

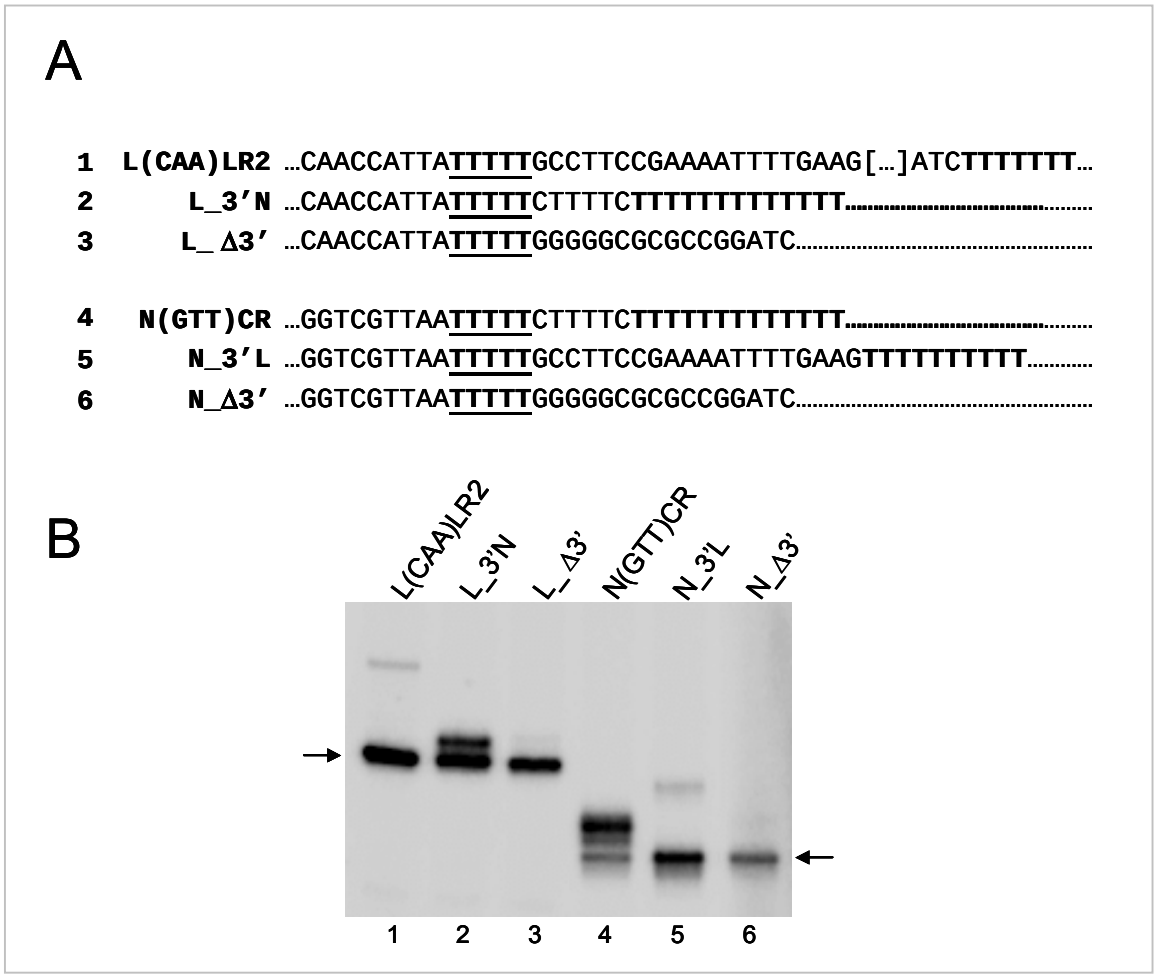



**Fig. 4.3: Proprietà di terminazione di varianti in 3' dei tDNA L(CAA)LR2 e N(GTT)CR.**
(A): Sequenza nucleotidica della regione del terminatore di diverse varianti in 3' di L(CAA)LR2 e N(GTT)CR. Il terminatore $T_5$ è in grassetto e sottolineato, l'eventuale secondo terminatore è in grassetto. I costrutti L_Δ3' e N_ Δ3' sono stati linearizzati 16 bp a valle di $T_5$. (B): Trascrizione *in vitro* dei templati riportati in *A*, condotta in presenza di UTP 250 µM. Le frecce indicano la posizione dei trascritti terminati al $T_5$ di L(CAA)LR2 (*sinistra*) e di N(GTT)CR (*destra*).

Come si può osservare in *Figura 4.3, B*, nel contesto L_3'N il terminatore $T_5$ risulta indebolito in modo non trascurabile, dal momento che, in presenza di una concentrazione di UTP 250 μM, mostra un'efficienza di terminazione soltanto del 55%, in confronto con l'efficienza del 98% che si osserva nel gene L(CAA)LR2 *wild-type* (cfr. corsie 1 e 2). Al contrario, in N_3'L la forza di $T_5$ come terminatore risulta fortemente aumentata (cfr. corsie 4 e 5).

Sempre a valle $T_5$, sui due *scaffold* L(CAA) e N(GTT) abbiamo posto una sequenza casuale, derivata dal plasmide nel quale sono stati clonati i tDNA templato, dando luogo ai due costrutti L_Δ3' e N_Δ3'. Dal momento che in questi ultimi due costrutti non è presente un terminatore *back up*, analogamente a quanto fatto in precedenza, il DNA templato è stato linearizzato a valle di $T_5$.

In trascrizione è risultato che, in L_Δ3', l'efficienza di terminazione alla sequenza $T_5$ resta praticamente inalterata rispetto a quanto osservato in L(CAA)LR2 *wild-type* (cfr. corsie 1 e 3); al contrario, la forza dell'elemento $T_5$ come terminatore risulta fortemente aumentata nel costrutto N_Δ3' rispetto al relativo gene *wild-type* (cfr. corsie 4 e 6).

Questo risultato suggerisce che $T_5$ *per se* costituisca, nei tDNA di *S. cerevisiae*, un efficiente terminatore trascrizionale e che la sequenza $CT_4CT_{13}$, presente naturalmente a valle del tDNA N(GTT)CR, sia, oppure contenga, un elemento capace di indebolire il terminatore quando posto a valle di esso. A supporto di questa conclusione si osservi che $T_5$ risulta essere un terminatore debole soltanto in due dei tDNA saggiati in *Figura 4.2*, e che in entrambi quei casi $T_5$ è seguito da sequenze simili e contenenti l'elemento CT ($CT_4CT_{13}$ in N(GTT)CR e $CTCT_8C$ in S(AGA)EL).

Come ulteriore controllo abbiamo confrontato i risultati di saggi di trascrizione in *single round* e *multiple round*. I rapporti tra trascritti terminati a $T_5$ e trascritti *readthrough* è risultato lo stesso in entrambe le condizioni; ciò significa che il riconoscimento del terminatore rimane inalterato durante cicli multipli di trascrizione ed esclude che il *pattern* di terminazione osservato possa derivare dal re-inzio facilitato della RNA polimerasi III.

Per capire meglio l'effetto di indebolimento del terminatore che abbiamo osservato, ci siamo concentrati sul costrutto chimerico L_3'N, la cui sequenza a valle di $T_5$, $CT_4CT_{13}$, abbiamo sottoposto ad analisi mutazionale.

In particolare abbiamo costruito mutanti in cui la sequenza $CT_4C$ è stata rimpiazzata da sequenze differenti, lasciando inalterato il secondo terminatore $T_{13}$ oppure raccorciandolo a $T_{10}$. Abbiamo infatti verificato che $T_{13}$ e $T_{10}$ possono essere considerati terminatori equivalenti in quanto capaci di indurre il 100% delle polimerasi in trascrizione a terminare.

Tutti i costrutti mutanti, schematizzati in *Figura 4.4, A*, sono stati testati *in vitro*, in un sistema di trascrizione ricostituito, in almeno tre esperimenti indipendenti. Il risultato di queste trascrizioni è riportato in *Figura 4.4, B*.

A livello quantitativo, il *readthrough* di $T_5$ da parte della RNA polimerasi III è stato inizialmente calcolato sulla base del rapporto tra trascritti terminati a $T_5$ e terminati a $T_{13}/T_{10}$. Per ogni esperimento, quindi, il superamento del terminatore $T_5$ osservato con il templato L_3'N (corsia 2) è stato posto uguale a 100, e l'effetto delle mutazioni è stato quantificato come *readthrough* relativo a L_3'N.

Questa analisi trascrizionale ci ha consentito di trarre alcune fondamentali conclusioni.

Quando abbiamo mutato i primi 3 nucleotidi a valle di $T_5$ (CTT) in GCC, in modo tale da rendere le 6 bp a valle di $T_5$ uguali a quelle che si trovano nel tDNA L(CAA)LR2 *wild-type*, l'effetto di indebolimento del terminatore è risultato fortemente diminuito (~5 volte) rispetto a quanto osservato in L_3'N (cfr. corsie 2 e 4). Questo suggerisce che, probabilmente, le prime basi che seguono il terminatore $T_5$ sono le maggiori responsabili dell'indebolimento dello stesso.

Dal momento che la sostituzione di CTT con GCC, che aumenta la stabilità predetta della doppia elica del DNA immediatamente a valle di $T_5$, è risultata influenzare pesantemente il potenziale di indebolimento del terminatore, ci siamo domandati se, rimpiazzando la sequenza a valle con altre sequenze che determinano una stabilità del *duplex* simile od inferiore, potevamo conservare od aumentare l'effetto di indebolimento, cioè il superamento trascrizionale di $T_5$. Una possibile spiegazione di tipo meccanicistico dell'effetto del DNA a valle del terminatore, infatti, ipotizza che la stabilità del *duplex* che deve essere aperto od il legame di un NTP ad una regione di DNA che deve essere denaturata possa influenzare la traslocazione della RNA polimerasi insieme con la bolla di trascrizione, e dunque anche la terminazione.

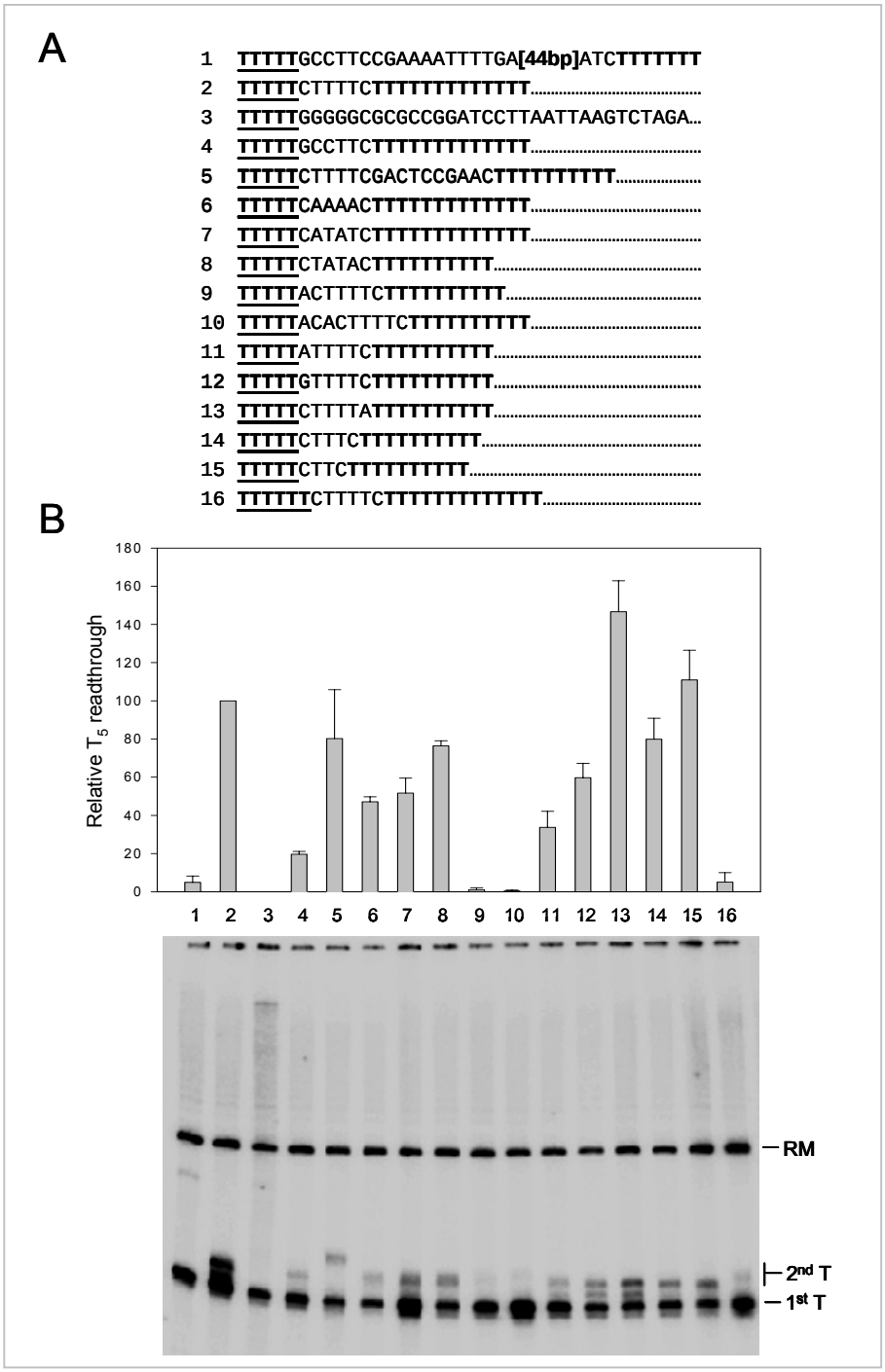



**Fig. 4.4: Proprietà del terminatore $T_5$ di L(CAA)LR2 fuso con diverse sequenze in 3'.**
(A): Sequenza nucleotidica della regione a valle del terminatore $T_5$ in diverse varianti di L(CAA)LR2. Il terminatore $T_5$ è riportato in grassetto e sottolineato, il terminatore *back up* in grassetto. I templati 1-3 sono stati utilizzati anche in *Fig. 4.3*. (B) *Sotto*: saggio di trascrizione *in vitro* dei templati elencati in *A*, identificati dai numeri sopra le corsie. Sulla destra è indicata la posizione dei trascritti terminati al primo o al secondo T-*run* e la posizione di un DNA di 415 nt usato come *recovery marker* (*RM*). *Sopra*: istogramma derivato dalla quantificazione di tre esperimenti indipendenti. Il superamento del terminatore $T_5$ da parte di Pol III è stato valutato sulla base del rapporto tra trascritti terminati al primo ed al secondo terminatore; in ogni esperimento, il *readthrough* osservato sul templato L_3'N (#*2*) è stato posto uguale a 100, e l'effetto delle mutazioni è stato calcolato come *readthrough* relativo rispetto a L_3'N. Le barre d'errore indicano l'errore *standard*.

Abbiamo quindi sostituito la serie $T_4$ della sequenza $CT_4C$ con TATA, ATAT oppure $A_4$. La stabilità predetta del *duplex*, secondo il sistema "*nearest neighbor thermodynamics*" risulta $A_4$=$T_4$>ATAT>TATA (Breslauer, KJ, 1986). Come si può osservare nelle corsie 6-8, una bassa stabilità del *duplex* nella regione immediatamente a valle di $T_5$, anche se in qualche misura contribuisce all'effetto di indebolimento, non è sufficiente a determinare in pieno tale effetto. Tra le tre sostituzioni, infatti, quella che causa il più alto grado di *readthrough* è anche quella che determina la più bassa $T_M$ nel DNA (sequenza TATA); la sequenza originale $T_4$, tuttavia, è in grado di indebolire maggiormente il terminatore rispetto a TATA. Questa conclusione è supportata anche dal paragone tra le proprietà di terminazione dei tDNA naturali indicati con i numeri 3 e 7 in *Figura 4.2*. Il terminatore $T_5$ del tDNA 7 è molto più forte rispetto a quello del tDNA 3, anche se la stabilità predicibile per le 6 bp immediatamente a valle ($CT_4C$ *vs* GAATAC) è molto simile.

In conclusione qualche ulteriore elemento, diverso dalla temperatura di *melting* del DNA, deve essere ricercato per dare piena spiegazione della funzionalità dell'elemento che indebolisce il terminatore.

Con un'altra serie di costrutti (n$^o$ 5, 9 e 10), abbiamo analizzato le richieste posizionali per l'indebolimento del terminatore $T_5$. Nel costrutto 5 la distanza tra le serie $T_4$ e $T_{10}$, che nel templato L_3'N è di solo 1 bp, è stata incrementata ad 11 bp tramite l'inserzione di sequenza *random*. In questo contesto l'indebolimento del terminatore è risultato molto simile a quello osservato in L_3'N. La sequenza $CT_4C$, pertanto, funziona come un elemento di indebolimento del terminatore autonomo, indipendente dal secondo terminatore più a valle. Questo risultato esclude anche che si possa pensare ad un meccanismo di tipo competitivo tra $T_5$ ed il terminatore forte posto più a valle.

L'indebolimento del terminatore, invece, è risultato fortemente compromesso in costrutti nei quali la distanza fra $T_5$ e la sequenza $CT_4C$ è stata aumentata dall'inserzione di tre o anche di una sola base (corsie 10 e 9). L'effetto sulla terminazione esercitato dall'elemento $CT_4C$, pertanto, dipende in modo critico dalla sua distanza dal terminatore $T_5$.

Con ulteriori costrutti mutanti ci siamo posti il fine di identificare l'elemento minimo in grado di determinare l'effetto sulla terminazione (costrutti 11-15).

Il residuo C che separa la serie $T_4$ dalla sequenza a valle $T_{10}$ può essere mutato senza perdita dell'effetto (corsia 13); al contrario, con questo mutante è stato osservato in modo riproducibile un aumento del *readthrough*. Nei templati 14 e 15, in cui la serie $T_4$ è stata

progressivamente ridotta, l'effetto di indebolimento è risultato mantenuto. La mutazione del residuo C che segue immediatamente $T_5$ con A è risultata ridurre l'efficienza di *readthrough* di circa il 65% rispetto a quanto osservato in L_3'N (cfr. corsie 2 e 11); mutando lo stesso residuo in G si è prodotto un più modesto ma tuttavia significativo effetto, una riduzione del *readthrough* di circa il 40% (cfr. corsie 2 e 12). Ne deriva che la presenza di un residuo C immediatamente a valle del terminatore $T_5$ è un requisito assolutamente necessario per ottenere un pieno effetto di indebolimento del terminatore.

I risultati ottenuti con questa analisi trascrizionale, insieme con i dati derivati dalla figura *4.2*, suggeriscono che il dinucleotide CT, posto immediatamente a valle di un terminatore $T_5$, costituisce il minimo elemento di sequenza capace di produrre un significativo indebolimento del terminatore. Questo elemento, tuttavia, risulta avere effetto soltanto su un terminatore costituito da 5 T: quando abbiamo incrementato la lunghezza del segnale di terminazione a $T_6$, infatti, l'effetto di indebolimento è stato perso del tutto (corsia 16).

Soltanto la forza di un terminatore limite ($T_5$ per *S. cerevisiae*), pertanto, può essere modulata dalla particolare sequenza CT posta a valle di esso.

E' possibile che la capacità di modulare la funzionalità di un terminatore $T_5$ non sia una caratteristica esclusiva dell'elemento di sequenza CT e che esistano altri brevi elementi con funzione simile. Nè l'analisi di sequenza sui tDNA naturali, né l'estensiva mutagenesi della cassetta di sequenza derivata dal gene N_(GTT)CR, tuttavia, ha consentito di individuare alcun elemento con simili caratteristiche.

I tratti oligo-(dT) che funzionano come terminatori per Pol III hanno la capacità di indurre sia la pausa della polimerasi che il rilascio dell'RNA trascritto. Queste due caratteristiche sono funzionalmente distinguibili e sono caratterizzate da diverse esigenze di sequenza (Campbell and Setzer, 1992) (Guffanti et al., 2004).

L'indebolimento del terminatore $T_5$ da parte della sequenza CT posta immediatamente a valle potrebbe coinvolgere un impedimento di entrambe oppure di una sola di queste abilità. Per esaminare più in profondità il meccanismo tramite il quale $T_5$ risulta indebolito, abbiamo analizzato la terminazione sui templati L_3'N (n$^o$ 2 in *Fig. 4.4*) e L_3'$GT_4$ (n$^o$ 12 in *Fig. 4.4*, nel quale la C che segue immediatamente $T_5$ è sostituita da G), in presenza di concentrazioni di CTP *standard* (500 μM) oppure 10 volte inferiori (50 μM). Abbassando

la concentrazione di CTP, poiché risulta ridotto il tasso di incorporazione del nucleotide successivo a $T_5$, si dovrebbe produrre un aumento della pausa di Pol III sul terminatore in L_3'N, ma non in L_3'$GT_4$. Allo stesso modo, l'abbassamento della concentrazione di GTP dovrebbe incrementare la pausa di Pol III in L_3'$GT_4$ ma non in L_3'N.

Come mostrato in *Figura 4.5*, in presenza di CTP 50 µM (corsia 5), il terminatore $T_5$ di L_3'N è risultato significativamente più forte che in presenza di CTP 500 µM (corsia 2), con una riduzione di ~3 volte della frazione di polimerasi che oltrepassa il terminatore (dal 45% al 14%). L'abbassamento della concentrazione di GTP, invece, non ha mostrato di avere effetto sul *readthrough* con questo templato (cfr. corsie 2 e 8). Con il templato L_3'$GT_4$, invece, l'abbassamento della concentrazione di CTP non ha influenzato il grado di superamento del terminatore $T_5$, mentre la frazione di polimerasi che facevano *readthrough* è risultata leggermente ma riproducibilmente ridotta a bassa concentrazione di GTP (dal 22% al 17%, cfr. corsie 3 e 9).

L'indebolimento del terminatore $T_5$, pertanto, può essere specificamente contrastato da condizioni di reazione che favoriscono la pausa della polimerasi a livello del $T_5$.

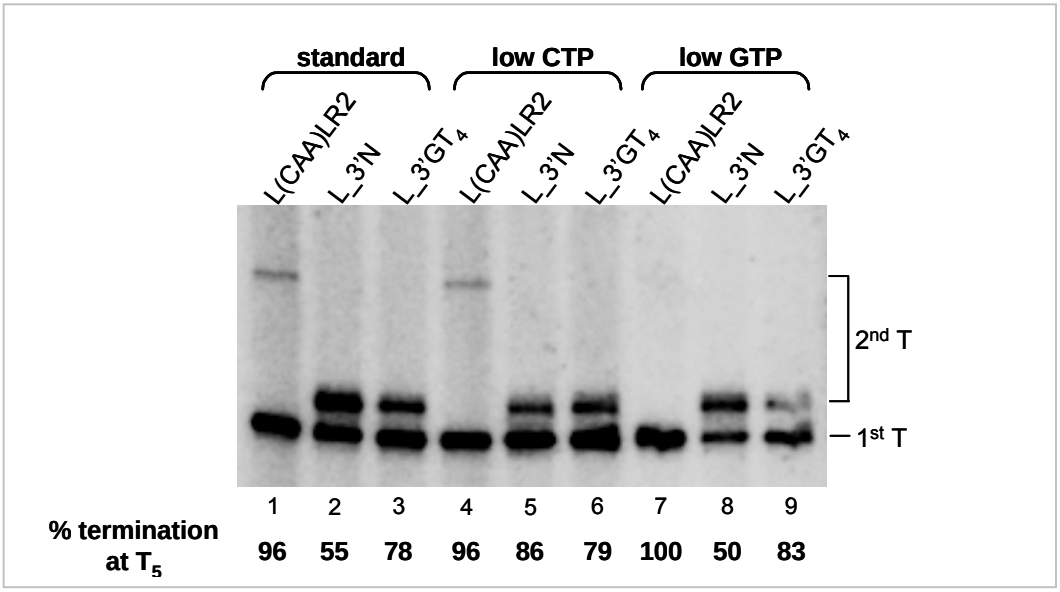



**Fig. 4.5: Il superamento del terminatore $T_5$ da parte di Pol III è influenzato dalla concentrazione del successivo nucleotide entrante.**
Saggio di trascrizione *in vitro* dei templati L(CAA)LR2, L_3'N e L_3'$GT_4$ (*#1*, *#2* e *#12* in *Fig. 4.4*) condotto alternativamente in presenza di una concentrazione *standard* di NTPs (500µM ATP, CTP, GTP; 250 µM UTP; corsie 1-3), di una bassa concentrazione di CTP (50 µM; corsie 4-6) o di una bassa concentrazione di GTP (50 µM; corsie 7-9). La posizione dei trascritti terminati al primo o al secondo T-*run* è indicata sulla destra. Sotto ad ogni corsia è riportata la percentuale di trascritti terminati al $T_5$.

Riguardo al rilascio del trascritto, abbiamo condotto saggi di trascrizione in cui abbiamo variato due condizioni che sono note influenzare quella fase nei terminatori rho-indipendenti dei batteri (Reynolds et al., 1992): abbiamo testato la sensibilità del terminatore $T_5$ a variazioni nella concentrazione di KCl (da 50 a 150 mM) e nella concentrazione di ioni $Mg^{++}$ (da 3 a 10 mM). I risultati di questi saggi non hanno rivelato alcuna variazione significativa nell'efficienza di terminazione; è dunque improbabile che gli effetti osservati possano dipendere dalla fase di rilascio dell'RNA trascritto.

Nel loro insieme, i nostri dati indicano che l'effetto dovuto alla sequenza a valle non dipende né dalla temperatura di *melting* del DNA né dal rilascio del trascritto, ma mostrano che l'elemento indebolitore di $T_5$ risulta alterare la pausa di Pol III a livello del terminatore.

E' possibile che l'elemento CT possa incrementare la processività di Pol III alla posizione immediatamente a monte attraverso un meccanismo che coinvolge il legame dei corrispondenti NTPs ad un sito allosterico presente sulla polimerasi, come recentemente proposto per spiegare gli effetti della sequenza a valle sull'allungamento sia da parte della RNA polimerasi di *E. coli* (Holmes and Erie, 2003) che della RNA polimerasi II di lievito (Nedialkov et al., 2003).

I dati sinora riportati hanno rivelato che il contesto di sequenza a valle è in grado di influenzare in modo non irrilevante la terminazione Pol III-dipendente, in modo indipendente dalla specifica sequenza tDNA che precede l'elemento $T_5$.

Se una sequenza a valle del terminatore può esercitare un simile effetto, a maggior ragione si può pensare che la regione a monte dello stesso, che è trascritta, possa influenzare la polimerasi in fase di terminazione.

Per verificare se esistono influenze dovute alla sequenza in 5' rispetto al terminatore, abbiamo generato un costrutto (SCR1Δ_3'L) in cui il terminatore $T_5$ + 20 bp di sequenza in 3' derivate dal tDNA L(CAA)LR2 + $T_{10}$ *back up* sono stati fusi alle prime 90 bp del gene *SCR1*.

*SCR1* è il gene di lievito che codifica per l'RNA 7SL, facente parte della particella di riconoscimento del segnale. Al di là della presenza di una A-*box* e di una B-*box*, che consentono di ottenere un'efficiente trascrizione *in vitro* (Dieci et al., 2002), questa

porzione del gene *SCR1* è sostanzialmente una sequenza non correlata con quella dei geni per tRNA.

Come mostrato in *Figura 4.6, A*, l'inserzione di un terminatore forte $T_7$ in posizione +90 è risultata in un templato (SCR1Δ_T7) la cui trascrizione è terminata efficientemente al $T_7$ e genera un trascritto lungo 90 nt; non si osservano, con questo templato, trascritti dovuti a *readthrough* (corsie 1, 3 e 5). Sorprendentemente, invece, il terminatore $T_5$ + 3' di L(CAA)LR2, che generava un'efficienza di terminazione del 97% nel templato originale, è risultato comportarsi come un terminatore debole nel costrutto SCR1Δ_3'L, consentendo al 40% circa delle polimerasi di terminare al $T_{10}$ posto più a valle in presenza di UTP 250 μM (corsia 2). Trascritti *readthrough* sono risultati visibili anche a più basse concentrazioni di UTP, anche se in proporzioni progressivamente inferiori con UTP 25 μM (corsia 4) e 2,5 μM (corsia 6). Questi dati suggeriscono che la sequenza a monte, derivata da *SCR1,* riduca la terminazione a livello del $T_5$ influenzando la pausa della RNA polimerasi.

E' possibile che l'effetto della sequenza a monte sulla terminazione sia dovuto ad un'interazione diretta polimerasi-DNA; non si può escludere, tuttavia, che anche l'RNA trascritto giochi un ruolo attivo nell'influenzare il processo di terminazione, in analogia con quanto si osserva nei terminatori rho-indipendenti dei batteri (von Hippel, 1998) (Nudler and Gottesman, 2002).

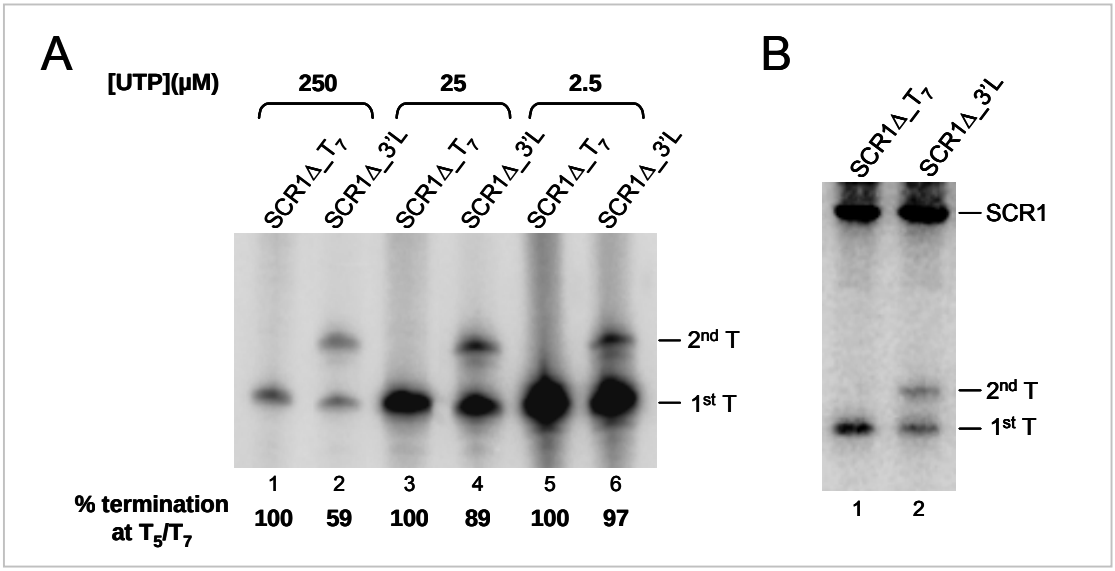



**Fig. 4.6: Effetto della sequenza a monte sulla forza del terminatore $T_5$ *in vitro* e *in vivo*.**
(A): Trascrizione *in vitro* dei templati SCR1Δ_T7 e SCR1Δ_3'L condotta in presenza di UTP 250 μM (corsie 1-2), 25 μM (corsie 3-4) o 2,5 μM (corsie 5-6). La posizione dei trascritti terminati al primo o

al secondo T-*run* è indcata sulla destra. (B): Profilo di espressione dei templati SCR1Δ_T7 e SCR1Δ_3'L *in vivo*. L'RNA totale estratto da cellule di lievito trasformate con i due templati, indicati sopra le corsie, stato ibridato con uno specifico oligonucleotide marcato radioattivamente che si ibrida entro le prime 90 bp dell'RNA prodotto da *SCR1*. Sulla destra è indicata la posizione dei trascritti derivati dalla copia cromosomica del gene *SCR1* e dai due templati plasmidici SCR1-derivati (*1st T, 2nd T*).

E' stato possibile monitorare la terminazione della trascrizione di SCR1Δ_3'L anche *in vivo*, dal momento che la distanza di 20 bp tra l'elemento $T_5$ ed il terminatore a valle $T_{10}$ consente la produzione di trascritti che differiscono tra loro per la taglia e che sono caratterizzati da una stabilità paragonabile *in vivo*. Lo stesso tipo di analisi non è invece stato possibile con i geni per tRNA perché l'elevata instabilità dei trascritti allungati in 3' rispetto a quelli maturati, insieme con il fatto che esistono diverse copie della maggior parte dei geni per tRNA (Percudani et al., 1997), rende piuttosto complicata la *detection* dei prodotti di terminazione su questi geni.

I templati SCR1Δ_T7 e SCR1Δ_3'L sono stati inseriti nel vettore ad alto numero di copie YEp352 e trasformati nel ceppo di *S. cerevisiae* YPH500. L'RNA totale estratto da cellule cresciute in fase logaritmica è stato quindi soggetto ad analisi *Northern* utilizzando una specifica sonda antisenso per SCR1.

Come mostrato in *Figura 4.6, B*, il 55% dei trascritti derivati dal costrutto SCR1Δ_3'L è terminato alla sequenza $T_5$, mentre il 45% dei trascritti è di 20 nt più lungo e deriva dal superamento del terminatore $T_5$ (corsia 2); al confronto, il terminatore $T_7$ nel contesto SCR1Δ_T7 è risultato pienamente efficiente anche *in vivo* (corsia 1).

Questo significa che le stesse proprietà di riconoscimento del terminatore viste *in vitro* si ritrovano anche *in vivo* e non sono pertanto un artificio dovuto all'uso di un sistema di trascrizione non naturale.

L'effetto del contesto di sequenza visto fino a questo momento potrebbe in linea di principio essere mediato da TFIIIC, fattore di trascrizione che generalmente copre una porzione notevole delle unità trascrizionali Pol III, che si estende fino alla regione del terminatore nei geni per tRNA (Klemenz et al., 1982) (Camper et al., 1985).

E' stato proposto in studi precedenti che attività legate a TFIIIC possano influenzare la terminazione Pol III-dipendente in uomo (Wang et al., 2000) e che, in lievito, TFIIIC faciliti il re-inizio trascrizionale da parte di Pol III sulle unità trascrizionali più lunghe (Ferrari et al., 2004). Il ruolo di TFIIIC, pertanto, non è limitato alla sola fase di assemblaggio del complesso di trascrizione.

Per verificare se TFIIIC influenza la terminazione al $T_5$ nei diversi contesti analizzati, abbiamo modificato, in modo tale da consentirne la trascrizione TFIIIC-indipendente, alcuni dei templati nei quali il terminatore risultava indebolito. A tal fine, le porzioni trascritte di L(CAA)LR2, L_3'N e SCR1Δ_3'L sono state fuse con la porzione in 5' del gene *SNR6* di *S. cerevisiae*, che è nota supportare l'assemblaggio di TFIIIB mediato dalla TATA-*box* e l'inizio della trascrizione in assenza di TFIIIC (Ferrari et al., 2004) (Moenne et al., 1990).

Come mostrato in *Figura 4.7, A*, anche in assenza di TFIIIC si osserva un pronunciato indebolimento del terminatore $T_5$ da parte sia della sequenza a valle 3'N che della sequenza a monte SCR1Δ (cfr. corsie 5 e 6 con corsia 4); questo dimostra che gli effetti di indebolimento non dipendono da questo fattore di trascrizione. Un confronto quantitativo fra l'efficienza di terminazione in presenza (corsie 1-3) o in assenza (corsie 4-6) di TFIIIC suggerisce piuttosto che TFIIIC favorisca il riconoscimento del terminatore da parte di Pol III, un effetto specialmente evidente con il costrutto 5'U6_L_3'N (cfr. corsie 2 e 5); questa osservazione può essere spiegata postulando che TFIIIC, legato al DNA, possa leggermente diminuire il tasso di allungamento della RNA polimerasi in trascrizione.

Per testare in modo più diretto se il distinguere tra segnali di terminazione $T_5$ forti oppure deboli è una proprietà intrinseca di Pol III, abbiamo sfruttato la riportata capacità di questo enzima di iniziare la trascrizione su templati lineari che contengono un'estremità 3'-protrudente (Bardeleben et al., 1994).

Abbiamo preparato versioni linearizzate dei templati L(CAA)LR2, L_3'N e SCR1Δ_3'L (vedi *Materiali e Metodi*), e condotto la trascrizione in assenza di fattori di trascrizione. L'inizio della trascrizione è stato fatto occorrere all'estremità 3'-protrudente generata per digestione con l'enzima di restrizione SacI, in una posizione che si trova 5 bp a monte rispetto al naturale sito di inizio della trascrizione.

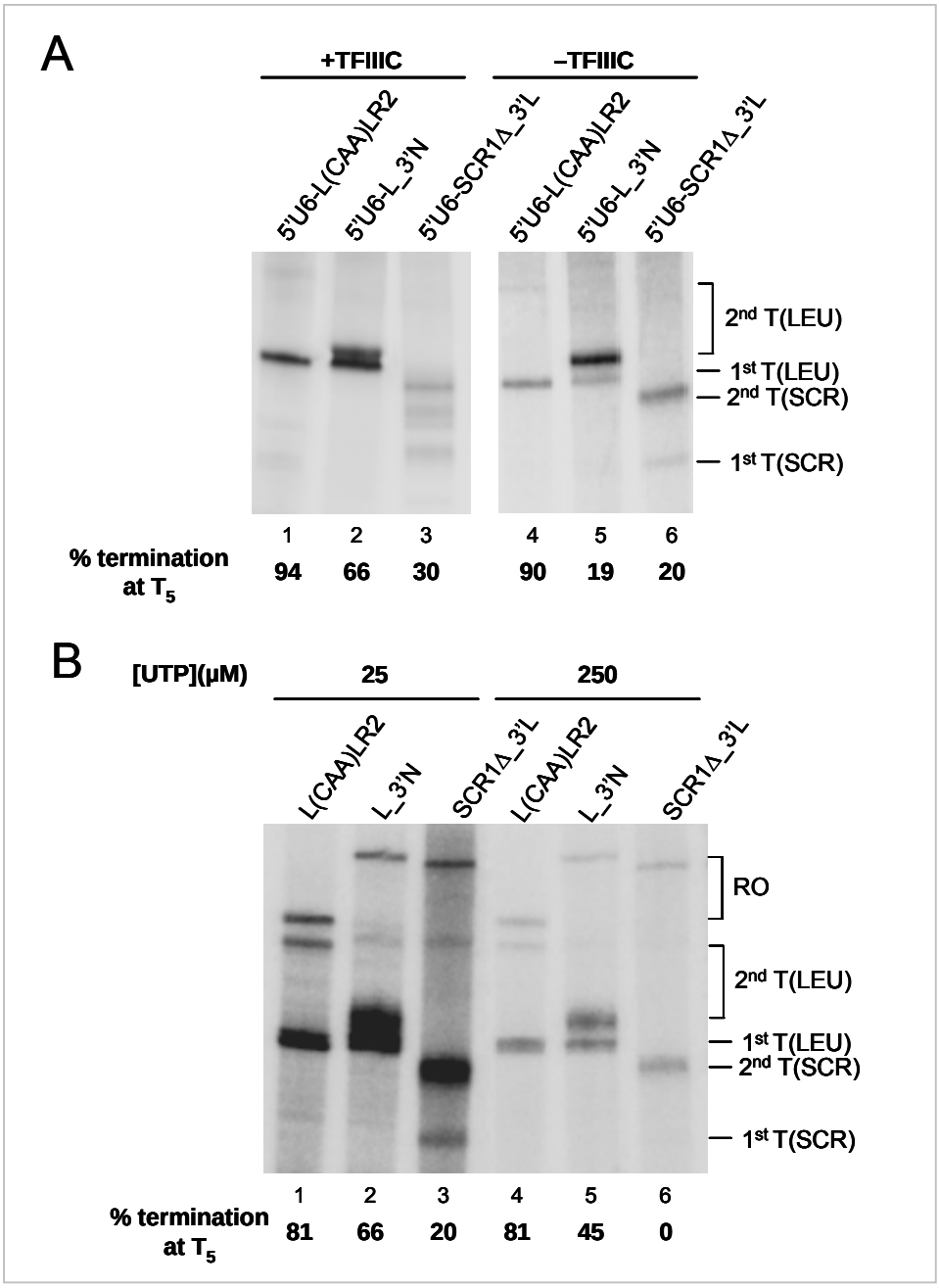



**Fig. 4.7: Il superamento del terminatore $T_5$ è una proprietà intrinseca di Pol III.**
(A): Trascrizione *in vitro* dei costrutti derivati dalla fusione della sequenza in 5' di *SNR6* con L(CAA)LR2, L_3'N e SCR1Δ_3'L. Le reazioni sono state assemblate in presenza di UTP 250 μM e alternativamente in presenza (corsie 1-3) o in assenza (corsie 4-6) di TFIIIC. Sotto ad ogni corsia è riportata la percentuale di trascritti terminati al $T_5$. Tutte le corsie mostrate fanno parte di uno stesso gel di trascrizione; le corsie 4-6, tuttavia, derivano da un'esposizione più lunga resasi necessaria per compensare il più basso *output* trascrizionale che si ottiene in assenza di TFIIIC. Sulla destra è indicata la posizione dei trascritti terminati al primo o al secondo T-*run* del templato SCR1-derivato (*1st T(SCR), 2nd T(SCR)*) e dei templati derivati da L(CAA)LR2 (*1st T(LEU), 2nd T(LEU)*). (B): Trascrizione *in vitro* di frammenti di restrizione, contenenti i templati indicati sopra le corsie, con un'estremità 3'-protrudente. Le reazioni sono state condotte in presenza della sola Pol III purificata e di UTP alternativamente 25 μM (corsie 1-3) o 250 μM (corsie 4-6). Sotto ogni corsia è riportata la

percentuale di trascritti terminati al $T_5$. Sulla destra come in *A* è indicata la posizione dei trascritti terminati al primo o al secondo T-*run*, insieme con la posizione dei trascritti terminati per *run-off* (*RO*).

Come mostrato in *Figura 4.7, B*, ed in accordo con uno studio precedente (Bardeleben et al., 1994), la RNA polimerasi III purificata trascrive specificamente ed efficientemente questi templati.

Come con i templati originali, anche in questo caso Pol III è risultata terminare la trascrizione sulla stessa regione $T_5$ + terminatore *back up* (localizzato, rispettivamente, 67, 7 e 21 bp a valle in L(CAA)LR2, L_3'N e SCR1Δ_3'L). Tuttavia, una piccola percentuale dei complessi di trascrizione (il 10% al massimo) ha mostrato di fallire il riconoscimento sia del terminatore $T_5$ che del terminatore più a valle, dando luogo a trascritti *run-off*. Questa piccola quantità di trascritti *run-off* deriva probabilmente da complessi di trascrizione che non sono in grado di spiazzare l'RNA dagli ibridi DNA/RNA; è stato riportato che un simile artefatto dipende dal fare iniziare la trascrizione su templati con estremità 3'-protrudenti o contenenti una coda poli-(dC) (Bardeleben et al., 1994; Campbell and Setzer, 1992). Per questa ragione i trascritti *run-off* non sono stati considerati ai fini della quantificazione.

Il segnale $T_5$ nel contesto *wild-type* di L(CAA)LR2 ha mostrato di funzionare come un terminatore piuttosto forte sia con UTP 250 μM che 25 μM, con soltanto il 19% dei complessi di trascrizione che oltrepassa il segnale $T_5$ per terminare più a valle (corsie 1 e 4). Al contrario, il segnale di terminazione $T_5$ è risultato ancora indebolito in modo considerevole nei contesti L_3'N e SCR1Δ_3'L, sia con UTP 25 (corsie 2 e 3) che 250 (corsie 5 e 6) μM. L'effetto di indebolimento è risultato drammatico nel caso di SCR1Δ_3'L, nel quale non è stato possibile visualizzare trascritti terminati a $T_5$ con una concentrazione di UTP 250 μM (corsia 6).

L'indebolimento del terminatore $T_5$ da parte sia della sequenza a valle che della sequenza a monte, dunque, riflette proprietà di riconoscimento del terminatore intrinseche di Pol III.

Confrontando il grado di indebolimento osservato nelle stesse condizioni in assenza di fattori di trascrizione o nel sistema di trascrizione completo, infine, si può notare che l'effetto di indebolimento risulta aumentato nel sistema privo di fattori. Questo aumento

del *readthrough*, almeno in parte, può essere attribuito all'assenza di TFIIIC (vedi osservazioni precedenti relative a *Fig. 4.7, A*).

Studi recenti di localizzazione su scala genomica del macchinario trascrizionale Pol III hanno identificato il gene *SNR52*, codificante uno snoRNA, come nuovo gene di classe III (Harismendy et al., 2003) (Roberts et al., 2003) (Moqtaderi and Struhl, 2004).

Analisi *Northern* e di *primer extension* hanno rivelato che il trascritto primario di *SNR52* è un RNA precursore di ~250 nt dal quale una sequenza *leader* di ~160 nt viene rimossa per generare lo snoRNA maturo, di ~90 nt; l'analisi di sequenza ha rivelato la presenza di putativi elementi A- e B-*box* entro la sequenza *leader* di *SNR52* (Harismendy et al., 2003) (Roberts et al., 2003).

Una caratteristica particolare di questa unità trascrizionale è il fatto che una serie di 6 T consecutive si trova all'interno della regione trascritta, poche bp a valle della A-*box* (vedi *Fig. 4.8*). Come già detto, la sequenza $T_6$ costituisce un segnale di terminazione efficiente per la RNA polimerasi III di lievito, come abbiamo verificato anche trascrivendo diversi tDNA naturali con un segnale di terminazione $T_6$.

```
      1          11         21         31         41         51
      |          |          |          |          |          |
  1 CTTTGAAAAG ATAATGTATG ATTATGCTTT CACTCATATT TATACAGAAA CTTGATGTTT
 61 TCTTTCGAGT ATATACAAGG TGATTACATG TACGTTTGAA GTACAACTCT AGATTTTGTA
121 GTGCCCTCTT GGGCTAGCGG TAAAGGTGCG CATTTTTTCA CACCCTACAA TGTTCTGTTC
181 AAAAGATTTT GGTCAAACGC TGTAGAAGTG AAAGTTGGTG CGCATGTTTC GGCGTTCGAA
241 ACTTCTCCGC AGTGAAAGAT AAATGATCTA CTATGATGAA TGACATTAGC GTGAACAATC
301 TCTGATACAA AATCGAAAGA TTTTAGGATT AGAAAAACTT ATGTTGCCTT CCTTCTGAAA
361 TTTATTTTTT GTCACTATTG TTATGTAAAA TGCCACCTCT GACAGTATGG AACGC
```

**Fig. 4.8: Sequenza nucleotidica dell'unità trascrizionale** *SNR52*.
La sequenza corrispondente al trascritto primario è riportata in grassetto, quella corrispondente allo snoRNA maturo in grassetto sottolineato. In verde sono evidenziate la A- e la B-*box*, in rosso le due serie $T_6$, localizzate la prima entro la regione trascritta, la seconda in posizione canonica, alla fine dell'unità trascrizionale.

La trascrizione *in vitro* del gene *SNR52*, mostrata in *Figura 4.9*, in presenza di UTP 25 µM produce due trascritti fondamentali. Quello più lungo è stato identificato come il trascritto

primario sulla base della sua taglia (250 nt); l'RNA più corto ha invece la taglia attesa per un trascritto terminato all'elemento $T_6$ intragenico.

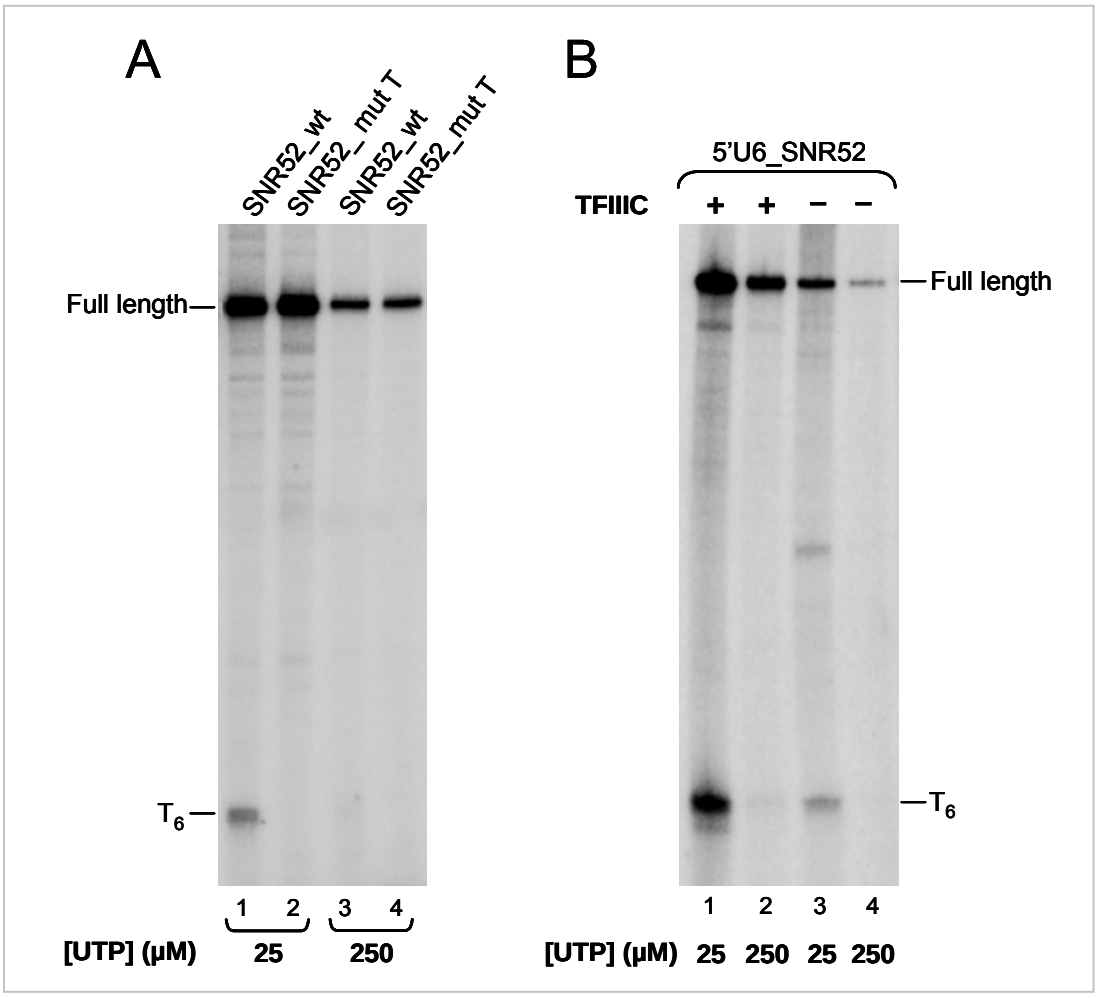



**Fig. 4.9: Superamento della serie $T_6$ entro l'unità trascrizionale naturale** *SNR52*.
(A): Trascrizione *in vitro* del gene *SNR52* e della sua variante SNR52_mutT (nella quale è distrutta la serie $T_6$ interna) condotta in presenza di UTP 25 µM (corsie 1-2) o 250 µM (corsie 3-4). Sulla sinistra è indicata la posizione del trascritto primario *full-length* e del trascritto terminato alla serie $T_6$. (B): Trascrizione *in vitro* del templato 5'U6-SNR52 condotta alternativamente in presenza (corsie 1-2) o in assenza (corsie 3-4) di TFIIIC, con UTP 25 µM (corsie 1 e 3) o 250 µM (corsie 2 e 4). Sulla destra, la posizione dei diversi trascritti è riportata come in *A*.

Infatti, la trascrizione di un templato *SNR52* mutante, in cui la serie $T_6$ è stata distrutta, ha prodotto più alti livelli del trascritto *full-length*, mentre il trascritto più corto è scomparso (cfr. corsie 1 e 2 in *Fig. 4.9, A*).

La quantificazione dei trascritti in corsia 1, corretta per il numero di U incorporate, ha portato alla stima che almeno il 70% delle polimerasi in trascrizione oltrepassa l'elemento $T_6$ in *SNR52*. $T_6$ dunque, nel contesto *SNR52*, si comporta in modo anomalo, come un terminatore debole, anche in condizioni (bassa concentrazione di UTP) che favoriscono la terminazione.

L'elemento $T_6$ scarsamente attivo di *SNR52* si trova appena a valle rispetto alla putativa A-*box*, all'interno della regione trascritta. Dal momento che TFIIIC contatta questa regione durante la trascrizione, questo fattore potrebbe essere richiesto affinché la RNA polimerasi oltrepassi l'elemento $T_6$. Per verificare una simile ipotesi, l'unità trascrizionale *SNR52* è stata fusa con la sequenza in 5' di *SNR6*, per consentirne la trascrizione in assenza di TFIIIC.

Come mostrato in *Figura 4.9, B*, la presenza o l'assenza di TFIIIC su *SNR52* non ha effetto sul comportamento di Pol III in fase di terminazione: in ogni caso $T_6$ si comporta come un elemento di terminazione altamente inefficiente, ad una concentrazione di UTP 250 μM (cfr. corsie 2 e 4).

La mancata terminazione sulla serie $T_6$ all'interno dell'unità trascrizionale *SNR52* ha un evidente significato biologico, dal momento che la regione codificante lo snoRNA si trova completamente a valle dell'elemento $T_6$. E' interessante notare che tale sequenza $T_6$, potenzialmente dannosa per la trascrizione del gene *SNR52*, in realtà si trova in posizione analoga nei genomi di almeno altri due *Saccharomyces* (*Saccharomyces paradoxus* e *Saccharomyces mikatae*) (Kellis et al., 2003); questo significa che l'elemento $T_6$ potrebbe avere qualche ruolo rilevante per la funzionalità dello snoRNA. In linea di principio, l'elemento $T_6$ potrebbe influenzare l'accessibilità e/o l'attività trascrizionale di *SNR52 in vivo*, o potrebbe giocare un ruolo nel processamento del trascritto primario.

Meno evidente è il ruolo biologico dell'indebolimento del terminatore $T_5$ nel caso dei geni per tRNA. Come risultato del *readthrough*, Pol III sintetizza tRNA precursori con estremità 3' più allungate le quali, in linea di principio, potrebbero stabilizzare il tRNA precursore, difendendolo dall'azione di 3'-esonucleasi.

Negli organismi eucariotici, dal lievito all'uomo, anche il legame della proteina La risulta in una protezione delle estremità 3' (vedi (Wolin and Cedervall, 2002) per review). La

proteina La, favorendo l'assunzione di una corretta conformazione, può influenzare la stabilità del pre-tRNA, il suo processamento, la localizzazione del tRNA maturo ed il suo caricamento (Wolin and Cedervall, 2002) (Maraia and Intine, 2002) (Intine et al., 2002) (Chakshusmathi et al., 2003). L'estensione in 3' causata dal *readthrough* potrebbe influenzare ciascuno di questi processi, direttamente o tramite la proteina La. A supporto di questa visione, è stato recentemente mostrato che un incremento nella lunghezza dell'oligo-(U) 3'-terminale dei pre-tRNA è in grado di influenzare il legame di La e la maturazione dei tRNA (Huang et al., 2005).

La possibilità di superare un oligo-(dT) da parte della RNA polimerasi III in trascrizione ha un'importanza ancora maggiore nei mammiferi. Come visto, infatti, nei genomi di topo e di uomo $T_4$ è di gran lunga il segnale di terminazione più frequente, ed in molti casi esso è l'unico segnale presente nella regione in 3' dei tDNA. Poiché serie di 4 T consecutive sono piuttosto frequenti all'interno delle regioni trascritte di questi geni, devono necessariamente esistere meccanismi o caratteristiche del contesto di sequenza capaci di sfavorire fortemente la terminazione su queste serie $T_4$. Tramite un'analisi statistica delle regioni fiancheggianti le serie $T_4$ intrageniche, tuttavia, non abbiamo individuato alcun elemento di sequenza conservato.

Se si pone il problema della presenza di potenziali terminatori intragenici, è lecito domandarsi come mai i terminatori per la RNA polimerasi III si sono evoluti verso serie di T progressivamente più corte. Una possibile spiegazione è che in questo modo, in presenza di terminatori più corti e quindi più deboli, sia lasciato più spazio per l'azione di fattori accessori, come NF1, che influenzano la terminazione da parte della RNA polimerasi III umana (Wang et al., 2000). Alternativamente, la predominanza di serie di T molto corte a valle dei tDNA di mammifero potrebbe riflettere l'importanza di avere estremità poli-(U) in 3' molto corte per il destino post-trascrizionale dei tRNA precursori.

## 4.2 INIZIO DELLA TRASCRIZIONE: SEQUENZA A MONTE DEI GENI PER TRNA ED EFFETTI DELLA PROTEINA NHP6

### *4.2.1 Sequenza a monte dei geni per tRNA*

Parte dei risultati riportati in questa sezione sono stati pubblicati nell'articolo riportato in appendice (Giuliodori et al., 2003).

Riguardo all'assemblaggio del fattore di trascrizione TFIIIB sui geni di classe III, in letteratura si trovano due serie di osservazioni apparentemente in contrasto: da una parte l'interazione TFIIIB-DNA, che coinvolge le prime 50 bp a monte del sito di inizio della trascrizione, è stata mostrata essere largamente sequenza-indipendente (Geiduschek and Kassavetis, 2001); dall'altra c'è un certo corpo di evidenze secondo il quale in lievito, così come in altri eucarioti, la regione in 5' di un particolare *subset* di geni per tRNA contiene elementi di sequenza conservati che danno un contributo alla funzionalità del promotore (Dieci et al., 2000; Hamada et al., 2001; Raymond et al., 1985).

Studi *in vitro* e *in vivo* su alcuni geni per tRNA di *Saccharomyces cerevisiae*, per esempio, hanno mostrato che mutazioni nella regione a monte, che è nota essere occupata da TFIIIB, non risultano mai in una completa abolizione della trascrizione ma spesso determinano una notevole diminuzione dell'efficienza di trascrizione (Shaw and Olson, 1984) (Joazeiro et al., 1996) (Lee et al., 1997). Questi dati sono a supporto dell'idea secondo la quale esistono alcuni contatti sequenza-specifici critici per il contatto fra TFIIIB e le regioni in 5' dei tDNA. In effetti pare piuttosto sorprendente che tali regioni a monte dei tDNA e degli altri geni di classe III non siano state soggette ad alcun tipo di pressione selettiva verso l'ottimizzazione del legame di TFIIIB, visto anche il notevole *output* trascrizionale della RNA polimerasi III.

Per dare ragione di queste osservazioni, abbiamo ipotizzato che gli elementi conservati nelle regioni in 5' non fossero stati identificati a causa della loro natura degenerata, che li renderebbe non visualizzabili tramite semplici allineamenti di sequenza, e della loro distribuzione eterogenea fra i diversi geni per tRNA.

Partendo da questa ipotesi abbiamo scelto di portare avanti un'analisi di tipo genomica sulle regioni in 5' dei geni per tRNA.

Il genoma di *S. cerevisiae* contiene 274 geni per tRNA, localizzati su tutti i 16 cromosomi (Percudani et al., 1997) (Hani and Feldmann, 1998); questo *set* ridondante di geni codifica per 42 specie di tRNA con diversa specificità di codone. Il numero di copie per ciascuna specie di tRNA varia da 1 a 16 e correla piuttosto bene sia con la frequenza dello specifico codone, sia con l'abbondanza intracellulare del particolare tRNA (Percudani et al., 1997).

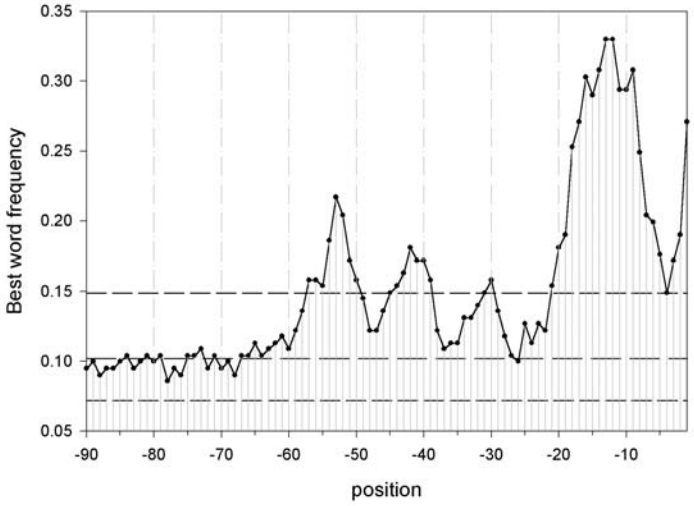
Poiché le A- e B-*box* hanno localizzazione intragenica, i geni che codificano per la stessa specie di tRNA condividono gli stessi elementi promotori, che legano TFIIIC, ma sono fiancheggiati da sequenze in 5', che interagiscono con TFIIIB, completamente differenti.

I tDNA di lievito spesso si trovano associati a retrotrasposoni Ty nel loro 5' (Hani and Feldmann, 1998). Dal momento che la conservazione in sequenza di questi elementi avrebbe potuto pregiudicare la nostra ricerca di caratteristiche conservate legate alla trascrizione, abbiamo escluso dalla nostra analisi 53 tDNA caratterizzati dalla presenza, entro le prime 100 basi a monte della sequenza codificante il tRNA maturo, di un retroelemento *sigma*, *delta* o *tau*.

Le regioni in 5' dei rimanenti 221 tDNA (fino a 90 bp a monte della posizione che corrisponde al primo nucleotide del tRNA maturo), sono state allineate ed analizzate per mezzo di un algoritmo disegnato specificamente che consente di ricercare la presenza di parole di n lettere entro date finestre. Cercando parole di 6 lettere entro finestre di 9 bp, consentendo 1 *mismatch*, abbiamo evidenziato quattro regioni che presentano parole di 6 lettere over-rappresentate, cioè parole che ricorrono con una frequenza sufficientemente alta da differenziarsi dal rumore locale. Una rappresentazione grafica dei risultati di questa analisi è mostrata in *Figura 4.10, A*.

La regione a monte della posizione –60 è caratterizzata da una completa casualità di sequenza, come è rivelato dalla sovrapposizione tra la frequenza di parole di 6 lettere a ciascuna posizione ed il valore di frequenza media che si deriva dall'analisi di *datasets* di sequenze casuali. In netto contrasto, le frequenze migliori di parole di 6 lettere nella regione del DNA a valle di –60 si discosta significativamente dal *background*, con picchi di conservazione chiaramente definiti attorno alle posizioni -53, -42, -30 e -13 rispetto al tRNA maturo. Le parole di 6 lettere che ricorrono con maggiore frequenza a queste posizioni ed a quelle immediatamente adiacenti sono riportate in *Figura 4.10, B*.

A




B

| POSITION | FREQUENCY | SEQUENCE |
|---|---|---|
| -54 | 0.186 | ATTTTT |
| **-53** | **0.217** | **TTTTTA** |
| -52 | 0.204 | TTTTTA |
| -43 | 0.163 | TATTAA |
| **-42** | **0.181** | **TATTAA** |
| -41 | 0.172 | ATATAG |
| -40 | 0.172 | TATATA |
| -39 | 0.158 | TATATT |
| -33 | 0.131 | ATTTTA |
| -32 | 0.140 | TCTTTA |
| -31 | 0.149 | ATTTTA |
| **-30** | **0.158** | **TTTTAT** |
| -29 | 0.136 | TTTAAT |
| -17 | 0.271 | TTTTCA |
| -16 | 0.303 | TTTCAA |
| -15 | 0.290 | TTTCAA |
| -14 | 0.308 | CAACAA |
| **-13** | **0.330** | **CAACAA** |
| -12 | 0.330 | AACAAA |
| -11 | 0.294 | AACAAA |

**Fig. 4.10: Elementi di sequenza conservati nella regione a monte dei tDNA.**
(A): In ascissa è riportata la posizione, compresa fra 0 (primo nucleotide codificante il tRNA maturo) e –90, delle regioni in 5' dei tDNA; in ordinata è riportata la frequenza con la quale ricorre la parola di 6 lettere più frequente in ogni posizione (*Best word frequency*). Le linee tratteggiate indicano, rispettivamente, la frequenza minima, media e massima ottenuta analizzando *datasets* di sequenze casuali. (B): E' riportata, per ogni posizione a monte (colonna di sinistra) la parola di 6 lettere che ricorre maggiormente (colonna di destra) e la corrispondente frequenza (colonna centrale). Le frecce sulla sinistra indicano i motivi di sequenza più frequenti in ciascuna delle quattro regioni parzialmente conservate.

L'elemento attorno alla posizione -42 è una sequenza TATA-*like*, mentre la conservazione di sequenza attorno a –53 e a –30, ai due lati del motivo TATA-*like*, rivela regioni ricche in T. I motivi identificati a -30, -42 e -53 definiscono una regione a monte che corrisponde a quella che è nota prendere contatto con TFIIIB (Kassavetis et al., 1989).

Le sequenze con più alta frequenza relativa sono quelle centrate attorno alla posizione –13 e corrispondono all'impronta dell'inizio della trascrizione Pol III-dipendente, come riportato da studi precedenti (Raymond et al., 1985) (Fruscoloni et al., 1995) (Zecherle et al., 1996).

Diversamente da quanto appena descritto per la sequenza a monte, un'analoga analisi di sequenza condotta sulle sequenze in 3' dei tDNA di lievito non ha rivelato alcuna regione con una significativa conservazione di sequenza, eccetto l'attesa regione ricca in T che corrisponde al terminatore per Pol III.

Abbiamo ritenuto utile, in vista delle successive analisi, classificare i 221 tDNA campione secondo la conformità delle loro regioni in 5' con il *pattern* riportato in *Figura 4.10*. Per ciascuno sono state considerate le regioni attorno a -53, -42, -30 e -13 e ad ogni singolo tDNA è stato attribuito un punteggio compreso tra 0 e 4.

La maggior parte delle famiglie multigeniche è costituita da membri che posseggono la stessa regione codificante e gli stessi promotori intragenici, ma che sono preceduti da regioni a monte con punteggi diversi.

Al fine di valutare l'effetto delle diverse regioni in 5' sulla trascrizione dei tDNA abbiamo scelto i geni che codificano per i $tRNA^{Pro}$(AGG), $tRNA^{Met}$(CAT) e $tRNA^{Asn}$(GTT) come rappresentativi di famiglie entro le quali il punteggio delle regioni a monte varia da 0 a 4.

Abbiamo isolato e testato trascrizionalmente *in vitro* i tDNA membri di ciascuna di queste famiglie; il risultato relativo ai $tDNA^{Asn}$(GTT) è riportato in *Figura 4.11*. Questa famiglia di tDNA comprende 10 membri uno dei quali però, tN(GTT)NL, abbiamo escluso dall'analisi in quanto fiancheggiato da un elemento Ty.

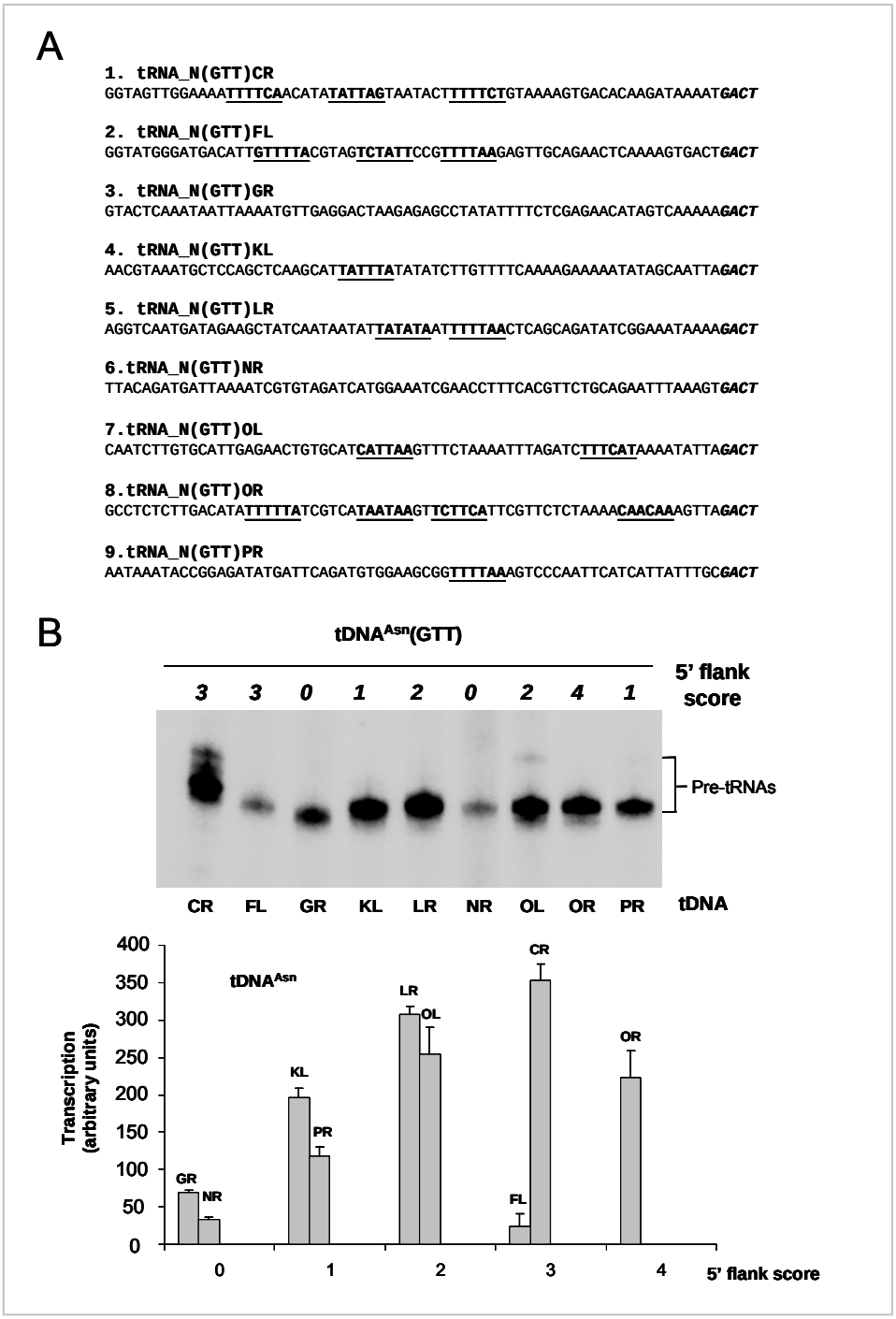



**Fig. 4.11: Trascrizione *in vitro* di geni per tRNA iso-codificanti, con sequenze in 5' diverse.**
(A): Sequenza nucleotidica (da –66 a +4) dei 9 tDNA iso-codificanti che costituiscono la famiglia genica $tRNA^{Asn}$(GTT). Le parole di 6 lettere che si conformano al *consensus* individuato entro le quattro regioni parzialmente conservate sono in grassetto e sottolineate. (B) *Sopra*: trascrizione *in vitro* dei templati riportati in *A*. Sotto ad ogni corsia è riportato il gene per tRNA utilizzato, identificato dalle ultime due lettere del suo nome sistematico; il corrispondente punteggio della regione in 5' è indicato sopra ad ogni corsia. Sulla destra è indicata la posizione dei diversi pre-tRNA prodotti. *Sotto*: istogramma derivato dalla quantificazione di due indipendenti saggi di trascrizione *in vitro*. I diversi $tDNA^{Asn}$ sono raggruppati, sull'asse delle x, in base al punteggio della loro sequenza in 5'; le barre d'errore indicano l'errore *standard*.

Come mostrato in figura, queste nove copie geniche mostrano capacità di trascrizione *in vitro* sostanzialmente diverse. La copia genica trascritta a più alto livello, tN(GTT)CR, è caratterizzata da un punteggio 3; due copie geniche con bassa efficienza trascrizionale, invece, hanno punteggio 0: tN(GTT)GR, il cui livello di trascrizione é 5 volte inferiore rispetto a tN(GTT)CR, e tN(GTT)NR, trascritto 10 volte meno efficientemente. L'unica eccezione, inaspettata, alla buona correlazione che esiste fra punteggio della regione in 5' ed efficienza di trascrizione *in vitro*, è rappresentata da tN(GTT)FL, che ha un punteggio pari a 3 ma è riproducibilmente trascritto ~15 volte meno efficientemente rispetto a tN(GTT)CR. Un'anomalia che si può notare in tN(GTT)FL, tuttavia, è il fatto che nel suo motivo -42 (di sequenza TCTATT) è presente un *mismatch* rispetto alla sequenza consenso che potrebbe inattivarne la funzione di TATA-*box*.

La prima conclusione che può essere tratta dai dati relativi a questa ed alle altre famiglie di tDNA testate, è che la regione fiancheggiante in 5' sembra influenzare in modo largamente prevedibile l'*output* trascrizionale *in vitro* dei diversi membri di tali famiglie. Più in specifico, pare esserci una stretta correlazione tra i livelli di trascrizione prodotti da ciascun tDNA iso-codificante ed il grado di conformità della sua regione in 5' al profilo di sequenza composito che abbiamo identificato. Come suggerito dal comportamento anomalo del gene tN(GTT)FL, tuttavia, non si può escludere la possibilità che altre caratteristiche trascrizionalmente importanti della regione in 5' possano sfuggire al sistema di classificazione usato sin qui.

Abbiamo quindi focalizzato la nostra attenzione sui due geni tN(GTT)CR e tN(GTT)NR, tDNA iso-codificanti che differiscono molto sia per il punteggio della loro regione in 5' (3 *vs* 0) che per la loro efficienza trascrizionale (vedi *Fig. 4.11*).

La delezione di tutta la sequenza che naturalmente si trova a monte della posizione -58 ha rivelato che la differenza nella capacità di trascrizione dei due geni non dipende da quella regione. Allo stesso modo si è rivelata ininfluente la regione in 3' rispetto alla sequenza codificante il tRNA maturo (dati non mostrati).

Esperimenti di mutagenesi hanno inoltre rivelato che i tre motivi centrati attorno alle posizioni -53, -42 e -30 di tN(GTT)CR possono essere singolarmente rimpiazzati dalle corrispondenti regioni di tN(GTT)NR senza che vi sia una riduzione significativa dell'efficienza di trascrizione (dati non mostrati). Il *pattern* di sequenza che abbiamo

identificato, pertanto, non deve essere considerato come una classica regione di controllo multipartita, costituita di *box* funzionalmente indipendenti e facilmente trapiantabili, ma piuttosto come una superficie di legame piuttosto larga e funzionalmente eterogenea che fornisce i siti di legame per almeno tre diverse proteine, i componenti di TFIIIB.

Attraverso esperimenti di ritardata migrazione su gel abbiamo verificato che, quando complessi TFIIIC-DNA, assemblati alternativamente sul tDNA forte oppure debole, vengono incubati con quantità crescenti di TFIIIB, i complessi assemblati sul gene col 5' forte possono essere *shiftati* verso l'alto da minori concentrazioni di TFIIIB (dati non mostrati). Inoltre, misurando il livello di trascrizione *in vitro* dei due tDNA tN(GTT)CR e tN(GTT)NR in presenza di diverse concentrazioni di TFIIIB, il tDNA con il 5'debole è risultato riproducibilmente trascritto da 5 a 6 volte meno efficientemente rispetto a tN(GTT)CR alle più basse concentrazioni di TFIIIB, mentre è risultato trascritto tanto quanto tN(GTT)CR in presenza di un largo eccesso di TFIIIB (dati non mostrati).

L'insieme di queste osservazioni suggerisce fortemente che, *in vitro*, la sequenza in 5' ad alto punteggio di tN(GTT)CR eserciti il suo effetto positivo sulla trascrizione facilitando l'assemblaggio del fattore TFIIIB.

Monitorare l'espressione dei singoli membri delle famiglie di tDNA *in* vivo è complicato dalla differenza di sequenza estremamente piccola che il più delle volte ne distingue i trascritti primari. Per esaminare l'effetto della regione a monte *in vivo*, dunque, abbiamo combinato le diverse sequenze in 5' con un gene *reporter* sintetico, originariamente disegnato dal gruppo di Feldmann (Krieg et al., 1991). Il gene reporter, $tDNA^{Syn2}$, contiene la sequenza codificante del $tRNA^{Glu}$(TTC) di *S. cerevisiae*, marcato tramite l'inserzione di una sequenza *intron-like* che non può essere rimossa dal tRNA precursore.

L'inserimento del *tag* pseudo-intronico ci ha permesso di identificare e quantificare i trascritti prodotti dai $tDNA^{Syn2}$ con le regioni a monte modificate tramite esperimenti di *Northern blot*, usando il $tRNA^{Glu}$(TTC) endogeno, senza introne, come *standard* interno (Krieg et al., 1991).

Abbiamo costruito due fusioni geniche, nelle quali la regione fiancheggiante in 5' tN(GTT)CR oppure tN(GTT)NR è stata fusa al tDNA Syn2.

La figura *4.12, A,* mostra il controllo del fatto che tali sequenze in 5' esercitano il loro distinto effetto trascrizionale *in vitro* anche quando poste a monte del $tDNA^{Syn2}$. Si noti inoltre che il templato [5'NR]Syn2 è trascritto con la stessa efficienza del $tDNA^{Syn2}$ controllo che porta in 5' la sequenza derivata dal plasmide YEp352 (cfr. corsie 1 e 3). Questa osservazione è fortemente a sostegno di un ruolo stimolatorio della regione a monte di tN(GTT)CR, piuttosto che di un effetto inibitorio della corrispondente regione a basso punteggio di tN(GTT)NR.

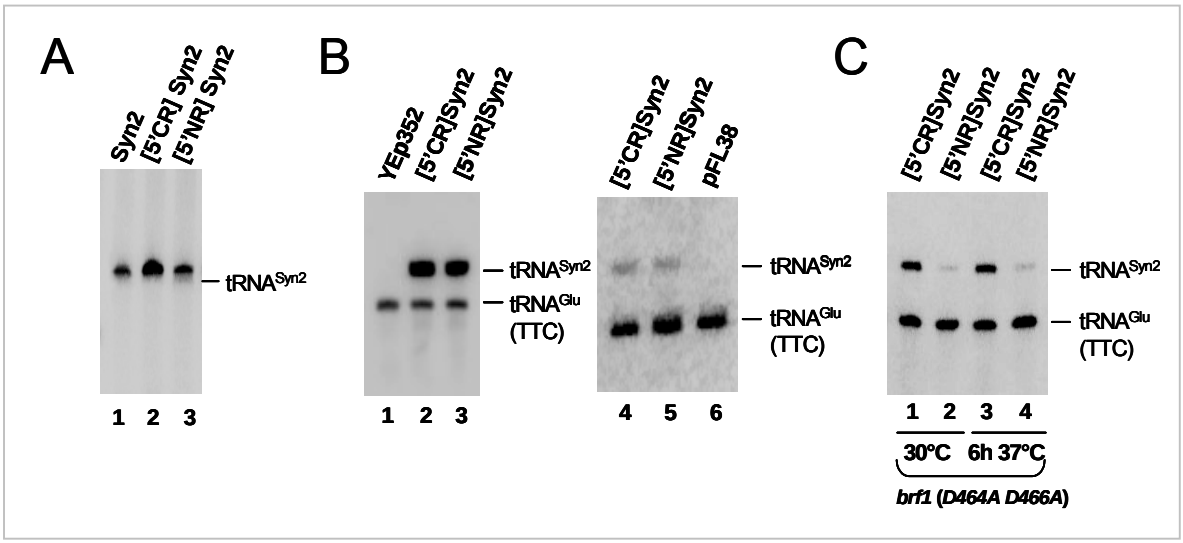



**Fig. 4.12: Analisi trascrizionale delle regioni in 5' di due $tDNA^{Asn}$(GTT)** ***in vivo*****.**
(A): Trascrizione *in vitro* del $tDNA^{Syn2}$ fuso con diverse sequenze a monte. I templati, indicati sopra le corsie, consistono nella sequenza codificante il $tRNA^{Syn2}$ fiancheggiata in 5' o dalla sequenza derivata dal plasmide YEp352 (*Syn2*) o da 58 bp derivate alternativamente da tN(GTT)CR (*[5'CR]Syn2*) o da tN(GTT)NR (*[5'NR]Syn2*). (B): Espressione dei due costrutti $tDNA^{Syn2}$ in ceppi di lievito *wild-type*. Nelle corsie 1-3 è stato usato il ceppo YPH500 trasformato con il vettore YEp352 (vuoto oppure contenente i costrutti); nelle corsie 4-6 è stato usato il ceppo W303 trasformato con pFL38. L'ibridazione è stata ottenuta con uno specifico oligonucleotide antisenso per i $tRNA^{Glu}$(TTC), marcato radioattivamente. Sulla destra è indicata la posizione dei trascritti derivati o dai $tDNA^{Glu}$(TTC) genomici (*$tRNA^{Glu}$(TTC)*) o dai due costrutti Syn2 plasmidici (*$tRNA^{Syn2}$*). (C): Espressione dei due costrutti $tDNA^{Syn2}$, portati dal vettore pFL38, nel ceppo mutante *brf1-II.6* (*D464A, D466A*). L'RNA totale estratto prima (corsie 1-2) o dopo (corsie 3-4) lo *shift* di 6h a 37°C è stato ibridato come in *B*. La posizione dei diversi trascritti è indicata sulla destra come in *B*.

Come mostrato in *Figura 4.12, B*, quando le fusioni [5'CR]Syn2 e [5'NR]Syn2, alternativamente entro il plasmide ad alto numero di copie YEp352 (corsie 1-3) o entro il plasmide centromerico pFL38 (corsie 4-6) sono state trasformate in due comuni ceppi di

lievito (YPH500 in corsie 1-3; W303 in corsie 4-6), non è stata osservata alcuna differenza nell'espressione del $tRNA^{Syn2}$ con le due diverse sequenze fiancheggianti (cfr. corsia 3 con 2 e corsia 5 con 4). I livelli *steady-state* del $tRNA^{Syn2}$ prodotti a partire dai costrutti clonati in pFL38 sono risultati ~60 volte più bassi rispetto a quelli ottenuti con YEp352, come atteso dal momento che pFL38 è un vettore centromerico.

Dal momento che la differenza nella capacità di trascrizione *in vitro* tra tN(GTT)CR e tN(GTT)NR dipende dalla concentrazione di TFIIIB (e specialmente della sua componente Brf1) fornita, abbiamo pensato di potere smascherare un corrispondente effetto della sequenza in 5' *in vivo* abbassando artificialmente i livelli cellulari di TFIIIB attivo. A tal fine, abbiamo misurato l'*output* trascrizionale dei costrutti [5'CR]Syn2 e [5'NR]Syn2 in un ceppo di *S. cerevisiae* che porta una mutazione condizionale termosensibile nel gene *BRF1*, mutazione che è nota influenzare la capacità di formare complessi TFIIIB-DNA. Tale allele mutante, *brf1-II.6*, è originato da una doppia sostituzione D464A/D466A nel dominio II di Brf1, che interagisce con TBP (Andrau et al., 1999; Juo et al., 2003).

Come mostrato in *Figura 4.12, C*, nel ceppo mutante la trascrizione del costrutto [5'NR]Syn2 è risultata fortemente ridotta (da 12 a 14 volte) rispetto a quella di [5'CR]Syn2, sia alla temperatura permissiva di 30°C che a quella non permissiva di 37°C. La regione a monte di tN(GTT)CR, pertanto, facilita il legame di TFIIIB e la trascrizione non soltanto *in vitro* ma anche *in vivo* entro cellule di lievito deficitarie nel componente chiave di TFIIIB, Brf1.

TFIIIB, legandosi al DNA, induce in esso un brusco ripiegamento (Leveillard et al., 1991) ed è stato mostrato che il legame di TFIIIB sulla regione a monte di un tDNA può essere potenziato artificialmente incrementando la flessibilità del DNA in particolari posizioni (Grove et al., 1999). E' dunque possibile che il *pattern* di sequenza da noi identificato nella regione in 5' dei tDNA rifletta la presenza di particolari caratteristiche strutturali, intrinseche in questa regione.

Al fine di esaminare questa possibilità, abbiamo analizzato il nostro insieme di sequenze a monte dei tDNA utilizzando un modello di propensione al ripiegamento del DNA basato sulla probabilità di taglio con DNasi I (Brukner et al., 1995). Questa analisi ha rivelato che una regione con propensione estremamente alta verso il ripiegamento è centrata attorno alla posizione –40, in buona sovrapposizione con l'elemento TATA-*like* in posizione -42;

tale regione è preceduta, intorno alla posizione –53, da una regione con una bassa deformabilità intrinseca. In modo analogo, la regione attorno al sito di inizio della trascrizione, che si sovrappone al motivo -13, presenta una marcata propensione verso il piegamento ed è fiancheggiata su entrambi i lati da regioni con bassa *bendability*.

Questa alternanza di segmenti di DNA adiacenti con opposta tendenza al ripiegamento suggerisce che esista un ben preciso *pattern* di deformabilità conservato nella regione a monte dei geni per tRNA.

In conclusione, le procedure sin qui descritte hanno consentito di individuare un *pattern* composito, di sequenza e strutturale, a monte dei geni per tRNA di lievito, costituito da motivi di sequenza parzialmente conservati. Le nostre analisi trascrizionali *in vitro* e *in vivo* sono a supporto dell'idea secondo la quale tali motivi rappresentano un'impronta *in cis* dell'interazione di TFIIIB; essi possono ottimizzare il legame di TFIIIB nella regione a monte dei tDNA ed in tal modo ne incrementano la trascrizione, in modo particolare in condizioni i cui TFIIIB è il fattore limitante.

### *4.2.2 Effetti di Nhp6 sui geni di classe III*

Come visto nell'introduzione, la proteina Nhp6 appartiene alla famiglia delle HMG *proteins*, è una proteina di architettura della cromatina che si lega al DNA in modo non sequenza-specifico ed è in grado di indurre un forte ripiegamento nel DNA.

Esiste una stretta correlazione tra la proteina Nhp6 ed il macchinario trascrizionale Pol III. In particolare è stato mostrato che Nhp6 è in grado di stimolare specificamente la trascrizione del gene *SNR6* (Kruppa et al., 2001; Lopez et al., 2001); non è mai stato osservato, finora, un effetto di Nhp6 sulla trascrizione dei geni per tRNA.

In tutti gli studi condotti, tuttavia, si è sempre osservato il livello di trascrizione dell'insieme dei geni per tRNA, o dei membri di una famiglia di tDNA (Kruppa et al., 2001; Lopez et al., 2001); questo fatto potrebbe avere mascherato un effetto trascrizionale più fine, a livello di singolo gene. L'identificazione dei motivi di sequenza conservati nella regione in 5' e la possibilità di classificare i diversi membri di una famiglia come tDNA forti oppure deboli, in base alla loro aderenza al *consensus* in 5', ci ha fornito uno

strumento per studiare più nel dettaglio i possibili effetti di Nhp6 sulla trascrizione dei geni per tRNA.

Il nostro lavoro prende spunto anche dall'osservazione che Nhp6 svolge un ruolo nell'incrementare la trascrizione di *SNR6*, gene nel quale è nota esserci un'anomala disposizione degli elementi promotori; si ritiene che proprio da questa disposizione sub-ottimale degli elementi promotori possano dipendere i diversi effetti trascrizionali osservati su *SNR6* (Martin et al., 2001; Shivaswamy et al., 2004). In modo analogo, si può considerare un gene per tRNA con una sequenza in 5' debole, come tN(GTT)NR, come unità trascrizionale con un promotore sub-ottimale per la trascrizione Pol III-dipendente. Si può immaginare che la struttura cromatinica, ed in particolare la proteina Nhp6, possa avere un'influenza sulla trascrizione dei tDNA più deboli, presumibilmente un'influenza positiva, in analogia con quanto osservato per *SNR6*.

Abbiamo scelto di testare l'effetto trascrizionale della proteina Nhp6 sui geni per tRNA utilizzando come geni *reporter* gli stessi costrutti $tRNA^{Syn2}$ (Krieg et al., 1991) tramite i quali abbiamo rivelato l'effetto della sequenza in 5'.

In un primo esperimento di *Northern blot* abbiamo trasformato i costrutti [5'CR]Syn2 (5' forte, 3 motivi) e [5'NR]Syn2 (5' debole, 0 motivi) alternativamente in un ceppo di lievito mutante *nhp6ΔΔ* (Y869) o nel suo corrispondente ceppo isogenico *wild-type* (Y865) (Costigan et al., 1994). Abbiamo quindi ibridato l'RNA totale, estratto da cellule cresciute in fase logaritmica alla temperatura permissiva di 30°C o *shiftate* per 2 h alla temperatura non permissiva di 37°C, con la specifica sonda antisenso per i $tRNA^{Glu}$(TTC) già utilizzata; tale sonda consente di distinguere i trascritti derivati dai costrutti $tDNA^{Syn2}$ da quelli derivati dalle copie genomiche del $tDNA^{Glu}$(TTC), che possono essere usati come *standard* interno.

Come si può osservare in *Figura 4.13*, mentre nel contesto *wild-type* i due costrutti Syn2 producono una quantità di trascritto sostanzialmente uguale, in accordo con quanto avevamo già osservato in due diversi ceppi *wild-type* (*Fig. 4.12, B*), nel ceppo *nhp6ΔΔ* il costrutto [5'NR]Syn2, che porta una sequenza in 5' debole, produce una quantità di trascritto marcatamente inferiore rispetto a [5'CR]Syn2, sia alla temperatura permissiva che a quella non permissiva.

Così come in condizioni di carenza di Brf1 si rivelava *in vivo* l'importanza della sequenza a monte per la trascrizione, così anche l'assenza della proteina Nhp6 ha rivelato il ruolo trascrizionale della sequenza a cui si lega il fattore TFIIIB.

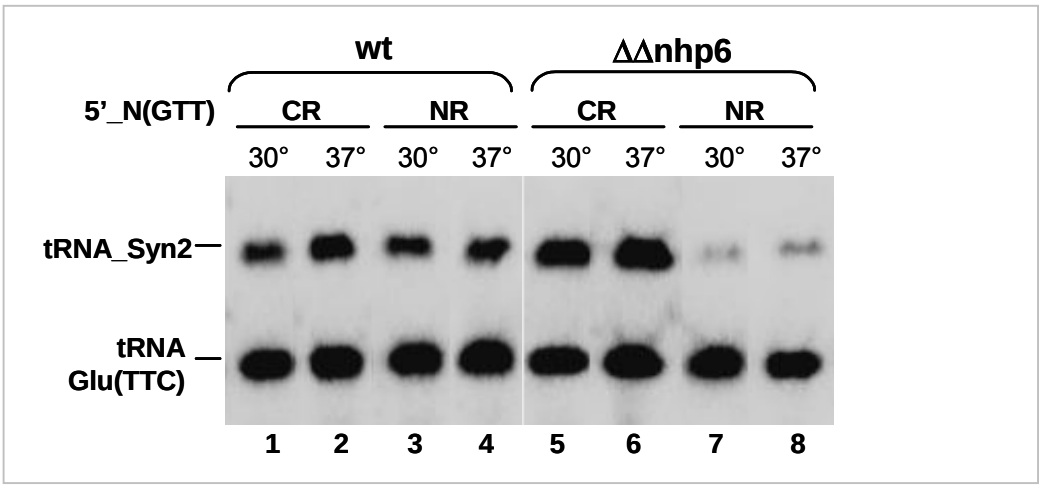



**Fig. 4.13: Espressione di due tDNA$^{Syn2}$, con diversa sequenza in 5', in lieviti privi di Nhp6.**
Analisi *Northern* condotta su RNA totale estratto da lieviti del ceppo Y865 (*wt*, corsie 1-4) o del ceppo Y869 (*ΔΔnhp6*, corsie 5-8), prima o dopo lo *shift* alla temperatura non permissiva di 37°C. I due costrutti, indicati con *CR* e *NR*, sono gli stessi usati in *Fig. 4.12*, trasformati nelle cellule tramite il vettore pFL46S. L'ibridazione è stata ottenuta con uno specifico oligonucleotide antisenso per i tRNA$^{Glu}$(TTC), marcato radioattivamente. Sulla sinistra è indicata la posizione dei trascritti derivati dalle copie genomiche di tDNA$^{Glu}$(TTC) (*tRNA_Glu(TTC)*) o dalle copie plasmidiche dei due costrutti Syn2 (*tRNA_Syn2*).

Per verificare questo risultato, abbiamo scelto di replicare l'esperimento precedente introducendo un secondo plasmide entro le cellule di lievito, il plasmide centromerico pFL39 vuoto oppure contenente il gene *NHP6A*. Insieme alla regione codificante abbiamo clonato anche le regioni regolatrici del gene, comprendendo 381 bp in 5' e 253 bp in 3'. Non è stato possibile introdurre il gene *NHP6A* entro un plasmide ad alto numero di copie perché tale manipolazione è risultata tossica per i lieviti che arrestavano quasi del tutto la loro crescita. Anche in letteratura è stato riportato un simile effetto dell'over-espressione di Nhp6A (Kolodrubetz et al., 2001).

Abbiamo fatto crescere i lieviti alla temperatura permissiva di 30°C dal momento che l'esperimento precedente non aveva rivelato effetti dovuti alla temperatura.

Il risultato dell'ibridazione è riportato in *Figura 4.14*.

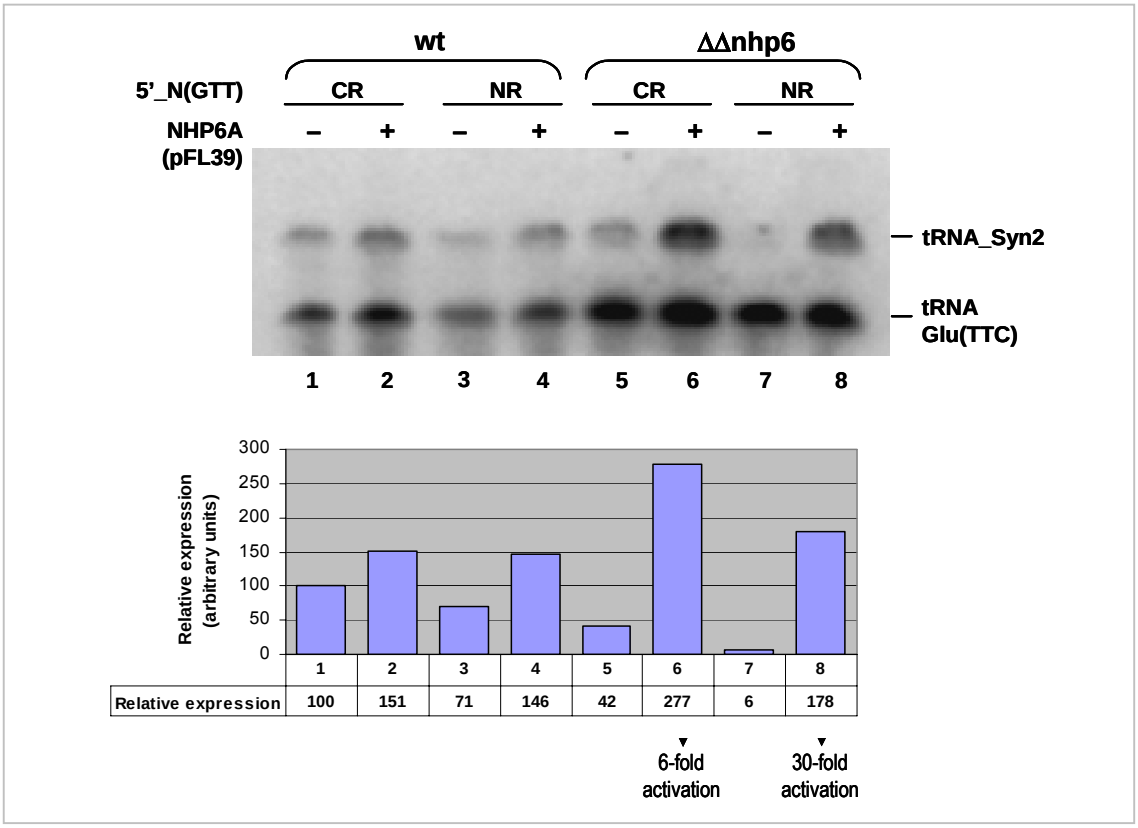



**Fig. 4.14: L'aggiunta di una copia del gene *NHP6* recupera il difetto trascrizionale di [5'NR]Syn2.**
*Sopra*: cellule di lievito dei ceppi Y865 o Y869 (*wt* o *ΔΔnhp6*) sono state trasformate con il plasmide pFL46S, contenete alternativamente il costrutto [5'CR]Syn2 o il costrutto [5'NR]Syn2, e con il plasmide pFL39, alternativamente vuoto o contenente il gene *NHP6A* (*coding sequence* + 381 bp in 5' + 253 bp in 3'), come indicato sopra le corsie. L'RNA totale estratto è stato ibridato come in *Fig. 4.13*. Sulla destra è indicata la posizione dei diversi trascritti, indicati come in *Fig. 4.13*. *Sotto*: istogramma derivato dalla quantificazione dell'immagine riportata sopra. Il segnale rilevato per ciascun trascritto tRNA_Syn2 è stato normalizzato rispetto al relativo trascritto tRNA_Glu(TTC) e riportato in ordinata e sotto ad ogni colonna; il valore ottenuto in corsia 1 è stato fissato a 100. Le frecce sotto al grafico indicano il diverso grado di attivazione della trascrizione ottenuto con i due costrutti [5'CR]Syn2 e [5'NR]Syn2 in corsia 6 e 8.

Come si può osservare nelle corsie 1-4 e nella relativa quantificazione, ottenuta utilizzando i trascritti $tRNA^{Glu}$(TTC) genomici come *standard* interno, nel contesto *wild-type* i due costrutti Syn2 sono risultati trascritti ad un livello paragonabile; l'aggiunta di una copia del gene *NHP6A* ha determinato per entrambi un incremento della trascrizione di ~1,5 volte. In assenza di Nhp6, invece, la trascrizione di [5'NR]Syn2 (corsia 7) è risultata fortemente compromessa; l'espressione di questo costrutto è risultata 7 volte inferiore rispetto a quella di [5'CR]Syn2 (corsia 5). Con l'aggiunta di *NHP6A*, si è determinato un notevole

incremento dell'espressione di questi costrutti, incremento che è quantitativamente molto più rilevante nel caso di [5'NR]Syn2; per quest'ultimo, infatti, si ha un recupero trascrizionale di ~30 volte (corsia 8), mentre per [5'CR]Syn2 il recupero trascrizionale è di sole 6 volte (corsia 6).

I risultati sin qui descritti indicano chiaramente che il basso livello di espressione del tDNA che porta la sequenza in 5' debole è dovuto all'assenza di Nhp6; l'aggiunta di una copia genica di *NHP6A*, infatti, è risultata sufficiente a recuperarne il difetto. Nhp6, inoltre, esercita la sua azione stimolatoria principalmente sul tDNA più debole. Nhp6 quindi funziona come una sorta di modulatore dell'espressione dei tRNA, che incrementa specificamente la trascrizione dei tDNA che presentano una regione di legame per TFIIIB non ottimale.

Si noti che l'aggiunta di una copia del gene *NHP6A* significa il recupero di tale attività nel mutante *nhp6ΔΔ*; è possibile, tuttavia, che il gene portato dal plasmide sia trascritto maggiormente rispetto alle copie genomiche dei geni *NHP6* del ceppo *wild-type*. Non è pertanto possibile, con questi dati, individuare una correlazione di tipo quantitativo tra livello intracellulare di proteina Nhp6 e trascrizione. Quello che resta chiaramente definito è lo specifico recupero nell'espressione dei $tDNA^{Syn2}$ nella condizione di presenza piuttosto che di assenza di Nhp6, ed il maggior grado di recupero osservato sul templato che porta il 5' debole.

Come ulteriore prova *in vivo* dell'interazione fra Nhp6 ed il complesso di trascrizione che si assembla sui geni per tRNA, abbiamo ritenuto opportuno verificare se il difetto trascrizionale nel mutante *nhp6ΔΔ* potesse essere recuperato potenziando il macchinario Pol III, in modo particolare aumentando il livello di espressione del componente limitante di TFIIIB, Brf1; osservazioni precedenti avevano infatti rivelato che il fenotipo termosensibile di lieviti mutanti *nhp6ΔΔ* può essere soppresso dall'over-espressione di *BRF1* (Kruppa et al., 2001). Abbiamo scelto Brf1 anche per il fatto che, *in vivo*, gli effetti della sequenza in 5' forte oppure debole si erano rivelati solo in condizioni di carenza di questo componente.

L'over-espressione di Brf1 è stata ottenuta introducendo entro le cellule di lievito un plasmide ad alto numero di copie contenete la sequenza codificante Brf1 + 194 bp in 5' e

619 bp in 3', le stesse regioni fiancheggianti già utilizzate in precedenza (Andrau et al., 1999).

Il risultato del saggio *Northern* dove come geni *reporter*, analogamente a quanto fatto in precedenza, sono stati utilizzati i due costrutti Syn2, è riportato in *Figura 4.15*.

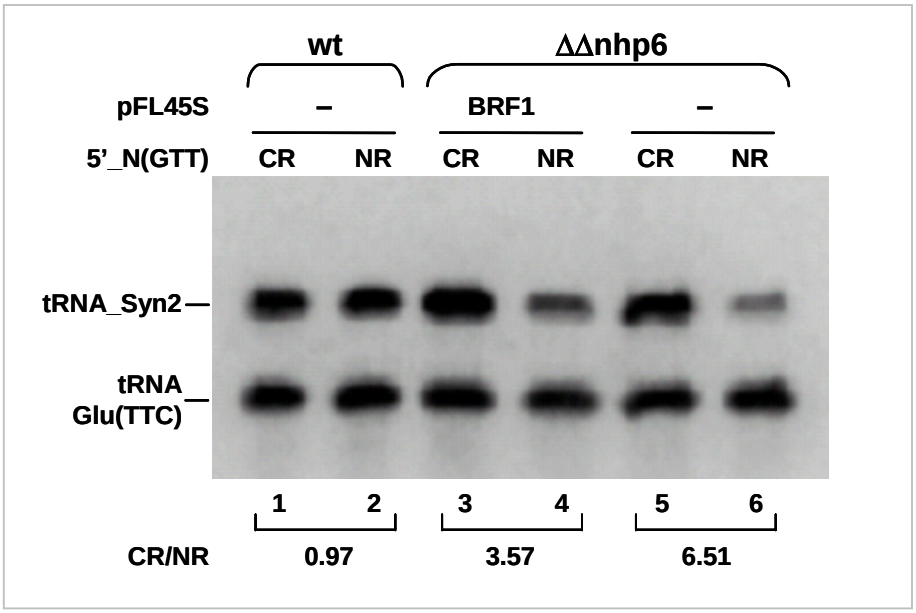



**Fig. 4.15: L'over-espressione di *BRF1* recupera parzialmente il difetto trascrizionale di [5'NR]Syn2**.
Cellule di lievito dei ceppi Y865 o Y869 (*wt* o *ΔΔnhp6*) sono state trasformate con il plasmide pFL46S, contenete alternativamente il costrutto [5'CR]Syn2 o il costrutto [5'NR]Syn2, e con il plasmide pFL45S, alternativamente vuoto o contenente il gene *BRF1* (*coding sequence* + 194 bp in 5' + 619 bp in 3'), come indicato sopra le corsie. L'RNA totale estratto è stato ibridato come in *Fig. 4.13*. Sulla sinistra è indicata la posizione dei diversi trascritti, indicati come in *Fig. 4.13*. Sotto ai numeri che indicano le corsie è riportato il rapporto ottenuto dalla quantificazione dei trascritti derivati da [5'CR]Syn2 e [5'NR]Syn2 (*CR/NR*) nelle tre condizioni testate.

Mentre rimangono confermati i dati già ottenuti nei precedenti esperimenti (rapporto ~1:1 fra [5'CR]Syn2 e [5'NR]Syn2 nel contesto *wild-type*, corsie 1 e 2; rapporto ~7:1 nel contesto *nhp6ΔΔ*, corsie 5 e 6), con l'over-espressione di Brf1 il livello di trascritto ottenuto dal templato [5'NR]Syn2 è risultato avvicinarsi a quello ottenuto con [5'CR]Syn2; si è ottenuto infatti un rapporto di circa 3:1 (corsie 3 e 4).

Questo risultato porta a concludere che l'azione di Nhp6 risiede probabilmente nel facilitare l'assemblaggio del fattore TFIIIB sulla regione a monte dei tDNA: aumentando

la disponibilità di questo fattore, infatti, le esigenze di sequenza *in cis* si rendono meno stringenti. Il fatto che si osservi un recupero trascrizionale soltanto parziale con l'over-espressione di Brf1 non è in realtà stupefacente, dal momento che Brf1 é soltanto uno dei componenti di TFIIIB e che l'assemblaggio del complesso di pre-inizio è un processo complesso e che coinvolge anche diversi fattori.

Visti i risultati sin qui ottenuti *in vivo*, ci siamo domandati se l'effetto di Nhp6 potesse essere dovuto ad un'interazione diretta fra Nhp6 e la cromatina in corrispondenza dei geni per tRNA.

A tal fine abbiamo allestito una *Chromatin Immunoprecipitation* (ChIP) utilizzando specifici ceppi di lievito commerciali che portano un TAP-*tag* su ciascuna delle due isoforme di Nhp6, Nhp6A e B, e su Brf1, utilizzato come controllo positivo; come controllo negativo abbiamo usato un ceppo in cui nessuna proteina è marcata tramite TAP-*tagging*.

Il TAP (*Tandem Affinity Purification*) *tag* è costituito da una cassetta di fusione che codifica per la CBP (*Calmodulin Binding Peptide*), un "*TEV protease cleavage site*" e due "*IgG binding domains*" della proteina A di *Staphylococcus aureus* (Prot A) (Rigaut et al., 1999). Le proteine di fusione ottenute con TAP-*tagging* sono funzionali e possono essere precipitate facendo uso di "*IgG beads*".

A seguito della immunoprecipitazione cromatinica abbiamo amplificato tramite PCR specifiche regioni corrispondenti a diversi trascritti Pol III; in particolare abbiamo scelto di verificare se Nhp6 si lega in corrispondenza del gene *SNR6*, che è noto essere influenzato da questa proteina, ed in corrispondenza dei geni per tRNA. Abbiamo scelto di amplificare, oltre alla regione codificante, anche una certa porzione di sequenza in 5' ed in 3', al fine di includere anche le regioni regolative di questi geni. I risultati relativi a *SNR6* ed a due dei tDNA testati sono riportati in *Figura 4.16*.

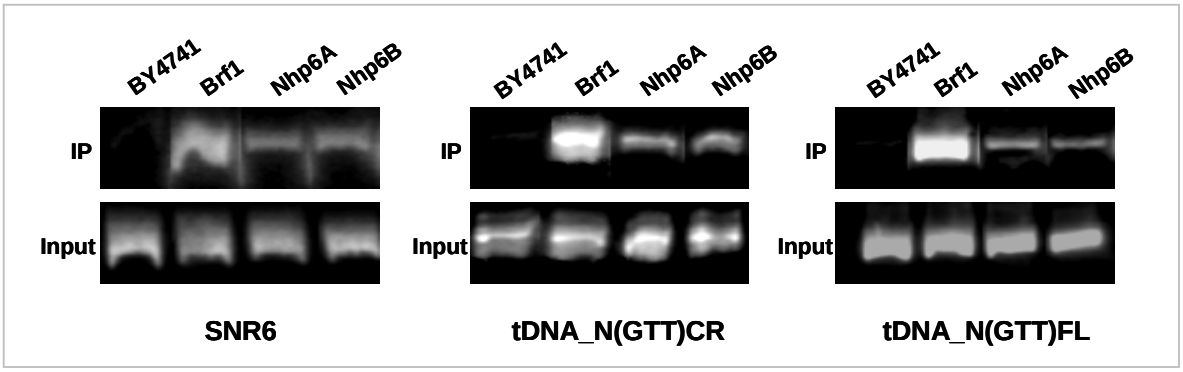



**Fig. 4.16: Nhp6 si lega sia su *SNR6* che sui geni per tRNA.**
*Chromatin immunoprecipitation* condotta su diversi ceppi di lievito in cui sono alternativamente marcate, tramite TAP-*tag,* le due isoforme di Nhp6 (*Nhp6A*, *Nhp6B*), Brf1 (controllo positivo, *Brf1*) o nessuna proteina (controllo negativo, *BY4741*). Le unità trascrizionali corrispondenti al gene *SNR6* ed a due geni per tRNA, riportate sotto i gel, sono state amplificate per PCR su aliquote equivalenti di cromatina immunoprecipitata (*IP*) oppure non trattata (*Input*) derivata dai diversi ceppi.

Dalla ChIP, Nhp6 è risultata essere associata alla cromatina sia in corrispondenza di *SNR6*, sia dei geni per tRNA. Questo risultato non esclude che Nhp6, che è una proteina abbondante in lievito e che è nota legarsi al DNA in modo aspecifico, non possa trovarsi legata anche in corrispondenza di altre regioni che non corrispondono a trascritti da Pol III; giustifica comunque l'ipotesi secondo la quale l'effetto di Nhp6 deriva da un'interazione diretta della proteina con il complesso di trascrizione Pol III.

Date le evidenze osservate *in vivo*, il modo più diretto per verificare se Nhp6 interviene direttamente nell'assemblaggio di un complesso Pol III attivo ci è sembrato quello di assemblare un saggio di trascrizione *in vitro* aggiungendo in reazione quantità crescenti della proteina Nhp6A ed usando come templato i due tDNA naturali, tN(GTT)CR e tN(GTT)NR, che recano, rispettivamente, la sequenza in 5' forte e debole.

Abbiamo prodotto e purificato la proteina Nhp6A tramite l'inserzione di un His-*tag* N-terminale.

Usando un sistema di trascrizione *standard*, comprendente la frazione parzialmente purificata TFIIIC, Pol III purificata, TBP e Brf1 ricombinanti e la frazione B", che contiene l'attività Bdp1 (Bartholomew et al., 1991) (Kassavetis et al., 1991), non è stato

possibile osservare alcun incremento nella trascrizione di nessuno dei due tDNA testati a seguito dell'addizione di rNhp6A, fino ad una quantità di 200 ng (dati non mostrati).

In questo sistema di trascrizione, tuttavia, si trovano due frazioni proteiche che, oltre alle attività TFIIIC e Bdp1, potrebbero contenere anche tracce della proteina Nhp6.

La spettrometria di massa sulla frazione TFIIIC utilizzata non ha rivelato la presenza di Nhp6. Per verificare se il mancato effetto di Nhp6 potesse derivare dal fatto questa proteina era già presente nella frazione B" utilizzata, abbiamo condotto un saggio di trascrizione simile al precedente, usando la proteina Bdp1 ricombinante in luogo della frazione B".

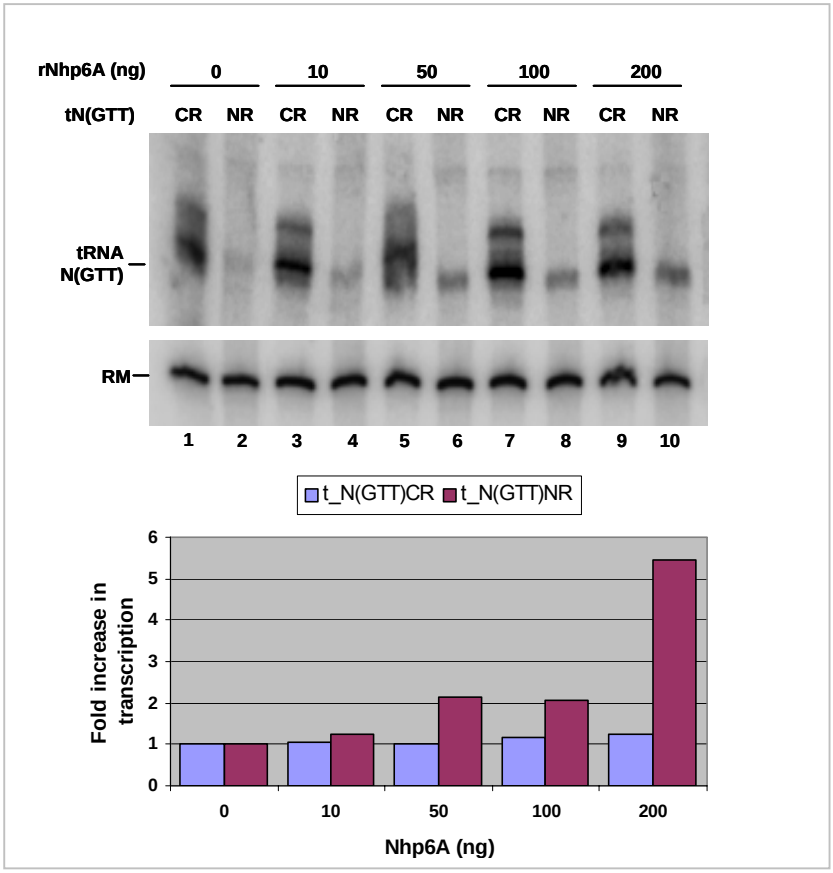



**Fig. 4.17: L'aggiunta di rNhp6A recupera specificamente la trascrizione del tDNA$^{Asn}$ con sequenza in 5' debole.**
*Sopra*: trascrizione *in vitro* dei tDNA templato tN(GTT)CR e tN(GTT)NR, condotta usando Pol III altamente purificata, TFIIIC parzialmente purificato e TFIIIB tutto ricombinante. Nelle diverse reazioni di trascrizione, durante la pre-incubazione del DNA con i fattori di trascrizione, sono state aggiunte quantità crescenti di rNhp6A (0-200 ng), indicate sopra il gel. Sulla sinistra è indicata la posizione dei tRNA trascritti ottenuti. Sullo stesso gel insieme ai trascritti sono stati caricati ~1000 cpm di un DNA di 283 bp usato come *recovery marker* (*RM*). *Sotto*: istogramma derivato dalla quantificazione del saggio in *A*. I segnali ottenuti sono stati normalizzati rispetto al *recovery marker*. Posto quindi uguale

a 1 il livello di trascrizione ottenuto per ciascuno dei due tDNA in assenza di Nhp6, l'incremento della trascrizione è stato calcolato come relativo a questa condizione e riportato sull'asse delle y; sull'asse delle x sono riportate le quantità di proteina ricombinante aggiunte.

Come mostrato in *Figura 4.17* e relativa quantificazione, la trascrizione di tN(GTT)CR (corsie dispari) è rimasta sostanzialmente invariata a seguito dell'aggiunta di rNhp6A; si osserva invece uno specifico incremento dell'espressione di tN(GTT)NR (corsie pari) al crescere della quantità di rNhp6A aggiunta, incremento che risulta massimo in presenza di 200 ng di Nhp6A. In tale condizione il livello di trascrizione del tDNA debole è risultato incrementato di oltre 5 volte rispetto alla condizione di assenza di Nhp6A.

Questo risultato porta alla conclusione che Nhp6 interviene direttamente nel facilitare l'assemblaggio del complesso di trascrizione Pol III e che specificamente aumenta la trascrizione dei tRNA che portano una sequenza in 5' debole. Nhp6 agisce come un livellatore verso l'alto della trascrizione dei tDNA: in assenza di Nhp6, sia *in vitro* che *in vivo*, si rivelano le differenze trascrizionali fra tDNA forti e deboli.

Una seconda conclusione che si può derivare dalle trascrizioni *in vitro* è che nella frazione B" si trova anche la proteina Nhp6; in effetti abbiamo verificato tale conclusione tramite *Western blot*, ottenendo un risultato positivo, in accordo con una simile osservazione riportata in letteratura (Kassavetis et al., 2005).

Con quale meccanismo Nhp6 agisce sulla trascrizione da parte della RNA polimerasi III?

Gli studi condotti sul gene *SNR6* tramite il sistema del doppio ibrido (Fields and Song, 1989), non hanno rivelato alcuna interazione diretta fra la proteina Nhp6 ed alcuno dei componenti di TFIIIB e TFIIIC (Lopez et al., 2001). E' stato quindi ipotizzato che Nhp6 funzioni nello stabilizzare la conformazione ripiegata del DNA nel complesso di pre-inizio della trascrizione sul gene *SNR6*. In modo analogo, è possibile che Nhp6 svolga la sua funzione di stimolazione sui geni per tRNA grazie alla sua capacità di legare il DNA e di indurre in esso un ripiegamento, che facilita il legame di TFIIIB.

Questo richiama il fatto che nella regione in 5' dei tDNA abbiamo osservato l'esistenza di un particolare *pattern* di *bendability* e che dalla presenza o assenza dei motivi in 5', in special modo del motivo TATA-*like* in posizione -42, può essere influenzata la tendenza al

ripiegamento del DNA. Sui tDNA che portano una sequenza in 5' debole, in cui sono assenti i motivi di sequenza conservata, si potrebbe rivelare maggiormante la stabilizzazione dovuta alla proteina strutturale Nhp6.

Recentemente è stato proposto che Nhp6 funzioni come fattore di selezione in grado di favorire la trascrizione specifica riducendo il grado di trascrizione non specifica (Kassavetis et al., 2005). Nhp6 e TBP, componente di TFIIIB, interagiscono entrambe con il solco minore ed inducono un netto ripiegamento del DNA ed è possibile che queste due proteine competano fra loro per il legame di TATA-*box* parziali.

Secondo questa ipotesi, sul tDNA debole che abbiamo testato si rivelerebbe l'effetto di Nhp6 poiché su questo templato, che possiede un promotore non ottimale per la trascrizione, si formerebbe una proporzione più alta di complessi di trascrizione non specifici; come fattore di selezione, Nhp6 consentirebbe di ottenere il più alto grado di trascritti specifici osservati *in vitro*.

Alcuni dati che abbiamo ottenuto in collaborazione con David Donze, dalla Louisiana State University, hanno rivelato che la proteina Nhp6 ha forti relazioni con l'apparato trascrizionale Pol III e con i geni per tRNA anche in un diverso contesto, quello del *locus HMR*-E di *Saccharomyces cerevisiae*.

I cromosomi eucariotici sono organizzati in domini strutturalmente e funzionalmente distinti; diverse regioni di DNA sono state identificate come elementi separatori di tali domini. In particolare un unico gene per $tRNA^{Thr}$, localizzato nella regione fiancheggiante il *locus HMR*-E in *S. cerevisiae*, è stato identificato come elemento barriera in grado di impedire il propagarsi dell'eterocromatina silente dal *locus HMR* (Donze and Kamakaka, 2001).

E' interessante notare che questo effetto barriera dipende in modo critico dall'attività trascrizionale del tDNA barriera: mutazioni negli elementi promotori A- e B-*box* o nei fattori di trascrizione TFIIIC e TFIIIB (compresa la mutazione *brf1-II.6* (Andrau and Werner, 2001)), sono risultate compromettere la funzionalità del $tDNA^{Thr}$ come barriera (Donze and Kamakaka, 2001). E' stato inoltre mostrato che le regioni fiancheggianti in 5' e in 3' giocano un ruolo chiave per la funzionalità dell'elemento barriera (Donze and Kamakaka, 2001).

E' possibile monitorare l'attività barriera di un dato elemento di DNA tramite il "saggio di silenziamento", basato su sistema plasmidico, già descritto in letteratura (Donze et al., 1999) e che abbiamo utilizzato per ottenere i nostri dati relativi a Nhp6.

I costrutti *reporter* utilizzati per misurare la funzione barriera sono schematizzati in *Figura 4.18, A*.

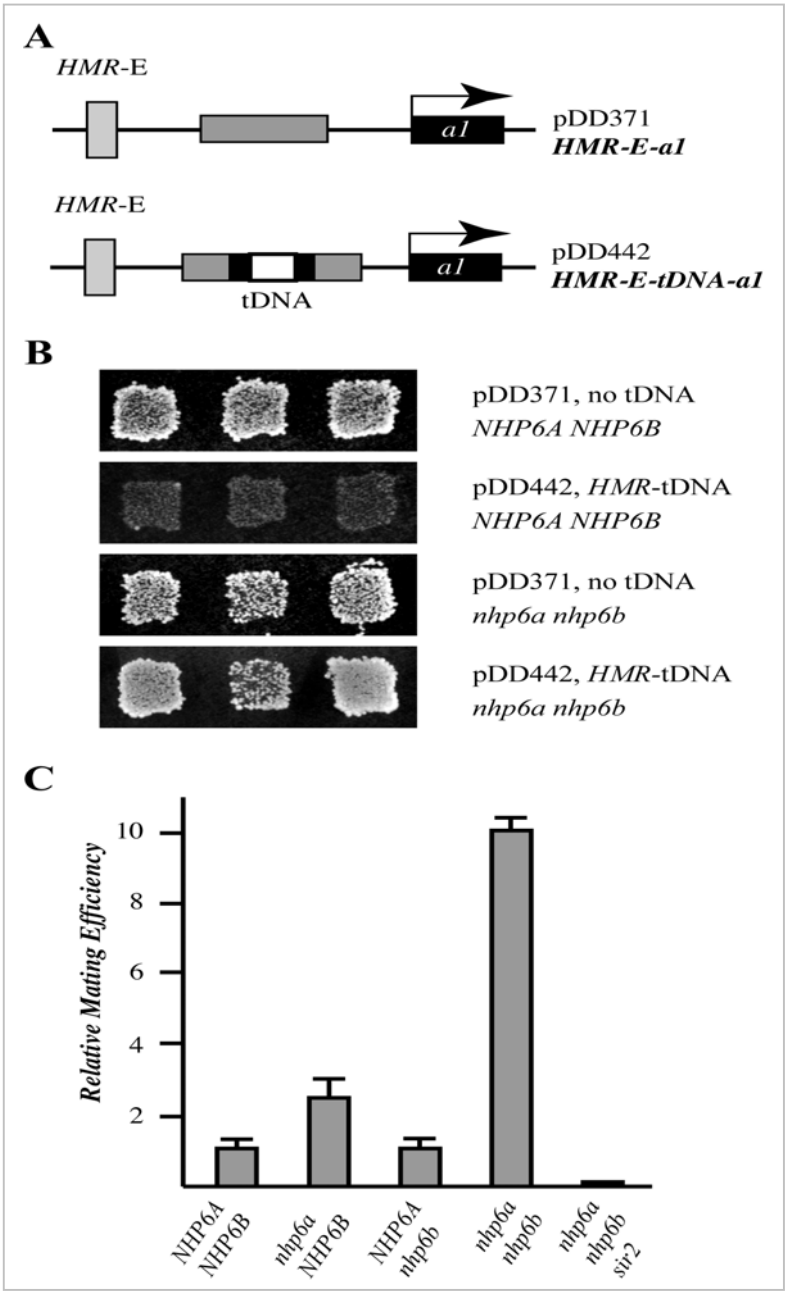



**Fig. 4.18: In assenza di Nhp6 si perde l'attività di barriera dell'*HMR*-tRNA$^{Thr}$.**
(A): Rappresentazione schematica dei plasmidi utilizzati per il saggio di *mating*. Nei due costrutti tra il *silencer HMR*-E ed il gene *a1* è alternativamente assente o presente il tDNA$^{Thr}$ barriera. (B): I ceppi di lievito DDY171 (*NHP6A NHP6B*) e DDY591 (*nhp6a nhp6b*) sono stati trasformati con i plasmidi

riportati in *A*. L'incremento del *mating* che si osserva nel ceppo DDY591 trasformato con pDD442 (quarto riquadro) indica che la capacità dell'*HMR*-tDNA di bloccare il silenziamento risulta compromessa nel *background nhp6a nhp6b*. (C): Analisi quantitativa degli effetti di mutazioni nei geni *nhp6* sull'efficienza di *mating*. I ceppi utilizzati contengono i costrutti *HMR*-E-tDNA integrati nuovamente entro il cromosoma III. Sull'asse delle y è riportata l'efficienza relativa di *mating*, sull'asse delle x sono riportate le caratteristiche dei diversi ceppi utilizzati. Il basso livello di *mating* ottenuto nel ceppo *nhp6a nhp6b sir2* dimostra che l'effetto di Nhp6 è dipendente dalle proteine Sir.

L'elemento *silencer* del *locus HMR*-E da solo (pDD371) è in grado di reprimere la trascrizione del gene *a1* posto a valle di esso. L'inserzione di un frammento di DNA che contiene un elemento barriera tra il *silencer* E ed il gene *a1* (pDD442), invece, blocca il diffondersi dell'eterocromatina silente e consente quindi l'espressione del gene *a1*.

Come si può osservare in *Figura 4.18, B*, trasformato in un ceppo di lievito *MATα hmrΔ*, il plasmide pDD371, in cui il gene *a1* è silenziato dall'elemento *HMR*-E, ha consentito alle cellule di mantenere il proprio fenotipo *α-mating*; questo è stato reso visibile dal fatto che queste cellule crescevano quando replicate su un appropriato *mating lawn* (primo riquadro). Quando invece il tDNA barriera è stato inserito tra l'elemento *HMR*-E ed il gene *a1* (plasmide pDD442), il diffondersi del *silencing* è risultato bloccato, consentendo l'espressione del gene *a1*; questo ha determinato il fenotipo *non-mating* (secondo riquadro).

Dato per noi interessante è che nel contesto di un ceppo di lievito *nhp6a nhp6b*, in cui entrambi i geni che codificano per Nhp6 sono deleti, l'abilità dell'*HMR*-tDNA di bloccare il propagarsi del *silencing* è risultato fortemente compromesso, come indicato dall'aumento del *mating* in cellule *nhp6a nhp6b* trasformate con il plasmide pDD442 (quarto riquadro).

La misura quantitativa del grado di *mating* di ceppi di lievito in cui la barriera *HMR*-tDNA è stata integrata nuovamente entro il cromosoma III (*Fig. 4.18, C*) ha mostrato, in accordo con i dati ottenuti nel sistema plasmidico, che il ceppo *nhp6a nhp6b* fa *mating* da 2,5 a 10 volte più efficientemente rispetto al ceppo *wild-type* per Nhp6; questo significa che in mancanza di Nhp6 risulta incrementato il diffondersi dell'eterocromatina anche quando l'elemento barriera si trova nel naturale contesto cromosomico. La delezione del gene per la proteina Sir2, inoltre, ha rivelato che tale effetto è dipendente dalla istone deacetilasi Sir;

questo indica che l'effetto barriera è mediato da variazioni nel grado di compattazione dei nucleosomi.

Si può immaginare che la perdita dell'effetto barriera in assenza di Nhp6 sia dovuta ad effetti sulla struttura cromatinica su larga scala, dal momento che Nhp6 è una proteina organizzatrice della cromatina. E' stato tuttavia osservato che l'attività del $tDNA^{Thr}$ come barriera è dipendente dalla trascrizione da parte di Pol III (Donze and Kamakaka, 2001). Un'ipotesi, plausibile anche se non provata, che dà ragione sia degli effetti trascrizionali di Nhp6, sia della sua influenza sull'attività barriera, è che la mancanza di Nhp6 possa determinare il diffondersi dell'eterocromatina come effetto secondario, dipendente dal mancato assemblaggio del complesso di trascrizione sullo specifico tDNA barriera.

Alcuni esperimenti di scambio delle regioni fiancheggianti in 5' ed in 3' fra il $tDNA^{Thr}$ barriera (tT(AGT)CR1) ed un secondo membro della stessa famiglia di tDNA (tT(AGT)NL1), sono risultati in un drastico calo della funzione barriera di tT(AGT)CR1, rivelando l'importanza di tali regioni (Donze and Kamakaka, 2001). Secondo il grado di aderenza al *consensus* che abbiamo identificato nella regione in 5' dei tDNA, tT(AGT)CR1 è classificato come gene con punteggio 2: sono presenti i due motivi in posizione –13 e –42; tT(AGT)NL1, che si è rivelato possedere un'attività di barriera fortemente ridotta, contiene invece soltanto il motivo in posizione –42. Su tT(AGT)NL1 il legame di TFIIIB e l'assemblaggio di un complesso di trascrizione attivo, pertanto, dovrebbe essere più difficile, in correlazione con la sua scarsa attività di barriera. Il gene tT(AGT)CR1 porta una sequenza a monte più forte ma comunque non ottimale ai fini dell'assemblaggio del complesso di trascrizione; questo gene infatti funziona come barriera in un contesto *wild-type*, ma non fa altrettanto in assenza di Nhp6.

Va comunque precisato che, oltre alla presenza dei motivi conservati in 5', deve essere ricercato qualche ulteriore elemento par dare ragione del diverso *output* trascrizionale dei tDNA isogenici, anche per il fatto che tali differenze si rivelano *in vivo* solo in condizioni particolari (carenza di Brf1; assenza di Nhp6). E' quindi necessario guardarsi dall'applicare in modo rigido le "regole" che riguardano il 5' dei tDNA.

A sfavore dell'ipotesi secondo la quale l'effetto di Nhp6 sull'attività barriera è un effetto secondario, mediato dalla trascrizione, l'over-espressione di Brf1, che in parte recuperava il difetto trascrizionale del mutante *nhp6ΔΔ* (*Fig. 4.15*), non è risultata sopprimere la

perdita della funzione barriera nello stesso mutante (dati non mostrati). Si tenga presente, tuttavia, che un simile sistema incrementa la disponibilità di soltanto uno dei diversi fattori che costituiscono l'apparato trascrizionale Pol III.

In ultima analisi sia l'effetto trascrizionale che l'effetto sul propagarsi dell'eterocromatina dipendono dalla proprietà di Nhp6 di legarsi al DNA ed indurre in esso una modificazione strutturale. Da un lato questa proteina determina condizioni ottimali per il legame di TFIIIB a monte dei geni di classe III con minore affinità verso esso (*SNR6* e tRNA deboli); dall'altro, come proteina non istonica della cromatina, Nhp6 interviene nella riorganizzazione della regione adiacente al $tDNA^{Thr}$ barriera. E' possibile che quest'ultimo processo sia mediato dall'assemblaggio dell'ingombrante complesso trascrizionale Pol III sul quel tDNA.

# 5. CONCLUSIONE

## 5.1 TERMINAZIONE DELLA TRASCRIZIONE: EFFETTI DEL CONTESTO SULLA FUNZIONALITÀ DEL TERMINATORE

Con i nostri dati abbiamo mostrato che il riconoscimento del segnale di terminazione da parte della RNA polimerasi III è in realtà più complesso della semplice "lettura" di una serie oligo-(dT) sul filamento codificante del DNA.

Abbiamo mostrato che l'evoluzione ha teso verso la selezione di terminatori oligo-(dT) progressivamente più brevi, $T_4$ nei mammiferi, osservazione, questa, che pone il problema di come tratti $T_4$, frequentemente presenti all'interno delle unità trascrizionali Pol III, possono essere oltrepassati dalla polimerasi per la produzione di trascritti finali *full-length*.

Abbiamo dimostrato che il riconoscimento di segnali oligo-(dT) da parte della polimerasi III di *Saccharomyces cerevisiae* può essere influenzato in modo considerevole dal contesto di sequenza entro il quale si trova la serie di T.

Tale influenza del contesto si rende evidente nei confronti dei segnali di terminazione più brevi, $T_5$ e $T_6$ in lievito; tali segnali minimi si trovano complessivamente nel 40% dei tDNA di *S. cerevisiae* che posseggono un singolo segnale di terminazione.

Abbiamo distinto due diversi tipi di effetto del contesto: il primo, che abbiamo esaminato più in dettaglio, è dovuto alle primissime basi che seguono un terminatore $T_5$; il secondo è dovuto alla sequenza che si trova a monte del terminatore ed è stato osservato quando una sequenza diversa da quella di un tDNA è stata fusa a monte di una regione-terminatore forte ($T_5$ + sequenza in 3' derivata da L(CAA)LR2).

Abbiamo inoltre osservato un esempio estremo di effetto del contesto nell'unità trascrizionale *SNR52*, entro la quale un elemento $T_6$, che si trova tra la A- e la B-*box*, non viene quasi riconosciuto come terminatore dalla RNA polimerasi III.

La principale conclusione che emerge dallo studio sull'effetto della sequenza a valle è che un elemento $T_5$, quando associato ad un gene per tRNA, è intrinsecamente forte come terminatore, dal momento che produce un'efficiente terminazione nella maggior parte dei

contesti analizzati. Esistono tuttavia particolari caratteristiche della sequenza a valle che possono influenzare la capacità di causare terminazione di un elemento $T_5$.

In particolare abbiamo identificato il di-nucleotide CT come elemento che, posto immediatamente a valle del terminatore, genera su di esso un effetto di "anti-terminazione". I nostri dati suggeriscono che l'elemento CT faciliti la traslocazione della polimerasi riducendo la pausa di quest'ultima sull'elemento $T_5$.

Una possibile spiegazione di questo effetto è che l'elemento CT incrementi la processività di Pol III alla posizione immediatamente a monte attraverso un meccanismo che coinvolge il legame dei corrispondenti NTPs ad un sito allosterico sulla stessa polimerasi, in analogia con quanto è stato proposto per la RNA polimerasi di *E. coli* e per la RNA polimerasi II di lievito (Holmes and Erie, 2003) (Nedialkov et al., 2003).

Costituisce un risultato piuttosto sorprendente il fatto che quello che si era rivelato un terminatore forte per Pol III nel contesto tDNA, è risultato un terminatore molto più debole nel diverso contesto generato dalla sua fusione con una porzione, contenente A- e B-*box*, del gene *SCR1*.

Ancora più sorprendente è l'osservazione che la serie di 6 T che si trova all'interno della regione trascritta di *SNR52*, serie che costituisce un segnale forte di terminazione in tutti i casi osservati e secondo i dati riportati in letteratura, risulta sostanzialmente inattiva come terminatore della trascrizione.

Anche in questi ultimi due casi gli effetti sulla terminazione sono risultati dipendere dalla pausa dell'enzima sull'oligo-(dT). In esperimenti in cui la trascrizione è stata condotta in assenza di TFIIIC o con la sola RNA polimerasi purificata, inoltre, abbiamo verificato che la distinzione tra terminatori forti e terminatori deboli è una proprietà intrinseca di Pol III, indipendente da fattori proteici accessori.

## 5.2 INIZIO DELLA TRASCRIZIONE: SEQUENZA A MONTE DEI GENI PER TRNA ED EFFETTI DELLA PROTEINA NHP6

Tramite un'analisi bioinformatica condotta sull'insieme dei tDNA di *Saccharomyces cerevisiae*, abbiamo individuato un *pattern* di sequenza conservato nella regione a monte

del sito di inizio della trascrizione, in quella porzione di DNA sulla quale ha luogo il legame di TFIIIB nella fase di inizio della trascrizione. Tale *pattern* consiste di quattro distinti "motivi" in cui particolari sequenze risultano over-rappresentate nell'insieme dei tDNA. Il *pattern* che abbiamo identificato, tuttavia, non deve essere considerato come una classica regione di controllo multipartita costituita di cassette di sequenza funzionalmente indipendenti e facilmente trapiantabili, piuttosto deve essere considerato come una superficie di contatto per il fattore di trascrizione TFIIIB piuttosto estesa e funzionalmente eterogenea.

Abbiamo mostrato sperimentalmente che la sequenza che si trova a monte dei geni per tRNA può influenzare il livello di espressione di tali geni e che il numero di motivi presenti a monte dei singoli tDNA correla piuttosto bene con il livello di trascrizione che essi possono supportare *in vitro*; membri di una stessa famiglia di tDNA, pertanto, che condividono la stessa sequenza codificante, e dunque gli stessi promotori intragenici, e che differiscono per le loro regioni in 5' e in 3', possono essere distinti e classificati come tRNA più o meno forti in base al loro grado di aderenza al *consensus* da noi individuato.

L'influenza trascrizionale esercitata dalla regione a monte, tuttavia, può essere rivelata *in vivo* soltanto in condizioni "limite" per la trascrizione Pol III-dipendente, come in lieviti nei quali è mutato il gene *BRF1*, che codifica per il componente limitante di TFIIIB, o in lieviti nei quali entrambi i geni che codificano per la HMG *protein* Nhp6 sono deleti. Abbiamo dimostrato, infatti, che la proteina Nhp6 è in grado di influenzare la trascrizione Pol III-dipendente non solo limitatamente al gene *SNR6*, come già riportato in letteratura (Kruppa et al., 2001; Lopez et al., 2001), ma anche sui tDNA di *S. cerevisiae*.

La trascrizione di un tDNA classificato come debole è risultata fortemente diminuita in assenza di Nhp6 e tale difetto trascrizionale è stato in parte recuperato potenziando il sistema di trascrizione Pol III tramite l'over-espressione di Brf1.

I dati che abbiamo ottenuto sono fortemente a favore della conclusione secondo la quale Nhp6 influenza l'espressione dei geni per tRNA intervenendo direttamente nella fase di assemblaggio del complesso di trascrizione, piuttosto che tramite effetti cromatinici su più larga scala. Abbiamo infatti mostrato che Nhp6 si trova associata alla cromatina in corrispondenza dei geni per tRNA; l'addizione di quantità crescenti di rNhp6A in saggi di trascrizione *in vitro*, inoltre, ha mostrato un corrispondente aumento dell'*output* trascrizionale del tRNA classificato come debole.

Riteniamo che Nhp6 possa esercitare la sua azione di stimolazione della trascrizione tramite la sua capacità di legarsi al DNA e di indurre in esso un profondo ripiegamento. E' noto infatti che il legame di TFIIIB, passaggio chiave per l'inizio della trascrizione, causa un ripiegamento del DNA; questa conformazione ripiegata potrebbe essere stabilizzata dal legame di Nhp6. In linea con questa ipotesi, abbiamo mostrato che dalla presenza dei motivi in 5' che abbiamo identificato dipende la presenza di un particolare *pattern* di *bendability* in quella regione. I tDNA deboli sono caratterizzati da una sequenza a monte che di per sé non determina una particolare propensione verso il ripiegamento; per questo l'effetto della proteina Nhp6 si rivelerebbe maggiormente sui tDNA deboli.

Secondo un'ipotesi alternativa, suggerita da recenti pubblicazioni (Kassavetis et al., 2005; Kassavetis and Steiner, 2006), Nhp6 potrebbe funzionare come fattore di selezione in grado di sfavorire, tramite una sorta di competizione, il legame di TBP a TATA-*box* imperfette, e quindi di impedire la produzione di trascritti non specifici. La produzione di trascritti aspecifici sarebbe più facile sul tDNA debole, che porta un sito di legame per TFIIIB non ottimale; l'effetto di Nhp6 risulterebbe quindi proporzionalmente maggiore su questo specifico templato.

Abbiamo quindi mostrato che la proteina Nhp6 ha un ruolo, precedentemente mai rivelato, sulla trascrizione dei tDNA di *Saccharomyces cerevisiae*; essa funziona come una sorta di modulatore dell'espressione dei tRNA, incrementando specificamente la trascrizione dei tRNA caratterizzati da una sequenza *in cis* non ottimale per la trascrizione. E' interessante notare a questo proposito che l'effetto di Nhp6 più immediatamente riconoscibile sulla trascrizione Pol III-dipendente riguarda l'espressione del gene *SNR6*, che, in modo analogo, è caratterizzato da una disposizione degli elementi promotori (A- e B-*box*) sub-ottimale ai fini della trascrizione.

Abbiamo infine mostrato che la proteina Nhp6, oltre ad avere un effetto trascrizionale, è in grado di influenzare la funzione non codificante del $tDNA^{Thr}$ che si trova al *locus HMR* di *Saccharomyces cerevisiae*. E' stato mostrato che tale tDNA funziona come barriera al propagarsi dell'eterocromatina al *locus HMR* (Donze and Kamakaka, 2001).

Tramite esperimenti *in vivo*, sia usando un sistema basato su plasmidi sia entro il contesto cromosomico, abbiamo mostrato che in un ceppo di lievito mutante, privo di Nhp6, si perde la funzione barriera dell'*HMR*-tRNA e che tale effetto è dipendente dalla presenza

della istone deacetilasi Sir2. Ciò significa che la funzione barriera è mediata da variazioni nel grado di compattazione dei nucleosomi.

Anche se non definitivamente dimostrata, un'ipotesi che lega i due effetti di Nhp6 sulla trascrizione dei geni per tRNA e sulla funzione di barriera del tDNA al locus *HMR*, è che quest'ultimo sia un effetto secondario, dipendente dall'assemblaggio del complesso di trascrizione sul $tDNA^{Thr}$.

Esperimenti di scambio delle regioni fiancheggianti in 5' e in 3' tra il tDNA barriera ed un altro tDNA appartenente alla stessa famiglia, hanno mostrato che le regioni fiancheggianti possono influenzare fortemente l'effetto barriera (Donze and Kamakaka, 2001).

Se in assenza della proteina Nhp6 risulta fortemente sfavorito il legame di TFIIIB sui tDNA con un 5' debole, e presumibilmente anche sull'*HMR*-tRNA, si può immaginare che dal mancato assemblaggio del complesso di trascrizione Pol III su quel gene dipenda la perdita dell'effetto barriera, in analogia con il chiaro risultato ottenuto in ceppi di lievito con mutazioni a carico delle A- e B-*box* del tDNA barriera o dei fattori di trascrizione TFIIIB e TFIIIC (Donze and Kamakaka, 2001).

Resta comunque possibile che Nhp6 influenzi l'effetto barriera al *locus HMR* semplicemente in quanto proteina organizzatrice della cromatina.

# 6. BIBLIOGRAFIA


Ahmad, K. and Henikoff, S. (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell*, **9**, 1191-1200.

Allain, F.H., Yen, Y.M., Masse, J.E., Schultze, P., Dieckmann, T., Johnson, R.C. and Feigon, J. (1999) Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for non-sequence-specific binding. *Embo J*, **18**, 2563-2579.

Allison, D.S., Goh, S.H. and Hall, B.D. (1983) The promoter sequence of a yeast tRNAtyr gene. *Cell*, **34**, 655-664.

Allison, D.S. and Hall, B.D. (1985) Effects of alterations in the 3' flanking sequence on in vivo and in vitro expression of the yeast SUP4-o tRNATyr gene. *Embo J*, **4**, 2657-2664.

Andrau, J.C., Sentenac, A. and Werner, M. (1999) Mutagenesis of yeast TFIIIB70 reveals C-terminal residues critical for interaction with TBP and C34. *J Mol Biol*, **288**, 511-520.

Andrau, J.C. and Werner, M. (2001) B"-associated factor(s) involved in RNA polymerase III preinitiation complex formation and start-site selection. *Eur J Biochem*, **268**, 5167-5175.

Baer, M., Nilsen, T.W., Costigan, C. and Altman, S. (1990) Structure and transcription of a human gene for H1 RNA, the RNA component of human RNase P. *Nucleic Acids Res*, **18**, 97-103.

Baker, R.E., Eigel, A., Vogtel, D. and Feldmann, H. (1982) *Embo J*, **1**, 291-295.

Bardeleben, C., Kassavetis, G.A. and Geiduschek, E.P. (1994) Encounters of Saccharomyces cerevisiae RNA polymerase III with its transcription factors during RNA chain elongation. *J Mol Biol*, **235**, 1193-1205.

Bargonetti, J., Friedman, P.N., Kern, S.E., Vogelstein, B. and Prives, C. (1991) Wild-type but not mutant p53 immunopurified proteins bind to sequences adjacent to the SV40 origin of replication. *Cell*, **65**, 1083-1091.

Bartholomew, B., Kassavetis, G.A. and Geiduschek, E.P. (1991) Two components of Saccharomyces cerevisiae transcription factor IIIB (TFIIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIIC. *Mol Cell Biol*, **11**, 5181-5189.

Bazett-Jones, D.P., Mendez, E., Czarnota, G.J., Ottensmeyer, F.P. and Allfrey, V.G. (1996) Visualization and analysis of unfolded nucleosomes associated with transcribing chromatin. *Nucleic Acids Res*, **24**, 321-329.

Bell, A.C., West, A.G. and Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, **98**, 387-396.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M. and Reinberg, D. (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science*, **301**, 1090-1093.

Bieker, J.J., Martin, P.L. and Roeder, R.G. (1985) Formation of a rate-limiting intermediate in 5S RNA gene transcription. *Cell*, **40**, 119-127.

Bobkova, E.V., Habib, N., Alexander, G. and Hall, B.D. (1999) Mutational analysis of the hydrolytic activity of yeast RNA polymerase III. *J Biol Chem*, **274**, 21342-21348.

Bogenhagen, D.F. (1985) The intragenic control region of the Xenopus 5 S RNA gene contains two factor A binding domains that must be aligned properly for efficient transcription initiation. *J Biol Chem*, **260**, 6466-6471.

Bogenhagen, D.F. and Brown, D.D. (1981) Nucleotide sequences in Xenopus 5S DNA required for transcription termination. *Cell*, **24**, 261-270.

Bogenhagen, D.F., Wormington, W.M. and Brown, D.D. (1982) Stable transcription complexes of Xenopus 5S RNA genes: a means to maintain the differentiated state. *Cell*, **28**, 413-421.

Bouvet, P., Dimitrov, S. and Wolffe, A.P. (1994) Specific regulation of Xenopus chromosomal 5S rRNA gene transcription in vivo by histone H1. *Genes Dev*, **8**, 1147-1159.

Braglia, P., Percudani, R. and Dieci, G. (2005) Sequence context effects on oligo(dT) termination signal recognition by Saccharomyces cerevisiae RNA polymerase III. *J Biol Chem*, **280**, 19551-19562.

Braun, B.R., Bartholomew, B., Kassavetis, G.A. and Geiduschek, E.P. (1992a) Topography of transcription factor complexes on the Saccharomyces cerevisiae 5 S RNA gene. *J Mol Biol*, **228**, 1063-1077.

Braun, B.R., Kassavetis, G.A. and Geiduschek, E.P. (1992b) Bending of the Saccharomyces cerevisiae 5S rRNA gene in transcription factor complexes. *J Biol Chem*, **267**, 22562-22569.

Brewster, N.K., Johnston, G.C. and Singer, R.A. (1998) Characterization of the CP complex, an abundant dimer of Cdc68 and Pob3 proteins that regulates yeast transcriptional activation and chromatin repression. *J Biol Chem*, **273**, 21972-21979.

Brewster, N.K., Johnston, G.C. and Singer, R.A. (2001) A bipartite yeast SSRP1 analog comprised of Pob3 and Nhp6 proteins modulates transcription. *Mol Cell Biol*, **21**, 3491-3502.

Brow, D.A. and Guthrie, C. (1990) Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. *Genes Dev*, **4**, 1345-1356.

Brukner, I., Sanchez, R., Suck, D. and Pongor, S. (1995) Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. *Embo J*, **14**, 1812-1818.

Bulger, M. and Groudine, M. (1999) Looping versus linking: toward a model for long-distance gene activation. *Genes Dev*, **13**, 2465-2477.

Burlingame, R.W., Love, W.E., Wang, B.C., Hamlin, R., Nguyen, H.X. and Moudrianakis, E.N. (1985) Crystallographic structure of the octameric histone core of the nucleosome at a resolution of 3.3 A. *Science*, **228**, 546-553.

Burnol, A.F., Margottin, F., Huet, J., Almouzni, G., Prioleau, M.N., Mechali, M. and Sentenac, A. (1993) TFIIIC relieves repression of U6 snRNA transcription by chromatin. *Nature*, **362**, 475-477.

Bustin, M. (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol*, **19**, 5237-5246.

Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B. and Kornberg, R.D. (1996) RSC, an essential, abundant chromatin-remodeling complex. *Cell*, **87**, 1249-1260.

Campbell, F.E., Jr. and Setzer, D.R. (1992) Transcription termination by RNA polymerase III: uncoupling of polymerase release from termination signal recognition. *Mol Cell Biol*, **12**, 2260-2272.

Camper, S.A., Yao, Y.A. and Rottman, F.M. (1985) Hormonal regulation of the bovine prolactin promoter in rat pituitary tumor cells. *J Biol Chem*, **260**, 12246-12251.

Cannon, R.E., Wu, G.J. and Railey, J.F. (1986) Functions of and interactions between the A and B blocks in adenovirus type 2-specific VARNA1 gene. *Proc Natl Acad Sci U S A*, **83**, 1285-1289.

Carey, M.F. and Singh, K. (1988) Enhanced B2 transcription in simian virus 40-transformed cells is mediated through the formation of RNA polymerase III transcription complexes on previously inactive genes. *Proc Natl Acad Sci U S A*, **85**, 7059-7063.

Chakshusmathi, G., Kim, S.D., Rubinson, D.A. and Wolin, S.L. (2003) A La protein requirement for efficient pre-tRNA folding. *Embo J*, **22**, 6562-6572.

Challice, J.M. and Segall, J. (1989) Transcription of the 5 S rRNA gene of Saccharomyces cerevisiae requires a promoter element at +1 and a 14-base pair internal control region. *J Biol Chem*, **264**, 20060-20067.

Chedin, S., Ferri, M.L., Peyroche, G., Andrau, J.C., Jourdain, S., Lefebvre, O., Werner, M., Carles, C. and Sentenac, A. (1998a) The yeast RNA polymerase III transcription machinery: a paradigm for eukaryotic gene activation. *Cold Spring Harb Symp Quant Biol*, **63**, 381-389.

Chedin, S., Riva, M., Schultz, P., Sentenac, A. and Carles, C. (1998b) The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. *Genes Dev*, **12**, 3857-3871.

Chernova, O.B., Chernov, M.V., Agarwal, M.L., Taylor, W.R. and Stark, G.R. (1995) The role of p53 in regulating genomic stability when DNA and RNA synthesis are inhibited. *Trends Biochem Sci*, **20**, 431-434.

Chesnokov, I., Chu, W.M., Botchan, M.R. and Schmid, C.W. (1996) p53 inhibits RNA polymerase III-directed transcription in a promoter-dependent manner. *Mol Cell Biol*, **16**, 7084-7088.

Chiang, C.M., Ge, H., Wang, Z., Hoffmann, A. and Roeder, R.G. (1993) Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *Embo J*, **12**, 2749-2762.

Chong, S.S., Hu, P. and Hernandez, N. (2001) Reconstitution of transcription from the human U6 small nuclear RNA promoter with eight recombinant polypeptides and a partially purified RNA polymerase III complex. *J Biol Chem*, **276**, 20727-20734.

Ciliberto, G., Raugei, G., Costanzo, F., Dente, L. and Cortese, R. (1983) Common and interchangeable elements in the promoters of genes transcribed by RNA polymerase iii. *Cell*, **32**, 725-733.

Clarke, E.M., Peterson, C.L., Brainard, A.V. and Riggs, D.L. (1996) Regulation of the RNA polymerase I and III transcription systems in response to growth conditions. *J Biol Chem*, **271**, 22189-22195.

Colbert, T. and Hahn, S. (1992) A yeast TFIIB-related factor involved in RNA polymerase III transcription. *Genes Dev*, **6**, 1940-1949.

Colbert, T., Lee, S., Schimmack, G. and Hahn, S. (1998) Architecture of protein and DNA contacts within the TFIIIB-DNA complex. *Mol Cell Biol*, **18**, 1682-1691.

Costigan, C., Kolodrubetz, D. and Snyder, M. (1994) NHP6A and NHP6B, which encode HMG1-like proteins, are candidates for downstream components of the yeast SLT2 mitogen-activated protein kinase pathway. *Mol Cell Biol*, **14**, 2391-2403.

Cox, L.S. and Lane, D.P. (1995) Tumour suppressors, kinases and clamps: how p53 regulates the cell cycle in response to DNA damage. *Bioessays*, **17**, 501-508.

Cozzarelli, N.R., Gerrard, S.P., Schlissel, M., Brown, D.D. and Bogenhagen, D.F. (1983) Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. *Cell*, **34**, 829-835.

Dammann, R., Lucchini, R., Koller, T. and Sogo, J.M. (1993) Chromatin structures and transcription of rDNA in yeast Saccharomyces cerevisiae. *Nucleic Acids Res*, **21**, 2331-2338.

Das, G., Henning, D., Wright, D. and Reddy, R. (1988) Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III. *Embo J*, **7**, 503-512.

Deininger, P.L. and Batzer, M.A. (1999) Alu repeats and human disease. *Mol Genet Metab*, **67**, 183-193.

Deshpande, A.M. and Newlon, C.S. (1996) DNA replication fork pause sites dependent on transcription. *Science*, **272**, 1030-1033.

Dieci, G., Giuliodori, S., Catellani, M., Percudani, R. and Ottonello, S. (2002) Intragenic promoter adaptation and facilitated RNA polymerase III recycling in the transcription of SCR1, the 7SL RNA gene of Saccharomyces cerevisiae. *J Biol Chem*, **277**, 6903-6914.

Dieci, G., Percudani, R., Giuliodori, S., Bottarelli, L. and Ottonello, S. (2000) TFIIIC-independent in vitro transcription of yeast tRNA genes. *J Mol Biol*, **299**, 601-613.

Dieci, G. and Sentenac, A. (1996) Facilitated recycling pathway for RNA polymerase III. *Cell*, **84**, 245-252.

Dingermann, T., Burke, D.J., Sharp, S., Schaack, J. and Soll, D. (1982) The 5- flanking sequences of Drosophila tRNAArg genes control their in vitro transcription in a Drosophila cell extract. *J Biol Chem*, **257**, 14738-14744.

Dingermann, T., Sharp, S., Schaack, J. and Soll, D. (1983) Stable transcription complex formation of eukaryotic tRNA genes is dependent on a limited separation of the two intragenic control regions. *J Biol Chem*, **258**, 10395-10402.

Donze, D., Adams, C.R., Rine, J. and Kamakaka, R.T. (1999) The boundaries of the silenced HMR domain in Saccharomyces cerevisiae. *Genes Dev*, **13**, 698-708.

Donze, D. and Kamakaka, R.T. (2001) RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae. *Embo J*, **20**, 520-531.

Eissenberg, J.C., Elgin, S.C.R., and Paro, R. (1995) Epigenetic regulation in *Drosophila*: a conspiracy of silence. In.

Emerson, B.M. and Roeder, R.G. (1984) DNA sequences and transcription factor interactions of active and inactive forms of mammalian 5 S RNA genes. *J Biol Chem*, **259**, 7926-7935.

Engelke, D.R., Ng, S.Y., Shastry, B.S. and Roeder, R.G. (1980) Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell*, **19**, 717-728.

Eriksson, P., Biswas, D., Yu, Y., Stewart, J.M. and Stillman, D.J. (2004) TATA-binding protein mutants that are lethal in the absence of the Nhp6 high-mobility-group protein. *Mol Cell Biol*, **24**, 6419-6429.

Eschenlauer, J.B., Kaiser, M.W., Gerlach, V.L. and Brow, D.A. (1993) Architecture of a yeast U6 RNA gene promoter. *Mol Cell Biol*, **13**, 3015-3026.

Fabrizio, P., Coppo, A., Fruscoloni, P., Benedetti, P., Di Segni, G. and Tocchini-Valentini, G.P. (1987) Comparative mutational analysis of wild-type and stretched tRNA3(Leu) gene promoters. *Proc Natl Acad Sci U S A*, **84**, 8763-8767.

Farmer, G., Bargonetti, J., Zhu, H., Friedman, P., Prywes, R. and Prives, C. (1992) Wild-type p53 activates transcription in vitro. *Nature*, **358**, 83-86.

Feng, H.P., Scherl, D.S. and Widom, J. (1993) Lifetime of the histone octamer studied by continuous-flow quasielastic light scattering: test of a model for nucleosome transcription. *Biochemistry*, **32**, 7824-7831.

Ferrari, R., Rivetti, C., Acker, J. and Dieci, G. (2004) Distinct roles of transcription factors TFIIIB and TFIIIC in RNA polymerase III transcription reinitiation. *Proc Natl Acad Sci U S A*, **101**, 13442-13447.

Fields, S. and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature*, **340**, 245-246.

Flaus, A. and Owen-Hughes, T. (2001) Mechanisms for ATP-dependent chromatin remodelling. *Curr Opin Genet Dev*, **11**, 148-154.

Formosa, T., Eriksson, P., Wittmeyer, J., Ginn, J., Yu, Y. and Stillman, D.J. (2001) Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN. *Embo J*, **20**, 3506-3517.

Formosa, T., Ruone, S., Adams, M.D., Olsen, A.E., Eriksson, P., Yu, Y., Rhoades, A.R., Kaufman, P.D. and Stillman, D.J. (2002) Defects in SPT16 or POB3 (yFACT) in Saccharomyces cerevisiae cause dependence on the Hir/Hpc pathway: polymerase passage may degrade chromatin structure. *Genetics*, **162**, 1557-1571.

Fruscoloni, P., Zamboni, M., Panetta, G., De Paolis, A. and Tocchini-Valentini, G.P. (1995) Mutational analysis of the transcription start site of the yeast tRNA(Leu3) gene. *Nucleic Acids Res*, **23**, 2914-2918.

Galli, G., Hofstetter, H. and Birnstiel, M.L. (1981) Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. *Nature*, **294**, 626-631.

Geiduschek, E.P. and Kassavetis, G.A. (2001) The RNA polymerase III transcription apparatus. *J Mol Biol*, **310**, 1-26.

Geiduschek, E.P. and Tocchini-Valentini, G.P. (1988) Transcription by RNA polymerase III. *Annu Rev Biochem*, **57**, 873-914.

Gerasimova, T.I. and Corces, V.G. (1996) Boundary and insulator elements in chromosomes. *Curr Opin Genet Dev*, **6**, 185-192.

Gerlach, V.L., Whitehall, S.K., Geiduschek, E.P. and Brow, D.A. (1995) TFIIIB placement on a yeast U6 RNA gene in vivo is directed primarily by TFIIIC rather than by sequence-specific DNA contacts. *Mol Cell Biol*, **15**, 1455-1466.

Geyer, P.K. (1997) The role of insulator elements in defining domains of gene expression. *Curr Opin Genet Dev*, **7**, 242-248.

Giuliodori, S., Percudani, R., Braglia, P., Ferrari, R., Guffanti, E., Ottonello, S. and Dieci, G. (2003) A composite upstream sequence motif potentiates tRNA gene transcription in yeast. *J Mol Biol*, **333**, 1-20.

Gottlieb, E. and Steitz, J.A. (1989a) Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III. *Embo J*, **8**, 851-861.

Gottlieb, E. and Steitz, J.A. (1989b) The RNA binding protein La influences both the accuracy and the efficiency of RNA polymerase III transcription in vitro. *Embo J*, **8**, 841-850.

Gottschling, D.E., Aparicio, O.M., Billington, B.L. and Zakian, V.A. (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. *Cell*, **63**, 751-762.

Gribnau, J., Diderich, K., Pruzina, S., Calzolari, R. and Fraser, P. (2000) Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. *Mol Cell*, **5**, 377-386.

Grove, A., Kassavetis, G.A., Johnson, T.E. and Geiduschek, E.P. (1999) The RNA polymerase III-recruiting factor TFIIIB induces a DNA bend between the TATA box and the transcriptional start site. *J Mol Biol*, **285**, 1429-1440.

Guffanti, E., Corradini, R., Ottonello, S. and Dieci, G. (2004) Functional dissection of RNA polymerase III termination using a peptide nucleic acid as a transcriptional roadblock. *J Biol Chem*, **279**, 20708-20716.

Gunnery, S., Ma, Y. and Mathews, M.B. (1999) Termination sequence requirements vary among genes transcribed by RNA polymerase III. *J Mol Biol*, **286**, 745-757.

Haber, J.E. (1998) Mating-type gene switching in Saccharomyces cerevisiae. *Annu Rev Genet*, **32**, 561-599.

Haffner, R. and Oren, M. (1995) Biochemical properties and biological effects of p53. *Curr Opin Genet Dev*, **5**, 84-90.

Hamada, M., Huang, Y., Lowe, T.M. and Maraia, R.J. (2001) Widespread use of TATA elements in the core promoters for RNA polymerases III, II, and I in fission yeast. *Mol Cell Biol*, **21**, 6870-6881.

Hamada, M., Sakulich, A.L., Koduru, S.B. and Maraia, R.J. (2000) Transcription termination by RNA polymerase III in fission yeast. A genetic and biochemically tractable model system. *J Biol Chem*, **275**, 29076-29081.

Han, M. and Grunstein, M. (1988) Nucleosome loss activates yeast downstream promoters in vivo. *Cell*, **55**, 1137-1145.

Hani, J. and Feldmann, H. (1998) tRNA genes and retroelements in the yeast genome. *Nucleic Acids Res*, **26**, 689-696.

Harismendy, O., Gendrel, C.G., Soularue, P., Gidrol, X., Sentenac, A., Werner, M. and Lefebvre, O. (2003) Genome-wide location of yeast RNA polymerase III transcription machinery. *Embo J*, **22**, 4738-4747.

Hartzog, G.A., Speer, J.L. and Lindstrom, D.L. (2002) Transcript elongation on a nucleoprotein template. *Biochim Biophys Acta*, **1577**, 276-286.

Hebbes, T.R., Clayton, A.L., Thorne, A.W. and Crane-Robinson, C. (1994) Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *Embo J*, **13**, 1823-1830.

Hernandez, N. and Lucito, R. (1988) Elements required for transcription initiation of the human U2 snRNA gene coincide with elements required for snRNA 3' end formation. *Embo J*, **7**, 3125-3134.

Hess, J., Perez-Stable, C., Wu, G.J., Weir, B., Tinoco, I., Jr. and Shen, C.K. (1985) End-to-end transcription of an Alu family repeat. A new type of polymerase-III-dependent terminator and its evolutionary implication. *J Mol Biol*, **184**, 7-21.

Higuchi, R. (1990) PCR Protocols:A Guide to Methods and Applications. In Academic Press, I. (ed.), San Diego, pp. 177-183.

Hofstetter, H., Kressman, A. and Birnstiel, M.L. (1981) A split promoter for a eucaryotic tRNA gene. *Cell*, **24**, 573-585.

Hollstein, M., Sidransky, D., Vogelstein, B. and Harris, C.C. (1991) p53 mutations in human cancers. *Science*, **253**, 49-53.

Holmes, S.F. and Erie, D.A. (2003) Downstream DNA sequence effects on transcription elongation. Allosteric binding of nucleoside triphosphates facilitates translocation via a ratchet motion. *J Biol Chem*, **278**, 35597-35608.

Huang, W., Cao, Z., Han, Z., Li, M. and Wang, D. (1996) [Monoclonal antibodies to human progesterone receptors: production, indentification and clinical application]. *Hua Xi Yi Ke Da Xue Xue Bao*, **27**, 40-45.

Huang, Y., Intine, R.V., Mozlin, A., Hasson, S. and Maraia, R.J. (2005) Mutations in the RNA polymerase III subunit Rpc11p that decrease RNA 3' cleavage activity increase 3'-terminal oligo(U) length and La-dependent tRNA processing. *Mol Cell Biol*, **25**, 621-636.

Huet, J., Manaud, N., Dieci, G., Peyroche, G., Conesa, C., Lefebvre, O., Ruet, A., Riva, M. and Sentenac, A. (1996) RNA polymerase III and class III transcription factors from Saccharomyces cerevisiae. *Methods Enzymol*, **273**, 249-267.

Huet, J. and Sentenac, A. (1992) The TATA-binding protein participates in TFIIIB assembly on tRNA genes. *Nucleic Acids Res*, **20**, 6451-6454.

Intine, R.V., Dundr, M., Misteli, T. and Maraia, R.J. (2002) Aberrant nuclear trafficking of La protein leads to disordered processing of associated precursor tRNAs. *Mol Cell*, **9**, 1113-1123.

Izban, M.G. and Luse, D.S. (1991) Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. *Genes Dev*, **5**, 683-696.

Jahn, D., Wingender, E. and Seifart, K.H. (1987) Transcription complexes for various class III genes differ in parameters of formation and stability towards salt. *J Mol Biol*, **193**, 303-313.

Jambunathan, N., Martinez, A.W., Robert, E.C., Agochukwu, N.B., Ibos, M.E., Dugas, S.L. and Donze, D. (2005) Multiple Bromodomain Genes Are Involved in Restricting the Spread of Heterochromatic Silencing at the Saccharomyces cerevisiae HMR-tRNA Boundary. *Genetics*, **171**, 913-922.

Joazeiro, C.A., Kassavetis, G.A. and Geiduschek, E.P. (1996) Alternative outcomes in assembly of promoter complexes: the roles of TBP and a flexible linker in placing TFIIIB on tRNA genes. *Genes Dev*, **10**, 725-739.

Johnson, J.D. and Raymond, G.J. (1984) Three regions of a yeast tRNALeu3 gene promote RNA polymerase III transcription. *J Biol Chem*, **259**, 5990-5994.

Juo, Z.S., Chiu, T.K., Leiberman, P.M., Baikalov, I., Berk, A.J. and Dickerson, R.E. (1996) How proteins recognize the TATA box. *J Mol Biol*, **261**, 239-254.

Juo, Z.S., Kassavetis, G.A., Wang, J., Geiduschek, E.P. and Sigler, P.B. (2003) Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature*, **422**, 534-539.

Kaiser, M.W. and Brow, D.A. (1995) Lethal mutations in a yeast U6 RNA gene B block promoter element identify essential contacts with transcription factor-IIIC. *J Biol Chem*, **270**, 11398-11405.

Kaplan, C.D., Laprade, L. and Winston, F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science*, **301**, 1096-1099.

Kassavetis, G.A., Bardeleben, C., Kumar, A., Ramirez, E. and Geiduschek, E.P. (1997) Domains of the Brf component of RNA polymerase III transcription factor IIIB (TFIIIB): functions in assembly of TFIIIB-DNA complexes and recruitment of RNA polymerase to the promoter. *Mol Cell Biol*, **17**, 5299-5306.

Kassavetis, G.A., Bartholomew, B., Blanco, J.A., Johnson, T.E. and Geiduschek, E.P. (1991) Two essential components of the Saccharomyces cerevisiae transcription factor TFIIIB: transcription and DNA-binding properties. *Proc Natl Acad Sci U S A*, **88**, 7308-7312.

Kassavetis, G.A., Blanco, J.A., Johnson, T.E. and Geiduschek, E.P. (1992a) Formation of open and elongating transcription complexes by RNA polymerase III. *J Mol Biol*, **226**, 47-58.

Kassavetis, G.A., Braun, B.R., Nguyen, L.H. and Geiduschek, E.P. (1990) S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell*, **60**, 235-245.

Kassavetis, G.A., Joazeiro, C.A., Pisano, M., Geiduschek, E.P., Colbert, T., Hahn, S. and Blanco, J.A. (1992b) The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIIB. *Cell*, **71**, 1055-1064.

Kassavetis, G.A., Kumar, A., Letts, G.A. and Geiduschek, E.P. (1998a) A post-recruitment function for the RNA polymerase III transcription-initiation factor IIIB. *Proc Natl Acad Sci U S A*, **95**, 9196-9201.

Kassavetis, G.A., Kumar, A., Ramirez, E. and Geiduschek, E.P. (1998b) Functional and structural organization of Brf, the TFIIB-related component of the RNA polymerase III transcription initiation complex. *Mol Cell Biol*, **18**, 5587-5599.

Kassavetis, G.A., Letts, G.A. and Geiduschek, E.P. (1999) A minimal RNA polymerase III transcription system. *Embo J*, **18**, 5042-5051.

Kassavetis, G.A., Nguyen, S.T., Kobayashi, R., Kumar, A., Geiduschek, E.P. and Pisano, M. (1995) Cloning, expression, and function of TFC5, the gene encoding the B" component of the Saccharomyces cerevisiae RNA polymerase III transcription factor TFIIIB. *Proc Natl Acad Sci U S A*, **92**, 9786-9790.

Kassavetis, G.A., Riggs, D.L., Negri, R., Nguyen, L.H. and Geiduschek, E.P. (1989) Transcription factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Mol Cell Biol*, **9**, 2551-2566.

Kassavetis, G.A., Soragni, E., Driscoll, R. and Geiduschek, E.P. (2005) Reconfiguring the connectivity of a multiprotein complex: fusions of yeast TATA-binding protein with Brf1, and the function of transcription factor IIIB. *Proc Natl Acad Sci U S A*, **102**, 15406-15411.

Kassavetis, G.A. and Steiner, D.F. (2006) NHP6 is a transcriptional initiation fidelity factor for RNA polymerase III transcription in vitro and in vivo. *J Biol Chem*.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B. and Lander, E.S. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*, **423**, 241-254.

Kern, S.E., Kinzler, K.W., Bruskin, A., Jarosz, D., Friedman, P., Prives, C. and Vogelstein, B. (1991) Identification of p53 as a sequence-specific DNA-binding protein. *Science*, **252**, 1708-1711.

Kireeva, M.L., Komissarova, N., Waugh, D.S. and Kashlev, M. (2000) The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. *J Biol Chem*, **275**, 6530-6536.

Kireeva, M.L., Walter, W., Tchernajenko, V., Bondarenko, V., Kashlev, M. and Studitsky, V.M. (2002) Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Mol Cell*, **9**, 541-552.

Klemenz, R., Stillman, D.J. and Geiduschek, E.P. (1982) Specific interactions of Saccharomyces cerevisiae proteins with a promoter region of eukaryotic tRNA genes. *Proc Natl Acad Sci U S A*, **79**, 6191-6195.

Ko, L.J. and Prives, C. (1996) p53: puzzle and paradigm. *Genes Dev*, **10**, 1054-1072.

Kolodrubetz, D., Kruppa, M. and Burgum, A. (2001) Gene dosage affects the expression of the duplicated NHP6 genes of Saccharomyces cerevisiae. *Gene*, **272**, 93-101.

Kornberg, R.D. and Lorch, Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, **98**, 285-294.

Krieg, R., Stucka, R., Clark, S. and Feldmann, H. (1991) The use of a synthetic tRNA gene as a novel approach to study in vivo transcription and chromatin structure in yeast. *Nucleic Acids Res*, **19**, 3849-3855.

Krol, A., Carbon, P., Ebel, J.P. and Appel, B. (1987) Xenopus tropicalis U6 snRNA genes transcribed by Pol III contain the upstream promoter elements used by Pol II dependent U snRNA genes. *Nucleic Acids Res*, **15**, 2463-2478.

Kruppa, M., Moir, R.D., Kolodrubetz, D. and Willis, I.M. (2001) Nhp6, an HMG1 protein, functions in SNR6 transcription by RNA polymerase III in S. cerevisiae. *Mol Cell*, **7**, 309-318.

Kundu, T.K., Wang, Z. and Roeder, R.G. (1999) Human TFIIIC relieves chromatin-mediated repression of RNA polymerase III transcription and contains an intrinsic histone acetyltransferase activity. *Mol Cell Biol*, **19**, 1605-1615.

Kunkel, G.R. and Pederson, T. (1988) Upstream elements required for efficient transcription of a human U6 RNA gene resemble those of U1 and U2 genes even though a different polymerase is used. *Genes Dev*, **2**, 196-204.

Kunkel, G.R. and Pederson, T. (1989) Transcription of a human U6 small nuclear RNA gene in vivo withstands deletion of intragenic sequences but not of an upstream TATATA box. *Nucleic Acids Res*, **17**, 7371-7379.

Lamond, A.I. and Earnshaw, W.C. (1998) Structure and function in the nucleus. *Science*, **280**, 547-553.

Landrieux, E., Alic, N., Ducrot, C., Acker, J., Riva, M. and Carles, C. (2005) A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation.

Lassar, A.B., Martin, P.L. and Roeder, R.G. (1983) Transcription of class III genes: formation of preinitiation complexes. *Science*, **222**, 740-748.

Lee, J.Y., Evans, C.F. and Engelke, D.R. (1991) Expression of RNase P RNA in Saccharomyces cerevisiae is controlled by an unusual RNA polymerase III promoter. *Proc Natl Acad Sci U S A*, **88**, 6986-6990.

Lee, Y., Wong, W.M., Guyer, D., Erkine, A.M. and Nazar, R.N. (1997) In vivo analyses of upstream promoter sequence elements in the 5 S rRNA gene from Saccharomyces cerevisiae. *J Mol Biol*, **269**, 676-683.

Leveillard, T., Kassavetis, G.A. and Geiduschek, E.P. (1991) Saccharomyces cerevisiae transcription factors IIIB and IIIC bend the DNA of a tRNA(Gln) gene. *J Biol Chem*, **266**, 5162-5168.

Lin, R., Leone, J.W., Cook, R.G. and Allis, C.D. (1989) Antibodies specific to acetylated histones document the existence of deposition- and transcription-related histone acetylation in Tetrahymena. *J Cell Biol*, **108**, 1577-1588.

Lobo, S.M. and Hernandez, N. (1989) A 7 bp mutation converts a human RNA polymerase II snRNA promoter into an RNA polymerase III promoter. *Cell*, **58**, 55-67.

Lobo, S.M., Lister, J., Sullivan, M.L. and Hernandez, N. (1991) The cloned RNA polymerase II transcription factor IID selects RNA polymerase III to transcribe the human U6 gene in vitro. *Genes Dev*, **5**, 1477-1489.

Lobo, S.M., Tanaka, M., Sullivan, M.L. and Hernandez, N. (1992) A TBP complex essential for transcription from TATA-less but not TATA-containing RNA polymerase III promoters is part of the TFIIIB fraction. *Cell*, **71**, 1029-1040.

Loo, S. and Rine, J. (1995) Silencing and heritable domains of gene expression. *Annu Rev Cell Dev Biol*, **11**, 519-548.

Lopez, S., Livingstone-Zatchej, M., Jourdain, S., Thoma, F., Sentenac, A. and Marsolier, M.C. (2001) High-mobility-group proteins NHP6A and NHP6B participate in activation of the RNA polymerase III SNR6 gene. *Mol Cell Biol*, **21**, 3096-3104.

Lopez-De-Leon, A., Librizzi, M., Puglia, K. and Willis, I.M. (1992) PCF4 encodes an RNA polymerase III transcription factor with homology to TFIIB. *Cell*, **71**, 211-220.

Lowe, T.M. and Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*, **25**, 955-964.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251-260.

Lustig, A.J. (1998) Mechanisms of silencing in Saccharomyces cerevisiae. *Curr Opin Genet Dev*, **8**, 233-239.

Ma, B. and Hernandez, N. (2001) A map of protein-protein contacts within the small nuclear RNA-activating protein complex SNAPc. *J Biol Chem*, **276**, 5027-5035.

Mahy, N.L., Perry, P.E. and Bickmore, W.A. (2002) Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol*, **159**, 753-763.

Mao, J., Schmidt, O. and Soll, D. (1980) Dimeric transfer RNA precursors in S. pombe. *Cell*, **21**, 509-516.

Maraia, R.J. and Intine, R.V. (2002) La protein and its associated small nuclear and nucleolar precursor RNAs. *Gene Expr*, **10**, 41-57.

Margottin, F., Dujardin, G., Gerard, M., Egly, J.M., Huet, J. and Sentenac, A. (1991) Participation of the TATA factor in transcription of the yeast U6 gene by RNA polymerase C. *Science*, **251**, 424-426.

Marsolier, M.C., Tanaka, S., Livingstone-Zatchej, M., Grunstein, M., Thoma, F. and Sentenac, A. (1995) Reciprocal interferences between nucleosomal organization and transcriptional activity of the yeast SNR6 gene. *Genes Dev*, **9**, 410-422.

Martin, M.P., Gerlach, V.L. and Brow, D.A. (2001) A novel upstream RNA polymerase III promoter element becomes essential when the chromatin structure of the yeast U6 RNA gene is altered. *Mol Cell Biol*, **21**, 6429-6439.

Marzouki, N., Camier, S., Ruet, A., Moenne, A. and Sentenac, A. (1986) Selective proteolysis defines two DNA binding domains in yeast transcription factor tau. *Nature*, **323**, 176-178.

Mathews, C.K. (1972) Biochemistry of deoxyribonucleic acid-defective amber mutants of bacteriophage T4. 3. Nucleotide pools. *J Biol Chem*, **247**, 7430-7438.

Matsuzaki, H., Kassavetis, G.A. and Geiduschek, E.P. (1994) Analysis of RNA chain elongation and termination by Saccharomyces cerevisiae RNA polymerase III. *J Mol Biol*, **235**, 1173-1192.

Mattaj, I.W., Dathan, N.A., Parry, H.D., Carbon, P. and Krol, A. (1988) Changing the RNA polymerase specificity of U snRNA gene promoters. *Cell*, **55**, 435-442.

Mazabraud, A., Scherly, D., Muller, F., Rungger, D. and Clarkson, S.G. (1987) Structure and transcription termination of a lysine tRNA gene from Xenopus laevis. *J Mol Biol*, **195**, 835-845.

Meyers, R.E. and Sharp, P.A. (1993) TATA-binding protein and associated factors in polymerase II and polymerase III transcription. *Mol Cell Biol*, **13**, 7953-7960.

Mitchell, M.T., Hobson, G.M. and Benfield, P.A. (1992) TATA box-mediated polymerase III transcription in vitro. *J Biol Chem*, **267**, 1995-2005.

Mittal, V. and Hernandez, N. (1997) Role for the amino-terminal region of human TBP in U6 snRNA transcription. *Science*, **275**, 1136-1140.

Moenne, A., Camier, S., Anderson, G., Margottin, F., Beggs, J. and Sentenac, A. (1990) The U6 gene of Saccharomyces cerevisiae is transcribed by RNA polymerase C (III) in vivo and in vitro. *Embo J*, **9**, 271-277.

Moir, R.D., Sethy-Coraci, I., Puglia, K., Librizzi, M.D. and Willis, I.M. (1997) A tetratricopeptide repeat mutation in yeast transcription factor IIIC131 (TFIIIC131) facilitates recruitment of TFIIB-related factor TFIIIB70. *Mol Cell Biol*, **17**, 7119-7125.

Moqtaderi, Z. and Struhl, K. (2004) Genome-wide occupancy profile of the RNA polymerase III machinery in Saccharomyces cerevisiae reveals loci with incomplete transcription complexes. *Mol Cell Biol*, **24**, 4118-4127.

Morse, R.H., Roth, S.Y. and Simpson, R.T. (1992) A transcriptionally active tRNA gene interferes with nucleosome positioning in vivo. *Mol Cell Biol*, **12**, 4015-4025.

Murphy, S., Tripodi, M. and Melli, M. (1986) A sequence upstream from the coding region is required for the transcription of the 7SK RNA genes. *Nucleic Acids Res*, **14**, 9243-9260.

Murphy, S., Yoon, J.B., Gerster, T. and Roeder, R.G. (1992) Oct-1 and Oct-2 potentiate functional interactions of a transcription factor with the proximal sequence element of small nuclear RNA genes. *Mol Cell Biol*, **12**, 3247-3261.

Nacheva, G.A., Guschin, D.Y., Preobrazhenskaya, O.V., Karpov, V.L., Ebralidse, K.K. and Mirzabekov, A.D. (1989) Change in the pattern of histone binding to DNA upon transcriptional activation. *Cell*, **58**, 27-36.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger, J.H., Yan, H. and Burton, Z.F. (2003) NTP-driven translocation by human RNA polymerase II. *J Biol Chem*, **278**, 18303-18312.

Ng, H.H., Robert, F., Young, R.A. and Struhl, K. (2002) Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. *Genes Dev*, **16**, 806-819.

Nguyen, V.T., Kiss, T., Michels, A.A. and Bensaude, O. (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*, **414**, 322-325.

Novy, R., Drott, D., Yaeger, K. and Mierendorf, R. (2001) *inNovations*, **12**, 1-3.

Nudler, E. and Gottesman, M.E. (2002) Transcription termination and anti-termination in E. coli. *Genes Cells*, **7**, 755-768.

O'Donohue, M.F., Duband-Goulet, I., Hamiche, A. and Prunell, A. (1994) Octamer displacement and redistribution in transcription of single nucleosomes. *Nucleic Acids Res*, **22**, 937-945.

Ong, W.C., Ibrahim, M., Town, M. and Johnson, J.D. (1997) Functional differences among the six Saccharomyces cerevisiae tRNATrp genes. *Yeast*, **13**, 1357-1362.

Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S. and Reinberg, D. (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell*, **92**, 105-116.

Orphanides, G., Wu, W.H., Lane, W.S., Hampsey, M. and Reinberg, D. (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature*, **400**, 284-288.

Ouyang, C., Martinez, M.J., Young, L.S. and Sprague, K.U. (2000) TATA-Binding protein-TATA interaction is a key determinant of differential transcription of silkworm constitutive and silk gland-specific tRNA(Ala) genes. *Mol Cell Biol*, **20**, 1329-1343.

Parsons, J.D. (1995) Improved tools for DNA comparison and clustering. *Comput Appl Biosci*, **11**, 603-613.

Partridge, J.F., Borgstrom, B. and Allshire, R.C. (2000) Distinct protein interaction domains and protein spreading in a complex centromere. *Genes Dev*, **14**, 783-791.

Percudani, R., Pavesi, A. and Ottonello, S. (1997) Transfer RNA gene redundancy and translational selection in Saccharomyces cerevisiae. *J Mol Biol*, **268**, 322-330.

Pieler, T., Appel, B., Oei, S.L., Mentzel, H. and Erdmann, V.A. (1985a) Point mutational analysis of the Xenopus laevis 5S gene promoter. *Embo J*, **4**, 1847-1853.

Pieler, T., Digweed, M. and Erdmann, V.A. (1985b) RNA structural dynamics: pre-melting and melting transitions in E. coli 5S rRNA. *J Biomol Struct Dyn*, **3**, 495-514.

Pieler, T., Oei, S.L., Hamm, J., Engelke, U. and Erdmann, V.A. (1985c) Functional domains of the Xenopus laevis 5S gene promoter. *Embo J*, **4**, 3751-3756.

Pruss, D., Hayes, J.J. and Wolffe, A.P. (1995) Nucleosomal anatomy--where are the histones? *Bioessays*, **17**, 161-170.

Rameau, G., Puglia, K., Crowe, A., Sethy, I. and Willis, I. (1994) A mutation in the second largest subunit of TFIIIC increases a rate-limiting step in transcription by RNA polymerase III. *Mol Cell Biol*, **14**, 822-830.

Ravindra, A., Weiss, K. and Simpson, R.T. (1999) High-resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating-type locus HMRa. *Mol Cell Biol*, **19**, 7944-7950.

Raymond, G.J. and Johnson, J.D. (1983) The role of non-coding DNA sequences in transcription and processing of a yeast tRNA. *Nucleic Acids Res*, **11**, 5969-5988.

Raymond, K.C., Raymond, G.J. and Johnson, J.D. (1985) In vivo modulation of yeast tRNA gene expression by 5'-flanking sequences. *Embo J*, **4**, 2649-2656.

Reynolds, R., Bermudez-Cruz, R.M. and Chamberlin, M.J. (1992) Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators. *J Mol Biol*, **224**, 31-51.

Rhoades, A.R., Ruone, S. and Formosa, T. (2004) Structural features of nucleosomes reorganized by yeast FACT and its HMG box component, Nhp6. *Mol Cell Biol*, **24**, 3907-3917.

Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. and Klug, A. (1984) Structure of the nucleosome core particle at 7 A resolution. *Nature*, **311**, 532-537.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol*, **17**, 1030-1032.

Roberts, D.N., Stewart, A.J., Huff, J.T. and Cairns, B.R. (2003) The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proc Natl Acad Sci U S A*, **100**, 14695-14700.

Roberts, S., Colbert, T. and Hahn, S. (1995) TFIIIC determines RNA polymerase III specificity at the TATA-containing yeast U6 promoter. *Genes Dev*, **9**, 832-842.

Ruone, S., Rhoades, A.R. and Formosa, T. (2003) Multiple Nhp6 molecules are required to recruit Spt16-Pob3 to form yFACT complexes and to reorganize nucleosomes. *J Biol Chem*, **278**, 45288-45295.

Ruth, J., Conesa, C., Dieci, G., Lefebvre, O., Dusterhoft, A., Ottonello, S. and Sentenac, A. (1996) A suppressor of mutations in the class III transcription system encodes a component of yeast TFIIIB. *Embo J*, **15**, 1941-1949.

Sadowski, C.L., Henry, R.W., Lobo, S.M. and Hernandez, N. (1993) Targeting TBP to a non-TATA box cis-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE. *Genes Dev*, **7**, 1535-1548.

Sathyanarayana, U.G., Freeman, L.A., Lee, M.S. and Garrard, W.T. (1999) RNA polymerase-specific nucleosome disruption by transcription in vivo. *J Biol Chem*, **274**, 16431-16436.

Saunders, A., Werner, J., Andrulis, E.D., Nakayama, T., Hirose, S., Reinberg, D. and Lis, J.T. (2003) Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. *Science*, **301**, 1094-1096.

Schmidt, O., Mao, J., Ogden, R., Beckmann, J., Sakano, H., Abelson, J. and Soll, D. (1980) Dimeric tRNA precursors in yeast. *Nature*, **287**, 750-752.

Schnell, R. and Rine, J. (1986) A position effect on the expression of a tRNA gene mediated by the SIR genes in Saccharomyces cerevisiae. *Mol Cell Biol*, **6**, 494-501.

Schramm, L. and Hernandez, N. (2002) Recruitment of RNA polymerase III to its target promoters. *Genes Dev*, **16**, 2593-2620.

Schramm, L., Pendergrast, P.S., Sun, Y. and Hernandez, N. (2000) Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes Dev*, **14**, 2650-2663.

Segall, J., Matsui, T. and Roeder, R.G. (1980) Multiple factors are required for the accurate transcription of purified genes by RNA polymerase III. *J Biol Chem*, **255**, 11986-11991.

Sethy, I., Moir, R.D., Librizzi, M. and Willis, I.M. (1995) In vitro evidence for growth regulation of tRNA gene transcription in yeast. A role for transcription factor (TF) IIIB70 and TFIIIC. *J Biol Chem*, **270**, 28463-28470.

Sethy-Coraci, I., Moir, R.D., Lopez-de-Leon, A. and Willis, I.M. (1998) A differential response of wild type and mutant promoters to TFIIIB70 overexpression in vivo and in vitro. *Nucleic Acids Res*, **26**, 2344-2352.

Setzer, D.R. and Brown, D.D. (1985) Formation and stability of the 5 S RNA transcription complex. *J Biol Chem*, **260**, 2483-2492.

Shaaban, S.A., Bobkova, E.V., Chudzik, D.M. and Hall, B.D. (1996) In vitro analysis of elongation and termination by mutant RNA polymerases with altered termination behavior. *Mol Cell Biol*, **16**, 6468-6476.

Sharp, S., DeFranco, D., Dingermann, T., Farrell, P. and Soll, D. (1981) Internal control regions for transcription of eukaryotic tRNA genes. *Proc Natl Acad Sci U S A*, **78**, 6657-6661.

Shaw, K.J. and Olson, M.V. (1984) Effects of altered 5'-flanking sequences on the in vivo expression of a Saccharomyces cerevisiae tRNATyr gene. *Mol Cell Biol*, **4**, 657-665.

Shivaswamy, S., Kassavetis, G.A. and Bhargava, P. (2004) High-level activation of transcription of the yeast U6 snRNA gene in chromatin by the basal RNA polymerase III transcription factor TFIIIC. *Mol Cell Biol*, **24**, 3596-3606.

Smith, J.S. and Boeke, J.D. (1997) An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev*, **11**, 241-254.

Sprague, K.U., Larson, D. and Morton, D. (1980) 5' flanking sequence signals are required for activity of silkworm alanine tRNA genes in homologous in vitro transcription systems. *Cell*, **22**, 171-178.

Stajich, J.E., Block, D., Boulez, K., Brenner, S.E., Chervitz, S.A., Dagdigian, C., Fuellen, G., Gilbert, J.G., Korf, I., Lapp, H., Lehvaslaiho, H., Matsalla, C., Mungall, C.J., Osborne, B.I., Pocock, M.R., Schattner, P., Senger, M., Stein, L.D., Stupka, E., Wilkinson, M.D. and Birney, E. (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res*, **12**, 1611-1618.

Sterner, D.E. and Berger, S.L. (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev*, **64**, 435-459.

Stone, E.M. and Pillus, L. (1998) Silent chromatin in yeast: an orchestrated medley featuring Sir3p [corrected]. *Bioessays*, **20**, 30-40.

Strubin, M. and Struhl, K. (1992) Yeast and human TFIID with altered DNA-binding specificity for TATA elements. *Cell*, **68**, 721-730.

Studitsky, V.M., Clark, D.J. and Felsenfeld, G. (1994) A histone octamer can step around a transcribing polymerase without leaving the template. *Cell*, **76**, 371-382.

Studitsky, V.M., Kassavetis, G.A., Geiduschek, E.P. and Felsenfeld, G. (1997) Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. *Science*, **278**, 1960-1963.

Sun, F.L., Cuaycong, M.H., Craig, C.A., Wallrath, L.L., Locke, J. and Elgin, S.C. (2000) The fourth chromosome of Drosophila melanogaster: interspersed euchromatic and heterochromatic domains. *Proc Natl Acad Sci U S A*, **97**, 5340-5345.

Sun, F.L. and Elgin, S.C. (1999) Putting boundaries on silence. *Cell*, **99**, 459-462.

Tackett, A.J., Dilworth, D.J., Davey, M.J., O'Donnell, M., Aitchison, J.D., Rout, M.P. and Chait, B.T. (2005) Proteomic and genomic characterization of chromatin complexes at a boundary. *J Cell Biol*, **169**, 35-47.

Taggart, A.K., Fisher, T.S. and Pugh, B.F. (1992) The TATA-binding protein and associated factors are components of pol III transcription factor TFIIIB. *Cell*, **71**, 1015-1028.

Teichmann, M., Wang, Z. and Roeder, R.G. (2000) A stable complex of a novel transcription factor IIB- related factor, human TFIIIB50, and associated proteins mediate selective transcription by RNA polymerase III of genes with upstream promoter elements. *Proc Natl Acad Sci U S A*, **97**, 14200-14205.

Thuillier, V., Brun, I., Sentenac, A. and Werner, M. (1996) Mutations in the alpha-amanitin conserved domain of the largest subunit of yeast RNA polymerase III affect pausing, RNA cleavage and transcriptional transitions. *Embo J*, **15**, 618-629.

Topper, J.N. and Clayton, D.A. (1990) Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. *Nucleic Acids Res*, **18**, 793-799.

van Holde, K.E.a.Y., T.D. (1985) Nucleosome motion: evidence and models. In.

von Hippel, P.H. (1998) An integrated model of the transcription complex in elongation, termination, and editing. *Science*, **281**, 660-665.

Wakimoto, B.T. and Hearn, M.G. (1990) The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2L of Drosophila melanogaster. *Genetics*, **125**, 141-154.

Waldschmidt, R., Wanandi, I. and Seifart, K.H. (1991) Identification of transcription factors required for the expression of mammalian U6 genes in vitro. *Embo J*, **10**, 2595-2603.

Walter, W., Kireeva, M.L., Studitsky, V.M. and Kashlev, M. (2003) Bacterial polymerase and yeast polymerase II use similar mechanisms for transcription through nucleosomes. *J Biol Chem*, **278**, 36148-36156.

Wang, Y. and Stumph, W.E. (1995) RNA polymerase II/III transcription specificity determined by TATA box orientation. *Proc Natl Acad Sci U S A*, **92**, 8606-8610.

Wang, Z., Bai, L., Hsieh, Y.J. and Roeder, R.G. (2000) Nuclear factor 1 (NF1) affects accurate termination and multiple-round transcription by human RNA polymerase III. *Embo J*, **19**, 6823-6832.

Wang, Z. and Roeder, R.G. (1996) TFIIIC1 acts through a downstream region to stabilize TFIIIC2 binding to RNA polymerase III promoters. *Mol Cell Biol*, **16**, 6841-6850.

Waterman, M.S. and Jones, R. (1990) Consensus methods for DNA and protein sequence alignment. *Methods Enzymol*, **183**, 221-237.

Weiler, K.S. and Wakimoto, B.T. (1995) Heterochromatin and gene expression in Drosophila. *Annu Rev Genet*, **29**, 577-605.

Weiss, K. and Simpson, R.T. (1998) High-resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating type locus HMLalpha. *Mol Cell Biol*, **18**, 5392-5403.

White, R.J. (1998) *RNA Polymerase III Transcription*.

White, R.J., Jackson, S.P. and Rigby, P.W. (1992) A role for the TATA-box-binding protein component of the transcription factor IID complex as a general RNA polymerase III transcription factor. *Proc Natl Acad Sci U S A*, **89**, 1949-1953.

Whitehall, S.K., Kassavetis, G.A. and Geiduschek, E.P. (1995) The symmetry of the yeast U6 RNA gene's TATA box and the orientation of the TATA-binding protein in yeast TFIIIB. *Genes Dev*, **9**, 2974-2985.

Willis, I.M. (2002) A universal nomenclature for subunits of the RNA polymerase III transcription initiation factor TFIIIB. *Genes Dev*, **16**, 1337-1338.

Willoughby, D.A., Vilalta, A. and Oshima, R.G. (2000) An Alu element from the K18 gene confers position-independent expression in transgenic mice. *J Biol Chem*, **275**, 759-768.

Wolffe, A.P. and Pruss, D. (1996) Deviant nucleosomes: the functional specialization of chromatin. *Trends Genet*, **12**, 58-62.

Wolin, S.L. and Cedervall, T. (2002) The La protein. *Annu Rev Biochem*, **71**, 375-403.

Yang, Z., Zhu, Q., Luo, K. and Zhou, Q. (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature*, **414**, 317-322.

Yen, Y.M., Wong, B. and Johnson, R.C. (1998) Determinants of DNA binding and bending by the Saccharomyces cerevisiae high mobility group protein NHP6A that are important for its biological activities. Role of the unique N terminus and putative intercalating methionine. *J Biol Chem*, **273**, 4424-4435.

Yoon, J.B., Murphy, S., Bai, L., Wang, Z. and Roeder, R.G. (1995) Proximal sequence element-binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. *Mol Cell Biol*, **15**, 2019-2027.

Young, L.S., Rivier, D.H. and Sprague, K.U. (1991) Sequences far downstream from the classical tRNA promoter elements bind RNA polymerase III transcription factors. *Mol Cell Biol*, **11**, 1382-1392.

Yukawa, Y., Sugita, M., Choisne, N., Small, I. and Sugiura, M. (2000) The TATA motif, the CAA motif and the poly(T) transcription termination motif are all important for transcription re-initiation on plant tRNA genes. *Plant J*, **22**, 439-447.

Zambetti, G.P., Bargonetti, J., Walker, K., Prives, C. and Levine, A.J. (1992) Wild-type p53 mediates positive regulation of gene expression through a specific DNA sequence element. *Genes Dev*, **6**, 1143-1152.

Zecherle, G.N., Whelen, S. and Hall, B.D. (1996) Purines are required at the 5' ends of newly initiated RNAs for optimal RNA polymerase III gene expression. *Mol Cell Biol*, **16**, 5801-5810.

# 7. APPENDICE

| Costrutto | Plasmide | Primer fw | Primer rev |
|---|---|---|---|
| P(AGG)CR | pGEM | Pro_26_fw | Pro_26_rev |
| N(GTT)CR | pBlueScript KS | Asn_19_plus | Asn_19_minus |
| N(GTT)KL | pGEM | Asn_167_plus | Asn_167_minus |
| N(GTT)OL | pGEM | Asn_239_fw | Asn_239_rev |
| V(AAC)MR2 | pY7 | | |
| L(CAA)LR2 | pGEM | Leu_186_fw | Leu_186_rev |
| S(AGA)EL | pGEM | Ser_aga66_fw | Ser_aga66_rev |
| M(CAT)E | pGEM | Met_cat67_fw | Met_cat67_rev |
| | | | |
| L_3'N | pNEB | Leu_186_fw | Leu_186+T4T13 |
| L_Δ3' | pNEB | Leu_186_fw | Leu_186_term_rev |
| N_3'L | pGEM | Asn_19_plus | Asn_TLeu |
| N_Δ3' | pNEB | Asn_19_plus | Asn_19_T5_rev |
| L(CAA)#4 in fig. 4.4 | pNEB | Leu_186_fw | Leu_186_T13 |
| L(CAA)#5 in fig. 4.4 | pGEM | Leu_186_fw | LeuT4_lont_T10 |
| L(CAA)#6 in fig. 4.4 | pGEM | Leu_186_fw | Leu_CA4CT13_rev |
| L(CAA)#7 in fig. 4.4 | pGEM | Leu_186_fw | Leu_CATATCT13_rev |
| L(CAA)#8 in fig. 4.4 | pGEM | Leu_186_fw | Leu_TATA |
| L(CAA)#9 in fig. 4.4 | pGEM | Leu_186_fw | Leu_+2_T4 |
| L(CAA)#10 in fig. 4.4 | pGEM | Leu_186_fw | Leu_+4_T4 |
| L(CAA)#11 in fig. 4.4 | pGEM | Leu_186_fw | Leu_AT4CT10 |
| L(CAA)#12 in fig. 4.4 | pGEM | Leu_186_fw | Leu_GTTTT |
| L(CAA)#13 in fig. 4.4 | pGEM | Leu_186_fw | Leu_CT4AT10 |
| L(CAA)#14 in fig. 4.4 | pGEM | Leu_186_fw | Leu_CT3CT10 |
| L(CAA)#15 in fig. 4.4 | pGEM | Leu_186_fw | Leu_CT2CT10 |
| L(CAA)#16 in fig. 4.4 | pGEM | Leu_186_fw | Leu_T6+Coda |
| | | | |
| SCR1Δ_T7 | YEp352 | SCR1_TATA | SCR1_del |
| SCR1Δ_3'L | YEp352 | SCR1_plus | SCR_SeqLEU |
| | | | |
| 5'U6-L(CAA)LR2 | pNEB193 -SacI | 5'U6_Sac_Leu186 | Leu_186_rev |
| 5'U6-L_3'N | pNEB193 -SacI | 5'U6_Sac_Leu186 | Leu_186+T4T13 |
| 5'U6-SCR1Δ_3'L | pNEB193 -SacI | U6_SCR1 | SCR_SeqLEU |
| 5'Sac-SCR1Δ_3'L | pNEB193 -SacI | 5'_Sac_SCR | SCR_SeqLEU |
| | | | |
| SNR52_wt | pNEB193 | SNR_52_fw | SNR_52_rev |
| SNR52_mutT | pNEB193 | SNR52_Tdown_fw | SNR52_Tdown_rev |
| 5'U6_SNR52 | pNEB193 | 5'U6_SNR52 | SNR_52_rev |

**Tab. 7.1: Elenco dei costrutti clonati e nome dei *primers* utilizzati, relativo alla sezione *4.1*.**

| Costrutto | Plasmide | Primer fw | Primer rev |
|---|---|---|---|
| N(GTT)CR | pBlueScript KS | Asn_19_plus | Asn_19_minus |
| N(GTT)FL | | Asn_95_plus | Asn_95_minus |
| N(GTT)GR | pBlueScript KS | Asn_107_plus | Asn_107_minus |
| N(GTT)KL | pGEM | Asn_167_plus | Asn_167_minus |
| N(GTT)LR | pGEM | Asn_195_fw | Asn_195_rev |
| N(GTT)NR | pBlueScript KS | Asn_gtt232_plus | Asn_gtt232_rev |
| N(GTT)OL | pGEM | Asn_239_fw | Asn_239_rev |
| N(GTT)OR | pBlueScript KS | Asn_246_plus | Asn_246_minus |
| N(GTT)PR | pGEM | Asn_262_plus | Asn_262_minus |
| | | | |
| tDNA_Syn2 | YEp352 | Glu(TTC)syn2_fw | Glu(TTC)syn2_rev |
| [5'CR]Syn2 | YEp352 ; pFL38 ; pFL46S | Asn19-58 | Asn19_flank_rev |
| [5'NR]Syn2 | YEp352 ; pFL38 ; pFL46S | Asn232_short | Asn232_flank_rev |
| BRF1 | pFL45S | | |
| | | | |
| NHP6A_gene | pFL45S; pFL39 | Nhp6A_fw | Nhp6A_rev |
| NHP6A_ORF | pET28-pme | Nhp6A_orf_fw | Nhp6A_orf_rev |

**Tab. 7.2: Elenco dei costrutti clonati e nome dei *primers* utilizzati, relativo alla sezione *4.2*.**

| Nome oligonucleotide | SEQUENZA |
|---|---|
| **5'_Sac_SCR** | CTGACGAGCTCTGCTAGGCTGTAATGGCTTTCTGG |
| **5'U6_Sac_Leu186** | TTTCGGCTACTATAAATAAATGTTTTTTTCGGAGCTCTTACAACAAATAAGTGG TTGTTTGG |
| **5'U6_SNR52** | TTTCGGCTACTATAAATAAATGTTTTTTTCGCAACTACTCTAGATTTTGTAGTG CCC |
| **Asn_107_minus** | GGAGGATTTCACAAATCTTCC |
| **Asn_107_plus** | GATTACCGGTTGGTGTACTC |
| **Asn_167_minus** | GGTTGATTTGAAGCCATTGGC |
| **Asn_167_plus** | CGTAAATGCTCCAGCTCAAGC |
| **Asn_19_minus** | GATTTTTCCATTCGCCATGC |
| **Asn_19_plus** | CATACTCGAAGGGTAGTTGG |
| **Asn_19_T5_rev** | AAAAATTAACGACCCCAGTGAGG |
| **Asn_195_fw** | GAAGGTCAATGATAGAAGCTATC |
| **Asn_195_rev** | CTATTTTAATGGAGTTTTTAGCTC |
| **Asn_239_fw** | CATTTCTCACAATCTTGTGCATTG |
| **Asn_239_rev** | GACATAGGATAAAAAACGCCATTG |
| **Asn_246_minus** | GAGATAACTCCGACATAGCAAC |
| **Asn_246_plus** | CATTCTTGGGCCTCTCTTGAC |
| **Asn_262_minus** | GCCATAATAATATTAAAAAACTTAACG |
| **Asn_262_plus** | CCGGAGATATGATTCAGATGTGG |
| **Asn_95_minus** | GGAAAAAGATATGATACAATAAAATC |
| **Asn_95_plus** | GGGATGACATTGTTTTACGTAG |
| **Asn_gtt232_plus** | CATGGCAAGGCTTACAGATG |
| **Asn_gtt232_rev** | CATGTAAGTGATAGCGATTTG |
| **Asn_TLeu** | AAAAAAAAAACTTCAAAATTTTCGGAAGGCAAAAATTAACGACCCCAGTGAGG |
| **Asn19_flank_rev** | TATTGCGGATCCTTATCTTGTGTCACTTTTACAG |
| **Asn19-58** | GAAAATTTTCAACATATATTAGTAATAC |
| **Asn232_flank_rev** | TATTGCGGATCCTTAAATTCTGCAGAACGTGAAAG |
| **Asn232_short** | GATTAAAATCGTGTAGATCATGG |
| **Glu(TTC)syn2_fw** | TATGCCGGATCCGATATAGTGTAACGGCTATCACATCACGCTTCAGCAAATAT GCTGAGCCGCAGTAATCTGC |
| **Glu(TTC)syn2_rev** | CGTCGACAAAAAATAACTCCGATACGGGGAGTCGAACCCCGGTCTCCACGGT CCGCAGATTACTGCGGCTCAGC |
| **Glu_ttc_synt_minus** | CGTCGACAAAAAATAACTCCGATACGGGGAGTCGAA |
| **Leu_+2_T4** | AAAAAAAAAAGAAAAGTAAAAATAATGGTTGCTAAGAGATTCG |
| **Leu_+4_T4** | AAAAAAAAAAGAAAAGTGTAAAAATAATGGTTGCTAAGAGATTCG |
| **Leu_186_fw** | GCAAATAAACGTTGGAAATTGGC |
| **Leu_186_rev** | GGTAGGAAAAAAAGATGCTGCAC |
| **Leu_186_T13** | AAAAAAAAAAAAAGAAGGCAAAAATAATGGTTGC |
| **Leu_186_term_rev** | AAAAATAATGGTTGCTAAGAGATTCG |
| **Leu_186+T4T13** | AAAAAAAAAAAAAGAAAAGAAAAATAATGGTTGCTAAGAGATTCG |
| **Leu_AT4CT10** | AAAAAAAAAAGAAAATAAAAATAATGGTTGCTAAGAGATTCG |
| **Leu_CA4CT13_rev** | AAAAAAAAAAAAAGTTTTGAAAAATAATGGTTGCTAAGAGATTCG |
| **Leu_CATATCT13_rev** | AAAAAAAAAAAAAGATATGAAAAATAATGGTTGCTAAGAGATTCG |

| | |
|---|---|
| **Leu_CT2CT10** | AAAAAAAAAAGAAGAAAAATAATGGTTGC |
| **Leu_CT3CT10** | AAAAAAAAAAGAAAGAAAAATAATGGTTGC |
| **Leu_CT4AT10** | AAAAAAAAAATAAAAGAAAAATAATGGTTGC |
| **Leu_GTTTT** | AAAAAAAAAAGAAAACAAAAATAATGGTTG |
| **Leu_T6+Coda** | AAAAAAAAAAAAAGAAAAGAAAAAATAATGGTTGC |
| **Leu_TATA** | AAAAAAAAAAGTATAGAAAAATAATGGTTG |
| **LeuT4_lont_T10** | AAAAAAAAAAGTTCGGAGTCGAAAAGAAAAATAATGGTTGC |
| **Met_cat67_fw** | GCGACATGGAGAGATAGAGATAG |
| **Met_cat67_rev** | CGCAGAAGTTAAAGCTTATAGTC |
| **Nhp6A_fw** | CTCAGAGCTCTGGCAAAAATCGTCCTCTGTTG |
| **Nhp6A_orf_fw** | ATGGTCACCCCAAGAGAACCTAAG |
| **Nhp6A_orf_rev** | CTACTAAGCCAAAGTGGCGTTATATAAC |
| **Nhp6A_rev** | TCATGGATCCAAGAGCTGCACTCGGTCTACG |
| **Pro_26_fw** | GGTCAATGACGATGGCAGGG |
| **Pro_26_rev** | CCTGAGTACACTTGTTTTACGTC |
| **SCR_prext_2** | CTGGCCGAGGAACAAATCCTTCCTCGCGGC |
| **SCR_SeqLEU** | AAAAAAAAAACTTCAAAATTTTCGGAAGGCAAAAACGTGCAATCCGTG |
| **SCR1_del** | AAAAAAACGTGCAATCCGTGTCTAGCCGCG |
| **SCR1_plus** | TGATCAACTTAGCCAGGACATCC |
| **SCR1_TATA** | GTATAAAATCGAAAGTTTATTCCAATTG |
| **Ser_aga66_fw** | GGGAACTCTCAGAACGGGGG |
| **Ser_aga66_rev** | CGAGTTTTTACCCATTTAGGAAAAG |
| **SNR_52_fw** | CTTTGAAAAGATAATGTATGATTATGC |
| **SNR_52_rev** | GCGTTCCATACTGTCAGAGGTG |
| **SNR52_Tdown_fw** | GGTGCGCATTCGTTCACACCCTAC |
| **SNR52_Tdown_rev** | GTAGGGTGTGAACGAATGCGCACC |
| **U6_SCR1** | TTTCGGCTACTATAAATAAATGTTTTTTTCCCAACTAGCTAGGTTGTA |

**Tab. 7.3: Sequenza (5'→3') degli oligonucleotidi utilizzati.**

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 20, Issue of May 20, pp. 19551–19562, 2005
Printed in U.S.A.
© 2005 by The American Society for Biochemistry and Molecular Biology, Inc.

# Sequence Context Effects on Oligo(dT) Termination Signal Recognition by *Saccharomyces cerevisiae* RNA Polymerase III*

Received for publication, October 28, 2004, and in revised form, March 18, 2005
Published, JBC Papers in Press, March 22, 2005, DOI 10.1074/jbc.M412238200

**Priscilla Braglia, Riccardo Percudani, and Giorgio Dieci‡**

*From the Dipartimento di Biochimica e Biologia Molecolare, Università degli Studi di Parma, Parco Area delle Scienze 23/A, 43100 Parma, Italy*

**Eukaryotic RNA polymerase (Pol) III terminates transcription at short runs of T residues in the coding DNA strand. By genomic analysis, we found that $T_5$ and $T_4$ are the shortest Pol III termination signals in yeasts and mammals, respectively, and that, at variance with yeast, oligo(dT) terminators longer than $T_5$ are very rare in mammals. In *Saccharomyces cerevisiae*, the strength of $T_5$ as a terminator was found to be largely influenced by both the upstream and the downstream sequence context. In particular, the CT sequence, which is naturally present downstream of $T_5$ in the 3′-flank of some tDNAs, was found to act as a terminator-weakening element that facilitates translocation by reducing Pol III pausing at $T_5$. In contrast, tDNA transcription termination was highly efficient when $T_5$ was followed by an A or G residue. Surprisingly, however, when a termination-proficient $T_5$ signal was taken out from the tDNA context and placed downstream of a fragment of the *SCR1* gene, its termination activity was compromised, both *in vitro* and *in vivo*. Even the $T_6$ sequence, acting as a strong terminator in tRNA gene contexts, was unexpectedly weak within the *SNR52* transcription unit, where it naturally occurs. The observed sequence context effects reflect intrinsic recognition properties of Pol III, because they were still observed in a simplified *in vitro* transcription system only consisting of purified RNA polymerase and template DNA. Our findings strengthen the notion that termination signal recognition by Pol III is influenced in a complex way by the region surrounding the T cluster and suggest that read-through transcription beyond T clusters might play a significant role in the biogenesis of class III gene products.**

Eukaryotic RNA polymerase (Pol)[1] III is unique among DNA-dependent RNA polymerases in recognizing a simple run of T residues on the coding strand as a termination signal, in the apparent absence of accessory factors. The minimal signal thought to be sufficient to provoke Pol III termination varies among different eukaryotes. $T_4$ suffices for termination by *Xenopus* and human Pol III (1, 2), $T_4/T_5$ for *Schizosaccharomyces pombe* Pol III (3), and $T_5/T_6$ for *Saccharomyces cerevisiae* Pol III (4). Pol III termination within T clusters is generally heterogeneous and progressive (5), and termination efficiency tends to increase with the length of the T run (4). Mechanistically, Pol III termination involves extensive pausing, dictated by the T cluster itself (5, 6). Even during elongation, the addition of three UMP residues in succession is particularly slow (5). Pol III pausing at T clusters sets cycles of hydrolytic RNA chain retraction (7). Both pausing and the associated hydrolytic RNA cleavage are affected by mutations in the C160, C128, and C11 subunits of yeast Pol III (7–9). In the case of C128 and C11, the same mutations have also been shown to affect the termination properties of Pol III, thus suggesting a functional interplay between hydrolytic RNA cleavage and transcription termination (9, 10). The termination-prone character of Pol III is consistent with class III genes being very short, a fact that reduces the probability of unwanted termination signals within the coding regions. At the same time, the high efficiency of Pol III termination, by favoring polymerase recycling and transcription reinitiation, is likely to contribute in a crucial way to the high level of transcript production typical of Pol III (11–13).

The facility of Pol III termination poses questions that have not yet been addressed exhaustively. One of them relates to the mechanism allowing for easy termination by Pol III. Pol I and Pol II transcription units very often contain long T clusters, yet these two RNA polymerases do not terminate at these sequences, although T clusters are thought to generally induce RNA polymerase pausing. On which molecular features thus resides the distinctive ability of Pol III to terminate efficiently at short T clusters? At least part of the Pol III termination proneness has been attributed to the presence in this enzyme of the Rpc11p subunit, which is endowed with RNA cleavage-stimulating activity (9). Such a feature, however, seems not to be unique to Pol III, since Pol I and Pol II both contain subunits that share sequence homology with Rpc11p and might similarly be involved in termination (14). A recent study has further suggested that the termination-prone character of Pol III may also come from a marked propensity of the enzyme toward RNA release, such that every T-rich sequence at which Pol III pauses for a sufficiently long time would become a termination site (15). Such a termination propensity takes us to another important question, arising from the observation that Pol III-transcribed genes often contain internal stretches of T residues that would be predicted to provoke transcription termination. Is intragenic termination impaired in such cases, and by which mechanisms? A previously characterized example related to this issue is that of a $tRNA^{Lys}$ gene from *Xenopus laevis*, in which two $T_4$ tracts, one located at a canonical downstream position and the other one within the gene itself, determine transcription termination to different extents, with the internal $T_4$ cluster behaving as a much weaker terminator than the external one (16). The idea that the sequence context can generally influence the efficiency of T clusters first emerged from

* This study was supported by Human Frontier Science Program Organization Grant RGY0011/2002-C and by the Italian Ministry of Education, University, and Research (MIUR, 2003 COFIN Program). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 39-0521-905649; Fax: 39-0521-905151; E-mail: giorgio.dieci@unipr.it.

[1] The abbreviations used are: Pol, RNA polymerase; TFIIIB and TFIIIC, transcription factor IIIB and IIIC, respectively; nt, nucleotide(s); SnoRNA, small nucleolar RNA.

early studies of Pol III termination (1) and was confirmed and extended by later studies of transcription termination on the adenovirus VA RNA genes (17, 18). The surrounding sequences seem to influence Pol III termination especially at short T clusters. In *Xenopus*, for example, only $T_4$ terminators display a context-dependent strength, whereas clusters of five or more T residues are highly efficient as terminators independently from the context (1). The few documented cases of surrounding sequence effects can hardly be accommodated into general rules of sequence specificity, because the same flanking nucleotides were found to produce different (or even opposite) effects on the terminators of different class III templates (1, 16, 18). Further adding to the complexity of the sequence requirements for Pol III termination, several observations also suggested the existence of Pol III termination signals that do not conform to the general "T-cluster" rule (18–20).

In this study, we show that termination at oligo(dT) tracts by *S. cerevisiae* Pol III can be influenced by the flanking sequences to a previously unsuspected extent, and we address the sequence requirements and molecular mechanisms of such effects.

## EXPERIMENTAL PROCEDURES

*Statistical Analysis of tDNA 3′-Flanking Regions*—Coordinates for the tDNA downstream regions were predicted directly from the genomic sequences of *S. cerevisiae*, *S. pombe*, *Mus musculus* and *Homo sapiens* by using Pol3scan (21) and tRNAscan-SE (22). Predicted tRNA genes were considered for the analysis if the tRNA boundaries and anticodon type were found to be identical by the two methods. Although this criterion does not guarantee the identification of all the genomic tDNAs, it strongly limits the occurrence of false positive results. The portions corresponding to the 100 nucleotides immediately downstream to the mature tRNA coding sequence were extracted as FASTA files and used for the analysis of terminator signals. A Perl script was designed for statistical analysis of T-runs of arbitrary length, making use of the BioPerl library (23).

*DNA Templates*—For tRNA gene nomenclature, we refer to MIPS (Munich Information Center for Protein Sequences; available on the World Wide Web at mips.gsf.de/proj/yeast/rna/trna.html). The P(AGG)CR, N(GTT)CR, N(GTT)KL, and N(GTT)OL tDNAs have been described (24). The V(AAC)MR2 tDNA is carried by the previously described pY7 plasmid (25). The *Saccharomyces cerevisiae* tDNA L(CAA)LR2 was PCR-amplified from yeast genomic DNA (strain S288C) by using the high fidelity Deep Vent DNA polymerase (New England Biolabs) and the oligonucleotide primers L_fw (5′-GCAAATAAACGTTGGAAATTGGC) and L_rev (5′-GGTAGGAAAAAAAGATGCTGCAC). The tDNAs S(AGA)EL and M(CAT)E were amplified by using the *Pfu* Turbo DNA polymerase (Stratagene) and the following pairs of oligonucleotide primers: S_fw (5′-GGGAACTCTCAGAACGGGGG) and S_rev (5′-CGAGTTTTTACCCATTTAGGAAAAG); M_fw (5′-GCGACATGGAGAGATAGAGATAG), and M_rev (5′-CGCAGAAGTTAAAGCTTATAGTC). The 3′-mutant forms of N(GTT)CR and L(CAA)LR2 were obtained by PCR using the cloned tDNAs as templates and the appropriate mutagenic reverse primers. All of the amplification products were inserted into pUC-derived plasmid vectors and sequence-verified by dideoxy chain termination sequencing. The SCR1Δ_T7 construct is the previously described 3′Δ+90 variant of *SCR1* (26). For *in vivo* analysis, this *SCR1* gene variant was subcloned into the YEp352 shuttle vector. The SCR1Δ_3′L template was obtained using the cloned wild type *SCR1* gene as a template, the oligonucleotide 5′-TGATCAACTTAGCCAGGACATCC as a forward primer, and the oligonucleotide 5′-AAAAAAAAAACTTCAAAATTTTCGGAAGGCAAAAACGTGCAATCCGG as a reverse primer. The amplified fragment, corresponding to the fusion of the first 90 bp of the *SCR1* gene with the $T_5$ terminator plus 20 bp of 3′-flank of L(CAA)LR2 followed by a $T_{10}$ terminator, was cloned into the YEp352 vector. The 5′U6-L(CAA)LR2 and 5′U6-L_3′N templates were created by PCR using the corresponding cloned constructs as templates and the forward primer 5′-TTTCGGCTACTATAAATAAATGTTTTTTTCGGAGCTCTTACAACAAATAAGTGGTTGTTTGG, carrying the 5′-flanking sequence of the *SNR6* gene of *S. cerevisiae* immediately followed by a SacI site (underlined). The amplified fragments were inserted into a modified pNEB193 plasmid whose polylinker SacI site was disrupted. The 5′U6-SCR1Δ_3′L construct was generated by PCR using the oligonucleotide 5′-TTTCGGCTACTATAAATAAATGTTTTTTTCGCAACTAGCTAGGCTGTAATGGCTTTCTG, carrying the 5′ sequence of the *SNR6* gene, as a forward primer. In a different construct, the SacI restriction site was placed 5 bp upstream of SCR1Δ_3′L; the SCR1Δ_3′L variants were both inserted into the modified pNEB193. The *SNR52* transcription unit was PCR-amplified from yeast genomic DNA (strain S288C) by using *Pfu* Turbo DNA polymerase and the following oligonucleotide primers: SNR52_fw (5′-TGATCAACTTAGCCAGGACATCC), SNR52_rev (5′-GTTCTAAGTATTCTCATTTTATCC). The SNR52_mutT mutant was constructed by recombinant PCR (27). Two overlapping PCR primary products were generated using the SNR52_fw oligonucleotide in combination with the mutT_rev primer (5′-GTAGGGTGTGAACGAATGCGCACC) and the SNR52_rev oligonucleotide in combination with the mutT_fw primer (5′-GGTGCGCATTCGTTCACACCCTAC). Mutated positions in mutT_rev and mutT_fw are underlined. After gel purification, primary amplification products were mixed and used as templates in a subsequent amplification reaction, employing SNR52_fw and SNR52_rev as "outside" primers, which yielded the desired full-length secondary product. The 5′U6_SNR52 variant of *SNR52*, carrying the 5′-flanking sequence of *SNR6*, was generated by PCR using cloned wild type *SNR52* as a template and the oligonucleotide 5′-TTTCGGCTACTATAAATAAATGTTTTTTTCGCAACTA as a forward primer. All of the *SNR52* variants were cloned into the pNEB193 plasmid vector.

*In Vitro Transcription Analyses*—*In vitro* transcription reaction conditions and procedures for RNA purification and analysis were essentially as described (13, 28). Reactions contained 40 fmol of plasmid-borne class III template and the following amounts of transcription proteins: 150 ng of TFIIIC purified up to the DEAE-Sephadex A-25 step (29); 40 ng of pure, recombinant TATA box-binding protein, and 80 ng of recombinant Brf1, both purified from overexpressing *Escherichia coli* cells (29); 0.5 μg of B″ fraction partially purified, up to the Bio-Rex 70 chromatography step, from chromatin pellets generated during yeast nuclear extract preparation (30). The last three components reconstitute a highly active TFIIIB factor. Unless otherwise indicated, the ATP, CTP, and GTP concentration was 500 μM, and UTP concentration was 250 μM. Transcripts were radiolabeled using [α-$^{32}$P]UTP. Template DNA was first incubated with TFIIIC and TFIIIB for 20 min at 20 °C, and then Pol III (10 ng) and NTPs were added, and transcription was allowed to proceed for 20 min at 20 °C. In Fig. 7, *A* and *B*, transcription complexes were assembled either in the presence or in the absence of TFIIIC, using increased amounts of the different components of TFIIIB: 400 ng of TATA box-binding protein, 240 ng of Brf1 and either 0.7 μg of B″ fraction (Fig. 8*B*), or 20 ng of recombinant yeast Bdp1 (13) (Fig. 7*A*). For the factorless transcription experiment in Fig. 7*B*, 3′-overhanged restriction fragments were obtained by digesting the above described SacI site-containing templates with SacI, cutting 5 bp upstream of the natural transcription start site, and either BamHI (for L(CAA)LR2) or NdeI (for L_3′N and SCR1Δ_3′L), cutting downstream of the cloned insert in the plasmid polylinker. Restriction fragments were gel-purified, treated once with phenol/chloroform, and ethanol-precipitated prior to use in transcription assays. General transcription conditions were the same as in factor-directed transcription. The 3′-overhanged fragment (30 ng) was first incubated with 20 ng of purified Pol III and 400 μM CpU dinucleotide primer in transcription buffer containing 160 μg/ml bovine serum albumin for 15 min at 20 °C in a total volume of 25 μl. NTPs and [α-$^{32}$P]UTP (10 μCi) were then added, and transcription was allowed to proceed for 30 min at 20 °C. In all transcription experiments, radioactive transcripts were quantified by phosphorimaging. For termination efficiency measurements, the signals corresponding to RNAs of different length were normalized for the number of incorporated U residues.

*In Vivo RNA Analysis*—Yeast cells (YPH500 strain) were transformed with the different YEp352-SCR1Δ constructs by the lithium acetate procedure (31), and the resulting transformants were selected for uracil auxotrophy. Total RNA was prepared according to a previously described procedure (32). RNA samples (10 μg) were electrophoresed on 6% polyacrylamide, 7 M urea gels and transferred to a GeneScreen Plus membrane (PerkinElmer Life Sciences), which was then probed with a 5′-labeled oligonucleotide (5′-CTGGCCGAGGAACAAATCCTTCCTCGCGGC) complementary to the coding region of *SCR1* between positions +43 and +73. Hybridization was carried out overnight at 28 °C in 5× SSC solution, 5× Denhardt's solution, 0.1 mg/ml denatured salmon sperm DNA, 0.5% (w/v) SDS, followed by one short washing in 2× SSC solution containing 0.1% SDS and two short washings in 1× SSC solution containing 0,1% SDS. Hybridization products were visualized and quantified by phosphorimaging.

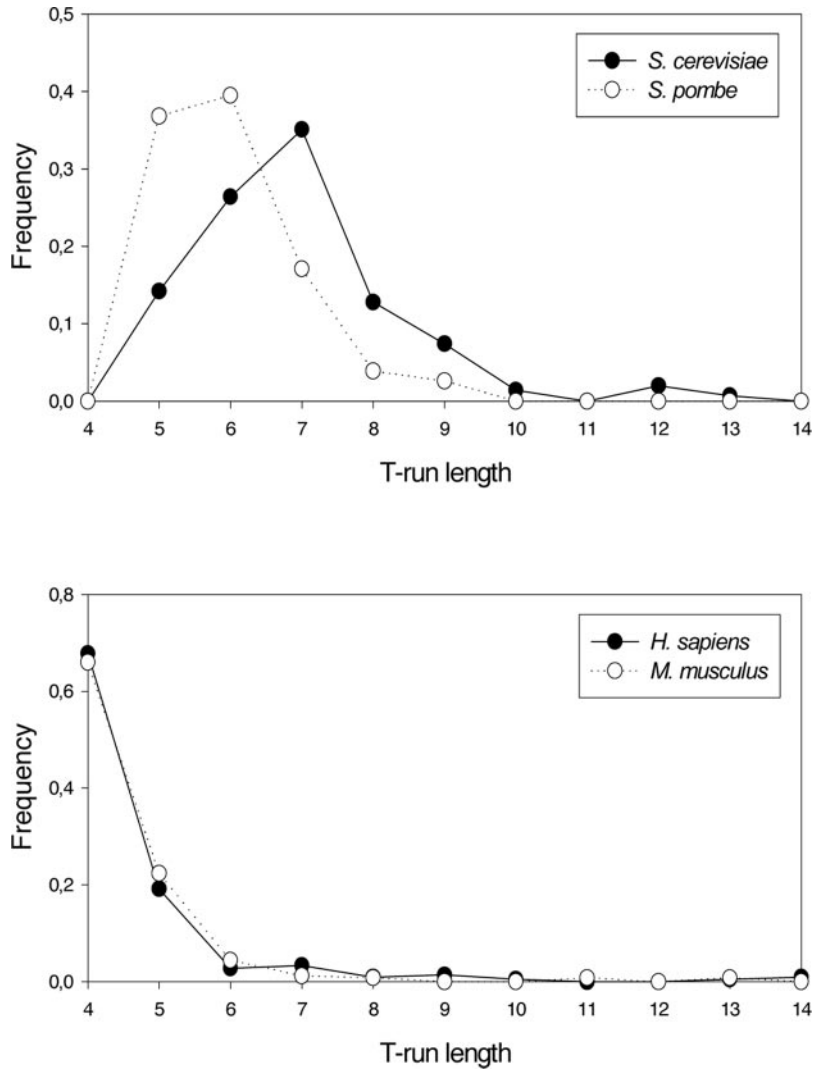



FIG. 1. **Distribution of tDNA terminator length among different eukaryotes.** Frequency distribution analysis of poly(dT) terminator length was conducted on the subsets of tDNAs (from *S. cerevisiae*, *S. pombe*, *M. musculus*, and *H. sapiens* genomes) possessing a single Pol III termination signal starting within the first 40 bp downstream of the mature tRNA coding region. Such subsets comprised 148, 76, 214, and 250 tDNAs in the case of *S. cerevisiae*, *S. pombe*, *H. sapiens*, and *M. musculus*, respectively. *Upper panel*, T-run length distribution in the yeasts *S. cerevisiae* and *S. pombe*. *Lower panel*, T-run length distribution in *H. sapiens* and *M. musculus*.

## RESULTS

*Oligo(dT) Termination Signals in Different Eukaryotic Genomes*—The flanking regions of eukaryotic tRNA genes present conserved sequence patterns that diverge considerably in going from yeast to mammalian genomes (24). As a further element of divergence, yeast and human RNA polymerase III enzymes respond differently to oligo(dT) terminators (3). To gain insight into Pol III termination signals on a genome-wide scale, we thus decided to analyze and compare the tDNA transcription terminators in two yeast (*S. cerevisiae* and *S. pombe*) and two mammalian (*M. musculus* and *H. sapiens*) genomes. Using the complete tRNA gene inventories from these genomes, we initially searched for the presence of runs of 4 or more T residues within the first 100 bp of the 3′-flanking region, because $T_4$ is the shortest known oligo(dT) tract causing Pol III termination (33). In *S. cerevisiae* and *S. pombe*, $T_4$ was never found as the sole potential termination signal: a run of 5 or more T residues was always present, and it always started within the first 40 bp following the end of the mature tRNA coding sequence.[2] This was taken as an indication that runs of less than 5 T residues are not used as Pol III terminators in yeast and that Pol III termination signals appear within the first 40 bp of 3′-flanking sequence. In contrast, many mammalian tDNAs had a run of 4 T residues as the sole potential termination signal in the same region, in agreement with the reported ability of human Pol III to terminate efficiently at $T_4$, and with the observation that yeast Pol III requires longer T stretches than vertebrate Pol III in order to terminate efficiently (3, 33). Again, most termination signals occurred within 40 bp of the 3′-flanking region. Therefore, we considered $T_5$ and $T_4$ as the minimal termination signals for yeast and mammalian tDNA transcription, respectively. By limiting further analysis to the first 40 bp of 3′-flanking sequence, we then selected, from each tDNA set, the tDNAs having a single T run within that region. Such "single terminator" tDNAs represented 56, 47, 69, and 67% of all tDNAs in *S. cerevisiae*, *S. pombe*, *H. sapiens*, and *M. musculus*, respectively. We then calculated, for the subsets of single terminator tDNAs from each genome, the frequency distribution of T run length. The aim of such analysis was to establish whether the T run length in the different tDNA sets correlates with lineage-specific differences in termination signal recognition by Pol III. The tDNAs with more than one termination signal within the first 40 bp of 3′-flank were excluded from analysis, because the presence of a potential "back up" terminator could have reduced the selective pressure on the upstream terminator. The results of this analysis are shown in Fig. 1. In *S. cerevisiae* and *S. pombe*, the most frequent T run lengths are 6 and 7, and (as already mentioned) there is no single terminator shorter than $T_5$. Remarkably, the tDNA terminators in fission yeast tend to be ~1 nt shorter than in *S. cerevisiae*, in agreement with a previous analysis showing

[2] The only exceptions were represented by dimeric tDNAs, in which the upstream tDNA unit has no termination signal (45, 58).

A

```
1.tRNA_i_L(CAA)LR2
CAACCATTATTTTTGCCTTCCGAAAATTTTGA[44bp]ATCTTTTTTTCCT…
2.tRNA_P(AGG)CR
CCCACATTGTTTTTACAATGTGTGGTGTCGCTACTGGCAGAAATGGTACGG…
3.tRNA_N(GTT)CR
GGTCGTTAATTTTTCTTTTCTTTTTTTTTTTTTATCTCCGACGGAGAAAAA…
4.tRNA_S(AGA)EL
TTGTCGTTATTTTTCTCTTTTTTTTCAATTTCCCTTGTTCGTCAGATCGAG…
5.tRNA_M(CAT)E
CGGCGCTAATTTTTCATTTCTTTTTGCCCGCAACAATAACACAGATCGCAA…
6.tRNA_N(GTT)KL
GGTCGCTTATTTTTAACTTTTTTACACTGAAGAAACAAATCAATTCTAATA…
7.tRNA_N(GTT)OL
GGGTCGTTATTTTTGAATACCCAATACTTTTTTTTCGGGCCATTTCACACA…
8.tRNA_V(AAC)MR2
GCGAAATCATTTTTATCTCAGCCGTTACGGAAAGTTTTTTAACGCCCGTCT…
```

B

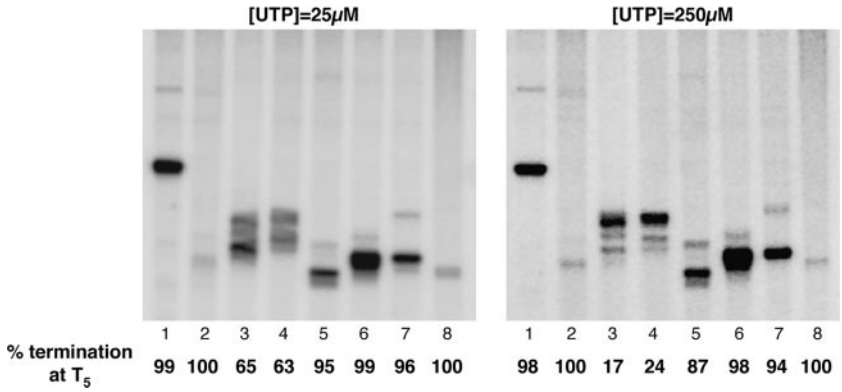



FIG. 2. ***In vitro* transcription of yeast tRNA genes displaying $T_5$ as the first termination signal.** *A*, nucleotide sequences of the terminator regions of eight yeast tRNA genes displaying $T_5$ as the first encountered termination signal. The sequences are aligned on the $T_5$ terminator (*underlined*). The sequence corresponding to the mature tRNA is in *boldface and italic type*. For tDNA 1, the sequence of a 44-bp tract (location indicated) has been omitted. *B*, *in vitro* transcription of the templates in *A*. *Lanes 1–8* refer to the templates identified by the same *number* in *A*. All of the tDNAs except for number 2 possess a downstream back-up terminator that allows to monitor polymerases reading through $T_5$; template 2 was linearized 121 bp downstream of $T_5$ in order to reveal read-through events as run-off transcripts. *Left*, the transcription reactions were carried out in the presence of 25 μM UTP; *right*, the transcription reactions were carried out in the presence of 250 μM UTP. The percentage of tDNA transcripts terminated at $T_5$ is reported *below* each *lane*.

that *S. cerevisiae* Pol III requires slightly longer T runs than the fission yeast enzyme in order to terminate efficiently (3). The T-run length distribution in mammals is completely different: $T_4$ is the most frequent tDNA terminator, and terminators longer than $T_5$ are rare. Again, this finding is in agreement with the reported ability of human Pol III (and, more generally, of vertebrate Pols III) to terminate efficiently at $T_4$ sequences (1, 3, 18, 33).

*Variable Strength of the $T_5$ Terminator in S. cerevisiae tDNAs*—Of the 148 "single terminator" tDNAs of *S. cerevisiae*, 22 bear a simple run of 5 Ts as a termination signal. For such tDNAs, the $T_5$ sequence is thus expected to be endowed with the ability to induce efficient Pol III termination. On the other hand, previous studies have shown that 5 consecutive T residues in the 3′-flanking sequence of yeast tDNAs are only partially effective in inducing Pol III transcription termination (3, 4). To better define the termination properties of the $T_5$ sequence, we initially analyzed transcription termination on eight yeast tRNA genes all displaying $T_5$ as the first termination signal encountered in the 3′-flanking region. These genes are listed in Fig. 2*A*, together with the features of their 3′-flanks. Two of them (numbers 1 and 2) are classified as "single terminator" tDNAs. In the other tRNA genes, $T_5$ is followed by a back up terminator located 3–20 bp downstream. The presence of a downstream T run allows to easily monitor polymerases reading through the $T_5$ terminator by their production of longer transcripts terminated at the downstream T cluster. This is also true in the case of tDNA 1 (L(CAA)LR2), since a back-up T run is present starting at 66 bp downstream of $T_5$. In the case of P(AGG)CR, where no back-up T run is present, the tDNA-containing plasmid was linearized 121 bp downstream of $T_5$, so as to be able to reveal reading-through polymerases by their production of longer, run-off transcripts. All of these templates were used to program *in vitro* transcription reactions in a reconstituted system from *S. cerevisiae* (see "Experimental Procedures"). Transcription was performed in the presence of either a low UTP concentration (25 μM; Fig. 2*B*, *left panel*) or a 10-fold higher UTP concentration (Fig. 2*B*, *right panel*) closer to the *in vivo* UTP concentration, estimated to be in the millimolar range (34, 35). Fig. 2*B* shows that the strength of $T_5$ as a terminator varies considerably in the different contexts and that $T_5$ read-through by Pol III is in some cases strongly favored at the higher UTP concentration (*cf.*, for example, *lanes 3* and *4* of the *left panel* with the *same lanes* of the *right panel*), in other words under conditions that increase the polymerization rate and thus decrease the dwell time of the enzyme at $T_5$. For six of the eight genes, the $T_5$ element acted as a strong terminator (*lanes 1*, *2*, and *5–8*). In contrast, for two tDNAs (*lanes 3* and *4*), a large fraction of polymerases read through the $T_5$ element and terminated at the back up terminator. The two $T_5$ elements behaving as weak terminators are both immediately followed by the sequence CT and, a few base pairs downstream, by a longer T run acting as a back up terminator (tDNA 3 and 4 in Fig. 2*A*). In contrast, the six $T_5$ elements acting as strong terminators are all immediately followed by A or G, with the exception of M(CAT)E (tDNA 5). The $T_5$ element of M(CAT)E is immediately followed by C and is characterized by a slightly lower strength as a terminator (*cf. lane 5* with *lanes 1*, *2*, *6*, *7*, and *8* in the *right panel* of Fig. 2*B*). It is interesting to note that $T_5$ is flanked by almost identical upstream sequences in the case of N(GTT)CR (tDNA 3), N(GTT)KL (tDNA 6), and N(GTT)OL (tDNA 7). Therefore, the remarkable difference in terminator strength between N(GTT)CR (*lane 3*) and the other two $tRNA^{Asn}$ genes (*lanes 6* and *7*)

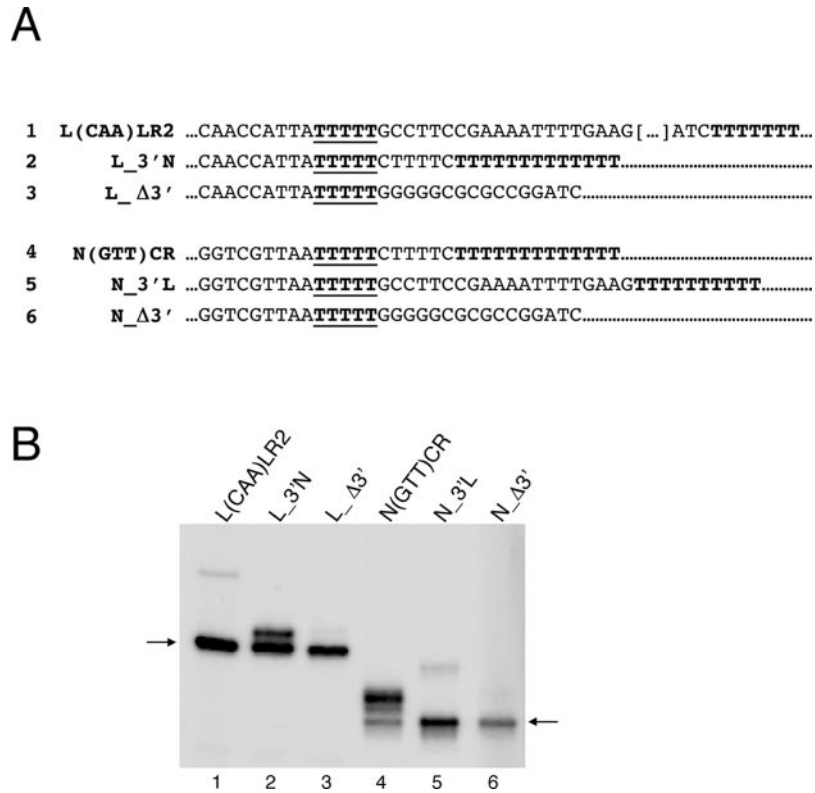



FIG. 3. **Termination properties of 3′-flank variants of L(CAA)LR2 and N(GTT)CR tDNAs.** *A*, nucleotide sequence of the terminator-surrounding regions of L(CAA)LR2 and N(GTT)CR tDNAs and their 3′-flank variants. The $T_5$ terminator is in *boldface type* and *underlined*. When present, the back-up terminator is in *boldface type*. The L_Δ3′ and N_Δ3′ constructs were both linearized 16 bp downstream of $T_5$ in order to reveal read-through events by the presence of run-off transcripts. *B*, *in vitro* transcription assay (conducted in the presence of 250 μM UTP) of the templates reported in *A*. The *arrows* on the *left* and on the *right* indicate the migration positions of transcripts terminated at the $T_5$ element of L(CAA)LR2 and N(GTT)CR, respectively.

can reasonably be attributed to the downstream sequence context.

*Influence of the Downstream Sequence on Pol III Termination at $T_5$*—To investigate in more detail the downstream sequence effect on $T_5$ terminator strength, we focused on two of the tDNAs tested in the previous experiment: L(CAA)LR2 (tDNA 1), whose $T_5$ is one of the strongest terminators, and N(GTT)CR (tDNA 3), whose $T_5$ is the weakest terminator observed in Fig. 2 and is flanked by a back up terminator 7 bp downstream. We initially constructed and tested two hybrid templates (Fig. 3*A*): the L_3′N template, in which the sequence downstream of the L(CAA)LR2 $T_5$ terminator was replaced by the $CT_4CT_{13}$ sequence naturally present downstream of $T_5$ in N(GTT)CR and the N_3′L template, in which the sequence downstream of the N(GTT)CR $T_5$ element was replaced by 20 bp of the sequence naturally present downstream of the L(CAA)LR2 $T_5$ terminator, followed by an artificially placed $T_{10}$ back-up terminator. In the L_3′N context, the $T_5$ terminator of L(CAA)LR2 turned out to be remarkably weakened, displaying (at 250 μM UTP) a termination efficiency of 55% as compared with 98% in the case of the wild type L(CAA)LR2 gene (Fig. 3*B*, *cf. lanes 1* and *2*). On the other hand, when N(GTT)CR with its $T_5$ element was fused to the 3′-flanking sequence derived from L(CAA)LR2, the strength of $T_5$ as a terminator was greatly increased (Fig. 3*B*, *cf. lanes 4* and *5*). There are two possible explanations for the above observations. The $T_5$ element might be intrinsically weak as a terminator, and the particular downstream context in L(CAA)LR2 would potentiate its termination proficiency; in this case, $T_5$ weakening in L_3′N would be simply due to the removal, and $T_5$ strengthening in N_3′L to the insertion, of a terminator-strengthening sequence. Alternatively, the $CT_4CT_{13}$ sequence naturally present 3′ to the $T_5$ element in N(GTT)CR might directly act as (or contain) a terminator-weakening element. According to the second hypothesis, the replacement of the sequence downstream of $T_5$ in N(GTT)CR with an unrelated sequence should increase the $T_5$ strength as a terminator, whereas the replacement of the sequence downstream of $T_5$ in L(CAA)LR2 with an unrelated sequence should not decrease the $T_5$ terminator efficiency. In the experiment in Fig. 3*B*, the termination properties of the unmodified L(CAA)LR2 and of the L_3′N construct were compared with those of a L(CAA)LR2 derivative (L_Δ3′) in which the 3′-flanking sequence downstream of $T_5$ was replaced by a plasmid sequence (*lane 3*). Similarly, the termination properties of unmodified N(GTT)CR and of the N_3′L construct were compared with those of a N(GTT)CR derivative (N_Δ3′) in which the $T_5$ 3′-flank was replaced by the same plasmid sequence as in L_Δ3′ (*lane 6*). Termination at $T_5$ appeared to be unaffected in the L_Δ3′ construct with respect to unmodified L(CAA)LR2 (*cf. lanes 1* and *3*), whereas the strength of the N(GTT)CR $T_5$ element was greatly increased in the N_Δ3′ construct (*cf. lanes 4* and *6*). (Since no rescue terminator is present downstream of $T_5$ in the L_Δ3′ and N_Δ3′ templates, they were linearized with BamHI at a position 16 bp downstream of the $T_5$ sequence). The $CT_4CT_{13}$ sequence thus behaves as (or appears to contain) a terminator-weakening element. This conclusion is further supported by the observation that $T_5$ is weak in only two of the eight tDNAs tested in the experiment in Fig. 2 (N(GTT)CR and S(AGA)EL) and that in both cases $T_5$ is followed by similar CT-containing sequences ($CT_4CT_{13}$ in the case of N(GTT)CR, $CTCT_8C$ in the case of S(AGA)EL).

*Sequence and Positional Requirements for $T_5$ Terminator Weakening*—To better understand the $T_5$ terminator weakening effect, we concentrated on the L_3′N chimeric template, whose downstream weakening sequence, $CT_4CT_{13}$, was subjected to mutational analysis. In particular, mutants were constructed in which the $CT_4C$ sequence was variously replaced, whereas the downstream $T_{13}$ tract was left unvaried (or shortened to $T_{10}$) and used as a back-up terminator. All of the mutant constructs, schematically illustrated in Fig. 4*A*, were tested in an *in vitro* reconstituted system in at least three independent experiments. The results are shown in Fig. 4*B*. Pol III transcriptional read-through at $T_5$ was first calculated on the basis of the ratio between $T_5$-terminated and $T_{13}/T_{10}$-ter-

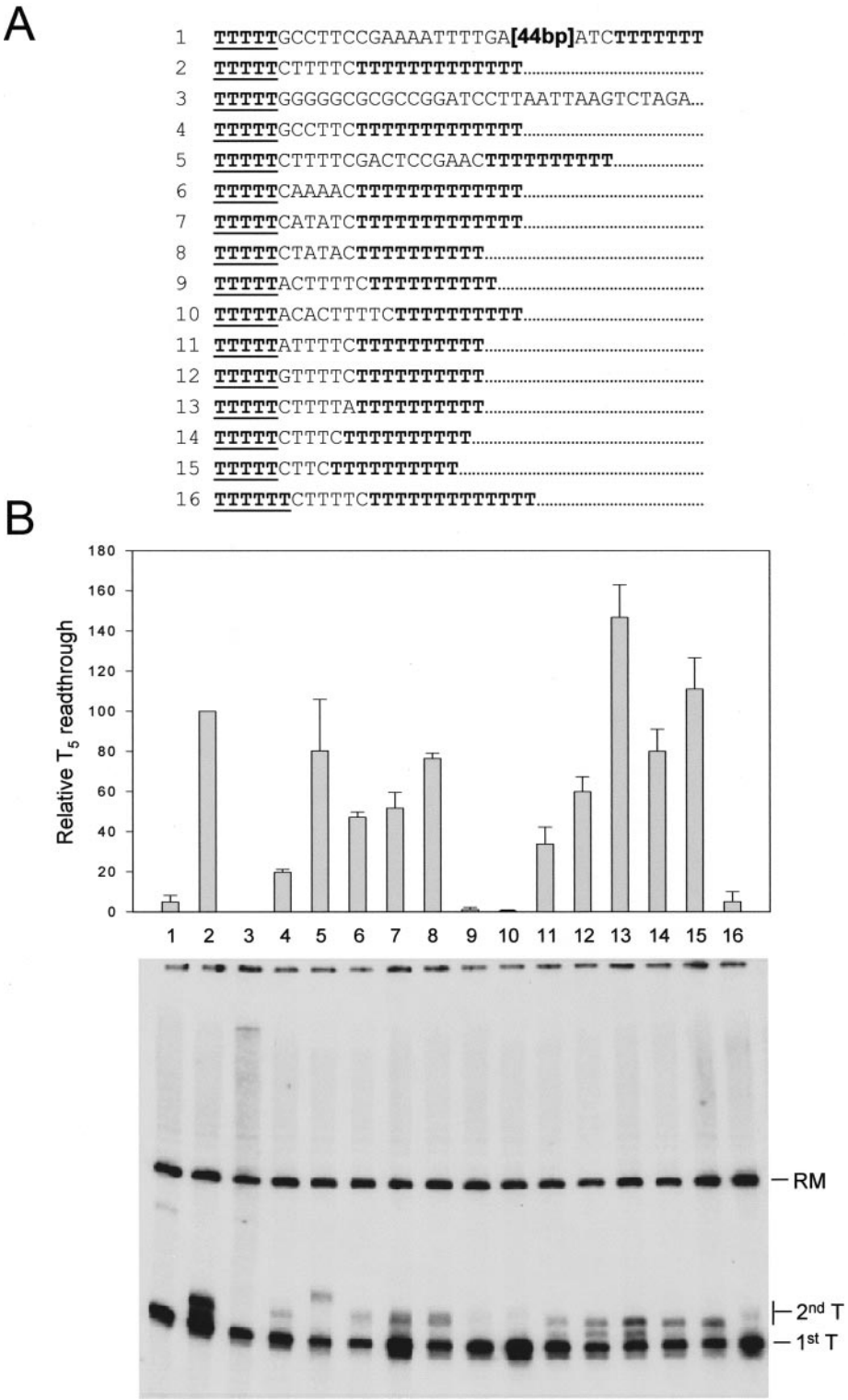



FIG. 4. **Properties of the L(CAA)LR2 $T_5$ terminator fused with different 3′-flanking sequences.** *A*, nucleotide sequence of the L(CAA)LR2 variants resulting from mutagenesis of the region downstream of $T_5$. Only the sequence 3′ to the $T_5$ terminator (in *boldface type* and *underlined*) is reported. The back-up terminator is in *boldface type*. Templates 1–3 are the same as in Fig. 3. *B*, the *lower part* shows the results of *in vitro* transcription assays of the templates listed in *A* (identified by the *numbers above* the *lanes*). The migration positions of the different tDNA transcription products, terminated at the first or the second T-run, are indicated on the *right*, together with the position of a 415-nt-long DNA used as a recovery marker (*RM*). Reported in the *upper part* is a *bar plot* deriving from quantification of three independent experiments. Pol III transcriptional read-through at the first terminator was calculated on the basis of the ratio between tDNA transcripts terminated at the first and second T-run. In each experiment, the transcriptional read-through observed with the L_3′N template (template 2) was set to 100, and the effect of the mutations was evaluated as relative $T_5$ read-through with respect to L_3′N and reported on the *y* axis, with *error bars* indicating the S.E.

minated transcripts. In each experiment, the extent of $T_5$ terminator read-through observed with the L_3′N template (*lane 2*) was arbitrarily set to 100, and the effect of the mutations was evaluated as relative $T_5$ read-through with respect to L_3′N. When the first 3 nucleotides downstream of $T_5$ (CTT) were changed to GCC, so as to render the 6 base pairs downstream of $T_5$ identical to the sequence found in the wild type L(CAA)LR2 tDNA, the weakening effect was strongly dimin-

ished (~5-fold) with respect to L_3′N (*cf. lanes 2* and *4*). Since the CTT to GCC substitution, which increases the predicted stability of the DNA double helix immediately downstream of $T_5$, strongly affected its weakening potential, we wondered whether replacing the downstream sequence with other sequences characterized by a similar (or even lower) DNA duplex stability would preserve (or increase) the weakening effect, and thus $T_5$ read-through. To address this point, the $T_4$ stretch in $CT_4C$ was replaced by TATA, ATAT, or $A_4$. As shown in Fig. 4*B*, *lanes 6–8*, and in the *bar plot above* the *gel*, such substitutions led to a partial loss of the weakening effect: a 40–50% loss in the case of $A_4$ and ATAT and only a 20% loss in the case of TATA. The prediction of DNA duplex stability on the basis on nearest neighbor thermodynamics gives $A_4 = T_4 >$ ATAT $>$ TATA (36). Therefore, among the three substitutions, the one causing the strongest read-through (TATA) is also the one displaying the lesser DNA duplex stability. However, despite its higher predicted stability, the original $T_4$ sequence causes more read-through than TATA. We conclude that a low DNA duplex stability in the region immediately downstream of $T_5$ is not sufficient for full $T_5$ terminator weakening, even if it somehow contributes to the weakening effect. This conclusion is also supported by the comparison of the termination properties of tDNAs 3 and 7 in Fig. 2. The $T_5$ terminator of tDNA 7 is much stronger than the $T_5$ terminator of tDNA 3, yet the predicted stability of the immediately downstream 6-bp sequence ($CT_4C$ for tDNA 3, GAATAC for tDNA 7) is very similar (36). With another series of constructs (templates 5, 9, and 10 in Fig. 4) we then analyzed the positional requirements for $T_5$ terminator weakening. In all of these constructs, the $T_{13}$ element was replaced by $T_{10}$, a sequence that works as efficiently as $T_{13}$ as a downstream terminator and does not interfere with the $T_5$ weakening effect (data not shown). In construct 5, the distance between the $T_4$ and $T_{10}$ stretches, which are separated by only 1 bp in the L_3′N template (*lane 2*), was increased to 11 bp. As shown in Fig. 4*B*, $T_5$ terminator weakening in this context was very similar to that observed with L_3′N (*cf. lanes 2* and *5*). The $CT_4C$ sequence thus appears to act as an autonomous, termination-weakening element. In contrast, the weakening effect was found to be compromised in constructs in which the distance between $T_5$ and the $CT_4C$ sequence was increased by 3 bp (construct 10, *lane 10*) or even by only 1 bp (construct 9, *lane 9*). The terminator weakening effect exerted by the $CT_4C$ sequence thus seems to depend in a critical way on its distance from the $T_5$ terminator. Having restricted the terminator-weakening function to the $CT_4C$ sequence, we constructed other mutants (constructs 11–15 in Fig. 4*A*) in order to identify the minimal sequence requirements for the terminator weakening effect. As shown in *lane 13*, the C residue separating the $T_4$ stretch from the downstream $T_{10}$ sequence could be mutated without any loss of terminator weakening (on the contrary, a 50% increase in $T_5$ read-through with respect to L_3′N was reproducibly observed with this mutant; *cf. lanes 2* and *13*). With templates 14 and 15, in which the length of the $T_4$ stretch was progressively reduced, the weakening effect was maintained (*lanes 14* and *15*). The mutation of the C residue, immediately following the $T_5$ element, to A reproducibly reduced the read-through efficiency by ~65% with respect to L_3′N (*cf. lanes 2* and *11*). The mutation of the same residue to G produced a more modest, but still significant effect, reducing $T_5$ read-through by ~40% with respect to L_3′N (*cf. lanes 2* and *12*). A C residue immediately downstream of $T_5$ is thus required for full weakening efficiency. The results of this mutational analysis, together with the data in Fig. 2, suggest that a CT dinucleotide immediately downstream of a $T_5$ terminator represents a minimal sequence element capable of producing a significant terminator weakening effect. That CT alone is sufficient to induce read-through is especially evident from the termination properties of tDNA 4 in Fig. 2 and template 8 in Fig. 4. This weakening element, however, appears to be only effective on the $T_5$ terminator. When the termination signal length was increased to $T_6$, the weakening effect was lost (Fig. 4*B*, *lane 16*).

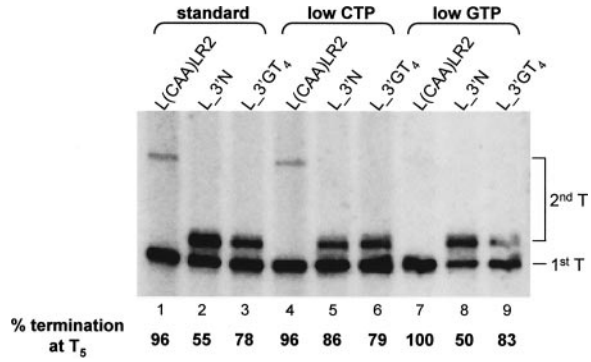



FIG. 5. **Transcriptional read-through at a $T_5$ terminator is influenced by the concentration of the next incoming nucleotide.** *In vitro* transcription assays of the L(CAA)LR2, L_3′N, and L_3′GT4 templates (template 1, 2, and 12 of Fig. 4, respectively) were carried out in the presence of either standard concentrations of NTPs (500 μM ATP, CTP, and GTP; 250 μM UTP; *lanes 1–3*), low CTP concentration (50 μM; *lanes 4–6*), or low GTP concentration (50 μM, *lanes 7–9*). The migration positions of the transcripts terminated at the first or second T-run are indicated on the *right*. The percentage of tDNA transcripts terminated at $T_5$ is reported *below* each *lane*.

*Mechanism of $T_5$ Terminator Weakening by Downstream Sequence*—The oligo(dT) tracts acting as Pol III terminators have the ability to induce both RNA polymerase pausing and transcript release. These two features are functionally distinguishable and are characterized by different sequence requirements (6, 15). Weakening of the $T_5$ terminator by a CT sequence placed immediately downstream may involve an impairment of both (or only one) of these abilities. To gain insight into the mechanism of $T_5$ terminator weakening, we first analyzed termination on the L_3′N template (construct 2 in Fig. 4*A*) and on the L_3′$GT_4$ template (template 12 of Fig. 4*A*, in which the C immediately downstream of $T_5$ is replaced by G) in the presence of a standard (500 μM) and a 10-fold lower concentration of CTP. Lowering CTP concentration is expected to increase the dwell time of Pol III across the $T_5$ sequence in L_3′N, but not in L_3′$GT_4$, by reducing the rate of incorporation of the first nucleotide downstream of $T_5$. In the same way, lowering the GTP concentration should increase the dwell time of Pol III onto the $T_5$ element in L_3′$GT_4$, but not in L_3′N. As shown in Fig. 5, in the presence of 50 μM CTP (*lane 5*) the $T_5$ terminator in L_3′N turned out to be significantly stronger than with 500 μM CTP (*lane 2*), with a ~3-fold reduction in the fraction of reading-through polymerases (from 45% in *lane 2* to 14% in *lane 5*). Lowering the GTP concentration had instead no effect on $T_5$ read-through with this template (*cf. lanes 2* and *8*). With the L_3′$GT_4$ template, lowering the CTP concentration did not influence the extent of $T_5$ read-through, whereas the fraction of reading-through polymerases was slightly (but reproducibly) decreased at low GTP concentration (from 22% in *lane 3* to 17% in *lane 9*). $T_5$ terminator weakening can thus be specifically counteracted by reaction conditions that favor polymerase pausing on $T_5$.

*Effect of the Upstream Sequence Context on $T_5$ Terminator Efficiency*—Altogether the above data show that termination efficiency at $T_5$ can be strongly influenced by the downstream sequence context, in a manner that is independent from the specific tRNA coding sequence preceding the $T_5$ element. To further test for possible effects of the upstream flanking

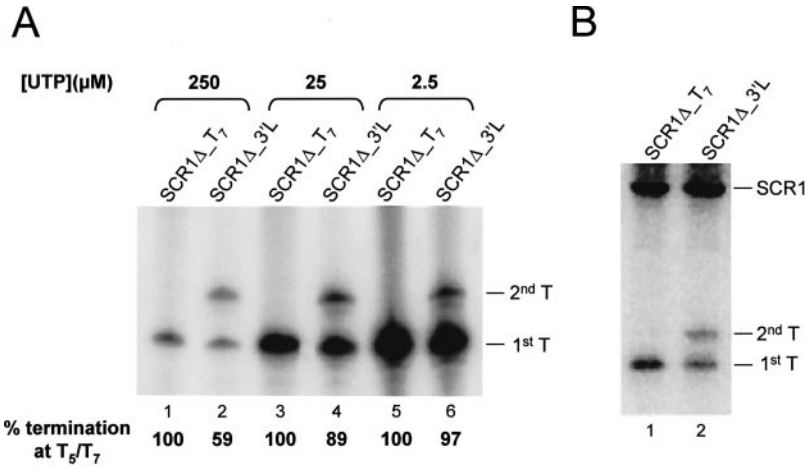



FIG. 6. **Upstream sequence effect on $T_5$ terminator strength *in vitro* and *in vivo*.** *A*, *in vitro* transcription of the SCR1Δ_$T_7$ and SCR1Δ_3′L templates was carried out in the presence of 250 μM UTP (*lanes 1* and *2*), 25 μM UTP (*lanes 3* and *4*), or 2.5 μM UTP (*lanes 5* and *6*). The migration positions of transcripts terminated at the first or second T-run of the SCR1-derived templates are indicated on the *right*. *B*, *in vivo* expression profiles of the SCR1Δ_$T_7$ and SCR1Δ_3′L templates. Total RNA was extracted from yeast cells (YPH500 strain) transformed with YEp352 carrying the template indicated *above* the *lanes*. Fractionated RNA samples (10 μg each) were probed with a radiolabeled oligonucleotide annealing within the first 90 bp of the *SCR1* RNA product. The migration positions of the transcripts derived from the chromosome-encoded *SCR1* gene and from the two plasmid-borne *SCR1*-derived templates (*1st T* and *2nd T*) are indicated on the *right*.

sequence on $T_5$-dependent termination, we generated a construct (SCR1Δ_3′L) in which the L(CAA)LR2 $T_5$ terminator, plus 20 bp of its natural 3′-flanking sequence followed by $T_{10}$, were fused to the first 90 bp of the *SCR1* gene, coding for the *S. cerevisiae* 7SL RNA. Apart from the presence of A- and B-blocks allowing for efficient *in vitro* transcription by the Pol III system (26), this portion of the *SCR1* gene is essentially unrelated to tRNA coding sequences. As shown in Fig. 6*A*, *lane 1*, the insertion of a $T_7$ terminator at position +90 of *SCR1* results in a template (SCR1Δ_$T_7$) whose transcription is efficiently terminated at $T_7$, thus generating 90-nt-long transcripts; no read-through transcription products were observed when this template was linearized 25 bp downstream of $T_7$ (data not shown). Surprisingly, the L(CAA)LR2-derived $T_5$ element plus 3′-flank, which produced a 97% termination efficiency in the case of the original template (see, for example, Fig. 2, *lane 1*), behaved as a poor terminator in the SCR1Δ_3′L construct, allowing ~40% of the polymerases to read through and terminate at the downstream $T_{10}$ terminator in the presence of 250 μM UTP (Fig. 6*A*, *lane 2*). Read-through was significantly less in the presence of 25 μM UTP (*lane 4*) and almost absent with 2.5 μM UTP (*lane 6*), thus suggesting that the *SCR1*-derived upstream sequence reduces termination at $T_5$ by affecting Pol III pausing. Transcription termination with the SCR1Δ_3′L construct could also be monitored *in vivo*, because the 20-bp distance between the $T_5$ element and the downstream $T_{10}$ terminator allows for the production of transcripts that differ from each other in size significantly and that are characterized by comparable *in vivo* stabilities.[3] The SCR1Δ $T_7$ and SCR1Δ_3′L templates were inserted into the high copy number YEp352 vector and transformed into the YPH500 *S. cerevisiae* strain. The RNAs derived from logarithmically growing cells were then subjected to Northern blot analysis with an *SCR1*-specific antisense oligonucleotide. As shown in Fig. 6*B* (*lane 2*) 55% of the plasmid-encoded, *SCR1*-derived transcripts were terminated at $T_5$, whereas 45% were 20 nt longer and derived from $T_5$ read-through and termination at the downstream $T_{10}$ sequence. By comparison, the $T_7$ terminator in the SCR1Δ_$T_7$ construct appeared to be fully efficient *in vivo* (*lane 1*).

*Distinguishing between Weak and Strong $T_5$ Terminator Sequences Is an Intrinsic Property of Pol III*—The observed sequence context effects on Pol III termination at $T_5$ might in principle be mediated by TFIIIC, a transcription factor that usually covers a large part of the transcription unit, extending up to the terminator region of tRNA genes (37, 38). TFIIIC-associated activities have previously been proposed to influence Pol III termination in humans (39), and yeast TFIIIC has recently been shown to facilitate transcription reinitiation by Pol III on long transcription units (13). The role of TFIIIC is thus not restricted to preinitiation complex assembly. To test whether TFIIIC influences termination at $T_5$ in different sequence contexts, some of the templates in which the $T_5$ terminator is weakened were modified so as to allow for their TFIIIC-independent transcription. To this end, the transcribed portions of L(CAA)LR2, L_3′N and SCR1Δ_3′L were fused to the 5′-flanking region of the *S. cerevisiae SNR6* gene, which is known to support TATA box-mediated TFIIIB assembly and transcription initiation in the absence of TFIIIC (13, 40). As shown in Fig. 7*A*, a pronounced weakening of $T_5$ by both the 3′N downstream context and the SCR1Δ upstream context was observed in the absence of TFIIIC (*cf. lanes 5* and *6* with *lane 4*), thus demonstrating that the weakening effects do not depend on this factor. A comparison of the termination efficiencies in the presence (*lanes 1–3*) or absence (*lanes 4–6*) of TFIIIC rather suggests that TFIIIC somehow facilitated $T_5$ termination signal recognition by Pol III, an effect that is especially evident with the 5′U6-L_3′N template (*cf. lanes 2* and *5* in Fig. 7*A*). To test in a more direct way whether distinguishing between weak and strong $T_5$ termination signals is an intrinsic property of Pol III, we exploited the reported ability of Pol III to initiate transcription on linear templates containing a 3′-overhanging end (41). Linear, 3′-overhanged versions of L(CAA)LR2, L_3′N and SCR1Δ_3′L templates were prepared as described under "Experimental Procedures" and transcribed *in vitro* with purified RNA polymerase III in the absence of any transcription factor. As shown in Fig. 7*B*, and in agreement with a previous study (41), purified Pol III specifically and efficiently transcribed these templates. Transcription initiation was made to occur at the 3′-overhanging end generated by SacI cleavage, at a position 5 bp upstream of the natural

[3] The very low stability of 3′-extended derivatives with respect to 3′-matured products of tRNA genes, together with the presence in multiple copies of most tRNA genes (21), generally complicates the detection of *in vivo* termination products in the case of natural tRNA genes.

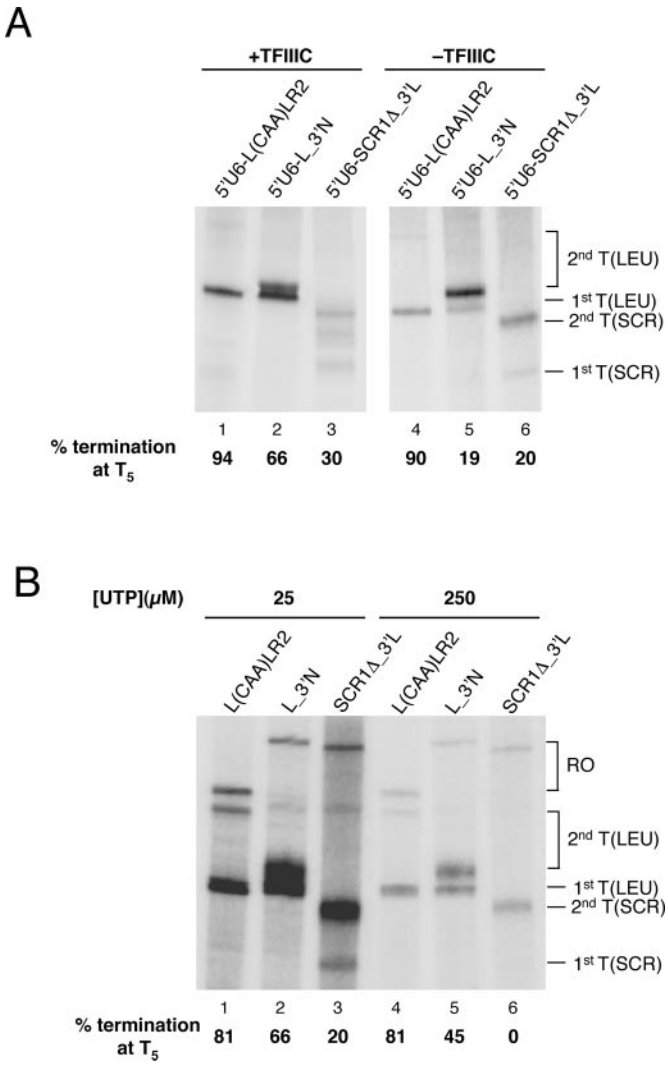



FIG. 7. **Transcriptional readthrough of the $T_5$ terminator is an intrinsic property of Pol III.** *A*, *in vitro* transcription of L(CAA)LR2, L_3′N, and SCR1Δ_3′L derivatives fused to the 5′-flanking region of *SNR6* was carried out at 250 μM UTP either in the presence (*lanes 1–3*) or in the absence (*lanes 4–6*) of TFIIIC. The percentage of transcripts terminated at the $T_5$ element is reported *below* each *lane*. All of the shown *lanes* were part of the same transcription gel. However, *lanes 4–6* derive from a ~10-fold darker exposure than *lanes 1–3*, to compensate for the lower transcription output obtained in the absence of TFIIIC. The migration positions of the transcripts terminated at the first or second T-run of either the *SCR1*-derived template (*1st T(SCR)* and *2nd T(SCR)*) or the L(CAA)LR2-derived templates (*1st T(LEU)* and *2nd T(LEU)*) are indicated on the *right*. *B*, 3′-overhanged restriction fragments containing the L(CAA)LR2, L_3′N and SCR1Δ_3′L templates were transcribed *in vitro* in the presence of purified Pol III and either 25 μM (*lanes 1–3*) or 250 μM (*lanes 4–6*) UTP. The percentage of transcripts terminated at the $T_5$ element is reported *below* each *lane*. The migration positions of the transcripts terminated at the first or second T-run of either the *SCR1*-derived template (*1st T(SCR)* and *2nd T(SCR)*) or the L(CAA)LR2-derived templates (*1st T(LEU)* and *2nd T(LEU)*) are indicated on the *right*, together with the migration position of run-off transcripts (*RO*).

transcription start site. Pol III terminated transcription at the same terminator region present in the original templates, in which a $T_5$ element of variable strength is followed by a back-up terminator (placed 67, 7, and 21 bp downstream of $T_5$ in L(CAA)LR2, L_3′N, and SCR1Δ_3′L, respectively). A small percentage of transcription complexes (10% at most) failed to recognize both the $T_5$ and back-up terminators, thus giving rise to run-off products (the DNA fragments used as templates ended 99, 238, and 249 bp downstream of $T_5$ in L(CAA)LR2, L_3′N, and SCR1Δ_3′L, respectively). Such a small amount of run-off transcripts probably derives from transcription complexes unable to displace the RNA from RNA/DNA hybrids, a reported artifact of initiating transcription at 3′-overhanged or dC-tailed templates (6, 41). Run-off transcripts were thus not considered for quantification purposes. The $T_5$ signal in the wild type context of L(CAA)LR2 acted as a rather strong terminator, both at 25 μM and at 250 μM UTP, with only 19% of transcription complexes reading through and terminating at the $T_7$ back up terminator (Fig. 7*B*, *lanes 1* and *4*). In contrast, the $T_5$ termination signal was remarkably weakened in the L_3′N and SCR1Δ_3′L contexts, both at 25 μM UTP (*lanes 2* and *3*) and at 250 μM UTP (*lanes 5* and *6*). The weakening effect was dramatic in the case of SCR1Δ_3′L, for which no transcripts terminated at $T_5$ could be detected at 250 μM UTP (*lane 6*). Weakening of $T_5$ by both the upstream and the downstream sequence contexts thus clearly reflect intrinsic recognition properties of Pol III. By comparing the extents of $T_5$ weakening observed at 250 μM UTP in the absence of transcription factors with those observed under the same conditions with a complete transcription system (*e.g.* see Fig. 5, *lane 2*, and Fig. 6, *lane 2*), we note that the weakening effect is significantly enhanced in the factor-free system. Such a $T_5$ read-through enhancement can at least in part be attributed to the absence of TFIIIC (see above in reference to Fig. 7*A*).

*Antitermination within a Natural Pol III Transcription Unit*—Recent studies of the genome-wide localization of the Pol III transcription machinery have identified the snoRNA encoding *SNR52* gene as a new class III gene (42–44). Northern blot and primer extension analyses suggested that the primary transcript of *SNR52* is a ~250-nt precursor RNA from which a leader sequence of 160 nt is cleaved to generate the ~90-nt mature snoRNA product, and sequence analysis revealed the presence of putative A- and B-block elements within the leader sequence (42, 43). An unusual feature of this sequence, however, is that a run of 6 T residues is present within the transcribed region, a few bp downstream of the putative A-block. The $T_6$ sequence is a highly efficient termination signal for yeast Pol III, as illustrated by the results shown in Fig. 4*B* (*lane 16*) and by the results of *in vitro* transcription experiments with several tDNAs bearing a $T_6$ terminator (data not shown). As shown in Fig. 8 (*lane 1*), *in vitro* transcription of *SNR52*, in the presence of 25 μM UTP, produced two main transcripts. The longer transcript was identified as the *SNR52* primary transcript on the basis of its size (250 nt). The shorter RNA had the size expected for a transcript terminated at the intragenic $T_6$ element. Indeed, transcription of a mutant *SNR52* template, in which $T_6$ was changed to TTCGTT, produced higher levels of the longer transcript, whereas the shorter transcript disappeared (*cf. lanes 1* and *2*). Quantification of the transcripts in *lane 1*, corrected for the number of incorporated U residues, led us to estimate that as much as 70% of the polymerases read through the $T_6$ element in *SNR52*. $T_6$ thus behaves as an unusually weak terminator in the *SNR52* context, even under conditions (low UTP concentration) that favor termination at T stretches. The inactive $T_6$ element of *SNR52* is located just downstream of a putative A-block promoter element within the transcribed region. Since TFIIIC is expected to contact such region during transcription, this factor might be required in order for Pol III to read through the internal $T_6$. To test this possibility, the *SNR52* transcription unit was fused with the 5′-flanking region of the *S. cerevisiae SNR6* gene to allow for TFIIIC-independent transcription of the resulting chimeric template. As shown in Fig. 8*B*, the presence or absence of TFIIIC on *SNR52* had no influence on the termination behavior of Pol III, characterized by highly inef-

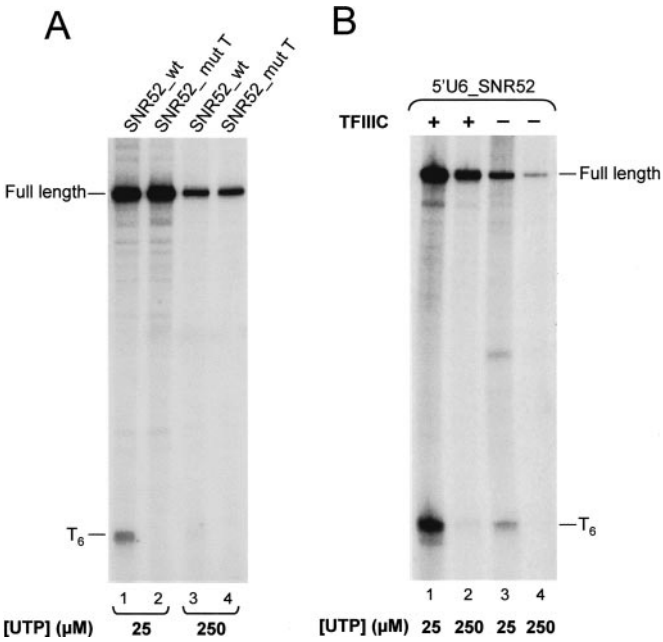



FIG. 8. **$T_6$ read-through within a natural Pol III transcription unit.** *A*, *in vitro* transcription of the *SNR52* gene and of the SNR52_mutT variant (in which the internal $T_6$ element is disrupted). The reactions in *lanes 1* and *2* were carried out in the presence of 25 μM UTP. The reactions in *lanes 3* and *4* were carried out in the presence of 250 μM UTP. The migration positions of the full-length primary transcript and of the $T_6$-terminated transcript are indicated on the *left*. *B*, *in vitro* transcription of the 5′U6-SNR52 template was carried out either in the presence (*lanes 1* and *2*) or in the absence (*lanes 3* and *4*) of TFIIIC. The reaction mixtures contained either 25 μM UTP (*lanes 1* and *3*) or 250 μM UTP (*lanes 2* and *4*).

ficient termination at the internal $T_6$ element at 250 μM UTP (*cf. lanes 2* and *4*).

## DISCUSSION

This work demonstrates that the recognition of oligo(dT) termination signals by *S. cerevisiae* RNA polymerase III can be strongly influenced by the sequence context in which the T stretch is embedded. Such influence becomes evident with the shortest termination signals encountered in yeast class III genes, namely $T_5$ and $T_6$, that together characterize ~40% of the yeast tDNAs possessing a single T run as a termination signal (see Fig. 1). We have been able to distinguish between two types of sequence context effect; one, that we have more extensively addressed, is due to the few base pairs immediately downstream of $T_5$; the other was observed when a non-tDNA sequence was fused upstream of a strong terminator region (composed of $T_5$ plus a particular 3′-flank) derived from L(CAA)LR2 tDNA. An extreme example of sequence context effect was observed with the *SNR52* transcription unit, in which a $T_6$ element placed between the A-block and the B-block promoter elements was barely recognized as a terminator by Pol III.

The main conclusion emerging from the study of the downstream sequence effect is that $T_5$, when associated with tRNA genes, is intrinsically strong as a terminator, because it induces highly efficient termination in most of the contexts analyzed; however, features of the downstream sequence exist that can significantly weaken the $T_5$ termination potential. A specific weakening element identified in this study seems to minimally consist of the CT dinucleotide placed immediately downstream of $T_5$ in the coding DNA strand. The tested natural tDNAs displaying such a feature were characterized by incomplete termination at $T_5$. The 3′-flank of N(GTT)CR, one of these tDNAs, could be transplanted 3′ to another tDNA, L(CAA)LR2, without losing its $T_5$ terminator weakening properties. By extensive mutagenesis, we were able to show that $T_5$ weakening absolutely requires a C residue immediately downstream of $T_5$ and is greatly favored by the presence of at least one T just after the C. With this respect, it is interesting to note that, in an early study of yeast Pol III termination, termination at $T_5$ was found to be incomplete, both *in vitro* and *in vivo*, for a template in which $T_5$ was followed by CT (4). $T_5$ terminator weakening might not be an exclusive property of the CT element, and other short weakening sequences may exist. However, neither natural tDNA analysis nor extensive mutagenesis of the N(GTT)CR-derived weakening element could reveal any other sequence with higher or similar weakening activity. By which mechanism does CT downstream of $T_5$ induce terminator read-through? Our data suggest that CT facilitates Pol III translocation by reducing its pause time at $T_5$. Indeed, read-through tends to be lesser under conditions (such as low UTP or CTP concentrations) that increase the dwell time of Pol III across the $T_5$ tract. We also tested the sensitivity of $T_5$ terminator weakening to variations in KCl concentration (50–150 mM range) and $Mg^{2+}$ ion concentration (3–10 mM range), which have the potential to influence the RNA release step of intrinsic bacterial termination (34), but we did not observe any significant variation in termination efficiency at $T_5$ either in the L(CAA)LR2 or in the L_3′N contexts (data not shown). Possible mechanistic explanations of the effect of downstream DNA on translocation include the stability of the duplex that must be melted or NTP binding to a melted region of downstream DNA, which would facilitate translocation (46–49). Our data tend to exclude a simple correlation between $T_5$ terminator weakening by downstream DNA and the ease of duplex melting. For example, template 6 in Fig. 4 displayed the same duplex stability as template 2 in the region immediately downstream of $T_5$ ($CA_4$ *versus* $CT_4$), yet the weakening effect was ~2-fold lower with $CA_4$ than with $CT_4$. Also, the introduction, in template 9 of Fig. 4, of an A residue before $CT_4$ completely abolished the weakening effect of $CT_4$ without changing significantly the duplex stability of this region. Our data rather suggest that the CT element might increase the processivity of Pol III at the position immediately upstream, possibly by a mechanism involving the binding of the corresponding NTPs to an allosteric NTP binding site, as recently proposed to explain downstream DNA effects on elongation by both *E. coli* RNA polymerase (49) and yeast RNA polymerase II (48).

Having established that $T_5$, followed by the sequence naturally present in the L(CAA)LR2 tDNA, behaves as a strong terminator for Pol III, even when fused 3′ to a completely different tDNA (see Fig. 3), we were surprised to find that the same sequence becomes a much weaker terminator when fused 3′ to an A- and B-block-containing segment of the *SCR1* gene. Even more surprising was the finding that the $T_6$ sequence, generally producing highly efficient Pol III termination, was substantially inactive as a terminator in the context of the *SNR52* gene, where it is naturally present between the A-block and B-block promoter elements preceding the snoRNA coding sequence. Although the sequence and positional requirements for both the *SCR1* and the *SNR52* effects were not investigated in detail, in both cases read-through was found to be favored at higher UTP concentrations, thus indicating that these effects involve an alteration of Pol III pausing. The results of experiments in which transcription initiation and termination were made to occur in the absence of TFIIIC further suggest that $T_5/T_6$ read-through in these two contexts does not require TFIIIC. Rather, in the absence of TFIIIC, read-through tended to be more pronounced, an observation that might be explained by postulating that DNA-bound TFIIIC slightly slows RNA

chain elongation, thus favoring termination at $T_5$. DNA-bound TFIIIC has previously been shown to exert no significant effect on RNA chain elongation by Pol III in the sense direction of the *SUP4* $tRNA^{Tyr}$ gene (41). It cannot be excluded, however, that in specific sequence contexts, TFIIIC contributes to Pol III pausing and termination. By *in vitro* transcription experiments conducted in the presence of purified Pol III and template DNA only, we showed that $T_5$ terminator weakening also occurs in the absence of any transcription factor and thus reflects intrinsic recognition properties of Pol III, a conclusion in agreement with that of an early study of termination by *Xenopus* Pol III (2). The $T_5$ weakening effect was especially evident when the strong L(CAA)LR2 terminator region was placed downstream of an A- and B-block-containing fragment of the *SCR1* gene. In this case, $T_5$ read-through enhancement can only depend on particular features of the upstream DNA region. This region might influence the Pol III behavior at $T_5$ through polymerase-DNA interactions, but it might also play an active role in the termination process through the corresponding RNA product, as is well known for the stem-loop coding portion of intrinsic bacterial terminators (50, 51).

The lack of termination at $T_6$ within the *SNR52* transcription unit is of evident biological significance, since the snoRNA coding portion of *SNR52* is entirely located downstream of the $T_6$ element. Curiously, such a sequence, which is potentially detrimental to *SNR52* gene transcription, is also present in the genomes of at least two other *Saccharomyces* species (*Saccharomyces paradoxus* and *Saccharomyces mikatae* (52)), thus indicating a possible relevant role in *SNR52* expression. In principle, the $T_6$ element might influence the accessibility and/or transcriptional activity of *SNR52 in vivo*, or it might play a role in pre-snoRNA processing. Less evident is the biological role of $T_5$ terminator weakening by short downstream sequences like CT in the case of tRNA genes. As a result of $T_5$ read-through, Pol III is expected to synthesize *in vivo* tRNA precursors with elongated 3′ trailers, that might in principle stabilize the pre-tRNA from 3′ exonucleolytic digestion. In organisms from yeast to humans, binding of La protein also protects the pre-tRNA 3′ trailer from exonucleolytic digestion (reviewed in Ref. 53). La protein binds the $UUU_{OH}$ at the 3′ end of nascent Pol III transcripts, and, by favoring the formation of correctly folded pre-tRNAs, it can affect pre-tRNA stability, processing, localization, and charging (53–56). The 3′ trailer extension caused by Pol III read-through might influence each of these processes either directly or by affecting La binding. In support of this view, it has recently been shown that an increase in 3′-terminal oligo(U) length of pre-tRNAs can influence association with La and tRNA maturation (57). The modulation of Pol III terminator readthrough might be of even higher significance in mammals. In the mouse and human genomes, $T_4$ is by far the most frequent termination signal for tRNA gene transcription, and in many cases, it is the sole potential termination signal encountered 3′ to the tDNA. Since runs of 4 T residues are often present within the internal transcribed region of tRNA genes, there must be features of the sequence context capable of strongly affecting Pol III termination at $T_4$. Such features, however, are not simply represented by conserved sequence elements flanking the $T_4$ terminator, as revealed by statistical analysis of such regions in mouse and human tDNA data sets.[4] Why did mammalian tDNA terminators evolve toward the shortest possible T runs? One possible answer is that, in the presence of weak T run terminators, more room is left for the modulation of termination by accessory factors, such as NF1, affecting transcription termination by human Pol III (39). Alternatively, the predominance of very short T runs downstream of mammalian tDNAs might reflect the importance of short 3′ polyuridylate tails for the post-transcriptional fate of mammalian pre-tRNAs.

[4] R. Percudani, unpublished observations.

*Acknowledgments*—We thank George Kassavetis, Christophe Carles, and Olivier Lefebvre for insightful discussions; Simone Ottonello for critical reading of the manuscript; Roberto Ferrari and Elisa Guffanti for *SNR52* constructs; and Joël Acker for recombinant Bdp1.

REFERENCES

1. Bogenhagen, D. F., and Brown, D. D. (1981) *Cell* **24,** 261–270
2. Cozzarelli, N. R., Gerrard, S. P., Schlissel, M., Brown, D. D., and Bogenhagen, D. F. (1983) *Cell* **34,** 829–835
3. Hamada, M., Sakulich, A. L., Koduru, S. B., and Maraia, R. J. (2000) *J. Biol. Chem.* **275,** 29076–29081
4. Allison, D. S., and Hall, B. D. (1985) *EMBO J.* **4,** 2657–2664
5. Matsuzaki, H., Kassavetis, G. A., and Geiduschek, E. P. (1994) *J. Mol. Biol.* **235,** 1173–1192
6. Campbell, F. E., Jr., and Setzer, D. R. (1992) *Mol. Cell. Biol.* **12,** 2260–2272
7. Bobkova, E. V., and Hall, B. D. (1997) *J. Biol. Chem.* **272,** 22832–22839
8. Thuillier, V., Brun, I., Sentenac, A., and Werner, M. (1996) *EMBO J.* **15,** 618–629
9. Chedin, S., Riva, M., Schultz, P., Sentenac, A., and Carles, C. (1998) *Genes Dev.* **12,** 3857–3871
10. Shaaban, S. A., Bobkova, E. V., Chudzik, D. M., and Hall, B. D. (1996) *Mol. Cell. Biol.* **16,** 6468–6476
11. Dieci, G., and Sentenac, A. (1996) *Cell* **84,** 245–252
12. Dieci, G., and Sentenac, A. (2003) *Trends. Biochem. Sci.* **28,** 202–209
13. Ferrari, R., Rivetti, C., Acker, J., and Dieci, G. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 13442–13447
14. Prescott, E. M., Osheim, Y. N., Jones, H. S., Alen, C. M., Roan, J. G., Reeder, R. H., Beyer, A. L., and Proudfoot, N. J. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 6068–6073
15. Guffanti, E., Corradini, R., Ottonello, S., and Dieci, G. (2004) *J. Biol. Chem.* **279,** 20708–20716
16. Mazabraud, A., Scherly, D., Muller, F., Rungger, D., and Clarkson, S. G. (1987) *J. Mol. Biol.* **195,** 835–845
17. Gunnery, S., and Mathews, M. B. (1995) *Mol. Cell. Biol.* **15,** 3597–3607
18. Gunnery, S., Ma, Y., and Mathews, M. B. (1999) *J. Mol. Biol.* **286,** 745–757
19. Emerson, B. M., and Roeder, R. G. (1984) *J. Biol. Chem.* **259,** 7926–7935
20. Hess, J., Perez-Stable, C., Wu, G. J., Weir, B., Tinoco, I., Jr., and Shen, C. K. (1985) *J. Mol. Biol.* **184,** 7–21
21. Percudani, R., Pavesi, A., and Ottonello, S. (1997) *J. Mol. Biol.* **268,** 322–330
22. Lowe, T. M., and Eddy, S. R. (1997) *Nucleic Acids Res.* **25,** 955–964
23. Stajich, J. E., Block, D., Boulez, K., Brenner, S. E., Chervitz, S. A., Dagdigian, C., Fuellen, G., Gilbert, J. G., Korf, I., Lapp, H., Lehvaslaiho, H., Matsalla, C., Mungall, C. J., Osborne, B. I., Pocock, M. R., Schattner, P., Senger, M., Stein, L. D., Stupka, E., Wilkinson, M. D., and Birney, E. (2002) *Genome Res.* **12,** 1611–1618
24. Giuliodori, S., Percudani, R., Braglia, P., Ferrari, R., Guffanti, E., Ottonello, S., and Dieci, G. (2003) *J. Mol. Biol.* **333,** 1–20
25. Baker, R. E., Eigel, A., Vögtel, D., and Feldmann, H. (1982) *EMBO J.* **1,** 291–295
26. Dieci, G., Giuliodori, S., Catellani, M., Percudani, R., and Ottonello, S. (2002) *J. Biol. Chem.* **277,** 6903–6914
27. Higuchi, R. (1990) in *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds) pp. 177–183, Academic Press, Inc., San Diego
28. Dieci, G., Percudani, R., Giuliodori, S., Bottarelli, L., and Ottonello, S. (2000) *J. Mol. Biol.* **299,** 601–613
29. Huet, J., Manaud, N., Dieci, G., Peyroche, G., Conesa, C., Lefebvre, O., Ruet, A., Riva, M., and Sentenac, A. (1996) *Methods Enzymol.* **273,** 249–267
30. Kassavetis, G. A., Joazeiro, C. A. P., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S., and Blanco, J. A. (1992) *Cell* **71,** 1055–1064
31. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J. Bacteriol.* **153,** 163–168
32. Wise, J. A. (1991) *Methods Enzymol.* **194,** 405–415
33. Geiduschek, E. P., and Tocchini-Valentini, G. P. (1988) *Annu. Rev. Biochem.* **57,** 873–914
34. Reynolds, R., Bermudez-Cruz, R. M., and Chamberlin, M. J. (1992) *J. Mol. Biol.* **224,** 31–51
35. Matthews, C. (1972) *J. Biol. Chem.* **247,** 7430–7438
36. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 3746–3750
37. Klemenz, R., Stillman, D. J., and Geiduschek, E. P. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 6191–6195
38. Camier, S., Gabrielsen, O., Baker, R., and Sentenac, A. (1985) *EMBO J.* **4,** 491–500
39. Wang, Z., Bai, L., Hsieh, Y. J., and Roeder, R. G. (2000) *EMBO J.* **19,** 6823–6832
40. Moenne, A., Camier, S., G., A., Margottin, F., Beggs, J., and Sentenac, A. (1990) *EMBO J.* **9,** 271–277
41. Bardeleben, C., Kassavetis, G. A., and Geiduschek, E. P. (1994) *J. Mol. Biol.* **235,** 1193–1205
42. Harismendy, O., Gendrel, C. G., Soularue, P., Gidrol, X., Sentenac, A., Werner, M., and Lefebvre, O. (2003) *EMBO J.* **22,** 4738–4747
43. Roberts, D. N., Stewart, A. J., Huff, J. T., and Cairns, B. R. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 14695–14700
44. Moqtaderi, Z., and Struhl, K. (2004) *Mol. Cell. Biol.* **24,** 4118–4127
45. Mao, J., Schmidt, O., and Soll, D. (1980) *Cell* **21,** 509–516

46. Palangat, M., Hittinger, C. T., and Landick, R. (2004) *J. Mol. Biol.* **341,** 429–442
47. Yarnell, W. S., and Roberts, J. W. (1999) *Science* **284,** 611–615
48. Nedialkov, Y. A., Gong, X. Q., Hovde, S. L., Yamaguchi, Y., Handa, H., Geiger, J. H., Yan, H., and Burton, Z. F. (2003) *J. Biol. Chem.* **278,** 18303–18312
49. Holmes, S. F., and Erie, D. A. (2003) *J. Biol. Chem.* **278,** 35597–35608
50. von Hippel, P. H. (1998) *Science* **281,** 660–665
51. Nudler, E., and Gottesman, M. E. (2002) *Genes Cells* **7,** 755–768
52. Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E. S. (2003) *Nature* **423,** 241–254
53. Wolin, S. L., and Cedervall, T. (2002) *Annu. Rev. Biochem.* **71,** 375–403
54. Maraia, R. J., and Intine, R. V. (2002) *Gene Expr.* **10,** 41–57
55. Intine, R. V., Dundr, M., Misteli, T., and Maraia, R. J. (2002) *Mol. Cell* **9,** 1113–1123
56. Chakshusmathi, G., Kim, S. D., Rubinson, D. A., and Wolin, S. L. (2003) *EMBO J.* **22,** 6562–6572
57. Huang, Y., Intine, R. V., Mozlin, A., Hasson, S., and Maraia, R. J. (2005) *Mol. Cell. Biol.* **25,** 621–636
58. Schmidt, O., Mao, J., Ogden, R., Beckmann, J., Sakano, H., Abelson, J., and Soll, D. (1980) *Nature* **287,** 750–752

doi:10.1016/j.jmb.2003.08.016 J. Mol. Biol. (2003) 333, 1–20

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

# A Composite Upstream Sequence Motif Potentiates tRNA Gene Transcription in Yeast

**Silvia Giuliodori, Riccardo Percudani, Priscilla Braglia, Roberto Ferrari Elisa Guffanti, Simone Ottonello* and Giorgio Dieci***

*Dipartimento di Biochimica e Biologia Molecolare, Università di Parma, Parco Area delle Scienze 23/A, 43100 Parma Italy*

*Corresponding authors*

Transcription of eukaryotic tRNA genes relies on the TFIIIC-dependent recruitment of TFIIIB on a ~50 bp region upstream of the transcription start site (TSS). TFIIIC specifically interacts with highly conserved, intragenic promoter elements, while the contacts between TFIIIB and the upstream DNA have long been considered as largely non-specific. Through a computer search procedure designed to detect shared, yet degenerate sequence features, we have identified a conserved sequence pattern upstream of *Saccharomyces cerevisiae* tDNAs. This pattern consists of four regions in which particular sequences are over-represented. The most downstream of these regions surrounds the TSS, while the other three districts of sequence conservation (appearing as a centrally located TATA-like sequence flanked by T-rich elements on both sides) are located across the DNA region known to interact with TFIIIB. Upstream regions whose sequence conforms to this pattern were found to potentiate tRNA gene transcription, both *in vitro* and *in vivo*, by enhancing TFIIIB binding. A conserved pattern of DNA bendability was also revealed, with peaks of bending propensity centered on the TATA-like and the TSS regions. Sequence analysis of other eukaryotic genomes further revealed the widespread occurrence of conserved sequence patterns upstream of tDNAs, with striking lineage-specific differences in the number and sequence of conserved motifs. Our data strongly support the notion that tRNA gene transcription in eukaryotes is modulated by composite TFIIIB binding sites that may confer responsiveness to variation in TFIIIB activity and/or concentration.

© 2003 Elsevier Ltd. All rights reserved.

*Keywords:* RNA polymerase III; TFIIIB; transcription; tRNA genes; 5′-flanking sequence

## Introduction

The promoters of eukaryotic tRNA genes (tDNAs) are essentially intragenic. Two highly conserved internal control regions, the A and B-blocks, form together a specific binding site for the multisubunit transcription factor (TF) IIIC. Promoter-bound TFIIIC provides an interaction platform for the productive assembly of TFIIIB on a ~50 bp region upstream of the transcription start site (TSS).[1] TFIIIB, which in budding yeast is made up by the three proteins TBP, Brf1 and Bdp1,[2] is considered as the true RNA polymerase (Pol) III initiation factor, because the TFIIIB–DNA complex is by itself capable of productively recruiting Pol III and supporting multiple rounds of transcription *in vitro*.[3] Despite the remarkable stability of the TFIIIB–DNA complex and its centrality in the transcription mechanism, TFIIIB contacts with its upstream DNA binding region have thus far been considered to be very poorly (or not at all) sequence-specific. This fact, together with the high level of sequence conservation of the internal control regions, has generated the idea that 5′-flanking sequences do not contribute to the promoter strength of tRNA genes, or do it only

Abbreviations used: Bdp1, TFIIIB double prime; Brf1, TFIIB-related factor 1; Pol III, RNA polymerase III; snRNA, small nuclear RNA; TBP, TATA box-binding protein; tDNA, tRNA gene; TF, transcription factor; TSS, transcription start site.

E-mail addresses of the corresponding authors: s.ottonello@unipr.it; giorgio.dieci@unipr.it

0022-2836/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.

sporadically. Evidence accumulated in the last two decades, however, argues against this simple assumption. The differential *in vitro* activity of silkgland-specific and constitutive $tRNA^{Ala}$ genes in the silkworm, *Bombyx mori*, has long been known to be due to sequence elements upstream of the TSS,[4] and it has later been shown to rely on the interaction of the TBP component of TFIIIB with a TATA-like element upstream of the constitutive gene.[5,6] Recent analyses have further revealed the presence of TATA-like elements upstream of many *Saccharomyces cerevisiae* tDNAs, and of the majority of *Arabidopsis thaliana* and *Schizosaccharomyces pombe* tDNAs.[7,8] The presence of upstream TATA elements is thus not restricted to particular class III genes, such as *SNR6* (coding for the U6 snRNA), that have long been known to support TFIIIC-independent TFIIIB recruitment *in vitro*.[9] Attesting to the functional significance of upstream TATA-like elements, a number of budding yeast tDNAs, bearing a consensus TATA box (TATAAATA) around position −30, were found to be transcribed *in vitro* in the absence of TFIIIC,[8] and the TATA box upstream of most *S. pombe* tDNAs was shown to be important for their expression *in vitro* and *in vivo*.[10] Besides the involvement of TATA-like sequences in tRNA gene transcription that is in agreement with the essential role of TBP in TFIIIB–DNA complex formation, many other examples of stimulatory (and, less often, inhibitory) effects of the upstream region on tRNA gene transcription have been reported, that cannot be simply attributed to the presence or absence of TATA-like elements.[11,12] In the case of budding yeast, *in vitro* and *in vivo* studies on selected tRNA genes have shown that mutations in the upstream region known to be occupied by TFIIIB in tRNA and 5 S rRNA genes never abolish transcription, but often result in a marked decrease in transcription efficiency.[13–15] This supports the idea that there are few critical, sequence-specific contacts between TFIIIB and 5′-flanking sequences, but that, nevertheless, TFIIIB binding is somehow influenced by as yet poorly defined sequence features of the 5′ flank. Other studies in yeast have shown that a 15 bp sequence element contiguous with the 5′ end of the mature tRNA coding sequence can act as a positive modulator of *in vitro* and *in vivo* transcription.[16] This region includes a conserved motif, surrounding the TSS, that influences transcription by increasing the rate of transcription initiation by Pol III, rather than the extent of TFIIIB–DNA complex formation.[17–19]

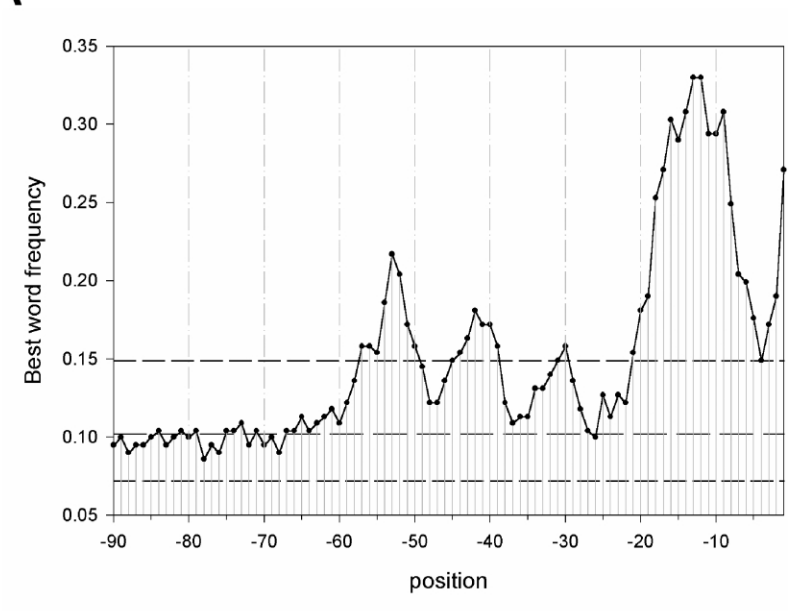
The molecular anatomy of the TFIIIB–DNA complex assembled on the yeast *SUP4* tRNA gene has been explored in great detail by systematic protein–DNA photocrosslinking analyses using both wild-type and mutant TFIIIB recombinant components.[2] The emerging picture is that of a tight nucleoprotein assembly, in which TBP occupies a key position and contacts a TATA-like element located approximately around the center of the TFIIIB-interacting segment. TBP, which is known to bend DNA when bound to the TATA box of Pol II promoters, supposedly does the same in the case of tRNA genes and, accordingly, the DNA was found to be bent sharply in TFIIIB–DNA complexes assembled upstream of tRNA (and 5 S rRNA) coding regions.[20,21] Of the other two TFIIIB components, Brf1 participates in TFIIIB–DNA complex formation by creating an extended connection between the opposite sides of the bent DNA, while Bdp1 generates an additional bend between the TSS and the TBP-binding site, extends the TFIIIB–DNA contacts upstream of the TATA box, and confers to the complex its distinctive tightness.[2] Very recently, the crystal structure of a ternary complex containing the TBP core domain, the Brf1 homology domain II (essential for TBP and Bdp1 anchoring) and 19 bp of U6 promoter DNA has been solved. The structure reveals an extensive interaction of this helical Brf1 domain with the convex surface of TBP, with an exceptionally large number of interactions between the two proteins.[22] Besides the function of TFIIIB in recruiting Pol III to the transcription initiation site, Brf1 and Bdp1 are further involved in post-recruitment steps of initiation.[2] Although many details of the TFIIIB–DNA interaction are still unknown, a simplified picture can be drawn in which three districts of upstream DNA are coordinately involved in TFIIIB binding. A central region, located at around −30 with respect to the TSS, provides the interaction surface for TBP, while the two adjacent regions serve as anchoring platforms for both Brf1 and Bdp1.[23,24] Even if all the distinctive physical properties of TFIIIB–tDNA complexes have been demonstrated to be essentially sequence-independent,[2] it seems quite surprising that the upstream region of tDNAs and other class III genes has not undergone any kind of selective pressure towards TFIIIB-binding optimization. With this in mind, we have carried out a genome-wide analysis of tDNA 5′-flanking regions in *S. cerevisiae*, and report here theoretical and experimental evidence for the existence and functional significance of a previously unidentified conserved sequence pattern. We further show that variants of such a pattern characterize the tDNA upstream regions in other eukaryotic genomes.

## Results

### Conserved sequence elements upstream of yeast tDNAs

The *S. cerevisiae* genome contains 274, mostly unclustered tRNA genes dispersed throughout its 16 chromosomes.[25,26] This redundant set of genes codes for 42 tRNA species with different codon specificities. The gene copy number for individual tRNA species ranges from 1 to 16 and correlates well with both the frequency of codon occurrence and the intracellular abundance of individual

A




B

| POSITION | FREQUENCY | SEQUENCE |
|---|---|---|
| -54 | 0.186 | ATTTTT |
| **-53** | **0.217** | **TTTTTA** |
| -52 | 0.204 | TTTTTA |
| -43 | 0.163 | TATTAA |
| **-42** | **0.181** | **TATTAA** |
| -41 | 0.172 | ATATAG |
| -40 | 0.172 | TATATA |
| -39 | 0.158 | TATATT |
| -33 | 0.131 | ATTTTA |
| -32 | 0.140 | TCTTTA |
| -31 | 0.149 | ATTTTA |
| **-30** | **0.158** | **TTTTAT** |
| -29 | 0.136 | TTTAAT |
| -17 | 0.271 | TTTTCA |
| -16 | 0.303 | TTTCAA |
| -15 | 0.290 | TTTCAA |
| -14 | 0.308 | CAACAA |
| **-13** | **0.330** | **CAACAA** |
| -12 | 0.330 | AACAAA |
| -11 | 0.294 | AACAAA |

**Figure 1**. Non-random sequence elements in the 5′-flanking regions of *S. cerevisiae* tDNAs. A, Individual tDNA upstream region positions, comprised between position 0 (first nucleotide of the mature tRNA) and −90, are indicated on the *x*-axis. The rates of occurrence of the most frequent six-letter words at each position (Best word frequency) are reported on the *y*-axis. The middle broken line indicates the average frequency value obtained from randomized datasets. The upper and lower lines correspond, respectively, to the maximum and minimum frequency values derived from the analysis of the same sets of randomly shuffled sequences. B, For each upstream region position (listed on the left column), the most frequent sequence (as specified in A) is reported on the right column, and the corresponding frequency is indicated in the middle column. Rows corresponding to the sequences most frequently occurring in each of the four partially conserved regions are shown in bold and are marked with arrows.

tRNAs.[25] Because of tDNA redundancy, and of the intragenic location of A and B-blocks, genes coding for the same tRNA species most often share identical TFIIIC-binding promoter elements, but are flanked by completely different TFIIIB-interacting sequences. Yeast tDNAs often carry a closely associated Ty retroelement in their 5′ flank.[26] Since sequence conservation of these elements might have biased our search for transcription-related conserved features, we excluded from our analysis the 53 tDNAs carrying, within the first 100 bp upstream of the mature tRNA coding sequence, a sigma, delta or tau retroelement. The 5′-flanking regions of the remaining 221 tDNAs (up to 90 bp upstream of the position corresponding to the starting nucleotide of the mature tRNA) were aligned and analyzed by means of a specifically designed algorithm that allows us to search for the presence of *n*-letter words within given windows. When six-letter words, within 9 bp sequence windows (allowing for one mismatch) were searched, four regions displaying over-represented six-letter words (i.e. words recurring at a sufficiently high frequency to make them noticeable against the local sequence noise) were evidenced. A graphical representation of the results of this analysis is shown in Figure 1A. Sequences upstream of −60 are characterized by complete randomness, as revealed by the close overlapping between the six-letter word frequency

at each position and an average frequency value derived from the analysis of randomized sequence sets (represented by the middle broken line in Figure 1A). In sharp contrast, the best six-letter word frequencies in the DNA region downstream of −60 significantly move away from the background, with clearly defined peaks of sequence conservation around positions −53, −42, −30, and −13 with respect to the first nucleotide of the mature tRNA. The six-letter words that occur most frequently at these and at immediately adjacent positions are reported in Figure 1B. Sequences around −42 are TATA-like, while sequence conservation around −53 and −30, on both sides of the TATA-like motif, tends to favor T-richness. The sequences with the highest relative frequencies are those around position −13, and correspond to the signature of the Pol III transcription initiation region, as reported by previous studies.[16,18,19] Considering that a 12 bp spacing between the TSS and the first nucleotide of the mature tRNA is most commonly found in yeast tDNAs, the −30, −42 and −53 motifs define an upstream segment, extending from −18 to −41 with respect to the TSS, that encompasses most of the region known to be contacted by TFIIIB.[1] By comparison, the same type of sequence analysis, conducted on the aligned 3′-flanking sequences of yeast tDNAs, failed to reveal any region of significant sequence conservation, apart from the expected T-rich region corresponding to the Pol III transcription terminator (data not shown).

Despite the degenerate character of upstream sequence conservation (especially within the TFIIIB-interacting region), we thought it useful, in view of subsequent analyses, to classify the 221 tDNAs in our sample according to the conformity of their 5′-flanking regions with the word frequency pattern reported in Figure 1. For each individual tDNA upstream region, the positions from −54 to −52 (" − 53" group), from −43 to −39 (" − 42" group), from −33 to −29 (" − 30" group) and from −17 to −11 (" − 13" group) were considered. Such positions cover the four districts displaying the higher-than-average best word frequencies reported in Figure 1B. Each of these positions is considered as the first letter of a six-letter word (read in the 5′ to 3′ direction). This word, and the three contiguous six-letter words read by moving three base-pairs in the downstream direction, is attributed to the position being considered. If at least one of these four words conforms (with one mismatch allowed) to the best frequency word reported in Figure 1B for the particular position under examination, the tDNA upstream sequence is considered as scoring at this position and, consequently, it is qualified as positive with respect to the group ( − 53, − 42, − 30 or − 13) to which such a position belongs, and it is assigned a 1-point-score. Since four groups are being considered, the upstream region point-score values can range from 0 to 4. As shown in Table 1, when the 221 tDNAs in our sample were classified according to the above described criterion, we found that most tDNA multigene families are made up by members sharing the same coding region and intragenic promoter, but preceded by upstream regions with very different score values often ranging from 0 to 3 or 4. The tDNAs classified as positive with respect to the − 13 group were the majority (54% of the entire set), in accord with the fact that the −13 frequency peak is the most clearly defined within the sequence conservation profile in Figure 1A. In all, 45% of the tDNAs turned out to be positive for the − 42 region, while those positive for the − 30 and − 53 regions were only 32% and 26% of the entire gene set, respectively. Interestingly, about one-third of the − 13 region-positive tDNAs belong to only four tDNA families: $tDNA^{Arg}$(TCT), $tDNA^{Leu}$ (CAA), $tDNA^{Lys}$(CTT) and $tDNA^{Val}$(AAC). Within each of these families, the vast majority of the members are positive for the − 13 region. It should be noted, however, that this type of classification (despite its apparent simplicity) presents some limitations. For example, an upstream region scoring at several positions within a certain group, without any mismatch, is assigned a 1-point-score for that particular group, just as an upstream sequence scoring at only one position of the same group with one mismatch. More stringent classification criteria, however, led to a generalized flattening of the score differences, and were thus less useful in view of functional studies addressing individual tDNAs.

### *In vitro* transcription properties of iso-coding tRNA genes with different upstream regions

The fact that most members of iso-coding tRNA gene families have heterogeneous 5′ flanks with distinct score values prompted us to evaluate the effect of different 5′ flanks on the transcription of natural tDNAs with identical internal control regions. The genes coding for $tRNA^{Pro}$(AGG), $tRNA^{Met}$(CAT), and $tRNA^{Asn}$(GTT) were taken as representative of tDNA families whose 5′ flank point-score values vary from 0 to 4. The 5′-flanking sequences of the tDNAs selected for transcriptional analysis are reported in Figure 2A. tDNAs 1 and 2 are the only two members of the small $tRNA^{Pro}$ (AGG) gene family. The 5′ flank of one of such genes, tP(AGG)CR, has a point-score of 0, while a 2-point-score value was assigned to the other gene, tP(AGG)NR, due to its consensus pattern conformity within the − 30 and − 13 groups of positions. When the two $tDNA^{Pro}$(AGG) gene copies were tested in a purified *in vitro* transcription system,[8] the low-score copy was reproducibly transcribed 12-fold less efficiently than the high-score copy (cf. the first two lanes in Figure 2B). A qualitatively similar effect was observed with two members of the $tRNA^{Met}$ gene family, tM(CAT)JL1, with a point-score of 0, and tM(CAT)MR, with a 4-point-score value (tDNAs 3 and 4 in Figure 2A). In this case, the high-score tDNA copy was

**Table 1.** 5′ Flank classification of *S. cerevisiae* tRNA genes

| tDNA | SCORE | -53 | -42 | -30 | -13 |
|---|---|---|---|---|---|
| tA(AGC)GR1 | 2 | | | • | • |
| tA(AGC)HR | - | | | | |
| tA(AGC)JL | - | | | | |
| tA(AGC)KL | 1 | | • | | |
| tA(AGC)KR | 0 | | | | |
| tA(AGC)LR2 | - | | | | |
| tA(AGC)MR2 | 3 | • | | • | • |
| tA(AGC)MR1 | 3 | | • | • | • |
| tA(AGC)PR | - | | | | |
| tA(AGC)DL | 1 | | | | • |
| tA(AGC)FR | 1 | | | | • |
| tA(TGC)GR2 | 0 | | | | |
| tA(TGC)AR | 0 | | | | |
| tA(TGC)LR1 | 1 | | • | | |
| tA(TGC)OR | - | | | | |
| tA(TGC)E | - | | | | |
| tR(ACG)JL3 | 2 | • | • | | |
| tR(AGC)KR | 0 | | | | |
| tR(ACG)LR1 | - | | | | |
| tR(ACG)OR | 1 | | | | • |
| tR(ACG)DR2 | 1 | | | | • |
| tR(ACG)ER | - | | | | |
| tR(CCG)LR2 | 1 | | • | | |
| tR(CCT)JR | 1 | | | | • |
| tR(TCT)GL | - | | | | |
| tR(TCT)GR1 | - | | | | |
| tR(TCT)BRR | 3 | • | | • | • |
| tR(TCT)GR2 | 2 | | • | | • |
| tR(TCT)JL2 | 2 | | | • | • |
| tR(TCT)JL1 | 1 | | | | • |
| tR(TCT)KL | 2 | • | | | • |
| tR(TCT)MR | 1 | | | | • |
| tR(TCT)ML | 2 | | • | | • |
| tR(TCT)DR1 | 3 | • | • | | • |
| tR(TCT)EL | 2 | | • | | • |
| **tN(GTT)GR** | **0** | | | | |
| **tN(GTT)KL** | **1** | | • | | |
| **tN(GTT)CR** | **3** | • | • | • | |
| **tN(GTT)LR** | **2** | | • | • | |
| tN(GTT)NL | - | | | | |
| **tN(GTT)NR** | **0** | | | | |
| **tN(GTT)OL** | **2** | | • | | • |
| **tN(GTT)OR** | **4** | • | • | • | • |
| **tN(GTT)PR** | **1** | | | • | |
| **tN(GTT)FL** | **3** | • | • | • | |
| tD(GTC)GR1 | 1 | | • | | |
| tD(GTC)GR2 | 2 | | • | | • |
| tD(GTC)BR | - | | | | |
| tD(GTC)JL3 | 2 | • | | | • |
| tD(GTC)JL2 | - | | | | |
| tD(GTC)JR | 2 | • | • | | |
| tD(GTC)JL1 | - | | | | |
| tD(GTC)KR | 2 | | • | | |
| tD(GTC)LR1 | 2 | | | • | • |
| tD(GTC)LR2 | - | | | | |
| tD(GTC)MR | 1 | | • | | |
| tD(GTC)OR | 1 | | | | • |
| tD(GTC)DR | - | | | | |
| tD(GTC)IL1 | 3 | • | • | | • |
| tD(GTC)IL2 | - | | | | |
| tC(GCA)GR | - | | | | |
| tC(GCA)BRR | 2 | | • | • | |
| tC(GCA)PR | 1 | | • | | |
| tC(GCA)PL | 1 | | | • | |
| tQ(CTG)MR | 2 | | | • | • |
| tQ(TTG)BRR | 2 | • | | | • |
| tQ(TTG)LR | 2 | | | • | • |
| tQ(TTG)CR | 3 | • | • | | • |
| tQ(TTG)DR1 | 0 | | | | |

| tDNA | SCORE | -53 | -42 | -30 | -13 |
|---|---|---|---|---|---|
| tQ(TTG)DR3 | 1 | | • | | |
| tQ(TTG)DR2 | - | | | | |
| tQ(TTG)ER | - | | | | |
| tQ(TTG)EL | - | | | | |
| tE(CTC)DR | 2 | | | • | • |
| tE(CTC)IL | 3 | • | • | • | |
| tE(TTC)GL2 | 4 | • | • | • | • |
| tE(TTC)GR | 3 | | • | • | • |
| tE(TTC)GL1 | - | | | | |
| tE(TTC)BR | 2 | | • | • | |
| tE(TTC)JL | 1 | • | | | |
| tE(TTC)KL | 1 | | | | • |
| tE(TTC)LR | 0 | | | | |
| tE(TTC)MR | 1 | | | • | |
| tE(TTC)CL | 3 | • | | • | • |
| tE(TTC)PL | 1 | | | • | |
| tE(TTC)IR | 0 | | | | |
| tE(TTC)ER1 | 2 | | • | | • |
| tE(TTC)ER2 | - | | | | |
| tE(TTC)ER3 | 1 | • | | | |
| tG(CCC)OR | 2 | • | • | | |
| G(CCC)DR2 | 1 | | • | | |
| tG(GCC)GR2 | - | | | | |
| tG(GCC)GR3 | 0 | | | | |
| tG(GCC)JR | 1 | | | | • |
| tG(GCC)JL | 2 | | • | | • |
| tG(GCC)BR | - | | | | |
| tG(GCC)ML | 2 | • | | | • |
| tG(GCC)OL1 | 2 | • | • | | |
| tG(GCC)CR | 2 | | | • | • |
| tG(GCC)OL2 | - | | | | |
| tG(GCC)PR1 | 1 | | • | | |
| tG(GCC)PR2 | 0 | | | | |
| tG(GCC)DL | 2 | | • | | • |
| tG(GCC)DR1 | - | | | | |
| tG(GCC)E | 2 | | | • | • |
| tG(GCC)FR1 | 2 | | • | | • |
| tG(GCC)FR2 | 0 | | | | |
| tG(TCC)GR1 | 2 | | • | | • |
| tG(TCC)NL | 0 | | | | |
| tG(TCC)OL3 | 1 | | | | • |
| tH(GTG)GL2 | 3 | • | • | | • |
| tH(GTG)HL | 0 | | | | |
| tH(GTG)KL | 1 | • | | | |
| tH(GTG)MR | 1 | | • | | |
| tH(GTG)ER1 | 1 | | • | | |
| tH(GTG)ER2 | - | | | | |
| tH(GTG)GL1 | - | | | | |
| tI(AAT)GR | 4 | • | • | • | • |
| tI(AAT)LR2 | 2 | | | • | • |
| tI(AAT)LR3 | 1 | • | | | |
| tI(AAT)NL1 | 1 | • | | | |
| tI(AAT)NL2 | 0 | | | | |
| tI(AAT)PR1 | 2 | | • | | • |
| tI(AAT)PR2 | - | | | | |
| tI(AAT)DR1 | - | | | | |
| tI(AAT)IL1 | - | | | | |
| tI(AAT)IL2 | 3 | | • | • | • |
| tI(AAT)BR | 1 | • | | | |
| tI(AAT)ER2 | 1 | | | • | |
| tI(AAT)ER1 | 2 | | | • | • |
| tI(TAT)LR1 | 1 | | • | | |
| tI(TAT)DR2 | 1 | | • | | |
| tL(CAA)GL1 | - | | | | |
| tL(CAA)GR2 | 1 | | | | • |
| tL(CAA)KR | 2 | | | • | • |
| tL(CAA)CL | 3 | | • | • | • |
| tL(CAA)LR2 | 3 | | • | • | • |
| tL(CAA)MR | 2 | | • | | • |

| tDNA | SCORE | -53 | -42 | -30 | -13 |
|---|---|---|---|---|---|
| tL(CAA)NL | 1 | | | | • |
| L(CAA)AR | 2 | | • | | • |
| tL(CAA)DR2 | 2 | | • | | • |
| tL(CAA)GL2 | 2 | | | • | • |
| tL(GAG)GR1 | 3 | | • | • | • |
| tL(TAA)BRR | 1 | | • | | |
| tL(TAA)JL | 2 | | • | | • |
| tL(TAA)KL | 0 | | | | |
| tL(TAA)LR5 | 0 | | | | |
| tL(TAA)NR | - | | | | |
| tL(TAA)DR1 | 1 | • | | | |
| tL(TAA)BRL | - | | | | |
| tL(TAG)JR | 2 | | | • | • |
| tL(TAG)LR4 | 4 | • | • | • | • |
| tL(TAG)LR1 | 2 | | | • | • |
| tK(CTT)GR2 | 2 | • | | | • |
| tK(CTT)GL1 | 4 | • | • | • | • |
| tK(CTT)JL | 3 | • | • | | • |
| tK(CTT)KL | - | | | | |
| tK(CTT)MR | - | | | | |
| tK(CTT)CR | 3 | • | • | | • |
| tK(CTT)PR1 | 2 | | • | | • |
| tK(CTT)DR1 | 3 | | • | • | • |
| tK(CTT)DR2 | 2 | • | | | • |
| tK(CTT)IL | 1 | | | | • |
| tK(CTT)ER | 1 | | | | • |
| tK(CTT)EL | 2 | | • | | • |
| tK(CTT)FR | 1 | | | • | |
| tK(CTT)GL2 | 2 | | • | | • |
| tK(TTT)GR1 | 1 | • | | | |
| tK(TTT)KR | 2 | | • | • | |
| tK(TTT)LR | 1 | | | • | |
| tK(TTT)OR | 2 | | • | • | |
| tK(TTT)PR2 | - | | | | |
| tK(TTT)DL | 1 | | | • | |
| tK(TTT)GL3 | 2 | | • | • | |
| **tM(CAT)JL1** | **0** | | | | |
| tM(CAT)JR | 3 | • | • | | • |
| tM(CAT)JL2 | 2 | | | • | • |
| tM(CAT)CR | 1 | | • | | |
| **tM(CAT)MR** | **4** | • | • | • | • |
| tM(CAT)OR1 | 2 | | • | | • |
| tM(CAT)OR2 | - | | | | |
| tM(CAT)PL | 0 | | | | |
| tM(CAT)DR | - | | | | |
| tM(CAT)E | 3 | • | • | | • |
| tF(GAA)GL | 0 | | | | |
| tF(GAA)HR2 | 1 | | | | • |
| tF(GAA)HR1 | 2 | | | • | • |
| tF(GAA)MR | 0 | | | | |
| tF(GAA)NL | 1 | | | | • |
| tF(GAA)PR2 | 2 | | • | | • |
| F(GAA)PR1 | 1 | | • | | |
| tF(GAA)DR | 1 | | • | | |
| tF(GAA)BR | 3 | • | • | • | |
| tF(GAA)FR | 3 | • | • | | • |
| **tP(AGG)NR** | **2** | | | • | • |
| **tP(AGG)CR** | **0** | | | | |
| tP(TGG)AL | 2 | | • | | • |
| tP(TGG)HR | 4 | • | • | • | • |
| tP(TGG)LL | 1 | | • | | |
| tP(TGG)ML | 2 | • | | • | |
| tP(TGG)NL2 | - | | | | |
| tP(TGG)NL1 | 1 | • | | | |
| tP(TGG)OL | 2 | | • | | • |
| tP(TGG)OR1 | 1 | | | | • |
| tP(TGG)OR2 | 2 | | | • | • |
| tP(TGG)FL | 2 | | • | | • |
| tP(TGG)OR2 | 2 | | | • | • |

| tDNA | SCORE | -53 | -42 | -30 | -13 |
|---|---|---|---|---|---|
| tP(TGG)FL | 2 | | • | | • |
| tS(AGA)GR | 2 | • | | • | |
| tS(AGA)HR | 1 | | | • | |
| tS(AGA)JR | 0 | | | | |
| tS(AGA)LR1 | - | | | | |
| tS(AGA)ML | 0 | | | | |
| tS(AGA)DL | 1 | | | • | |
| tS(AGA)AR | - | | | | |
| tS(AGA)DR3 | 1 | | | | • |
| tS(AGA)DR1 | - | | | | |
| tS(AGA)EL | - | | | | |
| tS(AGA)BR | 1 | | • | | |
| tS(CGA)CR | 3 | | • | • | • |
| tS(GCT)LR2 | - | | | | |
| tS(GCT)DR2 | - | | | | |
| tS(GCT)OL | 1 | • | | | |
| tS(GCT)FR | 1 | | | • | |
| tS(TGA)PR | 1 | | • | | |
| tS(TGA)IL | 2 | • | • | | |
| tS(TGA)ER | 1 | • | | | |
| tT(AGT)HR1 | 1 | | | • | |
| tT(AGT)JL | 2 | • | • | | |
| tT(AGT)CR | 2 | | • | | • |
| tT(AGT)NL2 | 3 | | • | • | • |
| tT(AGT)NL1 | 1 | | • | | |
| tT(AGT)OL | 1 | | | | • |
| tT(AGT)OR | - | | | | |
| tT(AGT)DL | 1 | | | | • |
| tT(AGT)IL2 | 2 | | • | | • |
| tT(AGT)IL1 | 1 | • | | | |
| tT(AGT)BR | 1 | | | | • |
| tT(CGT)KL | 0 | | | | |
| tT(TGT)GR1 | 1 | • | | | |
| tT(TGT)GR2 | 1 | | | | • |
| tT(TGT)HR2 | - | | | | |
| tT(TGT)PR | 0 | | | | |
| tW(CCA)GR | 0 | | | | |
| tW(CCA)GL | 0 | | | | |
| tW(CCA)JL | 3 | • | | • | • |
| tW(CCA)KL | 0 | | | | |
| tW(CCA)MR | 3 | • | • | | • |
| tW(CCA)PL | - | | | | |
| tY(GTA)JL | 0 | | | | |
| tY(GTA)JR | 2 | • | | | • |
| tY(GTA)ML | 1 | | | • | |
| tY(GTA)MR | 2 | | • | | • |
| tY(GTA)OL | 2 | | • | | • |
| tY(GTA)DR | - | | | | |
| tY(GTA)FR1 | 1 | | | | • |
| Y(GTA)FR2 | 3 | | • | • | • |
| tV(AAC)GR | 2 | | | • | • |
| tV(AAC)GL1 | 2 | • | | | • |
| tV(AAC)GL2 | 2 | | | • | • |
| tV(AAC)HL | - | | | | |
| tV(AAC)JL | 2 | | • | | • |
| tV(AAC)KL1 | 2 | | | • | • |
| tV(AAC)KL2 | 3 | | • | • | • |
| tV(AAC)LR | - | | | | |
| tV(AAC)MR3 | 3 | • | • | • | |
| tV(AAC)MR2 | 2 | • | | | • |
| tV(AAC)MR1 | 3 | | • | • | • |
| tV(ACC)OR | 2 | • | | | • |
| tV(AAC)ER1 | 3 | | • | • | • |
| tV(AAC)ER2 | 2 | | | • | • |
| tV(CAC)HR | 0 | | | | |
| tV(CAC)DR2 | 1 | | • | | |
| tV(TAC)BR | 0 | | | | |
| tV(TAC)DR1 | 3 | • | • | | • |

The 274 tDNAs are named according to MIPS. For each tDNA, the 5′-flank point-score value (ranging from 0 to 4) is reported (see the text for details). No value (–) is reported for the Ty-flanked tDNAs excluded from the analysis. Dots in the − 53, − 42, − 30 and − 13 columns indicate the conformity of tDNA 5′ flanks to the consensus profile with respect to the four districts of sequence conservation. The different iso-coding tDNA families are separated from each other. The tDNAs experimentally addressed in this work are shaded.

A

```
1. tRNA_P(AGG)CR                                                  +1
GATGGCAGGGAAAATAGGGCTGTATCCCAGACTTCAGGCGCTGTTCTGCCGGGATTTTTCAAAGGAGGGC

2. tRNA_P(AGG)NR
GTTGTGGTTGATGATCTATTGATGTTGAGGTACCATATTTTGAAATTCATTAAAACAACTTCTACCGGGC

3. tRNA_M(CAT)JL1
TAATAACCATGGAATCCGCAAATATTTGATGTAAAGTTTGTTTTTCTTCCCCAACATCCTTTGTATGCTT

4. tRNA_M(CAT)MR
TTATTCTCCATTGTTGCTTTTATTGAGTATTAAAGCATTTAATGCTAGAATCCTCCATAACAGACAGCTT

5. tRNA_N(GTT)CR
GGTAGTTGGAAAATTTTCAACATATATTAGTAATACTTTTTCTGTAAAAGTGACACAAGATAAAATGACT

6. tRNA_N(GTT)FL
GGTATGGGATGACATTGTTTTACGTAGTCTATTCCGTTTTAAGAGTTGCAGAACTCAAAAGTGACTGACT

7. tRNA_N(GTT)GR
GTACTCAAATAATTAAAATGTTGAGGACTAAGAGAGCCTATATTTTCTCGAGAACATAGTCAAAAAGACT

8. tRNA_N(GTT)KL
AACGTAAATGCTCCAGCTCAAGCATTATTTATATATCTTGTTTTCAAAAGAAAAATATAGCAATTAGACT

9. tRNA_N(GTT)LR
AGGTCAATGATAGAAGCTATCAATAATATTATATAATTTTTAACTCAGCAGATATCGGAAATAAAAGACT

10.tRNA_N(GTT)NR
TTACAGATGATTAAAATCGTGTAGATCATGGAAATCGAACCTTTCACGTTCTGCAGAATTTAAAGTGACT

11.tRNA_N(GTT)OL
CAATCTTGTGCATTGAGAACTGTGCATCATTAAGTTTCTAAAATTTAGATCTTTCATAAAATATTAGACT

12.tRNA_N(GTT)OR
GCCTCTCTTGACATATTTTTATCGTCATAATAAGTTCTTCATTCGTTCTCTAAAACAACAAAGTTAGACT

13.tRNA_N(GTT)PR
AATAAATACCGGAGATATGATTCAGATGTGGAAGCGGTTTTAAAGTCCCAATTCATCATTATTTGCGACT
```

B

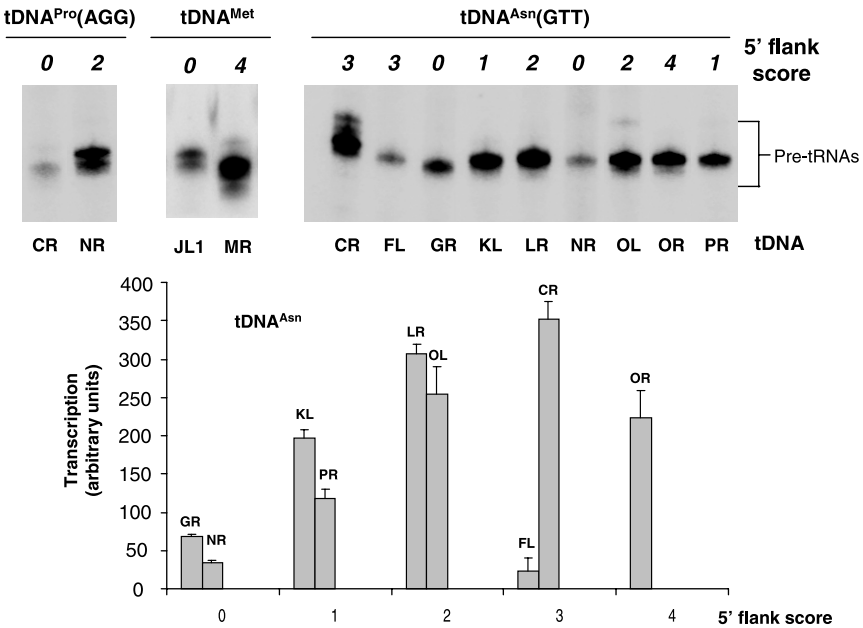



**Figure 2**. *In vitro* transcription of iso-coding tRNA genes with different upstream regions. A, Nucleotide sequences (from −66 to +4 with respect to the starting nucleotide of the mature tRNA) of three sets of isocoding tRNA genes that have been isolated and tested for their *in vitro* transcription properties. The first two tDNAs (tDNA 1 and 2) code for $tRNA^{Pro}$(AGG); tDNAs 3 and 4 code for $tRNA^{Met}$(CAT). Templates 5 to 13 represent the $tRNA^{Asn}$(GTT) gene family (only one member of this family, N(GTT)NL, was excluded from analysis due to the presence of a Ty element in its 5′ flank). In the listed 5′-flanking sequences, the six-letter words that have been recognized as conforming to the best frequency word within one of the four partially conserved districts are in bold and underlined. The presence of these words contributes to the final score of each upstream region, as explained in the text. B, The templates reported in A were used to program *in vitro* transcription reactions. The tRNA gene utilized for each reaction (identified by the last two letters of its systematic name) is indicated below each lane, while the corresponding 5′ flank score values are indicated above each lane. The migration position of the different pre-tRNA products is indicated by the bracket on the right. A bar plot derived from the quantification of $tDNA^{Asn}$(GTT) transcription is reported below the autoradiograms. The different $tDNA^{Asn}$ copies are grouped, on the *x-axis*, according to their 5′ flank score values; the average transcription levels measured in two independent experiments conducted and quantified in parallel are reported on the *y*-axis, with error bars indicating the standard error.

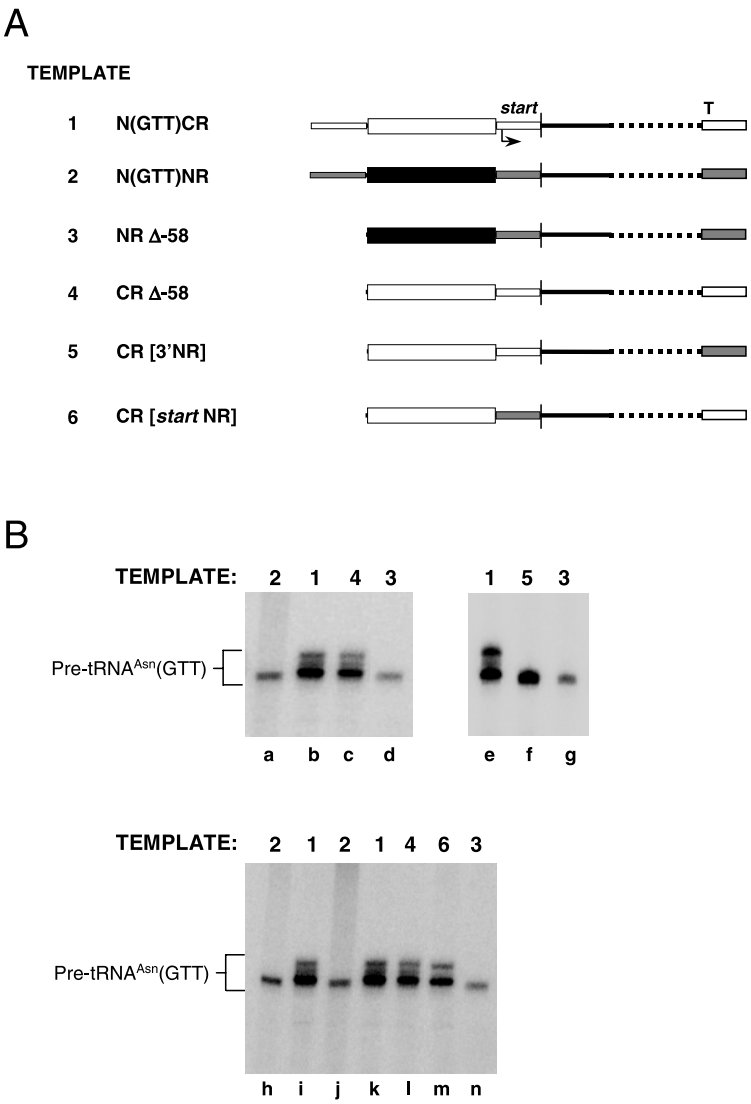



**Figure 3**. *In vitro* transcription of mutant tRNA$^{Asn}$(GTT) templates. A, Outline of the different mutant or hybrid variants of the tN(GTT)CR and tN(GTT)NR genes. The region from −53 to −14 is represented by a white or a black bar for the tN(GTT)CR or tN(GTT)NR genes, respectively. The position corresponding to the first nucleotide of the mature tRNA is marked by a vertical line. The bent arrow indicates the position of the transcription start site, while T indicates the position of the transcription termination site. The 3′-flanking sequences of tN(GTT)CR and tN(GTT)NR, as well as the sequences preceding the tRNA coding sequence up to −15, and thus including the transcription initiation site (start), are represented as white or gray thin horizontal bars, respectively. B, The templates listed in A were used to program *in vitro* transcription reactions. The template utilized for each reaction is indicated above each lane (same numbering as in A). The upper left, upper right and lower panels show the results of three independent *in vitro* transcription assays. The migration position of the pre-tRNA product is indicated on the left. Two main pre-tRNA species are produced from templates containing the bipartite tN(GTT)CR terminator and its associated 3′-flanking region.

transcribed fivefold more efficiently than the low-score copy (cf. the corresponding lanes in Figure 2B). To test for the occurrence of this 5′ flank effect on a more extended set of tDNAs, we isolated and transcriptionally tested *in vitro* the whole tRNA$^{Asn}$(GTT) gene family. This family comprises ten members, only one of which, tN(GTT)NL, was excluded from analysis because it is flanked by a Ty element. The other nine members of the tRNA$^{Asn}$(GTT) gene family (tDNAs 5–13 in Figure 2A) have 5′-flank point-score values ranging from 0 to 4. As shown in Figure 2B, these nine gene copies displayed substantially different *in vitro* transcription capacities. The most highly transcribed tRNA$^{Asn}$(GTT) gene copy was tN(GTT)CR, characterized by a 5′ flank point-score of 3. The tRNA$^{Asn}$(GTT) gene copies with a 2-point-score value (tN(GTT)LR and tN(GTT)OL), as well as the tN(GTT)OR gene copy with a 4-point score value, were all transcribed at slightly lower (less than twofold) levels than tN(GTT)CR. tN(GTT)PR and tN(GTT)KL, the two tRNA$^{Asn}$ gene copies with a 1-point-score value, were transcribed, respectively, about threefold and twofold less efficiently than tN(GTT)CR. Two of the three tRNA$^{Asn}$ gene copies with the lowest transcription efficiencies were those with a 5′ flank point-score of 0, tN(GTT)GR, whose transcription levels were fivefold lower than those of tN(GTT)CR, and tN(GTT)NR, that was transcribed tenfold less efficiently than tN(GTT)CR. The only, as yet unexplained, exception to this fairly close correlation trend between the 5′ flank point-score value and *in vitro* transcription capacity was the tN(GTT)FL gene, which has a 3-point score value, yet it was reproducibly transcribed ∼15-fold less efficiently than tN(GTT)CR. The first conclusion that can be drawn from these data is that, in all three tDNA gene families, the 5′ flank (perhaps in combination with different 3′ flanks, but see below) seems to influence in a largely predictable manner the *in vitro* transcriptional output of different members of such families, which share identical intragenic promoters, but have different upstream regions. More specifically, there seems to be a fairly close correlation between the *in vitro* transcription levels

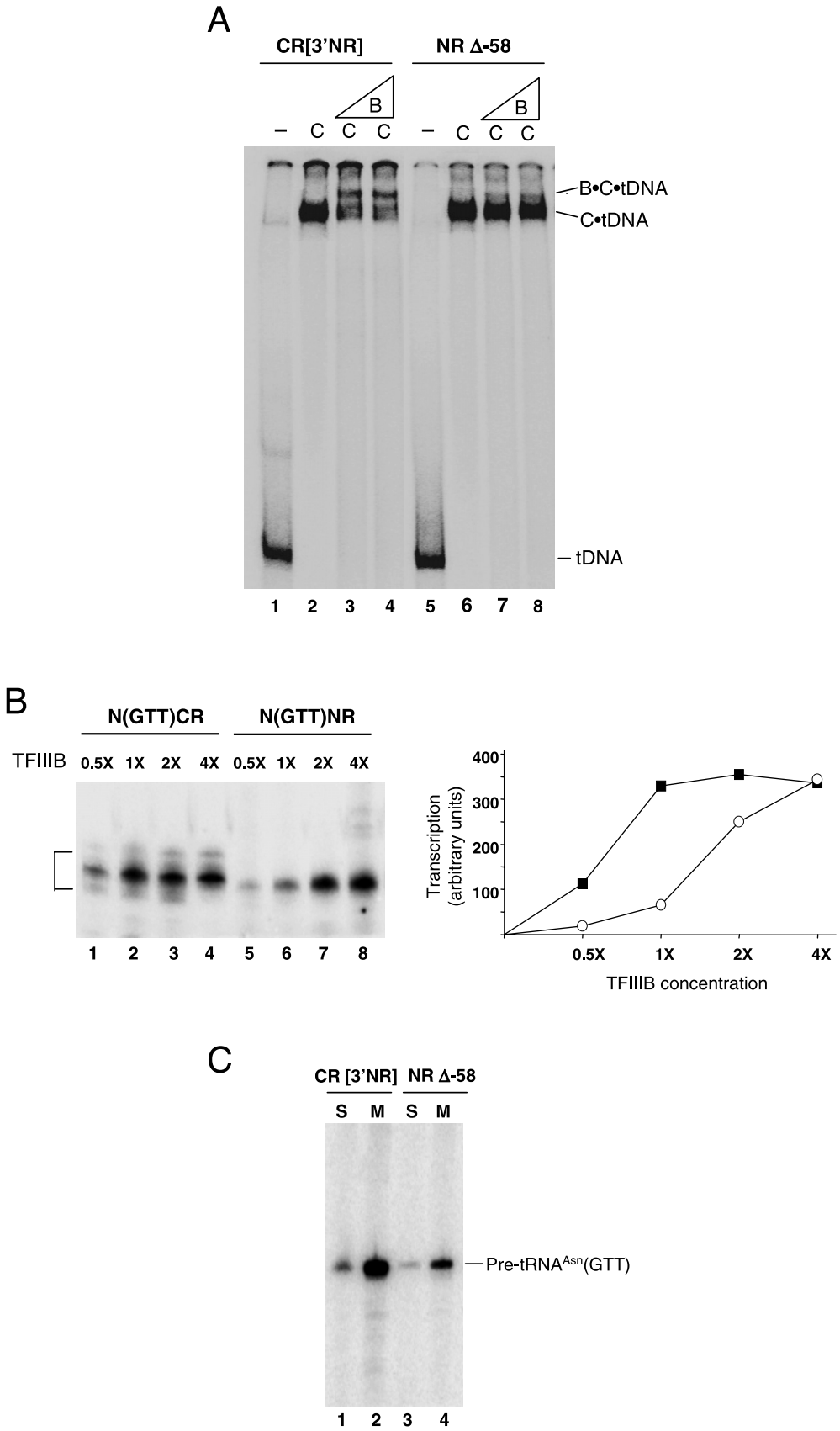



**Figure 4**. TFIIIB concentration responsiveness and transcription re-initiation properties of upstream region variants of tDNA$^{Asn}$(GTT). A, Radiolabeled DNA fragments containing either the CR[3′NR] (lanes 1–4) or the NR Δ-58 (lanes 5–8) tRNA$^{Asn}$(GTT) gene variants (only differing in the first 58 bp of their 5′-flanking regions) were incubated in the presence of either TFIIIC alone (lanes 2 and 6) or TFIIIC and increasing concentrations of reconstituted TFIIIB (lanes 3 and 7: 1 ng/μl of recombinant (r)TBP, 2 ng/μl of rBrf1 and 80 ng/μl of partially purified Bdp1 fraction; lanes

produced by individual iso-coding tDNAs and the degree of conformity of their 5′-flanking sequences to the composite upstream sequence consensus profile reported in Figure 1A. As suggested by the anomalous behaviour of the tN(GTT)FL gene, however, we cannot rule out the possibility that other transcriptionally important (yet largely idiosyncratic) features of the 5′-flanking region are somehow overlooked by our classification system.

### Boundaries of the stimulatory upstream region

Since the tDNA template constructs utilized for the previous experiments included about 80 bp of 5′-flanking and 50 bp of 3′-flanking sequences, we next tested the contribution to the observed transcriptional differences of the 5′-flanking region upstream of −53 (i.e. upstream of the region revealed as partially conserved by the sequence analysis reported in Figure 1A) and of the 3′-flank-flanking region, including the termination signal. This was investigated in detail for tN(GTT)CR and tN(GTT)NR, two iso-coding $tRNA^{Asn}$(GTT) genes that differ greatly in both 5′-flank score and transcription efficiency (see Figure 2). To this end, the mutant templates 3, 4 and 5 (outlined in Figure 3A) were constructed and transcriptionally assayed *in vitro*. In particular, templates 3 (NR Δ-58) and 4 (CR Δ-58) were derived from tN(GTT)NR and tN(GTT)CR by deletion of all the natural 5′-flanking sequence upstream of position −58. As shown in Figure 3B, whose three panels show the results of three independent transcription experiments, such a deletion only slightly reduced (1.5-fold) the *in vitro* transcription capacity both of tN(GTT)CR (cf. lanes b and c in the upper left panel, lanes i, k, and l in the lower panel) and of tN(GTT)NR (cf. lanes a and d in the upper left panel, lanes j and n in the lower panel). Importantly, however, the transcription efficiency difference between the two 5′-truncated templates remained exactly the same (i.e. sevenfold) as that of their 5′-extended counterparts (cf. lanes c and d with b and a in the upper left panel, and lanes l and n with h, i, j, and k in the lower panel). Therefore, the different transcription capacity of the high-score and low-score $tRNA^{Asn}$ gene copies does not depend on sequences upstream of −58. Similarly, the substitution of the 3′ flank of the highly transcribed tN(GTT)CR gene with the corresponding region of the poorly transcribed tN(GTT)NR (template 5 in Figure 3A) did not cause any decrease in tN(GTT)CR transcription, but only a change in the size distribution of full-length transcripts (cf. lanes e and f in Figure 3B, upper right panel), due to the introduction of the compact tN(GTT)NR terminator (a run of eight consecutive Ts) in place of the more complex and discontinuous tN(GTT)CR terminator region ($T_5CT_4CT_{13}$). The tN(GTT)CR variant carrying the 3′ flank of tN(GTT)NR was still transcribed seven times more efficiently than tN(GTT)NR (cf. lanes f and g in Figure 3B, upper right panel). The difference in transcription efficiency between tN(GTT)CR and tN(GTT)NR can thus be ascribed to differences in the 58 bp sequence immediately preceding the position corresponding to the 5′ terminus of mature tRNA. 5′-Flanking sequence swapping experiments conducted on the two members of the $tRNA^{Pro}$(AGG) gene family, tP(AGG)CR and tP(AGG)NR, similarly revealed that the transcriptional difference between these two genes is entirely due to the 5′ flank (data not shown). We also noticed the presence, at position −13 of tN(GTT)CR (see Figure 2A), of the sequence C<u>A</u>CAA (TSS underlined (data not shown)). This sequence is identical with the best frequency word at this position (CAACAA) except for a missing A residue. The corresponding sequence of tN(GTT)NR is CAGAA. Since the CAACAA element surrounding the TSS has been reported to positively influence transcription in other tRNA gene contexts,[16,18,19] the superior transcription capacity of tN(GTT)CR compared to tN(GTT)NR could, in principle, be due to differences in this region. To test this possibility, the

4 and 8: twofold higher concentrations of all three components). Protein–DNA complexes were resolved by native PAGE, and visualized by phosphorimaging. The migration positions of unbound DNA molecules (tDNA) and of tDNA-transcription factor complexes are indicated on the right. The total amount of complex in lanes 3 and 4 is 50% of the amount in lane 2; similarly, the amount of complex in lanes 7 and 8 is 70% with respect to that in lane 6. The reason for such a TFIIIB-dependent reduction in the total amount of complex has not been investigated in detail, but it is likely due to the presence of small, non-specific nucleic acid binding proteins (such as ribosomal proteins) in the Bdp1 fraction (G.D., unpublished results), that interfere with TFIIIC–DNA binding. B, *In vitro* transcription reactions programmed with either tN(GTT)CR (lanes 1–4) or tN(GTT)NR (lanes 5–8) were carried out in the presence of fixed concentrations of TFIIIC and RNA polymerase III, but increasing concentrations of TFIIIB, as indicated above each lane. The 1 × TFIIIB concentration corresponds to 0.8 ng/μl of rTBP, 1.6 ng/μl of rBrf1 and 10 ng/μl of a partially purified Bdp1 fraction. The migration position of the pre-$tRNA^{Asn}$ transcripts is indicated by a bracket on the left. The diagram on the right derives from phosphorimaging quantification of the gel shown on the left. Filled squares and open circles refer, respectively, to tN(GTT)CR and tN(GTT)NR transcription. C, Stalled ternary complexes, containing either a 6 nt or a 14 nt nascent RNA, were formed on either the CR[3′NR] (lanes 1, 2) or the NR Δ-58 (lanes 3, 4) template upon incubation in the presence of ATP, CTP, and GTP (CR[3′NR]) or GTP, ATP and UTP (NR Δ-58). Transcription was then resumed by the addition of the missing NTP alone (lanes 2 and 4) or together with 100 μg/ml of heparin (lanes 1 and 3), and allowed to proceed for seven minutes. The migration position of the pre-$tRNA^{Asn}$ transcript is indicated on the right.

CR[*start*NR] mutant was constructed, in which the tN(GTT)CR sequence extending from −15 to −1 is replaced by the corresponding sequence of tN(GTT)NR. As shown in Figure 3B (lower panel) the swapping of this region also did not cause any significant decrease in tN(GTT)CR transcription (cf. lanes l and m), thus indicating that the upstream region comprised between −58 and −15 is by itself sufficient to determine the higher transcription efficiency of tN(GTT)CR with respect to tN(GTT)NR (cf. lanes m and n).

### Mechanism of transcriptional stimulation by a high-score tDNA upstream region

Gel retardation experiments revealed that, as expected, TFIIIC binds with same affinity to the identical intragenic promoters of tN(GTT)CR and tN(GTT)NR (data not shown). However, when TFIIIC–tDNA complexes assembled on either gene (constructs CR[3′NR] and NR Δ-58, only differing for the first 58 bp of 5′-flanking region) were incubated with increasing amounts of reconstituted TFIIIB (Figure 4A), the CR[3′NR]-TFIIIC complexes (lanes 2–4) could be super-shifted in the presence of lower TFIIIB concentrations compared to NR Δ-58-TFIIIC complexes (lanes 6–8), whose super-shift products were barely visible even at the highest TFIIIB concentration. To further investigate whether the tN(GTT)CR and tN(GTT)NR upstream regions differ in their ability to assemble TFIIIB, we measured the *in vitro* transcription levels of tN(GTT)NR and tN(GTT)CR in the presence of varying concentrations of TFIIIB (Figure 4B). Transcription of the two tDNAs was carried out in the presence of either the same TFIIIB concentration utilized in previous experiments (lanes 2 and 6) or a twofold lower (lanes 1 and 5), a twofold higher (lanes 3 and 7), or a fourfold higher (lanes 4 and 8) TFIIIB concentration. As shown in Figure 4B, the tN(GTT)NR gene, whose transcription output was five- to sixfold lower than that of the tN(GTT)CR gene at the two lowest TFIIIB concentrations, was transcribed as efficiently as the high-score $tDNA^{Asn}$ at the highest TFIIIB concentration. Nearly identical results were obtained in two independent TFIIIB titration experiments (not shown). The above data strongly suggest that, *in vitro*, the high-score 5′-flanking sequence of tN(GTT)CR exerts its positive transcriptional effect by improving TFIIIB assembly. In experiments in which the concentrations of the three TFIIIB components were individually varied, the transcriptional difference between the two templates was found to be more sensitive to changes in Brf1, rather than TBP or Bdp1, concentration (data not shown).

Re-initiation is a crucial step in RNA polymerase III-directed transcription,[27] and TFIIIB stabilization during multiple rounds of transcription has been shown to be an important mechanism through which the 5′ flank of plant tRNA genes exerts its stimulatory effect.[28] We thus asked whether the same is true for the 5′ flanks of tN(GTT)CR and tN(GTT)NR. To address this point, we compared the transcription levels of NR Δ-58 and CR [3′NR] templates, only differing in their upstream regions, under conditions allowing for only one round of transcription.[1] In the experiment reported in Figure 4C, stable tDNA–TFIIIC–TFIIIB pre-initiation complexes were first assembled on the two genes. Stalled elongation complexes containing either a 6 nt (for CR [3′NR]) or a 14 nt (for NR Δ-58) nascent RNA were allowed to form through the addition of Pol III together with either ATP, CTP and GTP for CR [3′NR], or ATP, GTP and UTP for NR Δ-58. Such complexes were both resistant to 100 μg/ml of heparin (data not shown). The missing NTP was then added, alone (lanes 2 and 4) or with 100 μg/ml of heparin (lanes 1 and 3), and transcription was allowed to proceed for seven minutes. A comparison of lane 1 with 3 and of lane 2 with 4 shows that, when transcription is limited to a single round, the difference in transcriptional output from the two templates is very similar to that observed under multiple round transcription conditions (an approximately sixfold higher activity of CR Δ-58[3′NR] in both cases). Therefore, in keeping with the results of the gel retardation and TFIIIB titration experiments, a lower number of productive transcription complexes is formed on the tN(GTT)NR gene, but once formed these complexes support transcription re-initiation as efficiently as those assembled on the tN(GTT)CR gene.

### Influence of the $tDNA^{Asn}$(GTT) upstream regions on tRNA gene expression *in vivo*

Monitoring the *in vivo* expression of individual members of tDNA multigene families is complicated by the extremely small sequence differences that most often distinguish the primary transcripts produced by different members of such families. To monitor the *in vivo* transcriptional effect of the upstream region, we thus combined the different 5′-flanking regions with a synthetic reporter tRNA gene, as originally devised by Feldmann and co-workers.[29] The reporter gene, $tDNA^{Syn2}$, contains the *S. cerevisiae* $tRNA^{Glu}$(TTC) coding sequence, tagged by the insertion of an intron-like sequence that cannot be spliced out from the tRNA precursor. Pseudo-intron tagging allows the identification and quantification of transcripts synthesized from an upstream region-modified $tDNA^{Syn2}$ by Northern hybridization experiments, using the endogenous, intron-less $tRNA^{Glu}$(TTC) as an internal standard.[29] Two gene fusions were constructed, in which the 5′-flanking region of either tN(GTT)CR or tN(GTT)NR was fused to the Syn2 tDNA. The constructs were inserted both into the YEp352 multicopy vector and into the low-copy number, centromeric pFL38 plasmid, in order to detect possible gene dosage effects on tDNA transcription. The data reported in Figure

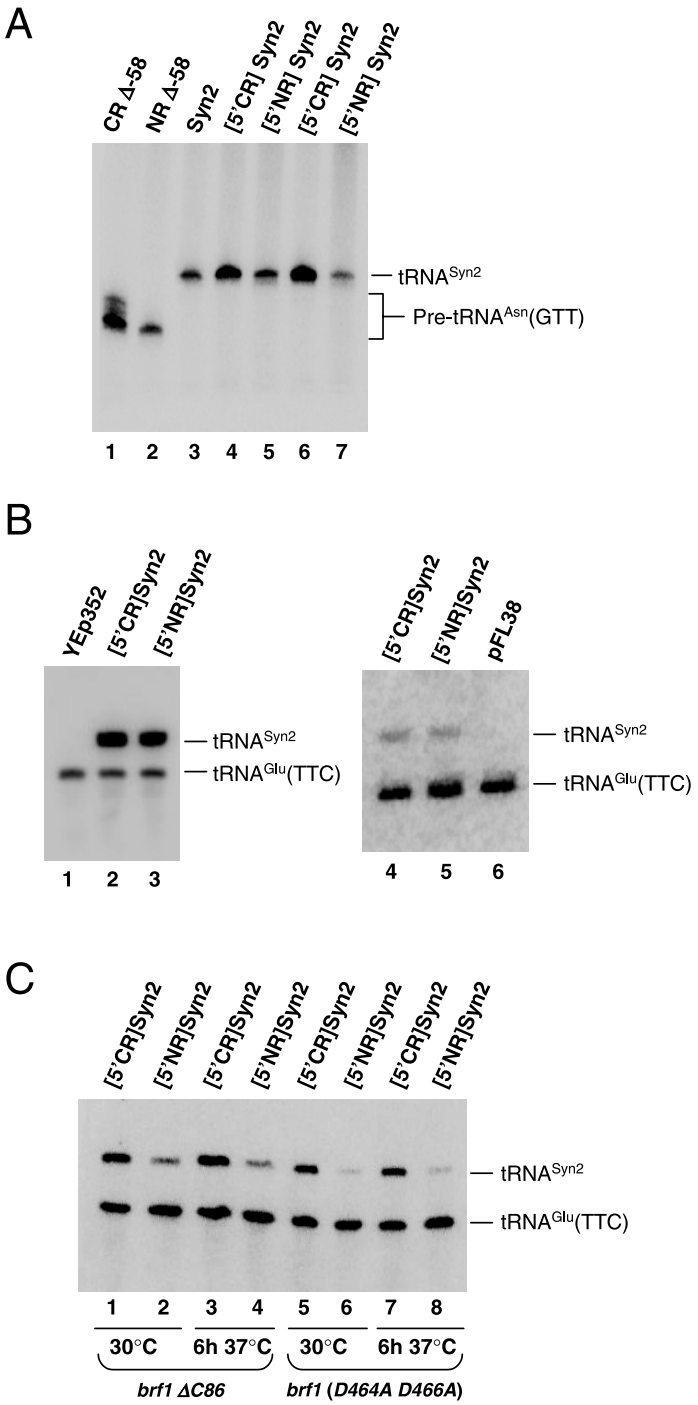



**Figure 5**. *In vivo* transcriptional analysis of the tDNA$^{Asn}$(GTT) upstream regions. A, *In vitro* transcription properties of tDNA$^{Syn2}$ fused to different upstream sequences. The templates indicated above each lane were used to program *in vitro* transcription reactions. The features of the CR Δ-58 and NR Δ-58 templates are specified in Figure 3A. The [5′CR]Syn2 and [5′NR]Syn2 templates consist of the tRNA$^{Syn2}$ coding sequence flanked by 58 base-pairs of 5′-flanking sequence derived from the tN(GTT)CR or the tN(GTT)NR gene, respectively. The Syn2 template consists of the tRNA$^{Syn2}$ coding sequence flanked by vector sequences provided by the YEp352 polylinker. The templates utilized for the reactions shown in lanes 4 and 5 were inserted into the YEp352 plasmid; templates for the reactions in lanes 6 and 7 were in the pFL38 plasmid. The migration positions of pre-tRNA$^{Asn}$ and tRNA$^{Syn2}$ products are indicated on the right. B, Expression of tDNA$^{Syn2}$ constructs in wild-type yeast strains. Total RNA was extracted from yeast cells (YPH500 strain for lanes 1–3, W303 strain for lanes 4–6) transformed with YEp352 (lane 1), YEp352 carrying [5′CR]Syn2 (lane 2), YEp352 carrying [5′NR]Syn2 (lane 3), pFL38 carrying [5′CR]Syn2 (lane 4), pFL38 carrying [5′NR]Syn2 (lane 5), or the empty pFL38 vector (lane 6). Fractionated RNA samples (5 μg each) were probed with a radiolabeled tRNA$^{Syn2}$ antisense oligonucletide. The migration positions of the genome-encoded, mature tRNA$^{Glu}$(TTC) and of plasmid-encoded tRNA$^{Syn2}$ transcripts are indicated on the right. C, Expression of tDNA$^{Syn2}$ constructs in *brf1* mutant yeast strains. Total RNA was extracted prior to (lanes 1, 2, 5, 6) or after (lanes 3, 4, 7, 8) a temperature shift (six hours at 37 °C) from mutant *S. cerevisiae* strains harboring either the *brf1* ΔC86 deletion[31] (lanes 1–4) or the *brf1* (D464A, D466A) double substitution mutation[30] (lanes 5–8). Both strains were transformed with pFL38 constructs carrying either the [5′CR]Syn2 template (lanes 1, 3, 5, 7) or the [5′NR]Syn2 template (lanes 2, 4, 6, 8). Gel-fractionated RNA samples were probed as described for A. The migration positions of the genome-encoded tRNA$^{Glu}$(TTC) and of plasmid-derived tRNA$^{Syn2}$ transcripts are indicated on the right.

5A show that the tN(GTT)CR and tN(GTT)NR 5′ flanks exert their distinctive transcriptional effects *in vitro* even when placed upstream of tDNA$^{Syn2}$. The [5′CR]Syn2 template was reproducibly transcribed about fivefold more efficiently than [5′NR]Syn2 (cf. lane 4 with 5 and lane 6 with 7). Also worth noting is the fact that the [5′NR]Syn2 template was transcribed as efficiently as a control tRNA$^{Syn2}$ gene bearing the sequence next to the YEp352 cloning site as its 5′ flank (cf. lanes 5 and 3). This observation argues strongly in favor of a stimulatory role of the high-score upstream region of tN(GTT)CR, rather than an inhibitory effect of the corresponding low-score region of tN(GTT)NR.

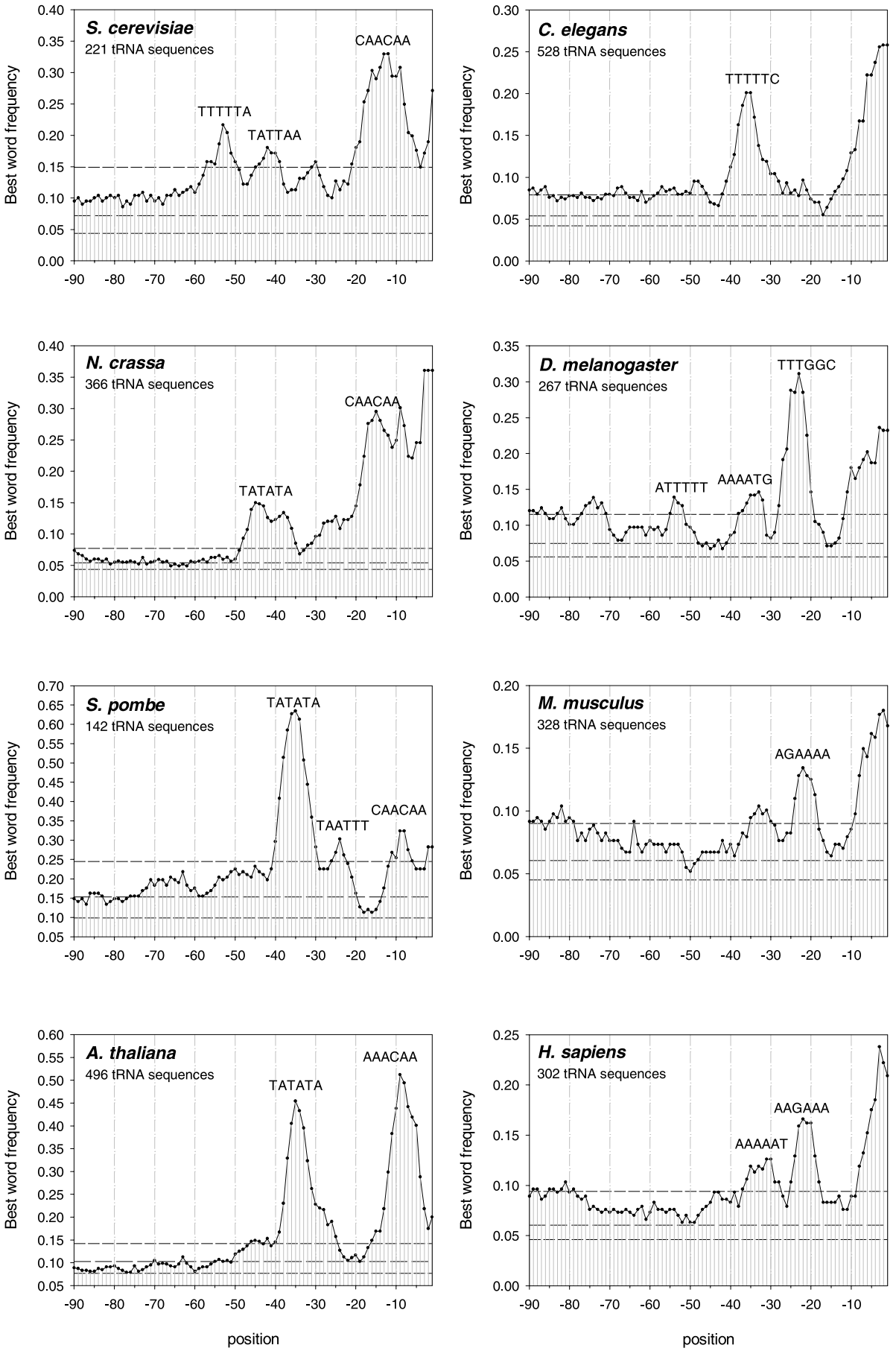



**Figure 6** (*legend opposite*)

When the tDNA$^{Syn2}$ fusions ([5′NR]Syn2 and [5′CR]Syn2, in either YEp352 (Figure 5(B), lanes 1–3) or pFL38 (lanes 4–6), were transformed into two common yeast laboratory strains (YPH500 for lanes 1–3; W303 for lanes 4–6), no difference in tRNA$^{Syn2}$ expression was observed when either type of 5′ flank was used as a tDNA$^{Syn2}$ fusion partner (cf. lane 3 with 2 and lane 5 with 4 in Figure 5(B)). The steady-state levels of the tRNA$^{Syn2}$ product synthesized from the pFL38-tRNA$^{Syn2}$ constructs were about 60-fold lower than those obtained with the YEp352-tRNA$^{Syn2}$ constructs, in keeping with an expected, correspondingly reduced *in vivo* level of the templates carried by the centromeric pFL38 vector. Despite this reduced template concentration, no difference in the *in vivo* transcriptional activities of the two genes could be detected. Since the difference in the *in vitro* transcription capacity of tN(GTT)CR and tN(GTT)NR depends on the concentration of TFIIIB (and especially of its Brf1 component), we reasoned that it might be possible to unmask a corresponding 5′ flank effect *in vivo* by artificially lowering the cellular levels of active TFIIIB. To this end, we measured the transcriptional output of the [5′CR]Syn2 and [5′NR]Syn2 constructs in two *S. cerevisiae* strains bearing conditional thermosensitive mutations in the *BRF1* gene that are known to affect TFIIIB–DNA complex formation. These two mutant strains harbored either the *brf1-II.6* mutant allele, originating from a double D464A/D466A substitution in the TBP-interacting homology domain II of Brf1,[22,30] or the *brf1-ΔC86* mutation, producing a truncated form of Brf1 lacking 86 amino acid residues at the C terminus.[31] As shown in Figure 5C, in both mutant strains, transcription of the [5′NR]Syn2 gene was considerably decreased (at least sevenfold) compared to that of [5′CR]Syn2, both at the permissive and non-permissive temperature. The effect of the 5′ flank was more pronounced in the D464A/D466A mutant: in this strain, [5′NR]Syn2 was transcribed 12–14 times less efficiently than [5′CR]Syn2 (cf. lane 6 with 5 and lane 8 with 7). The tN(GTT)CR upstream region thus seems to enhance TFIIIB binding and transcription not only *in vitro*, but also in yeast cells suffering from a deficit of the key TFIIIB component, Brf1.

## tDNA upstream regions in other eukaryotic genomes

Transcriptional modulation of particular subsets of tDNAs by their 5′-flanking regions has been documented previously in various eukaryotic systems besides yeast (see Introduction). We thus asked whether the presence of a conserved sequence pattern is a general feature of eukaryotic tRNA genes. The answer to this question is outlined in Figure 6, which summarizes the results of tDNA upstream sequence analyses conducted on eight different, wholly sequenced eukaryotic genomes. As expected on the basis of previous studies,[7,8,10] the upstream sequence profile in *S. pombe* and *A. thaliana* is dominated by a TATA box located at ~30 bp upstream of the TSS (*x*-axis positions in Figure 6 refer to the first nucleotide of the mature tRNA) and, especially in the case of plant tDNAs, by an A/CAACAA TSS motif. At variance with *S. cerevisiae*, in most organisms the tRNA gene TSS generally precedes by 7–8 bp the first nucleotide of the mature tRNA coding sequence. Accordingly, in both *S. pombe* and plant tDNAs, the TSS conserved motifs are centered around position −8, rather than position −12/−13 as in *S. cerevisiae*. The upstream sequence pattern of *S. pombe* and *A. thaliana* tDNAs resembles that of *S. cerevisiae*, and it also occurs in another fungus, the multicellular ascomycete *Neurospora crassa*. A very similar pattern appears to be maintained in the three *S. cerevisiae*-related yeast species (*S. paradoxus*, *S. mikatae* and *S. bayanus*) whose draft genome sequences have recently been published (data not shown).[32] Interestingly, such a fungal/plant profile disappears in metazoans such as *Caenorhabditis elegans*, *Drosophila melanogaster* and *Homo sapiens*. No significant TATA box conservation, nor any conserved TSS signature, was found upstream of tDNAs in the genomes of these organisms. Instead, upstream motifs mainly consisting of A or T-tracts were found at conserved positions corresponding to that of the TATA-like element in plants and fungi (TTTTTC in *C. elegans*; AAAATG in *D. melanogaster*; AAAAAT in *H. sapiens*). In addition, completely new motifs were identified near position −20 in *D. melanogaster* (a highly significant TTTGGC motif at position −23, possibly related to a previously identified upstream regulatory sequence[33]), *H. sapiens* (AAGAAA at −22) and *Mus musculus* (AGAAAA at −22). The close similarity of the latter motif, as well as of the general upstream sequence profile, in mouse and human tDNAs clearly points to the biological significance of upstream regulatory sequences in these organisms.[34] Altogether, these data suggest that TFIIIB assembly upstream of tRNA genes, although characterized by a generally low level of

**Figure 6**. Conserved sequence patterns in the tDNA upstream regions of different eukaryotes. Conserved sequence profiles were obtained and are represented as described for Figure 1. The organism and the total number of analyzed tDNA upstream sequences are indicated within each panel. The middle broken line indicates the average frequency value obtained from the analysis of multiple sets of randomly shuffled sequences. Maximum and minimum frequency values produced by such analysis are indicated by the upper and lower broken lines, respectively. The six-letter words most frequently occurring at significant peak positions are reported above the peaks.

sequence specificity (and thus, presumably, by a strong TFIIIC-dependence), is very likely influenced by partially conserved upstream elements in all eukaryotes.

## Discussion

We have addressed, on a genome-wide scale, the long-standing question of the occurrence, in the 5′-flanking regions of eukaryotic tRNA genes, of conserved sequence features that can influence their transcriptional properties. Such a matter is controversial, essentially because it relies on two apparently contrasting sets of observations. On one hand, the TFIIIB–tDNA interaction, which involves the 50 bp of DNA upstream of the TSS, has been shown in *S. cerevisiae* to be largely sequence-independent.[2] On the other hand, there is a large body of evidence indicating that in yeast, as well as in other eukaryotes, the 5′-flanking region of particular subsets of tRNA genes contains conserved sequence elements that contribute to promoter function.[8,10,11,16] The present work originated from the hypothesis that this apparent disagreement may be due to the degenerate nature of 5′-flanking elements, that makes them not easily detectable by standard sequence alignments, to their heterogeneous distribution within the different members of the tRNA gene complement, and to their variable transcriptional requirement, which may be strongly influenced, among other things, by the particular reaction conditions employed for transcription assays.[35]

Through a bioinformatic procedure specifically developed for the search of shared, yet degenerate elements in sets of aligned DNA sequences, we have unmasked a composite pattern made up by partially conserved motifs upstream of yeast tRNA genes. The results of *in vitro* and *in vivo* transcriptional analyses support the notion that such motifs represent a previously unidentified TFIIIB interaction signature that somehow optimizes TFIIIB binding and thus enhances tRNA gene transcription, especially under TFIIIB-limiting conditions.

### Sequence features and functional implications of tDNA upstream sequence conservation in *S. cerevisiae*

The upstream sequence pattern of *S. cerevisiae* tDNAs displays several interesting features, the most prominent of which is its composite nature. Indeed, this pattern comprises four regions in which particular sequences are over-represented over the random sequence background, thus appearing as conserved upstream motifs. The most downstream motif corresponds to a fairly conserved region previously identified near to the TSS of certain yeast tRNA genes.[16,36] Within this region, the sequence CAACAA is frequently found starting from the −14/−13 position. The first A of the CAACAA sequence is used as the TSS, and in the $tRNA^{Leu}$(CAA) gene family this motif significantly contributes to promoter strength.[16] It is also interesting to note that this motif is present in all members of this $tDNA^{Leu}$ family, as well as in most members of other three families coding for abundant tRNAs ($tRNA^{Arg}$(TCT), $tRNA^{Lys}$(CTT) and $tRNA^{Val}$(AAC))[25] (Table 1), but in only two out of the nine classified members of the $tDNA^{Asn}$(GTT) family mainly addressed in the present study. For the latter family, the −13 region does not seem to play a functionally dominant role in transcription. In fact, the region from −58 to −15, comprising only the three most upstream conserved elements, is sufficient to confer to tN(GTT)CR (the most highly transcribed member of the $tDNA^{Asn}$(GTT) family) an *in vitro* transcription efficiency that is seven to tenfold higher than that of tN(GTT)NR, whose 5′ flank (0-point-score) does not at all conform to the consensus pattern. It is thus possible that the TSS region (centered at −13) and the more upstream located conserved region (approximately extending from −30 to −53) differently affect transcription within different tDNA families.

The conserved region upstream of −13 is characterized by three positions, located at −30, −42 and −53 with respect to the first nucleotide of the mature tRNA, around which particular sequences are over-represented. Considering that −12 is the position where tRNA gene transcription most frequently starts in *S. cerevisiae*, this sequence conservation tends to be located at 18 bp, 30 bp and 41 bp upstream of the TSS, i.e. within the region that is known to be contacted by TFIIIB.[1,37] Indeed, functional analysis of the two above-mentioned $tRNA^{Asn}$(GTT) gene family members (both classified as negative with respect to the −13, TSS element) has shown that the presence of a high-score 5′-flanking region, conforming well to the consensus pattern upstream of the TSS region, potentiates TFIIIB assembly. This conclusion is supported by three lines of evidence. First, in gel retardation assays, a much more efficient TFIIIC-dependent recruitment of TFIIIB was observed with the high-score, tN(GTT)CR gene with respect to the low-score tN(GTT)NR iso-coding gene. Second, the transcription defect of tN(GTT)NR was abolished by increasing the concentration of TFIIIB added to *in vitro* transcription reaction mixtures. Third, the low-score 5′ flank of tN(GTT)NR strongly impaired the expression of a reporter tRNA gene *in vivo*, but only in the presence of reduced cellular levels of active TFIIIB.

The fact that the centrally located −42 motif is a TATA-like sequence is by itself quite remarkable, because TBP is thought to occupy a central position in the TFIIIB–tDNA complex, with the other two components (Brf1 and Bdp1) contacting the DNA on both sides of the TBP–DNA sub-complex. In the highly bent TFIIIB–tDNA complex,[20] TBP is thought to cause the main DNA distortion by inducing a sharp kink at the T–A base-pair steps

of the TATA element.[38] Thus, the regular spacing (~11 bp) between the −53, the −42 and the −30 districts of upstream sequence conservation might reflect an internal symmetry in the TFIIIB–DNA complex, with important protein–DNA contacts surrounding the TBP-mediated DNA bend and involving the same face of the DNA helix. Such contacts likely influence the half-life of the TFIIIB–DNA complex[39] rather than its rate of formation, because the *in vitro* transcriptional defect of low-score $tRNA^{Asn}$(GTT) and $tRNA^{Pro}$(AGG) genes was not rescued even upon prolonged pre-incubation with TFIIIC and TFIIIB (data not shown).

It is also worth noting that in the *SNR6* gene, as in some natural yeast tDNAs, whose transcription *in vitro* can take place in the absence of TFIIIC due to TATA-box-dependent TFIIIB assembly, a canonical TATA box is present at −30 with respect to the TSS,[8] exactly the same location of the conserved TATA-like element identified in this study. The other two-tDNA upstream motifs are characterized by T-richness and by a roughly symmetrical location with respect to the TATA-like element. These T-rich regions might be involved in specific contacts with Brf1 and Bdp1 that stabilize the intrinsically weak TBP–TATA interaction. Interestingly, the T-rich region we have identified downstream of the TATA-like element correlates both in sequence and position with a T-rich stretch upstream of the yeast *SNR6* gene, that becomes essential, as a promoter element, only when the chromatin structure of the gene is perturbed, either by a modification of the A to B-block spacing or by deletion of the gene coding for the chromatin architectural protein Nhp6.[40] In the case of the *SNR6* gene, this T-rich element was shown not to influence the autonomous (TFIIIC-independent) assembly of TFIIIB onto the upstream region, which thus seems to depend exclusively on the *SNR6* TATA box, at least *in vitro*. On the basis of our present results, it can only be concluded that the upstream region encompassing the TATA-like and the two T-rich elements influences TFIIIB assembly. One important difference between our TFIIIB binding and transcriptional studies and those of Martin *et al.*[40] is that our analysis was conducted in the context of a TFIIIC-dependent transcription mode, and with tRNA genes on which no TFIIIC-independent transcription, nor autonomous TFIIIB assembly takes place (data not shown). Thus, it is possible that the *SNR6* T-rich upstream element is without any effect in the context of an autonomous TFIIIB assembly pathway dominated by a strong TATA-box,[40] while it might affect TFIIIB assembly in the TFIIIC-directed transcription of tN(GTT)CR and other upstream-region positive class III genes.

The individual contribution of the three most upstream sequence conservation districts (−53, −42 and −30) to TFIIIB assembly and transcription was difficult to assess. Swapping mutagenesis revealed that the three elements upstream of tN(GTT)CR could be individually replaced by the corresponding regions of tN(GTT)NR without any significant reduction in transcription efficiency (data not shown). Thus, the conserved upstream sequence pattern identified by the present analysis should not be viewed as a classical multipartite control region made up by functionally independent and easily transplantable boxes, but rather as a fairly large and functionally heterogeneous binding surface providing docking sites for at least three distinct (yet interacting) proteins. The latter situation, which is reminiscent of complex Pol II nucleoprotein assemblies where highly flexible protein–protein interactions and non-specific DNA binding can mask the effect of localized sequence perturbations,[41] is clearly more difficult to address by simple dissection and DNA swapping experiments. We also have to note that the transcriptional enhancement brought about by different tDNA 5′ flanks is not fully predictable simply on the basis of their conformity to the upstream consensus pattern. For example, the N(GTT)FL tDNA was assigned a 5′ flank score of 3, yet it was the less efficiently transcribed member of its family†. On the same note, the different UGA suppression abilities of the six yeast $tRNA^{Trp}$ genes[42] cannot be explained simply on the basis of their 5′ flank score values (Table 1). Clearly, additional and perhaps more subtle (or hierarchically organized) upstream sequence features, not taken into account by our classification (e.g. intrinsic bending properties; see below), can influence tRNA gene transcription.

## Intrinsic bending properties of tDNA upstream regions

TFIIIB sharply bends DNA[20] and a previous study has shown that TFIIIB binding to a tDNA upstream region can be potentiated by artificially increasing DNA flexibility at particular positions.[43] The conserved sequence pattern upstream of tDNAs may thus also reflect the presence of peculiar intrinsic structural features in this region. To examine this possibility, our sample of tDNA upstream sequences was analyzed using a model of sequence-dependent DNA bending propensity of trinucleotides, based on the probability of DNase I cutting.[44] As shown in Figure 7, this analysis revealed a region of extremely high bending propensity centered around position −40 (28 bp upstream of the TSS), that quite precisely overlaps the −42, TATA-like upstream element (see Figure 1). Such a region is preceded, around position −53, by a region of low intrinsic bendability. This alternation of nearby DNA segments with opposite bending propensities suggests the

† The only anomaly we notice in the 5′ flank of the N(GTT)FL tDNA is the presence of a −42 motif (of sequence T<u>C</u>TATT) in which there is a single nucleotide mismatch that is expected to inactivate the TATA function.

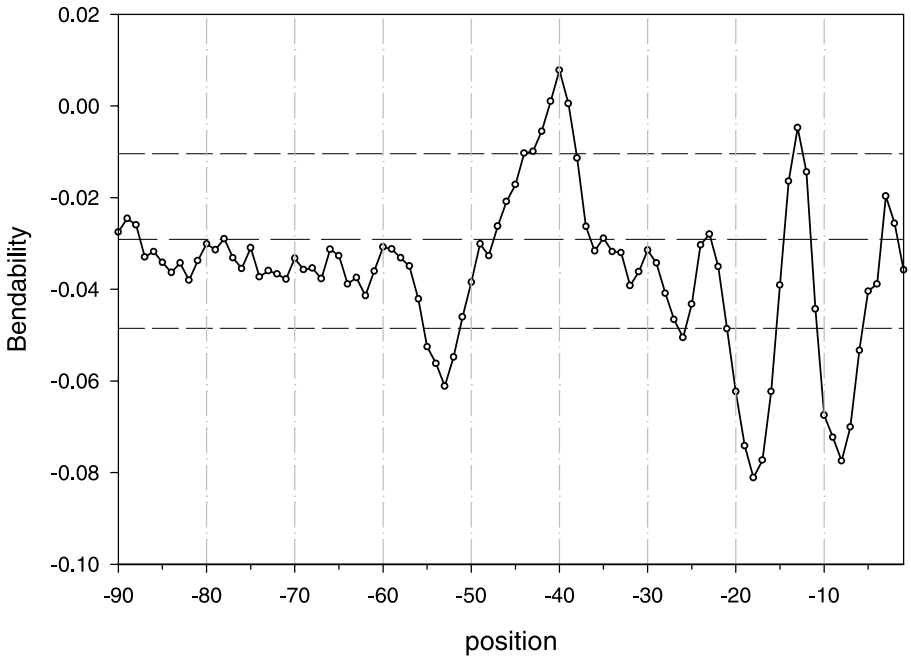



**Figure 7**. Bendability profiles of yeast tDNA upstream regions. The 0 position on the *x*-axis corresponds to the first nucleotide of the mature tRNA. Increasing values on the *y*-axis correspond to higher bendability. The average bendability for each position was calculated using DNase I-derived bendability parameters for trinucleotides.[44] The resulting profile was then smoothed by calculating a running average with a window size of 6. Due to this smoothing procedure, the bendability value reported for each position represents the average of the values for the first four triplets encountered when moving downstream. The middle broken line indicates the average frequency value obtained from the analysis of randomized datasets. The upper and lower broken lines correspond, respectively, to the maximum and minimum frequency values in the same sets of randomly shuffled sequences.

existence, in this region, of a conserved DNA deformability pattern, such that DNA bending by TBP might be favored by contacts between other TFIIIB components and a rigid upstream segment of DNA. A novel interesting feature revealed by the bendability analysis in Figure 7 is the marked propensity towards DNA bending of the region surrounding the tDNA TSS (and thus overlapping with the −13 motif in Figure 1A), that is flanked on both sides by low bendability regions. Such a feature suggests that the −13 region might promote transcription not only by specifying purine-rich nascent transcripts that stabilize the early ternary complex (as shown for the CAACAA element of the *SUP4* tRNA gene[19]) but also by providing a local region of high DNA deformability that facilitates Pol III binding and transcription initiation. It is tempting to speculate that DNA bending at the TSS might facilitate wrapping of the DNA around the polymerase, a mode of enzyme–DNA interaction that has been shown to occur in bacterial RNA polymerase open complexes[45] as well as in elongation complexes of bacterial RNA polymerase, yeast Pol III[46] and phage T7 RNA polymerase.[47]

## Possible biological significance of tDNA upstream sequence conservation

As already pointed out, the TFIIIB-binding signature identified in this work cannot be considered as a classical consensus sequence. Of the sequences that are over-represented within the districts of partial sequence conservation, only the −13, TSS motif is present in the majority (54%) of the analyzed tDNA 5′ flanks. Moreover, the upstream sequence scores attributed to individual tDNAs vary significantly depending on the stringency of the classification criteria. In spite of this limitation, we can conclude that individual tRNA gene families are highly heterogeneous with respect to upstream regions, comprising members whose 5′ flanks differ strongly in their conformity to the conserved sequence pattern. As a consequence, most iso-coding tRNA gene families include members that are expected to bind TFIIIB with quite different strengths. Our data suggest that such differences in TFIIIB binding will affect tDNA transcription most prominently under conditions of sub-optimal TFIIIB concentration or activity. This is particularly appealing, because TFIIIB down-regulation is the central strategy utilized in various systems to reduce class III gene transcription in response to external stimuli,[12,48] and the levels of active TFIIIB are known to fluctuate as a function of the nutritional status, growth phase and overall protein biosynthetic activity.[49–51] We anticipate that, within individual iso-coding tRNA gene families, different tDNAs will respond differently to a reduction in TFIIIB concentration and/or activity. The presence, within the same tRNA gene family, of gene copies with different sensitivities

to the concentration of a key transcription factor could thus serve a sort of homeostatic role aimed to guarantee the production of basal levels of tRNAs even under conditions of severely impaired metabolic activity and/or down-regulated transcription. On a different note, variations in TFIIIB occupancy and transcription rate within tRNA gene families might also be related to chromatin perturbation events that may change the accessibility or affinity of TFIIIB for its binding site as well as to extra-coding functions of tRNA genes, such as the recently demonstrated heterochromatin barrier activity of yeast tDNAs.[52]

### Diversity of tDNA upstream sequence patterns in eukaryotes

The statistical analysis of tDNA upstream regions in a number wholly sequenced eukaryotic genomes, besides *S. cerevisiae*, revealed the widespread occurrence of conserved sequence patterns in the 90 bp region immediately upstream of the tRNA coding sequence. Such patterns, however, display a remarkable diversity in terms of the number, sequence, and position of the over-represented six-letter words. The putative TFIIIB interaction signature found in plants and fungi is dominated by a TATA box or a TATA-like element. This motif is located at a fixed distance from a transcription initiation motif that is reminiscent of the Pol II initiator core promoter element.[53] Very different upstream signatures are present in the nematode, insect and mammalian genomes. In particular, TATA-like elements tend to be replaced by A or T-rich stretches placed at roughly equivalent positions (e.g. the −35 TTTTTC element in *C. elegans*), even though TATA-like sequences might be present and play important roles in special cases.[5,6] Also, completely new blocks of sequence conservation, located near position −20, appear in insect and mammals, as exemplified by the nearly identical A-rich motif found at position −22 upstream of mouse and human tDNAs. Because of its striking evolutionary conservation, such a motif is likely to play an important role in mammalian tRNA gene transcription.

From a broader perspective, a comparative view of Figure 6 strongly suggests that the mode of TFIIIB assembly upstream of tDNAs varies between lineages. In fungi and plants, the TFIIIB–tDNA interaction tends to be dominated by a TBP-recruitment step and/or by important TBP–TATA contacts. Mainly TATA-independent TFIIIB recruitment strategies are likely to operate, instead, in worms, insects, and mammals. Such strategies might involve specific TBP-related factors (TRFs), as it has been shown to be the case in *Drosophila*,[54] or TFIIIB components other than TBP/TRF. It is worth noting, in this regard, that no TRF, but only TATA-recognizing TBPs, are encoded in the genomes of fungi and plants, whose tDNAs are flanked by TATA (or TATA-like) elements. In contrast, TRFs have been identified in all the metazoans in which our analysis revealed the presence of non-TATA conserved sequence elements upstream of tRNA genes. It will thus be particularly interesting to find out whether non-TATA upstream motifs are specifically recognized by TRFs in these organisms. In addition, recent studies in the human system have revealed the existence of multiple forms of Brf1-like and Bdp1-like proteins, that might interact in various combinations, and with distinct specificities, with different class III genes.[55] In this case, upstream motifs might confer responsiveness not only to the levels, but also to the particular subunit composition of TFIIIB.

## Materials and Methods

### Statistical analysis of tDNA 5′-flanking regions

The 5′-flanking sequences of the 274 tRNA genes of *S. cerevisiae*, from position −1 up to position −90 with respect to the first nucleotide of the mature tRNA, were extracted from the yeast genome sequence retrieved from the MIPS web site†, using coordinates provided by Pol3scan.[25,56] Transfer RNA genes associated with Ty retroelements, identified on the basis of MIPS annotation and by similarity-based cluster analysis,[57] were eliminated from the set of tDNAs utilized for 5′ flank analysis.

The 5′-flanking regions of the remaining 221 tDNAs were analyzed with a specifically developed computer procedure, originally proposed by Waterman and Jones,[58] for the identification of short conserved motifs in unrelated sequences. Given a set of unaligned sequences, a fixed word size $k$, a number of possible mismatches $d$, and a window width $W$, the algorithm identifies, for each position, the best scoring word of length $k$. The sharpest signals above the background were obtained with $k = 6$, $d = 1$ and $W = 9$. The occurrence of consensus words with frequencies significantly deviating from random expectation was estimated with Monte Carlo simulations repeated 1000 times on randomized datasets obtained by shuffling individual sequences. An expected mean value was calculated from the resulting distribution. Minimum and maximum values, occurring in less than 1% of the sequences of randomized datasets, were chosen as significance thresholds.

The tDNA 5′-flanking sequences from other eukaryotic genomes were analysed with the same parameter values applied to *S. cerevisiae*. Coordinates for the tDNA upstream regions were predicted directly from raw genomic data by using Pol3scan and tRNAscan-SE.[59] Predicted elements were considered as tRNA genes only if the tRNA boundaries and anticodon type were found to be identical by the two methods. Redundant 5′-flanking sequences were excluded by similarity-based cluster analysis.[57] Expected values and significance of best frequency words were determined, for each genome, by computer simulations carried out on 100 randomized datasets.

The bendability profile of the tDNA 5′-flanking regions of *S. cerevisiae* was obtained using previously

† Munich Information center for protein sequences, http://mips.gsf.de

reported trinucleotide-based, DNAse I-derived parameters.[44] The average bendability of each consecutive, overlapping trinucleotide was calculated by moving from position −90 to position −1. Bendability values were further averaged by using a four nucleotide interval window. Expected values and significance were determined by computer simulations carried out on 1000 randomized datasets.

### PCR amplification, cloning and mutagenesis of tRNA genes

All the *S. cerevisiae* tRNA genes analyzed in this study were PCR-amplified from yeast genomic DNA (strain S288C) using the high-fidelity Deep Vent DNA polymerase (New England Biolabs) and gene-specific pairs of oligonucleotide primers. Amplification products, containing the tRNA coding region together with ~80 bp of 5′-flanking and ~50 bp of 3′-flanking sequences, were inserted into pUC-derived plasmid vectors and sequence-verified by dideoxy chain termination sequencing.

The CR Δ-58 and NR Δ-58 mutants were obtained by PCR using the tN(GTT)CR and tN(GTT)NR tDNAs as templates and gene-specific, "upstream deletion" primers. The CR [*start*NR] mutant was constructed by recombinant PCR.[60] Two overlapping PCR primary products were generated using the CR Δ-58 mutant as template and the oligonucleotide pairs N(GTT)CRfw (5′-CATA CTCG AAGG GTAG TTGG) with CR[*start*NR]rev (5′-CCAT GGAG TCAC TTTA AATT CTGC AACT TTTA CAG) and N(GTT)CRrev (5′-GATT TTTC CATT CGCC ATGC) with CR[*start*NR]fw (5′-CTGT AAAA GTTG CAGA ATTT AAAG TGAC TCCA TGG). The two primary amplification products were then mixed and used as templates for a subsequent amplification reaction, employing N(GTT)CRfw and N(GTT)CRrev as "outside" primers. The CR [3′NR] mutant was obtained by PCR using the N(GTT)CR tDNA as a template and a reverse oligonucleotide primer of sequence 5′GAAA TAAG AAAA AAAA CGAC CCCA GTAG GG, that fuses the tN(GTT)NR terminator, plus 8 bp of 3′-flanking sequence, to tN(GTT)CR.

### Electrophoretic mobility shift assays

Gel retardation assays were performed essentially as described.[61,62] Two 200 bp long DNA fragments, containing either the CR[3′NR] or the NR Δ-58 templates, only differing for 58 bp of 5′-flanking region, were prepared by *Eco*RI–*Hin*dIII digestion of the corresponding pUC-derived plasmids, gel-purified and end-labeled with [(α-$^{32}$P]dCTP and the Klenow fragment of DNA polymerase I. Reaction mixtures (20 μl) contained 24 mM Tris–HCl (pH 8.0), 10% (v/v) glycerol, 100 mM KCl, 1 mg/ml of ultrapure BSA, 5–10 fmol of radiolabeled DNA fragment, varying amounts of TFIIIC and TFIIIB, and pBlueScript-KS plasmid DNA at a 1:1 (w/w) DNA/protein ratio. Reaction mixtures were incubated for 25 minutes at 20 °C. Native gel electrophoresis and subsequent analysis were carried out as described.[62]

### *In vitro* and *in vivo* transcription analyses

Multiple and single round *in vitro* transcription assays, using recombinant or purified components of the Pol III transcription system, were carried out as described.[8]

For *in vivo* transcription analyses, the tRNA$^{Syn2}$ gene (an artificial variant of the tRNA$^{Glu}$(TTC) gene containing a pseudo-intron[29]) was synthesized by direct PCR amplification using two partially complementary oligonucleotides as primers:

Glu(TTC)syn2_fw (5′-TATG CCGG ATCC GATA TAGT GTAA CGGC TATC ACAT CACG CTTC AGCA AATA TGCT GAGC CGCA GTAA TCTGC), that generates a *Bam*HI site in the 5′ side of the amplicon; Glu(TTC)syn2_rev (5′CGTC GACA AAAA ATAA CTCC GATA CGGG GAGT CGAA CCCC GGTC TCCA CGGT CCGC AGAT TACT GCGG CTCAGC). The resulting, synthetic tDNA was then fused to the 5′-flanking sequences of either tN(GTT)CR or tN(GTT)NR. For this purpose, the two 5′-flanking regions (up to −58) were individually amplified. The tN(GTT)CR 5′ flank was amplified using the CR Δ-58 construct as template and the CR_*Bam*HI_rev (5′-TATT GCGG ATCC TTAT CTTG TGTC ACTT TTA CAG) and CR Δ-58fw oligonucleotides as primers. The tN(GTT)NR 5′ flank was amplified using the NR Δ-58 construct as template and NR_*Bam*HI_rev (5′-TATT GCGG ATCC TTAA ATTCT GCAG AACG TGA AAG) and NR Δ-58fw as primers. Each of the two 5′ flank amplicons was cut with *Bam*HI and ligated to the *Bam*HI-restricted tDNA$^{Syn2}$ fragment. The resulting, blunt-ended ligation products (CR_Syn2 and NR_Syn2) were inserted into the *Sma*I site of the multicopy plasmid YEp352, and subsequently subcloned into the centromeric plasmid pFL38. The tDNA$^{Syn2}$ amplicon, without any 5′ flank addition (Syn2), was also inserted into YEp352. The YEp352-CR_syn2, YEp352-NR_syn2 and YEp352-Syn2 plasmids were transformed into the W303-1b strain, the pFL38-CR_syn2 and pFL38-NR_syn2 plasmids into the YPH500 strain. Yeast total RNA extraction and RNA gel blot analysis were carried out as described[61] using the oligonucleotide Glu(TTC)-syn2_rev as a probe for both modified and natural tRNA$^{Glu}$(TTC).

Yeast mutant strains in which the chromosomal copy of *BRF1* is deleted[63] and a mutant copy of the same gene is provided by a centromeric plasmid were used for *in vivo* analyses carried out in the context *brf1* mutations. Two different *brf1* plasmids were utilized: the YpCC7 plasmid, carrying the *brf1ΔC86* mutant gene,[31] and the pRSM3*II.6 plasmid, which carries the *brf1* (*D464A D466A*) double mutation.[30] The two *brf1* thermosensitive strains were transformed with either pFL38-CR_syn2 or the pFL38-NR_syn2 construct. Total RNA was extracted before and after a six hour shift to 37 °C and subjected to gel blot analysis.

## Acknowledgements

We thank Christine Conesa and Michel Werner for plasmids and strains, Claudio Rivetti for insightful discussions, and Chiara Bertogalli for help in tDNA cloning. This study was supported by the Italian Ministry of Education, University and Research (COFIN 2001 Program) and by the Human Frontier Science Program Organization (grant RGY0011/2002-C to G.D.).

## References

1. Kassavetis, G. A., Riggs, D. L., Negri, R., Nguyen,

L. H. & Geiduschek, E. P. (1989). Transcription factor IIIB generates extended DNA interactions in RNA polymerase transcription complexes on tRNA genes. *Mol. Cell. Biol.* **9**, 2551–2566.
2. Geiduschek, E. P. & Kassavetis, G. A. (2001). The RNA polymerase III transcription apparatus. *J. Mol. Biol.* **310**, 1–26.
3. Kassavetis, G. A., Braun, B. R., Nguyen, L. H. & Geiduschek, P. E. (1990). *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell*, **60**, 235–245.
4. Young, L. S., Takahashi, N. & Sprague, K. U. (1986). Upstream sequences confer distinctive transcriptional properties on genes encoding silkgland-specific tRNAAla. *Proc. Natl Acad. Sci. USA*, **83**, 374–378.
5. Ouyang, C., Martinez, M. J., Young, L. S. & Sprague, K. U. (2000). TATA-binding protein–TATA interaction is a key determinant of differential transcription of silkworm constitutive and silk gland-specific tRNA(Ala) genes. *Mol. Cell. Biol.* **20**, 1329–1343.
6. Trivedi, A., Young, L. S., Ouyang, C., Johnson, D. L. & Sprague, K. U. (1999). A TATA element is required for tRNA promoter activity and confers TATA-binding protein responsiveness in *Drosophila* Schneider-2 cells. *J. Biol. Chem.* **274**, 11369–11375.
7. Choisne, N., Carneiro, V. T., Pelletier, G. & Small, I. (1997). Implication of 5′-flanking sequence elements in expression of a plant tRNA(Leu) gene. *Plant. Mol. Biol.* **36**, 113–123.
8. Dieci, G., Percudani, R., Giuliodori, S., Bottarelli, L. & Ottonello, S. (2000). TFIIIC-independent *in vitro* transcription of yeast tRNA genes. *J. Mol. Biol.* **299**, 601–613.
9. Moenne, A., Camier, S., Anderson, G., Margottin, F., Beggs, J. & Sentenac, A. (1990). The U6 gene of *Saccharomyces cerevisiae* is transcribed by RNA polymerase C (III) *in vivo* and *in vitro*. *EMBO J.* **9**, 271–277.
10. Hamada, M., Huang, Y., Lowe, T. M. & Maraia, R. J. (2001). Widespread use of TATA elements in the core promoters for RNA polymerases III, II, and I in fission yeast. *Mol. Cell. Biol.* **21**, 6870–6881.
11. Sprague, K. U. (1995). Transcription of eukaryotic tRNA genes. In *tRNA: Structure, Biosynthesis and Function* (Söll, D. & RajBhandary, U. L., eds), pp. 31–50, American Society for Microbiology, Washington.
12. White, R. J. (1998). *RNA Polymerase III Transcription*, 2nd edit., Springer, Berlin.
13. Shaw, K. J. & Olson, M. V. (1984). Effects of altered 5′-flanking sequences on the in vivo expression of a *Saccharomyces cerevisiae* tRNATyr gene. *Mol. Cell. Biol.* **4**, 657–665.
14. Joazeiro, C. A., Kassavetis, G. A. & Geiduschek, E. P. (1996). Alternative outcomes in assembly of promoter complexes: the roles of TBP and a flexible linker in placing TFIIIB on tRNA genes. *Genes Dev.* **10**, 725–739.
15. Lee, Y., Wong, W. M., Guyer, D., Erkine, A. M. & Nazar, R. N. (1997). *In vivo* analyses of upstream promoter sequence elements in the 5 S rRNA gene from *Saccharomyces cerevisiae*. *J. Mol. Biol.* **269**, 676–683.
16. Raymond, K. C., Raymond, G. J. & Johnson, J. D. (1985). *In vivo* modulation of yeast tRNA gene expression by 5′-flanking sequences. *EMBO J.* **4**, 2649–2656.
17. Raymond, G. J. & Johnson, J. D. (1987). The 5′-flanking sequence of yeast tRNA(Leu3) genes enhances the rate of transcription from stable pre-initiation complexes. *Nucl. Acids Res.* **15**, 9881–9894.
18. Fruscoloni, P., Zamboni, M., Panetta, G., De-Paolis, A. & Tocchini-Valentini, G. P. (1995). Mutational analysis of the transcription start site of the yeast tRNA(Leu3) gene. *Nucl. Acids Res.* **23**, 2914–2918.
19. Zecherle, G. N., Whelen, S. & Hall, B. D. (1996). Purines are required at the 5′ ends of newly initiated RNAs for optimal RNA polymerase III gene expression. *Mol. Cell. Biol.* **16**, 5801–5810.
20. Leveillard, T., Kassavetis, G. A. & Geiduschek, E. P. (1991). *Saccharomyces cerevisiae* transcription factors IIIB and IIIC bend the DNA of a tRNA(Gln) gene. *J. Biol. Chem.* **266**, 5162–5168.
21. Braun, B. R., Kassavetis, G. A. & Geidushek, E. P. (1992). Bending of the *Saccaromyces cerevisiae* 5 S rRNA gene in transcription factor complexes. *J. Biol. Chem.* **267**, 22562–22569.
22. Juo, Z. S., Kassavetis, G. A., Wang, J., Geiduschek, E. P. & Sigler, P. B. (2003). Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature*, **422**, 534–539.
23. Colbert, T., Lee, S., Schimmack, G. & Hahn, S. (1998). Architecture of protein and DNA contacts within the TFIIIB–DNA complex. *Mol. Cell. Biol.* **18**, 1682–1691.
24. Persinger, J., Sengupta, S. M. & Bartholomew, B. (1999). Spatial organization of the core region of yeast TFIIIB&DNA complexes. *Mol. Cell. Biol.* **19**, 5218–5234.
25. Percudani, R., Pavesi, A. & Ottonello, S. (1997). Transfer RNA gene redundancy and translational selection in *Saccharomyces cerevisiae*. *J. Mol. Biol.* **268**, 322–330.
26. Hani, J. & Feldmann, H. (1998). tRNA genes and retroelements in the yeast genome. *Nucl. Acids Res.* **26**, 689–696.
27. Dieci, G. & Sentenac, A. (2003). Detours and shortcuts to transcription reinitiation. *Trends. Biochem. Sci.* **28**, 202–209.
28. Yukawa, Y., Sugita, M., Choisne, N., Small, I. & Sugiura, M. (2000). The TATA motif, the CAA motif and the poly(T) transcription termination motif are all important for transcription re-initiation on plant tRNA genes. *Plant J.* **22**, 439–447.
29. Krieg, R., Stucka, R., Clark, S. & Feldmann, H. (1991). The use of a synthetic tRNA gene as a novel approach to study *in vivo* transcription and chromatin structure in yeast. *Nucl. Acids Res.* **19**, 3849–3855.
30. Andrau, J. C., Sentenac, A. & Werner, M. (1999). Mutagenesis of yeast TFIIIB70 reveals C-terminal residues critical for interaction with TBP and C34. *J. Mol. Biol.* **288**, 511–520.
31. Huet, J., Conesa, C., Carles, C. & Sentenac, A. (1997). A cryptic DNA binding domain at the COOH terminus of TFIIIB70 affects formation, stability, and function of preinitiation complexes. *J. Biol. Chem.* **272**, 18341–18349.
32. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*, **423**, 241–254.
33. DeFranco, D., Sharp, S. & Soll, D. (1981). Identification of regulatory sequences contained in the 5′-flanking region of *Drosophila* lysine tRNA2 genes. *J. Biol. Chem.* **256**, 12424–12429.
34. Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R. *et al.*

(2002). Initial sequencing and comparative analysis of the mouse genome. *Nature*, **420**, 520–562.
35. Wilson, E. T., Larson, D., Young, L. S. & Sprague, K. U. (1985). A large region controls tRNA gene transcription. *J. Mol. Biol.* **183**, 153–163.
36. Johnson, J. D. & Raymond, G. J. (1984). Three regions of a yeast tRNALeu3 gene promote RNA polymerase III transcription. *J. Biol. Chem.* **259**, 5990–5994.
37. Huibregtse, J. M. & Engelke, D. R. (1989). Genomic footprinting of a yeast tRNA gene reveals stable complexes over the 5′-flanking region. *Mol. Cell. Biol.* **9**, 3244–3252.
38. Kumar, A., Grove, A., Kassavetis, G. A. & Geiduschek, E. P. (1998). Transcription factor IIIB: the architecture of its DNA complex, and its roles in initiation of transcription by RNA polymerase III. *Cold Spring Harbor Symp. Quant. Biol.* **63**, 121–129.
39. Cloutier, T. E., Librizzi, M. D., Mollah, A. K., Brenowitz, M. & Willis, I. M. (2001). Kinetic trapping of DNA by transaction factor IIIB. *Proc. Natl Acad. Sci. USA*, **98**, 9581–9586.
40. Martin, M. P., Gerlach, V. L. & Brow, D. A. (2001). A novel upstream RNA polymerase III promoter element becomes essential when the chromatin structure of the yeast U6 RNA gene is altered. *Mol. Cell. Biol.* **21**, 6429–6439.
41. Fiering, S., Whitelaw, E. & Martin, D. I. (2000). To be or not to be active: the stochastic nature of enhancer action. *Bioessays*, **22**, 381–387.
42. Ong, W. C., Ibrahim, M., Town, M. & Johnson, J. D. (1997). Functional differences among the six *Saccharomyces cerevisiae* tRNATrp genes. *Yeast*, **13**, 1357–1362.
43. Grove, A., Kassavetis, G. A., Johnson, T. E. & Geiduschek, E. P. (1999). The RNA polymerase III-recruiting factor TFIIIB induces a DNA bend between the TATA box and the transcriptional start site. *J. Mol. Biol.* **285**, 1429–1440.
44. Brukner, I., Sanchez, R., Suck, D. & Pongor, S. (1995). Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. *EMBO J.* **14**, 1812–1818.
45. Rivetti, C., Guthold, M. & Bustamante, C. (1999). Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex. *EMBO J.* **18**, 4464–4475.
46. Rivetti, C., Codeluppi, S., Dieci, G. & Bustamante, C. (2003). Visualizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases. *J. Mol. Biol.* **326**, 1413–1426.
47. Tahirov, T. H., Temiakov, D., Anikin, M., Patlan, V., McAllister, W. T., Vassylyev, D. G. & Yokoyama, S. (2002). Structure of a T7 RNA polymerase elongation complex at 2.9 Å resolution. *Nature*, **420**, 43–50.
48. Upadhya, R., Lee, J. & Willis, I. M. (2002). Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. *Mol. Cell*, **10**, 1489–1494.
49. Dieci, G., Duimio, L., Peracchia, G. & Ottonello, S. (1995). Selective inactivation of two components of the multiprotein transcription factor TFIIIB in cycloheximide growth-arrested yeast cells. *J. Biol. Chem.* **270**, 13476–13482.
50. Sethy, I., Moir, R. D., Librizzi, M. & Willis, I. M. (1995). *In vitro* evidence for growth regulation of tRNA gene transcription in yeast. A role for transcription factor (TF) IIIB70 and TFIIIC. *J. Biol. Chem.* **270**, 28463–28470.
51. Clarke, E. M., Peterson, C. L., Brainard, A. V. & Riggs, D. L. (1996). Regulation of the RNA polymerase I and III transcription systems in response to growth conditions. *J. Biol. Chem.* **271**, 22189–22195.
52. Donze, D. & Kamakaka, R. T. (2001). RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in *Saccharomyces cerevisiae*. *EMBO J.* **20**, 520–531.
53. Smale, S. T., Jain, A., Kaufmann, J., Emami, K. H., Lo, K. & Garraway, I. P. (1998). The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes. *Cold Spring Harbor Symp. Quant. Biol.* **63**, 21–31.
54. Takada, S., Lis, J. T., Zhou, S. & Tjian, R. (2000). A TRF1:BRF complex directs *Drosophila* RNA polymerase III transcription. *Cell*, **101**, 459–469.
55. Schramm, L. & Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* **16**, 2593–2620.
56. Pavesi, A., Conterio, F., Bolchi, A., Dieci, G. & Ottonello, S. (1994). Identification of new eukaryotic tRNA genes in genomic DNA databases by a multistep weight matrix analysis of transcriptional control regions. *Nucl. Acids Res.* **22**, 1247–1256.
57. Parsons, J. D. (1995). Improved tools for DNA comparison and clustering. *Comput. Appl. Biosci.* **11**, 603–613.
58. Waterman, M. S. & Jones, R. (1990). Consensus methods for DNA and protein sequence alignment. *Methods Enzymol.* **183**, 221–237.
59. Lowe, T. M. & Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl. Acids Res.* **25**, 955–964.
60. Higuchi, R. (1990). Recombinant PCR. In *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J., eds), pp. 177–183, Academic Press, Inc., San Diego, CA.
61. Dieci, G., Giuliodori, S., Catellani, M., Percudani, R. & Ottonello, S. (2002). Intragenic promoter adaptation and facilitated RNA polymerase III recycling in the transcription of SCR1, the 7SL RNA gene of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**, 6903–6914.
62. Pizzi, S., Dieci, G., Frigeri, P., Piccoli, G., Stocchi, V. & Ottonello, S. (1999). Domain organization and functional properties of yeast transcription factor IIIA species with different zinc stoichiometries. *J. Biol. Chem.* **274**, 2539–2548.
63. Colbert, T. & Hahn, S. (1992). A yeast TFIIB-related factor involved in RNA polymerase III transcription. *Genes Dev.* **6**, 1940–1949.

***Edited by M. Yaniv***

*(Received 14 May 2003; received in revised form 5 August 2003; accepted 11 August 2003)*