

UNIVERSITÀ DEGLI STUDI DI TRIESTE

---

DIPARTIMENTO DI BIOCHIMICA, BIOFISICA E CHIMICA DELLE MACROMOLECOLE

DOTTORATO DI RICERCA IN BIOCHIMICA  
XI CICLO

TESI DI DOTTORATO

I FATTORI ARCHITETTURALI HMGI, HMGY E HMGI-C:  
STUDIO DELLA LORO ASSOCIAZIONE CON IL DNA E DELLA  
REGOLAZIONE DELLA LORO ESPRESSIONE

Dottoranda:

Dr. *Alessandra RUSTIGHI*

Relatore:

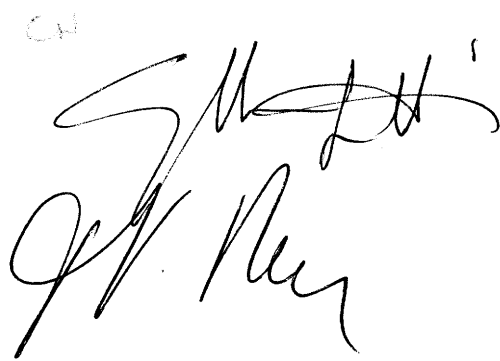
*Prof. Vincenzo GIANCOTTI*  
Università degli Studi di Trieste

Correlatore:

*Dr. Guidalberto MANFIOLETTI*  
Università degli Studi di Trieste

Coordinatore:

*Prof. Franco VITTUR*  
Università degli Studi di Trieste

CH  


---

ANNO ACCADEMICO 1998-1999

## Riassunto

È noto che la regolazione della trascrizione prevede la formazione di complessi multiproteici a livello dei promotori ed *enhancer* dei relativi geni. Nella costituzione di questi complessi nucleoproteici entrano in gioco numerose proteine, tra le quali alcune non hanno una funzione nell'attivazione trascrizionale propriamente detta, bensì sono essenziali come fattori architettonici nella strutturazione della cromatina attiva. Fra queste si collocano per prime le proteine HMGI (High Mobility Group I), un gruppo composto da tre proteine: HMGI ed HMGY derivanti da splicing alternativo dello stesso gene ed HMGI-C che invece è codificata da un gene diverso.

La loro attività funzionale si esplica tramite complesse interazioni col DNA e con altri fattori trascrizionali modulando in questo modo la trascrizione di numerosi geni. Il modello meglio caratterizzato è quello dell'*enhanceosoma* dell'Interferone  $\beta$  umano, ma si è dimostrato che un ampio numero di geni codificanti per citochine, molecole di adesione cellulare, fattori di crescita e fattori trascrizionali è regolato, sia positivamente che negativamente, dalle proteine HMGI, dimostrando il ruolo specifico delle stesse nella regolazione trascrizionale.

L'espressione di tali proteine in tessuti normali adulti è bassissima o non rilevabile, mentre elevati livelli delle stesse sono strettamente correlati a fenotipi indifferenziati o trasformati in senso neoplastico. In particolare per l'HMGI-C è stato dimostrato, con metodologie antisense, un ruolo causale della sua espressione nella trasformazione neoplastica indotta da virus oncogeni.

Inoltre estesi studi di tipo citogenetico sostengono un coinvolgimento della regione genica codificante per HMGI-C in numerosi tumori benigni di origine mesenchimale, quali leiomiomi uterini, adenomi pleiomorfici, amartomi polmonari e soprattutto lipomi. Correlati a questi dati si situano gli studi su topi nani, derivanti da inattivazione funzionale di entrambi gli alleli per l'HMGI-C, a cui si contrappongono fenotipicamente i topi transgenici "giganti" dovuti ad iperespressione di una forma di HMGI-C alterata.

Tutto ciò depone a favore di un ruolo della HMGI-C nella proliferazione e/o differenziamento cellulare, con massima rilevanza soprattutto per cellule di origine mesenchimale.

L'espressione fisiologica di *Hmgi-c* è ristretta agli stadi precoci dell'embriogenesi e decresce con l'inizio dell'organogenesi nell'embrione del topo spegnendosi subito dopo il 16° giorno

di età embrionale. Si è dimostrato che la riespressione di questo fattore in concomitanza con la trasformazione neoplastica avviene a livello trascrizionale, ma i processi molecolari coinvolti sono ancora oscuri.

Il lavoro svolto durante il dottorato di ricerca si è incentrato sullo studio dei meccanismi molecolari inerenti al funzionamento delle proteine HMGI come regolatori trascrizionali in diversi modelli cellulari, e alla regolazione della loro espressione. In particolare, si è visto che le proteine HMGY ed HMGI-C interferiscono con la regolazione trascrizionale di geni modulati da proteine omeobox, fattori indispensabili durante lo sviluppo embrionale. Si è proposto che le proteine HMGI in generale possano intervenire durante gli *switch* regolatori, tipici delle proteine contenenti un omeodominio. Inoltre si è stabilito per la prima volta che anche la proteina HMGI-C, in analogia ad HMGI(Y), agisce come fattore architetturale a livello del promotore dell'Interferone  $\beta$  umano.

Data l'associazione delle proteine HMGI con il fenotipo neoplastico è stato sviluppato un anticorpo contro il peptide N-terminale della HMGI(Y) che si è dimostrato un ottimo marcatore tumorale di lesioni già a livelli precoci. Inoltre si è stabilita una correlazione tra immunoreattività all'anticorpo e il grado di displasia in carcinomi della cervice uterina.

Focalizzando l'attenzione sul'HMGI-C, avente un ruolo d'elezione nella trasformazione cellulare, si è voluto comprendere i meccanismi trascrizionali che portano ad una riespressione di questo fattore. È stato quindi isolato e caratterizzato il promotore minimo del gene murino *Hmgi-c*, la cui attività basale è regolata da fattori trascrizionali appartenenti alle famiglie Sp1 e CTF/NF-1. Si è dimostrato che Sp1 ed Sp3 interagiscono con un sito multiplo del promotore, caratterizzato da una sequenza polipirimidinica a monte del sito di inizio della trascrizione, rilevante per l'attività genica. Sp1 e CTF/NF-1 hanno proprietà attivanti sul gene *Hmgi-c*, mentre Sp3 non sembra influenzarlo. L'elemento polipirimidinico è sotto attiva indagine per la sua capacità intrinseca di adottare strutture non-B del DNA, che possono oltresì influenzare l'attività del promotore.

La stessa regione del promotore minimo di *Hmgi-c* responsabile dell'attività basale è anche più attiva in cellule trasformate rispetto alla controparte normale. In relazione a ciò si è identificato un complesso proteico potenzialmente coinvolto in questa attività differenziale. Si presume un ruolo di repressione della trascrizione di *Hmgi-c* in cellule normali, dato che il complesso sopra citato risulta essere più abbondante o attivo in queste rispetto alle relative cellule trasformate. A supporto di quanto detto si è osservato che l'interazione tra il promotore ed il complesso avviene a livello del sito di inizio della trascrizione, permettendo

di ipotizzare un meccanismo di attivazione di *Hmgi-c* in cellule trasformate dovuta a derepressione.

Allo scopo di analizzare la completezza e la funzionalità della sequenza promotore del gene *Hmgi-c* da noi isolate, sono stati creati topi transgenici recanti un gene esogeno costituito da 4 Kb di sequenza 5' del gene *Hmgi-c* a monte del gene marcatore *lacZ* di *E. Coli* codificante per una  $\beta$ -galattosidasi. Attraverso il confronto tra espressione fisiologica nel topo normale durante l'embriogenesi e la colorazione blu degli embrioni transgenici derivante dall'espressione della  $\beta$ -galattosidasi, ovunque vi siano fattori in grado di attivare *Hmgi-c* si potrà stabilire se nelle 4 Kb di sequenza *Hmgi-c* sono custodite tutte le informazioni necessarie per la corretta espressione *in vivo*.

Durante il dottorato di ricerca ho avuto l'opportunità di acquisire tecniche biochimiche e di biologia molecolare di base ed avanzate, riguardanti la produzione e purificazione come anche l'analisi di proteine e DNA, e delle loro associazioni *in vitro*. Inoltre ho svolto ricerche usando svariati tipi di colture cellulari, effettuando saggi di attività *in vivo*.

Il lavoro è stato svolto presso il Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole.

# INDICE

<b>Capitolo 1. Introduzione</b>	pag. 1
1.1 Regolazione trascrizionale	pag. 1
Trascrizione basale	pag. 1
Comunicazione tra fattori basali della trascrizione e regolatori	pag. 3
Attivazione trascrizionale	pag. 4
Cromatina e regolazione trascrizionale	pag. 5
1.2 La famiglia HMGI	pag. 9
1.2.1 I geni <i>HMGI</i>	pag. 9
1.2.2 Le proteine HMGI	pag. 12
1.2.3 Il dominio di legame delle proteine HMGI: <i>AT-hook</i>	pag. 13
Motivi HMGI-simili in altre proteine	pag. 18
1.2.4 Modificazioni post-traduzionali delle HMGI: effetti biologici	pag. 21
Fosforilazione	pag. 21
Acetilazione	pag. 27
ADP-ribosilazioni	pag. 28
1.2.5 Funzioni biologiche delle proteine HMGI	pag. 31
HMGI e cromatina	pag. 31
HMGI e Controllo trascrizionale	pag. 37
1.2.6 Espressione delle HMGI	pag. 53
1.2.7 HMGI e trasformazione neoplastica	pag. 59
1.2.8 HMGI e tumori benigni	pag. 66
1.2.9 HMGI, mesenchima ed adipogenesi	pag. 74
<b>Capitolo 2. Scopo della tesi</b>	pag. 77
<b>Capitolo 3. Materiali e metodi</b>	pag. 79
3.1 Analisi del DNA	pag. 79

3.1.1	Estrazione di DNA plasmidico	pag. 79
3.1.2	Elettroforesi in gel d'agarosio	pag. 79
3.1.3	Manipolazioni del DNA	pag. 81
3.1.4	Trasferimento di DNA su filtro mediante <i>Southern Blotting</i>	pag. 86
3.1.5	Marcatura di DNA mediante Random Primed Radioabelling	pag. 86
3.1.6	Ibridazione di acidi nucleici su filtro di nylon	pag. 87
3.1.7	Trasformazione batterica	pag. 88
3.1.8	Sequenziamento del DNA	pag. 89
3.2	Analisi dell'RNA	pag. 92
3.2.1	Estrazione dell'RNA	pag. 92
3.2.2	Selezione dell'RNA poly(A) <sup>+</sup>	pag. 93
3.2.3	Elettroforesi di RNA su gel d'agarosio	pag. 93
3.2.4	Trasferimento su filtro di RNA mediante <i>Northern blotting</i>	pag. 94
3.2.5	Caratterizzazione dell'estremità 5' del filamento di RNA	pag. 94
3.3	Analisi delle proteine	pag. 95
3.3.1	Espressione di proteine ricombinanti	pag. 95
3.3.2	Estrazione di proteine con acido perclorico	pag. 96
3.3.3	Analisi elettroforetiche	pag. 96
3.3.4	Purificazione proteica mediante HPLC	pag. 97
3.3.5	Analisi delle proteine mediante spettrometria di massa	pag. 98
3.3.6	Trasferimento delle proteine su membrana ( <i>Western blotting</i> )	pag. 99
3.3.6	Riconoscimento con anticorpi	pag. 100
3.3.7	Interazioni proteina-DNA	pag. 101
3.4	Linee cellulari e condizioni di coltura	pag. 107
3.4.1	Linee cellulari	pag. 107
3.4.2	Condizioni di coltura	pag. 107
3.4.3	Trasfezione di cellule eucariotiche in adesione	pag. 108
3.5	Analisi di topi transgenici	pag. 112
3.5.1	Estrazione di DNA da pezzi di coda o dita	pag. 112
3.5.2	Estrazione del DNA genomico dal sacco vitellino	pag. 113
3.5.3	<i>Southern blot</i> genomico	pag. 113
3.5.4	Colorazione e disidratazione degli embrioni	pag. 114

3.6	Composizione delle soluzioni	pag. 115
-----	------------------------------	----------

## **Capitolo 4. Risultati e Discussione**

4.1	Le proteine HMGI in associazione con il DNA	pag. 122
4.1.1	Le proteine HMGI interferiscono con gli <i>omeodomini</i>	pag. 122
4.1.2	Il fattore architetturale HMGI-C aumenta l'attivazione trascrizionale mediata da NF- $\kappa$ B	pag. 127
4.2	Studio della regolazione dell'espressione dei fattori HMGI	pag. 130
4.2.1	Utilizzo delle proteine HMGI come marcatori nella diagnosi tumorale	pag. 130
4.2.2	Caratterizzazione del promotore <i>Hmgi-c</i>	pag. 133
	I fattori Sp1 e CTF/NF-1 sono coinvolti nell'espressione basale del promotore prossimale di <i>Hmgi-c</i>	pag. 133
	Il promotore <i>Hmgi-c</i> è più attivo in cellule trasformate che nella loro controparte normale	pag. 141
	Il tratto omopirimidinico/omopurinico del promotore minimo di <i>Hmgi-c</i> è una caratteristica comune di molti geni correlati con la crescita	pag. 146
	Analisi del promotore <i>in vivo</i>	pag. 156

<b>Conclusioni</b>	pag. 158
--------------------	----------

<b>Bibliografia</b>	pag. 160
---------------------	----------

<b>Appendice</b>	pag. 166
------------------	----------

# Capitolo 1.

## Introduzione

### 1.1 Regolazione trascrizionale

La regolazione coordinata e temporale dell'espressione genica è essenziale per la corretta crescita e lo sviluppo di tutti gli organismi, non ci sorprende quindi che la trascrizione di geni che codificano per proteine è un processo finemente accordato ed altamente controllato. Ricerche negli ultimi dieci anni hanno contribuito molto nella comprensione dell'apparato molecolare responsabile per la sintesi di RNA messaggero in cellule eucariotiche. Più di 40 diverse subunità proteiche funzionano di concerto per regolare la trascrizione operata dalla RNA polimerasi II a livello di specifici promotori di cromosomi eucariotici. Esperimenti recenti indicano che molti di questi fattori di trascrizione possono interagire fra di loro in modi altamente specifici per influenzare l'attività di geni individuali. (1)

#### Trascrizione basale

L'inizio della trascrizione è uno dei punti più importanti nella regolazione dell'espressione genica. Il meccanismo attraverso il quale una molecola di RNA viene accuratamente sintetizzata è altamente ordinato in tutti gli organismi viventi, sia procarioti che eucarioti. Per studiare la trascrizione eucariotica è istruttivo considerare i sistemi analoghi ma meno complessi della trascrizione procariotica. È ormai ben noto che il livello di trascrizione di qualsiasi operone procariotico è intimamente correlato alla sequenza del DNA promotoriale. L'inizio della trascrizione non regolata a livello di parecchi promotori procariotici richiede solo l'oloenzima della RNA polimerasi, costituita da 4 subunità centrali ed un fattore  $\sigma$  dissociabile. Il nucleo dell'enzima non può iniziare la trascrizione da promotori senza un



fattore  $\sigma$ , ma può allungare i trascritti di RNA messaggero dopo l'iniziazione e la dissociazione di  $\sigma$ . Sono stati identificati molti e diversi fattori  $\sigma$ , ciascuno dei quali programma il nucleo dell'enzima per trascrivere a partire da differenti classi di promotori e per rispondere a tipi diversi di segnali regolativi trascrizionali. Quindi i fattori  $\sigma$  sono fattori di specificità nella trascrizione procariotica, ed un oloenzima contenente un specifico fattore  $\sigma$  può funzionare solo su di un sottogruppo di promotori.

È probabile che le RNA polimerasi eucariotiche condividano alcuni aspetti di questa architettura del nucleo enzimatico/fattore  $\sigma$ . Infatti, le subunità centrali della RNA polimerasi II eucariotica sono capaci di catalizzare la sintesi di RNA, ma sono incapaci di attuare una trascrizione gene specifica. Invece una moltitudine di fattori di trascrizione accessori generali (denominati GTFs, costituiti da TFIIA, TFIIB, TFIID, TFIIE, TFIIF e TFIIH) lavorano assieme alla RNA polimerasi II per adempiere al riconoscimento del promotore e all'accurato inizio della trascrizione. (2) La gamma completa di questi GTFs assieme alla RNA polimerasi II sono in genere sufficienti per dirigere livelli basali di trascrizione *in vitro*, a partire da promotori forti, quali ad esempio promotori contenenti la TATA *box*. Quindi l'apparato trascrizionale eucariotico è caratterizzato da una RNA polimerasi centrale più numerosi fattori accessori, che nel loro insieme funzionerebbero come un fattore  $\sigma$ .

Un aspetto fondamentale che merita una considerazione speciale è la diversità e l'unicità dei diversi promotori utilizzati dalla RNA polimerasi seconda. Ad esempio non tutti i promotori di classe II contengono la TATA *box* e l'elemento *iniziatore*, (3) che sono i due elementi centrali di un classico promotore identificati finora. Inoltre la sequenza di questi elementi e la spaziatura tra di loro può variare in modo significativo tra diversi promotori dell'RNA polimerasi II. Il fatto che esistano i nuclei promotoriali così diversi indica che anche le interazioni tra l'apparato generale della trascrizione ed il DNA del promotore variano considerevolmente da promotore a promotore. Dall'analisi dei contatti del macchinario basale della trascrizione con molti promotori diversi tra di loro sono emersi alcuni meccanismi generali con cui i diversi elementi centrali dei promotori vengono riconosciuti ed utilizzati per dare inizio alla trascrizione. I più importanti si possono descrivere come meccanismi che avvengono a livello di promotori con la TATA *box* e quelli senza, definiti *TATA-less*. Tra tutti questi alcuni possiedono anche l'*iniziatore*. Comunque la TATA *box*, posta ad una distanza classica a monte del sito di inizio della trascrizione (-25 fino a -30) di molti geni viene specificamente riconosciuta da un fattore, denominato TBP (TATA box binding protein) (4, 5) che in seguito è stato identificato come subunità del fattore TFIID. Comunque molti promotori

eucariotici non possiedono né la TATA *box* né una sequenza alternativa che potrebbe servire come attacco per l'apparato trascrizionale in assenza della TATA *box*. In esperimenti di *footprinting* i promotori che possiedono un iniziatore mostrano una ampia protezione da parte di TFIID, ma non TBP, indicando quindi che un singolo GTF, ossia TFIID, è in grado di riconoscere due elementi promotoriali distinti e molto diversi tra di loro, come una TATA *box* ed un *iniziatore*.(6) Una spiegazione possibile per queste proprietà di legame multifunzionali al DNA è data dalla natura del fattore stesso, che è dato da un complesso grande e stabile di almeno nove subunità. (7). Comunque TFIID non sembra essere l'unico GTF a contattare il DNA durante l'inizio della trascrizione. È evidente innanzitutto che alcune subunità dell'RNA polimerasi II devono contattare il DNA per poter trascrivere, inoltre vi sono evidenze sperimentali che anche TFIIH, di cui una subunità ha attività elicastica, contatta lo stampo da trascrivere. Infine alcune subunità di TFIIA, TFIIB e la piccola subunità di TFIIF, RAP 30, sono state riscontrate in stretta vicinanza con il DNA del promotore tardivo dell'adenovirus Ad2. In conclusione molti, se non tutti i GTFs, possiedono il potenziale di contattare il DNA durante alcuni passaggi della reazione della trascrizione.

### **Comunicazione tra fattori basali della trascrizione e regolatori**

La trascrizione a partire da promotori naturali della RNA polimerasi II è strettamente controllata da azioni combinate di fattori di regolazione positivi e negativi, inclusi attivatori sito specifici, repressori e proteine associate alla cromatina. Affinché sia trascritto un qualsiasi gene una serie complessa di segnali deve essere reclutata in corrispondenza del promotore per stabilire il livello di produzione di RNA. In questo senso si sono rivelati importantissimi i contatti proteina-proteina tra attivatori, o piuttosto tra i loro domini di attivazione, ed i componenti dell'apparato generale della trascrizione. Ad oggi è stata dimostrata l'interazione di alcuni attivatori con diversi GTFs, incluse alcune subunità di TFIID, TFIIB, TFIIF e TFIIH. (2, 8) Per molte interazioni proteina-proteina però rimane da verificare un'effettiva funzione nell'attivazione trascrizionale.

Esistono coattivatori e cofattori, come ad esempio i TAFs, PC4, Dr2 ed altri (2, 8) i quali sono più importanti per l'attivazione trascrizionale piuttosto che per la trascrizione basale. La scoperta di questa caratteristica assieme all'uso di proteine ricombinanti ha permesso di sperimentare le interazioni tra attivatori e coattivatori nella trascrizione *in vitro*, e ha portato

alla conclusione, almeno per quel che riguarda i TAFs di TFIID, che non tutte le otto subunità sono richieste per l'attivazione attuata da attivatori individuali, ma che i contatti tra attivatore e TAFs specifici sono un passaggio essenziale durante il processo dell'attivazione trascrizionale.

Quindi il complesso TFIID può venire considerato come un trasmettitore sfaccettato che riceve segnali da differenti classi di attivatori e in qualche modo li ripone sul macchinario basale della trascrizione che induce l'attivazione a partire dai promotori bersaglio.

### **Attivazione trascrizionale**

L'attivazione trascrizionale di geni eucariotici durante lo sviluppo o in risposta a segnali extracellulari coinvolge l'assemblaggio regolato di complessi multiproteici a livello di "enhancers" e promotori. La natura complessa di questi processi fornisce potenzialmente innumerevoli possibilità di regolazione, ma soprattutto risulta in un meccanismo elaborato e infallibile per controllare l'espressione genica. Gli elementi cruciali in questi processi sono fattori di trascrizione sequenza-specifici che selezionano i geni da attivare e dirigono il montaggio del macchinario trascrizionale a livello del sito di inizio della sintesi di RNA messaggeri dalla RNA polimerasi II. (9)

Ma cos'è veramente l'attivazione trascrizionale? È altamente probabile che un legame diretto od un contatto tra un attivatore ed un GTF costituisca uno dei passi più precoci nel processo di attivazione trascrizionale. E come possono i contatti proteina-proteina produrre livelli di trascrizione aumentati a partire dai promotori bersaglio? Tre sono i meccanismi ipotizzati attraverso i quali contatti tra attivatori con il macchinario basale della trascrizione portano all'aumento della trascrizione. In primo luogo è possibile che un attivatore funzioni reclutando direttamente un GTF a livello del promotore bersaglio (10, 11), oppure come seconda ipotesi, il contatto attivatore-GTF potrebbe provocare un'alterazione conformazionale di uno o più elementi di un complesso di pre-iniziazione della trascrizione in via di formazione o parzialmente assemblato, o anche del DNA del promotore o di qualche componente di un complesso nucleoproteico montato a livello del promotore. Si parla anche di un cosiddetto montaggio indotto (o *induced fit*), in quanto i domini bersaglio degli attivatori spesso non sono strutturati, e lo diventano appena con il contatto con l'attivatore, evitando così interazioni improduttive e ottenendo specificità e flessibilità nel riconoscimento proteina-proteina. (12)

Una terza ipotesi invece mette in causa modificazioni covalenti dei GTFs che potrebbero avere effetti sia positivi che negativi sui livelli trascrizionali dei promotori in questione. Non serve nemmeno che gli attivatori abbiano attività enzimatica, ma basta che essi amplifichino l'attività di enzimi come ad esempio protein-chinasi già associate con i complessi d'inizio. Tre fattori con attività serin-chinasica, le cui attività potrebbero essere inficiate da attivatori trascrizionali, sono le subunità di TFIIF, denominate Cdk7/MO15, (13) la coppia chinasi/ciclina associata alla RNA polimerasi II in levito (14) ed infine la subunità TAF<sub>II</sub>250 del complesso TFIID. (15)

Un interrogativo che oggi si pongono gli esperti di trascrizione genica è quello di capire come un numero relativamente modesto di diversi fattori di trascrizione possano essere usati per ottenere gli altissimi livelli di specificità richiesti per regolare gli aspetti più complessi della trascrizione negli eucarioti superiori. Una parziale risposta è data dal fatto che fattori di trascrizione ed *enhancers* sono composti da elementi modulari. Ciò significa ad esempio che un tipico fattore di trascrizione contiene un dominio specifico per legarsi al DNA che contatta direttamente il DNA, un dominio di multimerizzazione che permette la formazione di omo- e eteromultimeri ed infine un dominio di attivazione trascrizionale. Inoltre questi domini possono essere combinati in modo modulare a formare nuovi fattori pienamente funzionali. (16) Analogamente gli *enhancers* contengono raggruppamenti distinti di siti di legame per fattori di trascrizione, e variazioni nella disposizione di tali siti forniscono il potenziale per creare complessi nucleoproteici unici (12).

Un ulteriore livello di specificità dell'attivazione trascrizionale di singoli geni è data dall'assemblaggio di complessi stereospecifici che si formano a livello di *enhancers*. Essenziali durante la formazione di questi complessi sono proteine che legano il DNA in modo sequenza specifico, ma che funzionano come componenti architetturali, che rendono possibili interazioni sinergiche tra proteine legate ai diversi siti a livello dello stesso complesso multifattoriale, risultando in specificità ed alti livelli di trascrizione. (Vedi sezione seguente e sezione 1.2.5 HMGI e trascrizione genica)

## **Cromatina e regolazione trascrizionale**

Anche se il determinante fondamentale dell'espressione genica differenziale risiede in interazioni combinatorie a livello di regioni regolative tra fattori specifici e proteine

ubiquitarie, crescenti evidenze indicano che anche le componenti cromatiniche rivestono un ruolo essenziale nei meccanismi di regolazione trascrizionale.

In primo luogo c'è il controllo dell'accesso dei fattori trascrizionali ai rispettivi siti da parte degli istoni che compattano il DNA nucleare in nucleosomi e in strutture cromatiniche d'ordine superiore. Ciò ha un duplice effetto. In alcuni casi, infatti, il posizionamento caratteristico di certe strutture cromatiniche ha un effetto positivo, in quanto porta a contatto elementi regolativi dispersi, fornendo l'infrastruttura necessaria all'instaurarsi di una conformazione tridimensionale utile a garantire un'azione efficiente del macchinario trascrizionale. (17) In altri, numerose evidenze indicano che la cromatina attiva ha una struttura "aperta", in cui condizione necessaria (anche se non sufficiente) all'attività trascrizionale è la decondensazione del DNA: l'impaccamento di elementi regolativi in nucleosomi consente quindi di mediare lo stato represso, dato che sia gli istoni *linkers* (H1 e H5) che proteine strutturalmente e funzionalmente simili sono in grado di contribuire, insieme alla stessa sequenza del DNA, al posizionamento specifico degli ottameri istonici. La competizione tra fattori trascrizionali ed istoni può modulare questo tipo di controllo trascrizionale.

Per alcuni gruppi di geni l'attivazione trascrizionale risulta strettamente dipendente dalla replicazione del DNA: in cellule quiescenti le regioni regolative risultano inaccessibili ai fattori trans-agenti e solo durante la fase di sintesi del DNA, l'assemblamento sequenziale della fibra cromatinica offre a queste molecole una finestra di opportunità di accedere ai rispettivi siti. In altri casi invece si verifica una competizione dinamica: l'attivazione trascrizionale indotta da un fattore deriva non solo dalla capacità di incrementare l'efficienza con cui il meccanismo di trascrizione basale utilizza la TATA *box*, ma soprattutto dalla possibilità di disassemblare strutture cromatiniche repressive, spiazzando direttamente i nucleosomi o reclutando enzimi specifici in grado di mediare la decondensazione, attraverso modificazioni covalenti degli istoni quali l'acetilazione (che riduce la tendenza dei nucleosomi ad aggregarsi e destabilizza l'ottamero stesso) e l'ubiquitinazione (che incrementa il tasso di degradazione proteolitica delle proteine bersaglio). (18)

Gli istoni non sono gli unici componenti della cromatina con il duplice ruolo di supporto fisico e di strumento di controllo dell'espressione genica. Infatti un livello addizionale di complessità strutturale è rappresentato dalle numerose proteine non istoniche, che compongono la fibra cromatinica o le sono comunque associate. Fra di esse il gruppo più abbondante e meglio caratterizzato, denominato HMG (*High Mobility Group*), è costituito da elementi cosiddetti architetturali che, modulando la struttura cromatinica, facilitano l'ordinata

progressione di processi fondamentali del metabolismo del DNA, fra cui anche la trascrizione. In generale, si possono evidenziare diverse modalità di controllo: alcune di queste proteine interagiscono principalmente con i nucleosomi e, competendo con gli istoni H1, ne rimuovono l'effetto di repressione e influenzano positivamente il potenziale trascrizionale della cromatina. Contrariamente a questo meccanismo che non coinvolge interazioni sequenza specifiche col DNA, altri membri di questa famiglia sono essenziali nell'assemblamento di strutture tridimensionali stereospecifiche su particolari regioni regolative. Infatti anche se non tutti i promotori ed *enhancers* richiedono per la loro attività precisi requisiti di posizionamento reciproco e di fase fra i siti componenti, un insieme di queste regioni inducibili ha evoluto questo particolare meccanismo per raggiungere elevati livelli di specificità e di attivazione genica.

Recentemente sono stati descritti nuovi fattori come elementi di assemblamento della cromatina, che si collegano direttamente al processo della trascrizione mediante un rimodellamento cromatinico. Questo processo comprende qualsiasi tipo di cambiamento osservabile che avviene sulla struttura cromatinica o del mononucleosoma ed è importantissimo affinché la cromatina sia una struttura dinamica. (19) Infatti è importante che vi sia un rimodellamento dei nucleosomi a livello e accanto alle sequenze trascrizionalmente attive. Questi rimaneggiamenti sono ATP-dipendenti e avvengono per opera di un apparato di assemblaggio cromatinico costituito da un complesso detto ACF (*ATP utilizing chromatin assembly and remodeling factor*) assieme ad un *chaperon* specifico per il *core* istonico, detto CAF-1 o NAP-1 (*Chromatin/nucleosome assembly factor/protein-1*). (20, 21) La reazione di assemblaggio cromatinico mediata da questi due complessi non sembra coinvolgere né acetilazione né deacetilazione dei componenti del *core* istonico, pur avendo attività enzimatiche di questo tipo. Interagiscono sì con istoni acetilati, ma tali modificazioni post-traduzionali sembrano essere più importanti per il trasporto degli istoni e per la loro stabilità, piuttosto che per il processo intrinseco di posizionamento.

Il gruppo di proteine non-istoniche HMG, molto abbondante nei nuclei di cellule degli eucarioti superiori, è caratterizzato da un elevato contenuto in aminoacidi carichi, che determina la peculiare elevata mobilità elettroforetica su gel in acido acetico urea, da cui deriva la loro denominazione *High Mobility Group*. Sebbene siano state definite le strutture di tutti i membri componenti, la loro funzione cellulare non è ancora completamente chiarita; infatti vengono tuttora definiti, in base a caratteristiche chimico fisiche, come componenti della cromatina estraibili da nuclei con soluzioni di NaCl 0.35 M, solubili in acido perclorico al 5%, con peso molecolare inferiore a 30 KDa ed elevato contenuto in aminoacidi carichi.

(22) All'interno di questo gruppo si possono evidenziare, in base all'omologia di sequenza aminoacidica, tre famiglie distinte.

Le proteine del gruppo HMG 1 e 2, molto abbondante e conservato, sono caratterizzate dalla presenza di un specifico dominio di legame al DNA, denominati *HMG-box*, e da un dominio carbossi-terminale acido, in grado di regolare l'affinità di legame per il DNA e di interagire con altre proteine, fra cui l'ottamero istonico.(23) Questi fattori non presentano specificità di sequenza nel legare il DNA; al contrario riconoscono regioni strutturalmente atipiche (DNA cruciforme o zone di transizione tra forme B e Z) (24) e sono a loro volta in grado di modulare la conformazione topologica del DNA, visto che sono in grado di introdurre curvature, superavvolgimenti e strutture a forcina.

Il loro ruolo, che tutt'ora resta elusivo, risulta intimamente legato alla regolazione generale della struttura e funzione della cromatina; in particolare, anche se l'effetto sul controllo trascrizionale è ancora controverso, è stata ampiamente comprovata la loro capacità di competere con gli istoni H1 per il legame di nucleosomi a livello di regioni ristrette, promuovendo l'accessibilità di domini locali della cromatina coinvolti nell'espressione genica.

(25) Un'altra funzione ipotizzabile deriva dal fatto che l'*HMG-box* è condiviso anche da una famiglia di fattori trascrizionali sequenza specifici (fra cui LEF1 di cellule linfoidi (26) SRY ed altri fattori coinvolti nella determinazione e nello sviluppo sessuale (27) Dato che questo dominio si lega preferenzialmente a regioni di DNA conformazionalmente distorto, il verificarsi in vivo di situazioni transitorie che producono, per svariati motivi, un incremento di queste strutture anomale, potrebbe intrappolare in modo aspecifico questi fattori, riducendone notevolmente la concentrazione. In questo contesto le HMG 1 e 2 potrebbero fornire la soluzione biologica a questo problema di competizione differenziale, saturando preferenzialmente questi siti aspecifici, visto che la loro quantità normalmente supera di diversi ordini di grandezza quella dei fattori trascrizionali specifici. (28)

La seconda famiglia, HMG 14 e 17, è data da un gruppo di proteine presente nella maggior parte degli eucarioti superiori, è estremamente conservato. Nonostante l'elevata omologia strutturale i due membri sono compresenti in tutti i tessuti: ne deriva che probabilmente il ruolo di entrambi è fondamentale nel metabolismo cellulare.

La loro struttura primaria, caratterizzata da una distribuzione di carica nettamente asimmetrica, con un dominio centrale basico e con una regione acida al carbossi-terminale, è quasi interamente essenziale alla loro funzione di elementi architettonici della fibra cromatinica. Sono infatti in grado di legare e di riconoscere specificamente l'ottamero istonico associato al DNA, senza requisiti di sequenza nucleotidica. (29) Nonostante ciò, queste

interazioni non sono casuali, visto che sono accompagnate dalla formazione unicamente di omodimeri fra questi fattori (30); è stato inoltre dimostrato che nella cromatina sono raggruppate ed associate a specifiche regioni, dove possono influenzare il potenziale trascrizionale rimuovendo l'effetto negativo dell'istone H1 (31) svolgendo strutture cromatiniche di ordine superiore. A riprova della loro capacità di modulare la struttura cromatinica dei geni trascritti, oltre alla colocalizzazione nucleare con le regioni trascrizionalmente attive (29), vi sono numerose evidenze *in vitro* di un'azione di attivazione trascrizionale da sequenze organizzate in strutture cromatiniche e non da stampi di DNA nudo. (32)

La terza famiglia è data dal gruppo delle proteine HMGI in mammifero ed è costituita da tre membri, HMGI, HMGY ed HMGI-C. HMGI dà il nome a tutta la famiglia, ma esistono diverse nomenclature, come ad esempio HMG-I, HMG I oppure anche HMG-I(Y) che spesso creano confusione.

## 1.2 La famiglia HMGI

Questa famiglia di proteine nucleari nonistoniche è strettamente associata alla regolazione della cromatina trascrizionalmente attiva. Tre sono le proteine ad oggi descritte, due cui due, HMGI e HMGY, derivano da *splicing* alternativo dello stesso gene e differiscono per un tratto di 11 amminoacidi presente tipicamente solo in HMGI. (33, 34) Ci si riferisce a loro come HMGI(Y), dato che fin'ora non si è trovata nessuna differenza funzionale tra le due isoforme, tranne in un caso descritto nella sezione 1.2.7. (35) La terza proteina, HMGI-C, è invece codificata da un gene diverso. (36-38)

### 1.2.1 I geni *HMGI*

Il gene umano *HMGI(Y)* si situa a livello del braccio corto del cromosoma 6 (6p21), un locus coinvolto in molti riarrangiamenti cromosomici associati ad una serie di tumori benigni (vedi sezione 1.2.8) Esso è costituito da sette introni (numerati 1-7) ed otto esoni (I-VIII) (Figura 1



in basso), che occupano una regione di circa 10 Kbasi. (39) Esso codifica per due proteine, HMGI ed HMGY, che si generano mediante *splicing* alternativo.(33)

Sono stati individuati quattro diversi siti di inizio della trascrizione posizionati in prossimità dei primi esoni della 5'UTR; questi ultimi, associati ad un complesso processo di *splicing* alternativo tra i primi sei esoni, possono spiegare tutti i diversi mRNA rilevabili a livello cellulare. Inoltre, grazie ad analisi di sequenza, si sono evidenziati nei quattro potenziali promotori siti di legame per numerosi fattori trascrizionali inducibili, tra cui AP-1, AP-2, PEA3 (indotti da esteri del forbolo, siero, acido retinoico e mitogeni), CREB (indotto dal calcio ed cAMP) ed NF- $\kappa$ B (vedi sezione 1.2.6).

Ogni dominio funzionale proteico è codificato da un esone diverso:

- ognuno dei tre DBD (DNA binding domain) presenti nella proteina HMGI(Y) è collocato rispettivamente negli esoni V (in cui si trova anche il sito di inizio della traduzione), VI e VII;
- l'esone VIII, invece, codifica per i 17 amminoacidi finali della coda COOH-terminale acida;
- il V esone di 33 pb, codifica per 11 residui amminoacidici esclusivi di HMGI. (39)

Ogni esone V-VII corrisponde ad un dominio di legame al DNA formato dalla sequenza tipo BBXRGRXBB (dove B indica lisina o arginina ed X una glicina o prolina). Quest'ultima è in grado di riconoscere e legarsi al solco minore di zone di DNA ricche in AT.

Il gene murino *Hmgi(y)* è posto a livello della regione "t-complex" sul cromosoma 17, la quale contiene numerosi geni che, se mutati, possono causare letalità embrionale. (34)

Il gene umano dell'HMGI-C è posto all'interno di una regione di 170 Kbasi sul braccio lungo del cromosoma 12, caratterizzata da numerosissimi riarrangiamenti cromosomici tipici di molte neoplasie benigne di origine mesenchimale. (40) Esso è composto da cinque esoni (Figura 1 in alto) di cui

- il primo contiene tutta la regione 5' non tradotta e la sequenza che codifica per il primo dominio all'N-terminale, contenente il primo dominio basico che ritiene la capacità di legarsi al DNA (DBD);
- gli esoni II e III codificano per gli altri due domini di legame al DNA,
- mentre il IV codifica per una regione definita spaziatrice, collocata fra il terzo DBD e l'estremità C-terminale, e
- l'ultimo esone contiene l'informazione genetica per la coda COOH-terminale acida della proteina e tutto il 3' non tradotto. È d'interesse notare la presenza di otto motivi AUUUA

solitamente legati alla promozione di una efficiente deadenilazione e quindi ad un rapido decadimento del relativo mRNA. (41)

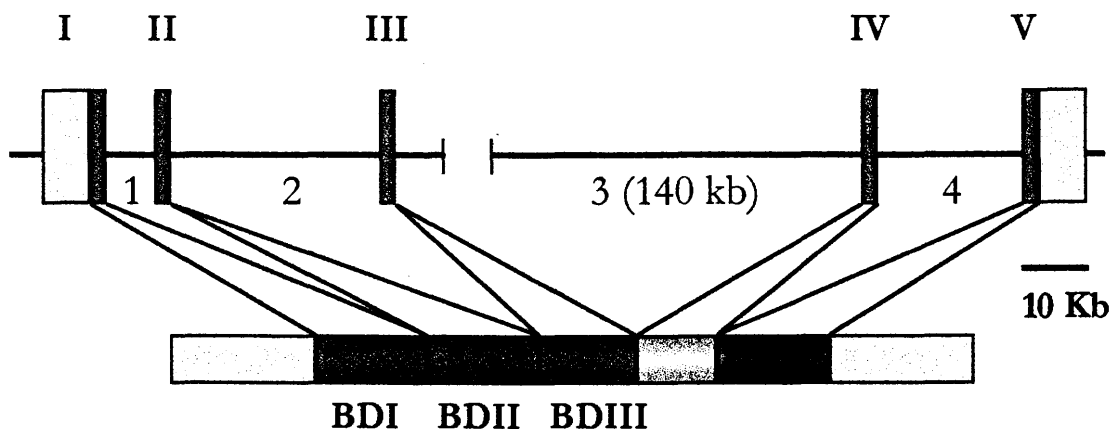
Quattro sono gli introni (1-4) di cui il terzo, che separa i tre DBD dal C-terminale, ha un'estensione ingente (circa 140 Kbase) ed è bersaglio della maggior parte dei punti di rottura nei riarrangiamenti cromosomici.

Studi di mutazioni inserzionali in topi hanno localizzato il gene murino per HMGI-C nel locus denominato *pygmy* sul cromosoma 10, (42) ed è del tutto analogo al gene umano. (37)

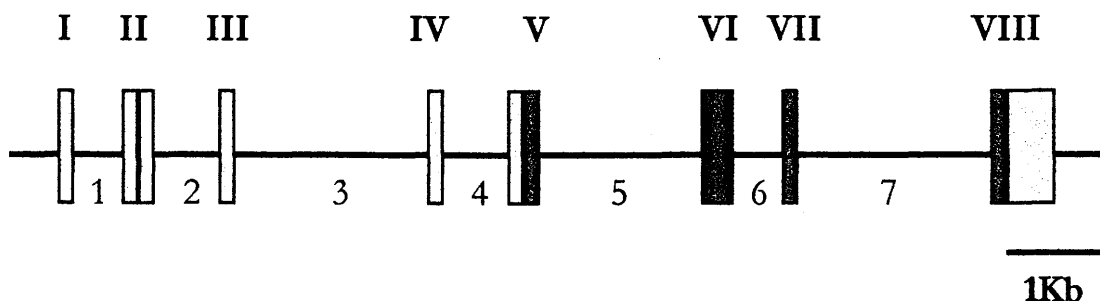
Grazie a comparazioni dirette è risultato che i geni HMGI hanno un elevato grado di conservazione tra diverse specie come il pollo, il topo, il ratto e l'uomo. Ad esempio nel gene per HMGI-C vi è identità nucleotidica tra le sequenze esoniche del topo e del uomo variano tra 82% ed il 96%.(43) Inoltre una eguale distribuzione esonica ed intronica tra le due specie suggerisce una limitazione nella divergenza evolutiva del gene dovuta a costrizioni funzionali. (37).

Nonostante la differenza consistente nelle dimensioni, i due geni *HMGI(Y)* ed *HMGI-C* presentano una struttura complessiva omologa (Figura 1), specialmente escludendo la presenza di siti multipli di inizio trascrizione al 5' del gene per *HMGI(Y)* ed il quarto esone caratteristico del gene per *HMGI-C*, tanto da ipotizzare che derivino da un'unità ancestrale per duplicazione e successivo accumulo di varie mutazioni puntiformi. Inoltre, se si considera la particolare organizzazione un esone-un dominio funzionale, si può allargare l'ipotesi della duplicazione e successiva divergenza agli stessi DBD, supportando così l'idea di uno "scivolamento evolutivo" anche verso altre strutture geniche (*exon-shuffling*).

## STRUTTURA GENE *HMGI-C*



## STRUTTURA GENE *HMGI(Y)*



**Figura 1:**

Rappresentazione schematica dell'organizzazione strutturale dei geni *HMGI-C* e *HMGI(Y)*. I rettangoli in grigio sono sequenze di DNA non tradotte, mentre quelli in blu rappresentano sequenze codificanti.

Per il gene *HMGI-C* viene rappresentato anche il trascritto, dove in grigio sono segnate le regioni 5' (a sinistra) e 3' (a destra) non tradotte, mentre in blu sono segnati i tre domini di legame al DNA (BDI-III), in giallo la regione spaziatrice ed in rosso la coda C-terminale acida.

## 1.2.2 Le proteine HMGI

L'estrazione di nuclei con acido perclorico al 5% porta ad una solubilizzazione selettiva dell'istone H1 e di un gruppo di proteine denominate High Mobility Group (denominazione derivante dalla loro tipica migrazione in elettroforesi a basso pH), tra le quali si distingue una sottofamiglia di tre elementi:

HMGI (107 amminoacidi; PM ~ 12.0 kDa) (44)

HMGY (96 amminoacidi; PM ~10.6 kDa) (33)

HMGI-C (109 amminoacidi; PM ~12.0 kDa) (36, 38)

collettivamente denominati HMGI. (22) L'omologia di sequenza amminoacidica tra le due isoforme HMGI ed HMGY e l'HMGI-C si aggira attorno al 50%, rendendo conto della loro appartenenza ad un'unica famiglia. (Figura 2)

### Struttura

Tutte e tre le proteine sono caratterizzate da una distinzione in domini funzionali (Figura 2):

- un dominio N-terminale a funzione sconosciuta e composizione amminoacidica poco conservata;
- tre domini basici altamente conservati, specifici per il legame con il DNA ricco in AT, denominati "AT-hooks";
- un dominio C-terminale ricco in amminoacidi acidi, caratterizzati dalla presenza di consensus di fosforilazione per la casein chinasi II;
- l'HMGI-C possiede un ulteriore tratto amminoacidico codificato dal IV esone, che funge da ponte tra la parte della proteina in grado di contattare il DNA e la coda C-terminale.

Le proteine HMGI sono estremamente stabili, con tempi di dimezzamento calcolati che vanno da 6h fino a 60 a seconda del tipo cellulare e delle situazioni. (45-47)

**Figura 2**

**Confronto della struttura primaria delle proteine della famiglia HMGI provenienti da diversi organismi.**

---

hHMGI	:	1	MSESSKSSQPLASKQEKDGT-----EKRGRGRPRKQPPVSPGTALVGSQKEPSEVPTP	54
mHMGI	:	1	MSESGSKSSQPLASKQEKDGT-----EKRGRGRPRKQPP??????????KEPSEVPTP	54
hHMGY	:	1	MSESSKSSQPLASKQEKSGT-----EKRGRGRPRKQPP-----KEPSEVPTP	43
mHMGY	:	1	MSESGSKSSQPLASKQEKSGT-----EKRGRGRPRKQPP-----KEPSEVPTP	43
hHMGI-C	:	1	MSARGEGAGQPSTSAQGQPAAPA-----POKRGRGRPRKQOQ-----EPTGEPSP	45
mHMGI-C	:	1	MSARGEGAGQPSTSAQGQPAAPV-----POKRGRGRPRKQOQ-----EPTCEPSP	45
gHMGI-C	:	1	MSAQGEGPGQSSTAAPQPAAAE-----POKRGRGRPRKQOQ-----EPTGEPSP	45
rHMGI-C	:	1	MSARGEGAGQPSTSAQGQPSTSAQGQPATPAPOKRGRGRPRKQOQ-----EPTCEPSP	53
hHMGI	:	55	KRPRGRPKGSKNKGAAKT--RKTTPGPKPRGRPKK-----LEK----EEEGISQESSEEEQ	107
mHMGI	:	55	KRPRGRPKGSKNKGAAKT--RKTTPGPKPRGRPKK-----LEK----EEEGISQESSEEEQ	107
hHMGY	:	44	KRPRGRPKGSKNKGAAKT--RKTTPGPKPRGRPKK-----LEK----EEEGISQESSEEEQ	96
mHMGY	:	44	KRPRGRPKGSKNKGAAKT--RKTTPGPKPRGRPKK-----LEK----EEEGISQESSEEEQ	96
hHMGI-C	:	46	KRPRGRPKGSKNKSPSKAAQKKAETGEKRPRGRPRKWPQQVVQKKPAQEETEETSSQESAEED	109
mHMGI-C	:	46	KRPRGRPKGSKNKSPSKAAQKKAETIGEKRPRGRPRKWPQQVVQKKPAQ-ETEETSSQESAEED	108
gHMGI-C	:	46	KRPRGRPKGSKNKSPSKAAQKKAETGEKRPRGRPRKWPQQVVQKKPAQEETEETSSQESAEED	109
rHMGI-C	:	54	KRPRGRPKGSKNKSPSKAAQKKAETIGEKRPRGRPRKWPQQIVQK-PAQ-ETEETSSQESAEED	115

---

h, umano; m, murino; g, di pollo; r, di ratto; Le sequenze rispettive di DNA sono depositate nelle banche dati Genbank/EMBL sotto i n° di accesso AF058287 (HMGI-C di pollo), L41617-41622 (HMGI-C di topo), U89695 (HMGI-C di ratto), J04179 (HMGI(Y) di topo), XXX (HMGI(Y) umane. In grigio sono evidenziati i domini AT-hook, conservati quasi al 100% tra le 4 diverse specie e le tre proteine.

### 1.2.3. Il dominio di legame tipico delle proteine HMGI: “AT-hook”

I tre domini di legame al DNA delle proteine HMGI riconoscono una grande varietà di tratti ricchi in AT, lunghi da quattro fino ad otto paia di basi. Nonostante questa blanda specificità, ad esempio HMGI lega il DNA con affinità nanomolare e studi di interferenza della metilazione hanno dimostrato che il sito di legame è posto a livello del solco minore di questi tratti di DNA. (48) Siti ad alta affinità per le proteine HMGI sono i cosiddetti siti multivalenti, costituiti da due o tre tratti ricchi in AT appropriatamente spazati, in modo tale da permettere un vero e proprio “attracco” di più di un dominio di legame al DNA contemporaneamente. (49) Questo permette di ipotizzare due modi di legarsi al DNA: un modello prevede una certa flessibilità della proteina, per cui due domini proteici possono contattare siti con spaziatura differente e con posizioni diverse rispetto all’asse dell’elica, mentre l’altro modello prevede l’uso alternato di due coppie di domini proteici (I+II, II+III, I+III) a seconda della spaziatura e posizione rispetto all’asse dell’elica del DNA. In entrambi i casi però una spaziatura maggiore di 10 pb riduce l’affinità di legame, suggerendo che questa distanza sia al disopra dell’apertura massima dei tre domini di legame al DNA delle proteine HMGI. (49)

A tale riguardo vale la pena annoverare un fatto descritto ancora agli albori della scoperta delle proteine HMGI. Infatti la proteina HMGI(Y) fu allora descritta come una proteina che lega il DNA satellite. Si evidenziò che tre siti di legame in questo DNA altamente ripetitivo potevano collocarsi in stretta vicinanza quando si trovava avvolto attorno ad un nucleosoma. (50) Se si estendono le proprietà di legame multivalenti descritte qui ad un sistema tridimensionale del nucleosoma, allora tre tratti di AT del DNA avvolto attorno al nucleo istonico potrebbero essere legati da una sola proteina HMGI. (49)

Studi di *binding* ed NMR dimostrano che la sequenza amminoacidica minima per formare questo dominio è data dal peptide PRGRP, inserito in un contesto di residui basici, capaci di prendere contatto con i gruppi fosfato della doppia elica, contribuendo così alla stabilità del complesso. Le interazioni centrali col DNA coinvolgono il motivo R-G-R, che segue esattamente la curvatura del DNA. Distinguendo tra vari tipi di siti di DNA è interessante notare che se il sito è un AAATTT il tripeptide può legare e/o scivolare, mentre con un sito AATT rimane fermo al centro. Le catene laterali delle Arginine vanno in direzioni opposte l’una dall’altra e lo scheletro peptidico assieme alla Glicina poggiano direttamente nel fondo del solco minore. Modelli molecolari propongono legami idrogeno tra i protoni dei gruppi

guanidinio delle Arginine e le basi e possibilmente anche con gli ossigeni dei fosfati del DNA. Tale nozione è supportata dal fatto che la sostituzione delle Arginine con Lisine fa sì che il peptide non interagisca più allo stesso modo. (51) Il peptide TPKRPRGRPCK sottoposto ad analisi computerizzate di simulazione e *model building* utilizzando programmi di minimizzazione energetica suggeriscono una struttura secondaria planare con una curvatura intrinseca denominata *Asx bend* o *hook* data la sporgenza dal piano delle catene laterali delle Arginine o Lisine, che sono componenti essenziali di questo tipo di dominio proteico. Da qui nasce il nome *AT-hook*. (52)

La struttura planare di questa unità modulare presenta molte analogie con quelle di farmaci antitumorali, quali la Netropsina e la Distamicina A e dell'intercalante Hoechst 33258. Anch'essi legano nel solco minore di sequenze di DNA ricche in AT e sono in grado di competere per tale legame con le proteine HMGI, motivo per cui probabilmente esplicano il loro potere antitumorale. La netropsina è stata co-cristallizzata con un dodecanucleotide a doppio filamento con struttura di B-DNA con la seguente sequenza: C-G-C-G-A-A-T-T-C-G-C-G. Molto delle proprietà di legame dell'*AT-hook* si deducono anche dal cristallo della netropsina, come ad esempio il fatto che sequenze di A alternate con T non sono dei siti di legame ottimi, dato che producono un elica di DNA-B alternato. (52, 53)

In assenza di DNA, lo spettro NMR di un frammento della proteina contenente il secondo (DBD2) ed il terzo (DBD3) *AT-hook* dell'HMGI è indicativo di *un random coil*. Solo dopo l'interazione col DNA i due DBD si strutturano ed adottano una conformazione ben definita nel solco minore. Le variazioni che questo legame induce alla struttura del DNA sono minime. Si è riscontrata una differenza notevole di affinità con il DNA tra il secondo e terzo *AT-hook*: il secondo ha un'affinità maggiore rispetto al terzo, dovuto alla presenza di sei amminoacidi carbossi-terminali rispetto al centro del DBD, che prendono contatto con lo scheletro zucchero-fostato da entrambi gli angoli del solco minore. Questo fatto fa sì che la superficie di contatto col DNA aumenti, contribuendo ad una maggiore stabilità. (53) Questo è l'aspetto che distingue i cosiddetti motivi HMGI di tipo I (DBD2) e di tipo II (DBD3). Esperimenti di footprinting proteico suggeriscono che anche il primo *AT-hook* (DBD1) appartenga al motivo HMGI di tipo II. Moduli addizionali nella proteina intera, che spesso non vengono considerati nell'analisi degli *AT-hook*, sono determinanti cruciali per l'affinità e la specificità, dato che le affinità del DBD2 e anche di DBD3 nella proteina intera, pur avendo lo stesso motivo di legame, sono maggiori di quattro e due ordini di grandezza in confronto ai peptidi, rispettivamente. (54)

Un confronto di molteplici proteine che possiedono domini del tipo *AT-hook*, descritti nel capitolo successivo, ha portato alla definizione di un *AT-hook* tipo III. In questi motivi vi è una posizione basica altamente conservata (lisina) e la presenza di alcuni residui amminoacidici polari a valle del motivo centrale che sono simili al *AT-hook* di tipo I. Mentre, come nel II tipo, il III tipo possiede una lisina predominante in seconda posizione a valle di GRP. Inoltre una lisina conservata a 4 residui in direzione carbossi-terminale rispetto al GRP centrale è caratteristico di questo tipo di *AT-hook*. Si prevede che legni con una maggiore affinità per il DNA che non il tipo II, ma meno del tipo I, che ha una superficie polare più estesa. (54)

Analisi di legame sono state effettuate anche con la proteina HMGI intera di *Chironomus tentans*, altamente omologa a quella umana, mediante la tecnica del "protein-footprinting". Con questa tecnica si analizzano le zone di protezione che il DNA crea sulla proteina marcata, protezione dal taglio chimico con radicali idrossilici. Incubando la proteina con differenti sequenze di DNA, rappresentanti diversi siti di legame delle HMGI, prima del taglio chimico, si sono ottenuti diversi tipi di protezione. Il legame di cHMGI ad una sequenza poly(dA-dT)·(dA-dT), oppure ad una sequenza cruciforme oppure al sito di legame tipico sull'*enhancer* dell'INF  $\beta$  (PRDII/NRDI), ha coinvolto, in tutti e tre i casi, il primo ed il secondo motivo *AT-hook*, lasciando invece la sequenza tra i due motivi leganti e tutto il carbossi-terminale esposti al taglio. Usando un oligonucleotide poly(dA-dT)·(dA-dT) più lungo di 30 pb, anche il terzo motivo *AT-hook* è stato protetto, suggerendo che il terzo motivo ha minore affinità di legame del primo e che l'intera proteina occupa più di 30 pb. (53)

Questi risultati sono simili anche per la HMGI umana. I dati di *footprinting* sull'elemento PRDII qui utilizzato (34 pb) indicano che tutti e tre i DBD contattano il DNA e che vi sono contatti più estesi, che coinvolgono probabilmente specificamente alcune basi o una specifica struttura al di fuori dei tratti ricchi in AT. Al contrario, nel complesso col DNA cruciforme il terzo DBD rimane scoperto, probabilmente perché questo DNA non riesce ad assestare tutta la proteina oppure perché la proteina si struttura in modo particolare quando contatta DNA cruciforme. (53) Dato che è stato dimostrato che a livello del DNA 4H (vedi sezione 1.2.5.1) si possono sistemare fino a due proteine HMGI, la seconda ipotesi sembra più plausibile. In questo caso però il secondo ed il terzo *AT-hook* rimanevano protetti. (56) Un fatto interessante invece è che nell'interazione tra HMGI di *Chironomus* e il DNA cruciforme rimane protetto un motivo speciale, denominato motivo PKRKK, che non è caratteristico solo per le proteine della famiglia HMGI, ma anche delle HMG 14 e 17, ed HMG 1 e 2. Peptidi contenenti questo



motivo interagiscono specificamente col solco minore del DNA. Dato che tutte queste famiglie legano preferenzialmente il DNA cruciforme questo potrebbe essere un motivo proteico che gioca un ruolo nel riconoscimento specifico di questo tipo di DNA. (53)

Istoni e proteine HMG sono i componenti proteici più abbondanti della cromatina. Nella struttura degli istoni e delle proteine HMG 1 e 2 si possono osservare domini ben caratterizzati e strutturati in posizioni centrali. Un'ulteriore caratteristica comune sono invece delle regioni non definite, cosiddetti "terminali", o "code". Invece i due gruppi HMGI 14 e 17 e HMGI sono composti principalmente da regioni flessibili senza struttura definita. È il legame col DNA che induce un assetto ordinato nello spazio delle regioni coinvolte nell'interazione col DNA e quindi della proteina intera, mentre la stessa interazione a sua volta può indurre un cambiamento conformazionale del DNA. È veramente un adattamento molecolare incredibile. (53)

Studi NMR sulla proteina intera dimostrano che i due DBDs di HMGI legano preferenzialmente a siti sequenziali ricchi in AT, piuttosto che siti su frammenti di DNA distinti. Questi dati sono concordi con quelli ottenuti dallo studio di interazione proteina-DNA che dimostrano un legame multivalente di HMGI all'elemento PRDII del promotore dell'interferone  $\beta$  umano (INF $\beta$ ). (49) (Vedi sezione 1.2.5.2). Inoltre mutazioni nel tratto AT di 5 pb dell'elemento PRDII hanno effetto maggiore sul legame di HMGI che non le mutazioni nel tratto AT vicino di 4 pb dell'elemento NRDI, suggerendo che PRDII contiene il sito di legame per HMGI a più alta affinità.

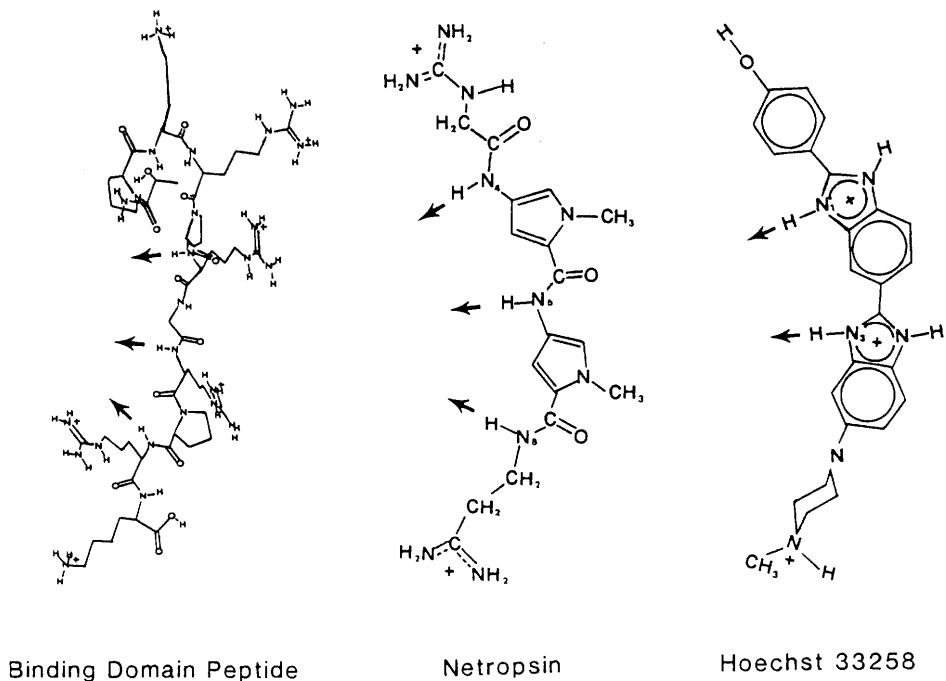
È probabile che DBD2 leghi il sito ad alta affinità in PRDII, mentre DBD3 aumenta l'affinità generale del complesso legandosi al sito più corto ed adiacente di NRDI. Analogamente i due siti in PRDIV sono legati dal 1° e 2° dominio DBD della proteina HMGI. (57) Sempre nel modello del promotore dell'INF $\beta$  viene descritto che due molecole di HMGI(Y) legano l'enhancer in modo cooperativo, usando sempre due DBDs ciascuna, per riconoscere le regioni ricche in AT poste in tandem; un'alta cooperatività intermolecolare è ottenuta se legano entrambe le proteine sullo stesso lato del DNA. Si dimostra ulteriormente che solo il DBD2, quello centrale, è responsabile del legame specifico al DNA, ma con un ulteriore AT-hook della stessa proteina che si lega ad un tratto di AT adiacente si ottiene cooperatività intramolecolare. (57)

È stato osservato che l'enhancer dell' INF $\beta$  esibisce una piccola curvatura intrinseca di circa 20° verso il solco minore, che è ripristinata dal legame con HMGI. Tali curvature infatti frequentemente risultano in una compressione dello spazio nel solco minore, mentre una

curvatura in direzione del solco maggiore provocherebbe un'espansione del solco minore. L'ampiezza del solco minore (7-7,5 Å) nei complessi degli HMGI DBDs 2 e 3 col DNA è comparabile a quella del DNA B (~6 Å). Quindi il ruolo architetturale principale delle proteine HMGI probabilmente è quello di revertere e prevenire distorsioni intrinseche della conformazione del DNA, come curvature e piegature, legandosi al solco minore, e contemporaneamente facilita il riconoscimento specifico dell'opposto solco maggiore da parte di altri fattori di trascrizione come ad esempio NF-κB. (54)

Al contrario delle proteine contenenti una HMG-box, come i fattori della famiglia HMG 1 e 2, che inducono curvature molto forti al DNA, gli AT-hook, presi singolarmente, hanno un effetto conformazionale modesto sul DNA. Piuttosto attraverso legami multivalenti di più di un AT-hook e di più di una proteina HMGI con siti multipli all'interno di molti promotori bersaglio si possono indurre cambiamenti estesi nell'architettura chromatinica, potenzialmente indispensabile per l'accesso e l'interazione di altri fattori trascrizionali col DNA. (58)

Figura 3



## Motivi HMGI-simili in altre proteine

Ricerche in banche dati non ridondanti di proteine (NCBI) hanno rivelato l'esistenza di motivi HMGI di tipo I, II ed uno nuovo definito III, in diverse proteine di cui sono riportate nella Tabella I i nomi, l'organismo da cui proviene, ulteriori domini presenti e le funzioni proposte. Moltissime evidenze sperimentali sulle proteine HMGI hanno dimostrato che queste proteine sono capaci di revertire curvature intrinseche del DNA, presumibilmente per permettere un legame efficace di fattori di trascrizione. (59) Analogamente proteine sintetiche con molti motivi AT-hook hanno un effetto fortissimo sull'organizzazione cromatinica quando sono aggiunti alla cromatina spermatica. (60) Proteine relate con tre o quattro motivi AT-hook si sono trovate in *Caenorhabditis elegans* e *Dictyostelium discoideum*, suggerendo che fattori architetturali simili alle HMGI erano già funzionali precocemente nell'evoluzione. Le piante tipicamente possiedono gruppi di proteine che possiedono un dominio H1 globulare all'N-terminale seguito da alcuni AT-hooks. La combinazione di due domini diversi suggerisce che tali proteine possano operare come una classe speciale di istoni *linker* coinvolti nell'organizzazione cromatinica. Tale ipotesi è avvalorata dal ritrovamento di motivi AT-hook all'interno di proteine istone-simili anche in *C. elegans* che potrebbero formare strutture atipiche simili a nucleosomi coinvolti nella regolazione trascrizionale o architetturale. (55)

Singoli o multipli AT-hooks si trovano anche in molte proteine con molti domini che si associano alla cromatina, come ad esempio HRX, TAFII250, ASH1, SWI2, ISWI, ENBP1, M33, doom/Mod(Mdg4) e *tramtrack*. (Referenze all'interno della ref. 55, e Tabella I) Alcune sono coinvolte nella organizzazione della cromatina durante stadi specifici. HRX (ALL-1) ed ASH1 sono membri del gruppo *trithorax* di geni regolativi coinvolti nella decondensazione cromatinica. HRX tra l'altro viene traslocato in una leucemia acuta linfoblastoide umana, che coinvolge aberrazioni dei geni codificanti per HRX ed AF17 ed HRX con AF10. Ne risultano proteine chimeriche, nelle quali i tre domini HMGI di tipo I di HRX si uniscono o rimpiazzano quelli di tipo II appartenenti alle proteine AF17 o AF10. Nessuno ha ancora stabilito l'effetto che lo scambio dei domini HMGI può provocare, ma si potrebbe ipotizzare che il passaggio da un dominio di tipo II ad uno di tipo I porti alla perdita della funzione corretta, (54) come verrà descritto nei capitoli successivi. SWI2 ed ISWI sono fattori di rimodellamento della cromatina ATP-dipendenti, mentre TAFII250 è un fattore basale della trascrizione che possiede anche un'attività istone-acetiltrasferasica. Le proteine Mdg4 e *tramtrack* che possiedono un dominio POZ operano invece da regolatori dell'isolamento

trascrizionale e nella repressione genica, rispettivamente. La presenza del motivo AT-hook in queste proteine indica che questi fattori regolatori dell'architettura cromatinica probabilmente usano l'informazione del DNA in forma di tratti ricchi in AT, per trovare regioni specifiche sui cromosomi. Ad esempio, la ricchezza di tratti ricchi in AT nelle SAR ed un loro presunto coinvolgimento nell'attivazione trascrizionale renderebbe valida l'ipotesi che gli AT-hook servono per legare queste proteine a regioni cromosomiche specifiche. In effetti molte proteine elencate nella Tabella I sono proteine associate alle SAR con in più domini proteici che funzionano ad esempio come metilasi, oppure omeodomini e *zinc-fingers*. Molto probabilmente in questi casi il motivo AT-hook serve o come attacco specifico al DNA per poter compiere funzioni come la metilazione, oppure come contatto peptidico addizionale nel solco minore accanto all'interazione del dominio di legame al DNA più grosso e sul solco maggiore. L'AT-hook agirebbe come un cofattore incorporato che sottilmente altera l'affinità e la specificità del dominio più importante, riconoscendo siti nelle vicinanze del sito di legame di DNA principale. Un simile modo di funzionare si conosce già dagli *omeodomini* e dalla *GRIP box* dei recettori ormonali nucleari recentemente descritta.

Al di fuori del mondo eucariotico il motivo AT-hook non è frequente ed in archebatteri è inesistente, mentre in alcuni batteri si possono trovare a livello delle trasposasi. La presenza quasi esclusiva in eucarioti suggerisce che questo motivo è stato selezionato attivamente solo dopo l'evoluzione della cromatina eucariotica. Inoltre si tratta di un elemento evolutivamente mobile, in quanto, in analogia alle traslocazioni che lo coinvolgono in molte neoplasie umane (vedi sezione 1.2.8), confrontando a coppie geni ortologi o paraloghi si possono spesso notare come unica differenza la presenza o assenza di un AT-hook. Per esempio il gene HRX umano ne possiede alcuni, mentre il suo ortologo trithorax in *D. melanogaster* non ce li ha. Evidenze analoghe si hanno anche confrontando i geni TAFII250, polycomb e SWI2 tra diverse specie tassonomiche. Chiaramente, essendo ciascun dominio di legame al DNA delle proteine HMGI codificato da un esone distinto, alla base di questo meccanismo evolutivo vi è sicuramente l'*exon-shuffling*.

In conclusione il dominio AT-hook è un piccolo motivo proteico che interagisce col DNA in molte proteine che hanno funzioni nel nucleo. Tale dominio serve ad ancorare molte proteine che modificano la cromatina nel solco minore del DNA, permettendo così il loro spostamento verso siti ricchi in AT della cromatina. Viene inoltre proposto che quelle proteine che oltre all'AT-hook possiedono un dominio H1 globulare funzionano come istoni non-canonicali in un'organizzazione cromatinica atipica nella regolazione della trascrizione. Sembra quindi in

---

generale che l'AT-hook serve come motivo proteico ausiliario cooperando con altre attività di legame al DNA oppure facilitando cambiamenti nella struttura del DNA. (55)

**Tabella I****Proteine contenenti un motivo AT-hook**

<b>Tabella I.</b>					
<b>Nome della proteina</b>	<b>Specie n° gi della proteina</b>	<b>n</b>	<b>Domini accessori</b>	<b>Funzioni proposte</b>	<b>ref.</b>
AT1	Os_1362172	15	domini globulare dell'istone H1		55
D1	Dm_117239	10		lega il DNA satellite D1 ricco in AT.	55
pabf	Nt_555655	8	2 domini globulari dell'istone H1	Regolatore trascrizionale tessuto-specifico.	55
ENBP1	Vs_1360637	6	dominio JOR (dominio "β-helix" modificato) + dominio ricco in Cisteine	Proteina che lega il DNA a livello di elementi ricchi in AT del gene ENOD.	55
ORF T05A7.4	Ce_1055146	4			55
AAC-rich mRNA	Dd_112945	4			55
HMG-Y-related A	Gm_1708261	4	dominio globulare dell'istone H1		55
HMG-Y-related B	Gm_1708262	4	dominio globulare dell'istone H1		55
PF1	Os_453692	4	dominio globulare dell'istone H1	lega i tratti ricchi in AT presenti nel promotore del fitocromo A3 del riso.	55
HMG-I(Y)	At_1808590	4	dominio globulare dell'istone H1		55
HMG-I(Y)	Ca_2129885	4	dominio globulare dell'istone H1		55
CarD	Mx_1022328	4	domino C-terminale globulare simile ad altri domini batterici	Regolatore trascrizionale di geni che rispondono alla luce ed al digiuno.	55
Proteina omeotica relata a patogenesi (PRHP)	Pc_1346791	4	PHD finger + dominio omeo	Lega specificamente "elicitor-RE" in risposta a patogeni del fungo.	55
HMG-I(C)	Hs_1708263	3		Fattore cromatinico architetturale, cofattore per diverse proteine che legano il DNA	36, 37, 55
HMG-I(Y)	Hs_123377	3		Fattore cromatinico architetturale, cofattore per diverse proteine che legano il DNA.	33, 34, 55
HMG-I(Y)	Hs_123383	3		Fattore cromatinico architetturale, cofattore per diverse proteine che legano il DNA.	33, 34, 55

**Tabella I.bis Continuazione**

ALL-1	Hs_1490271	3	dominio SET + Zn finger tipo metilasi	Ortologo di Trithorax, regolatore positivo della decondensazione cromatinica ed espressione dei geni del cluster <i>Hox</i> .	55, 172
F1N21.16	At_2760331	3	Transposasi della famiglia Mudra con C2HC finger ed un HTH		55
ORFT18B16.30	At_2828281	3	dominio ATPasico (AP)		55
Sum1p	Sc_1711597	2		Coinvolto nella regolazione trascrizionale del silenziamento attraverso i geni Sir; mutazioni dominanti di questo gene fanno perdere la dipendenza dai geni Sir per il silenziamento.	55
Swi2p/Snf2p	Sc_134589	2	elicasi simile alla superfamiglia 2 Swi/Snf + bromodominio	Componente ATPasico del complesso SWI di rimodellamento cromatinico.	55
Lin-15B	Ce_1078881	2	Trasposone modificato tipo Hermit	Regolatore negativo della via di trasduzione del segnale ras/Tirosinachinasi dipendente nello sviluppo vuvale.	55
C11G6.1	Ce-1229035	2	Prodos/ZK1320.7-type histone fold protein	Nuova proteina cromosomiale.	55
TAFII250	Dm_1335892	2	2 bromodomini	Acetiltransferasi istonica, fattore trascrizionale basale dell'RNAPolimerasi II. (Alcuni ceppi del moscerino hanno solo 1 AT-hook).	55
ASH1	Dm_1335892	2	PHD fingers, dominio SET e BAM	Membro del gruppo Trithorax di proteine cromatiniche, funzione proposta nella regolazione positiva della decondensazione della cromatina.	55
zhx-1	Mm_2137401	2	2 domini C <sub>2</sub> H <sub>2</sub> Zn finger + 5 domini omeo	Regolatore trascrizionale	55
Proteina nucleare SA-1	Mm_2204230	2	dominio SA1/2-rec11	Proteina nucleare espressa in organi ematopoietici	55
ORF KIAA0437	Hs_2879925	2			
TTF-I interacting peptide 5	Hs_2183083	2	Bromodominio + dominio TAM di legame al DNA	Forma un complesso con il fattore di terminazione I della RNA polikmerasi I; proposta funzione architetturale.	51
PD1	At_2213536	2	dominio globulare PD1 trovato anche in proteine arcaiche MTH1230, AF0141	Lega il DNA	
SC1	Hs_809489	1	dominio FHA	Regolatore della progressione del ciclo cellulare.	52
X box-binding regulatory factor	Hs_1350587	1	dominio RFX	Regolatore dello sviluppo di cellule B.	53

**Tabella I**  
**Proteine contenenti un motivo AT-hook**

<b>Tabella I.</b>					
<b>Nome della proteina</b>	<b>Specie n° gi della proteina</b>	<b>n</b>	<b>Domini accessori</b>	<b>Funzioni proposte</b>	<b>ref.</b>
AT1	Os_1362172	15	domini globulare dell'istone H1		55
D1	Dm_117239	10		lega il DNA satellite D1 ricco in AT.	55
pabf	Nt_555655	8	2 domini globulari dell'istone H1	Regolatore trascrizionale tessuto-specifico.	55
ENBP1	Vs_1360637	6	dominio JOR (dominio "β-helix" modificato) + dominio ricco in Cisteine	Proteina che lega il DNA a livello di elementi ricchi in AT del gene ENOD.	55
ORF T05A7.4	Ce_1055146	4			55
AAC-rich mRNA	Dd_112945	4			55
HMG-Y-related A	Gm_1708261	4	dominio globulare dell'istone H1		55
HMG-Y-related B	Gm_1708262	4	dominio globulare dell'istone H1		55
PF1	Os_453692	4	dominio globulare dell'istone H1	lega i tratti ricchi in AT presenti nel promotore del fitocromo A3 del riso.	55
HMG-I(Y)	At_1808590	4	dominio globulare dell'istone H1		55
HMG-I(Y)	Ca_2129885	4	dominio globulare dell'istone H1		55
CarD	Mx_1022328	4	domino C-terminale globulare simile ad altri domini batterici	Regolatore trascrizionale di geni che rispondono alla luce ed al digiuno.	55
Proteina omeotica relata a patogenesi (PRHP)	Pc_1346791	4	PHD finger + dominio omeo	Lega specificamente "elicitor-RE" in risposta a patogeni del fungo.	55
HMG-I(C)	Hs_1708263	3		Fattore cromatinico architetturale, cofattore per diverse proteine che legano il DNA	36, 37, 55
HMG-I(Y)	Hs_123377	3		Fattore cromatinico architetturale, cofattore per diverse proteine che legano il DNA.	33, 34, 55
HMG-I(Y)	Hs_123383	3		Fattore cromatinico architetturale, cofattore per diverse proteine che legano il DNA.	33, 34, 55



**Tabella I.bis Continuazione**

ALL-1	Hs_1490271	3	dominio SET + Zn finger tipo metilasi	Ortologo di Trithorax, regolatore positivo della decondensazione cromatinica ed espressione dei geni del cluster <i>Hox</i> .	55, 172
F1N21.16	At_2760331	3	Transposasi della famiglia Mudra con C2HC finger ed un HTH		55
ORFT18B16.30	At_2828281	3	dominio ATPasico (AP)		55
Sum1p	Sc_1711597	2		Coinvolto nella regolazione trascrizionale del silenziamento attraverso i geni Sir; mutazioni dominanti di questo gene fanno perdere la dipendenza dai geni Sir per il silenziamento.	55
Swi2p/Snf2p	Sc_134589	2	elicasi simile alla superfamiglia 2 Swi/Snf + bromodominio	Componente ATPasico del complesso SWI di rimodellamento cromatinico.	55
Lin-15B	Ce_1078881	2	Trasposone modificato tipo Hermit	Regolatore negativo della via di trasduzione del segnale ras/Tirosinachinasi dipendente nello sviluppo vuvale.	55
C11G6.1	Ce-1229035	2	Prodos/ZK1320.7-type histone fold protein	Nuova proteina cromosomiale.	55
TAFII250	Dm_1335892	2	2 bromodomini	Acetiltransferasi istonica, fattore trascrizionale basale dell'RNAPolimerasi II. (Alcuni ceppi del moscerino hanno solo 1 AT-hook).	55
ASH1	Dm_1335892	2	PHD fingers, dominio SET e BAM	Membro del gruppo Trithorax di proteine cromatiniche, funzione proposta nella regolazione positiva della decondensazione della cromatina.	55
zhx-1	Mm_2137401	2	2 domini C <sub>2</sub> H <sub>2</sub> Zn finger + 5 domini omeo	Regolatore trascrizionale	55
Proteina nucleare SA-1	Mm_2204230	2	dominio SA1/2-rec11	Proteina nucleare espressa in organi ematopoietici	55
ORF KIAA0437	Hs_2879925	2			
TTF-I interacting peptide 5	Hs_2183083	2	Bromodominio + dominio TAM di legame al DNA	Forma un complesso con il fattore di terminazione I della RNA polikmerasi I; proposta funzione architeturale.	51
PD1	At_2213536	2	dominio globulare PD1 trovato anche in proteine arcaiche MTH1230, AF0141	Lega il DNA	
SC1	Hs_809489	1	dominio FHA	Regolatore della progressione del ciclo cellulare.	52
X box-binding regulatory factor	Hs_1350587	1	dominio RFX	Regolatore dello sviluppo di cellule B.	53

**Tabella I.ter Continuazione**

LIM/Homeodomain protein LH-2	Hs_423900	1	dominio LIM e dominio omeo	Regolatore dello sviluppo neurale e linfoide nell'uomo.	55
Retinoblastoma-binding protein 1, RBP1	Hs_1710030	1	Chromodomain + dominio JOR ( $\beta$ -helix domain modificato) + Bright DNA binding domain	Proteina fosforilata che lega la proteina del Retinoblastoma; proposta funzione di fattore architetturale.	
ESX (epithelial-restricted with serine box)	Hs_1754538	1	dominio ETS	Fattore trascrizionale, regola l'espressione del recettore II del TGF $\beta$ .	55
Autoantigene DFS	Hs_1945445	1		Autoantigene nella dermatite atopica	
Putative transcription factor CR53	Hs_2407245	1	domini SCAN + Zn fingers (?)	Regolatore trascrizionale	
ORF H0601	Hs_2662163	1	dominio POZ + C <sub>2</sub> H <sub>2</sub> Zn fingers		
ORF R29381	Hs_2995577	1	dominio metiltransferasico unico, trovato anche in <i>C. elegans</i>		
Peregrin	Hs_1705500	1	bromodominio + PHD fingers * regione ricca in Cisteine	Regolatore trascrizionale; proposto ruolo come fattore architetturale.	55
Methyl-CpG-binding protein 2	Hs_420279	1	dominio TAM di legame al DNA	L'ortologo aviari è stato dimostrato essere una proteina che lega le SAR.	55
AF-10	Hs_1703190	1	2 PHD fingers	Proposta proteina cromosomiale, forma proteine di fusione con gli AT-hooks della proteina HRX (ALL-1) in leucemie.	55
AF-17	Hs_2134713	1	2 PHD fingers	Proposta proteina cromosomiale, forma proteine di fusione con gli AT-hooks della proteina HRX (ALL-1) in leucemie.	55
Barx 1	Mm_1550783	1	dominio omeo	Espressa in precursori dello sviluppo neurale e craniale	55
CTCF1	Mm_1256410	1	11 C <sub>2</sub> H <sub>2</sub> Zn fingers	Regolatore dello sviluppo neurale	55
M33 polycomb-like protein	Mm_2266988	1	Chromodomain	Coinvolto nella condensazione cromatinica	55
ORF epitope	Mm_2384711	1			
Castor protein	Dm_419965	1	Peculiar modified C2H2 Zn fingers	Regolatore trascrizionale nello sviluppo neurale.	55
Alternatively spliced tramtrack	Dm_578319	1	dominio POZ + 2 Zn fingers	Repressore nello sviluppo neurale.	55

**Tabella I.querter Continuazione**

Doom-modifier of mdg4	Dm_942610	1	dominio POZ	Proteina/Enhancer interagente con l'elemento "Insulator" della variegazione con imprinting; un'isoforma, Doom, è un regolatore pro-apoptotico.	55
ISWI	Dm_1362615	1	elicasi simile alla superfamiglia 2 Swi/Snf + dominio SANT	Componente ATPasico del complesso NURF di rimodellamento cromatinico.	55
Apterous protein	Dm_231559	1	domini LIM + dominio omeo	Regolatore dello sviluppo neurale e della polarità dorsoventrale dell'ala.	55
Zinc finger 30C	Dm_2653645	1	12 C <sub>2</sub> H <sub>2</sub> Zn fingers		55
Zinc finger AT-hook protein	Dm_2982495	1	11 C <sub>2</sub> H <sub>2</sub> Zn fingers		55
ORF YN06	Ce_465974	1			55
ORF C27B7.4	Ce_1000467	1	elicasi simile alla superfamiglia 2 Swi/Snf	Ortologo del gene umano <i>ATR-X</i> .	55
ORF F40F9.7	Ce_1262959	1	dominio Histone fold	Proposta proteina cromosomiale.	55
ORF C44F1.2	Ce_780182	1	dominio KDWK	Proteina che lega il DNA.	55
ORF C25G4.4	Ce_1256454	1	dominio KDWK	Proteina che lega il DNA.	55
ORF T11A5.1	Ce_1301722	1			55
ORF YG23	Sc_1723670	1	2 bromodomini e 1 dominio BAM		55
ORF F18E2.3	Ce_1418507	1	dominio SA1/2-rec11		55
ORF F16B12.6	Ce_1673468	1	dominio globulare ristretto a <i>C. elegans</i>		55
ORF C43E11.3	Ce_1703572	1	dominio SET		55
ORF C34F6.9	Ce_264801	1	Idrolasi del C-terminale dell'ubiquitina		55
ORF F15E6.1	Ce_2702442	1	SET PHD finger + dominio modificato		55
Orc2p	Sc_464317	1		Comonente del complesso di riconoscimento dell'origine di replicazione.	55
Swi5p	Sc_135077	1	3 C <sub>2</sub> H <sub>2</sub> Zn fingers	Regolatore della trascrizione ciclo cellulare-specifica in lievito.	55
Mif2p	Sc_462601	1	AraC type $\beta$ -helix motif	Proteina legante il centromero richiesta per l'integrità del fuso mitotico.	55
ORF YDR36	Sc_2131446	1			55
ARS-binding protein 2	Sp_1724085	1		proteina che lega ARS in <i>S. pombe</i> .	55
Orc 1 <sup>p</sup>	Sp_1163108	1	dominio BAM + dominio AAA ATPasico	Subunità del complesso di idrolisi dell'ATP che riconosce l'origine.	55

**Tabella I. Continuazione**

AMT1	Cg_1168451	1	Fungal-type metal-binding cluster	Proteina legante il DNA che regola la risposta ai metalli pesanti.	55
Manx	Mo_308967	1	dominio Tudor + dominio C-terminale ricco in Cisteine	Regolatore "Master" dello sviluppo della coda negli ascidi.	55
ICP4	GHV_643430	1		Attivatore trascrizionale attivo in <i>trans</i> dell'herpesvirus della malattia di Marek.	55
RF1	SHV_279827	1		Attivatore trascrizionale di geni precoci tardivi.	55
Tuple1/HirA	Fr_2352031	1	ripetizioni WD40	Regolatore dell'espressione genica degli istoni.	55
MADSbox homolog Umc1	Um_2708784	1	MADS box	Regolatore trascrizionale di molti geni ferormone-responsivi in <i>Ustilago</i> .	55
C	To_2894102	1			55
Tc1	An_1911486	1	Transposasi + HTH	Famiglia Tc1 di transposoni con domini HTH e dell'integrasi.	55

Le proteine contenenti un motivo AT-hook sono elencate nella colonna intitolata "Nome della proteina". Nella seconda colonna sono elencate le abbreviazioni dei nomi delle specie (vedi sotto) associate con l'inserto nella banca dati, seguite dal numero *ig* per la sequenza proteica della Genbank (usando questi numeri si estraggono le sequenze dalla Genbank all'indirizzo: <http://ncbi.nlm.nih.gov/Entrez/protein.html>). La colonna intitolata "n" mostra il numero di motivi AT-hook presenti in ciascuna proteina. La quarta colonna elenca altri domini proteici presenti in ciascun polipeptide; se non c'è scritto nulla significa che al momento non si sono identificati. Alcuni di questi sono domini ricorrenti anche in altre proteine, ma non ancora ben definiti.

An, *Aspergillus nidulans*; At, *Arabidopsis thaliana*; Ca, *Carassius auratus*; Ce, *Caenorhabditis elegans*; Cg, *Candida glabrata*; Dd, *Dictyostelium discoideum*; Dm, *Drosophila melanogaster*; Fr, *Fugu rubripes*; GHV, *Gallid herpesvirus*; Gm, *Glycine max*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Mo, *Molgula oculata*; Mx, *Myxococcus xanthus*; Nt, *Nicotiana tabacum*; Os, *Oryza sativa*; Pc, *Petrostelium crispum*; Sc, *Saccaromyces cerevisiae*; SHV, *Saimiriine herpesvirus 1*; Sp, *Schizosaccharomyces pombe*; To, *Torulasporea delbrueckii*; Um, *Ustilago maydis*; Vs, *Vicia sativa*.

## 1.2.4 Modificazioni post-traduzionali delle HMGI: effetti biologici

Le modificazioni post-traduzionali delle proteine in generale sono un mezzo che rende idonea una proteina per adempiere alle sue funzioni nel distretto cellulare corretto, oppure per modulare le funzioni delle stesse in modo reversibile.

Finora le modificazioni osservate sulle HMGI sono fosforilazioni, adenilazioni, acetilazioni, metilazioni, glicosilazioni ed ADP-ribosilazioni, ma non tutte sono state studiate approfonditamente. (61, 62)

### Fosforilazione

La fosforilazione è un meccanismo biochimico usato dalla cellula per modificare reversibilmente l'attività delle proteine ed ha un ruolo essenziale nella regolazione di numerosi processi in cellule eucariotiche. Si calcola che circa il 30% delle proteine cellulari siano fosforilate. In particolare la fosforilazione modula l'attività di fattori coinvolti in molteplici vie di trasduzione del segnale e nella trascrizione genica. (63) L'identificazione di chinasi che fosforilano le proteine HMGI potrebbe contribuire a comprendere come sia regolata l'azione di queste proteine.

Esistono diverse proteinchinasi che differiscono per la specificità dell'amminoacido che fosforilano. Fin'ora sono stati scoperti 3 tipi di fosforilazioni che possono avvenire sulle proteine HMGI(Y). La Casein-chinasi di tipo II (CK-II) è una serina/treonina chinasi ubiquitaria, altamente conservata che si trova nel citoplasma e nel nucleo di cellule eucariotiche. Essa sembra essere correlata alla trasduzione del segnale tra mitogeni extracellulari e risposte nucleari. Molte fosfoproteine nucleari implicate nella proliferazione, come ad esempio *c-myc*, *c-myb*, *c-fos*, *E1A* ed il *large T antigen* possiedono potenziali siti di fosforilazione da parte della CK-II, e per alcuni di questi è stata dimostrata l'effettiva fosforilazione sia *in vitro* che *in vivo*. (64) Il consensus per questa chinasi è dato da una serina o treonina seguita da un gruppo di residui di acido glutammico o aspartico, con speciale riguardo al residuo in posizione +3 rispetto all'amminoacido da fosforilare (di norma S/TXXD/E). Inoltre residui di fosfoserina possono sostituire gli acidi carbossilici in posizione +3 come determinanti di specificità. Nella proteina HMGI umana le serine in posizione 102 e

103 corrispondono a tali criteri, e anche la serina in posizione 99, nel caso che la Ser-102 sia fosforilata. È stato dimostrato che sia *in vitro* che *in vivo* la CK-II fosforila la proteina HMGI a livello delle serine 102 e 103, e anche 99, ma solo in seguito alla modificazione avvenuta sulla Ser-102. (64) Tecniche sofisticate di spettrometria di massa hanno permesso di concludere che anche l'isoforma HMGY, estratta dal tessuto di carcinoma polmonare di Lewis (LLC), un carcinoma spontaneo nel topo, si presenta come due specie molecolari iperfosforilate: una bi-ed una tri-fosforilata. Entrambe possiedono una fosfato a livello delle serine che corrispondono alle serine 102 e 103 dell'HMGI umana, mentre differiscono per una terza fosforilazione, che corrisponderebbe alla Ser-99 in HMGI. (Vedi Figura 2).(65)

Esistono però forme HMGI(Y) più estesamente fosforilate. Ad esempio si sono osservati ulteriori siti fosforilati a livello delle treonine 21, 53, 77 e 78, due dei quali, Thr-53 e Thr-78 (presente solo nelle proteine umane) sono chinasi  $p34^{cdc2}$ -dipendenti. Questa chinasi è molto importante in quanto in cellule eucariotiche il segnale iniziale per la fase M mitotica e meiotica è dato dalla sua attivazione. La  $p34^{cdc2}$  rimane poi attiva durante tutto il periodo del processo mitotico. È noto che avviene una diffusa riorganizzazione della cellula quando dall'interfase si procede verso la mitosi: gli aspetti più manifesti sono la dissoluzione del nucleo, la formazione del fuso mitotico, la condensazione cromosomica e il *rounding up* della cellula nel momento della divisione. Processi biochimici come la trascrizione genica, la traduzione e traffico intermembrana sono transientemente inibiti. Alla fine della mitosi tutti questi riarrangiamenti e cambiamenti revertano e la cellula torna all'interfase. (66) Il fatto che le HMGI siano fosforilate dalla  $p34^{cdc2}$  significa che è sottoposta a modificazioni ciclo cellulare dipendenti. Tali fosforilazioni sono rese evidenti analizzando cellule in coltura sincronizzate ed arrestate in metafase e confrontandole con cellule non sincronizzate. Ad esempio un esperimento di questo tipo in cellule HeLa mostrava la presenza, in cellule arrestate in metafase, di forme di HMGI ed Y fosforilate con una migrazione in elettroforesi diversa dalle forme estratte da cellule non sincronizzate. Risultati analoghi sono stati ottenuti da linfociti attivati con la PHA (fitoemagglutinina) e cellule CHO sincronizzate, indicando quindi l'avvenuta modificazione ciclo cellulare specifica. (67)

Proprio con cellule HeLa sincronizzate è stato dimostrato per la prima volta che le proteine HMGI ed Y sono fosforilate specificamente dalla  $p34^{cdc2}$  (o chinasi istone H1-specifica) durante la metafase a livello delle treonine 53 e 78 (mancante in HMGY), poste all'interno del consensus (Zaa)-Ser/Thr\*-Pro-(Xaa)-Zaa tipico della chinasi  $p34^{cdc2}$ , dove Zaa è generalmente una Lisina o Arginina e Xaa un residuo polare. I residui tra parentesi non sono sempre presenti. (68-70) È importante notare che all'aumentare e diminuire dell'attività di questa

chinesi durante il ciclo cellulare si nota un andamento parallelo dell'entità della fosforilazione della proteina HMGI.

L'effetto delle modificazioni qui elencate sulle funzioni di queste proteine sono oggetto di attive ricerche. Il ruolo delle fosforilazioni operate dalla CK-II a livello della coda C-terminale acida è tuttora oscuro, in primo luogo perché la funzione della coda carbossi-terminale stessa non è ancora nota. Si sa che non funge da dominio di attivazione trascrizionale (48), come si poteva supporre (71), ma si pensa che abbia piuttosto un ruolo nel modulare l'affinità di legame della proteina al DNA (osservazioni nostre). Il fatto di aggiungere ulteriori 4/6 cariche negative nel contesto di un dominio già negativo, con carica netta di  $-8$ , costituisce un evento la cui funzione non è facilmente comprensibile. Era stato proposto che queste fosforilazioni potrebbero aumentare l'interazione tra HMGI e i domini basici dell'istone H1, facendo così partecipare le HMGI nella condensazione cromatinica durante la mitosi. (64)

L'attività modulatrice delle proteine HMGI(Y) nella trascrizione di molti geni inducibili è ben assodata. Un esempio è quello del promotore  $\epsilon$  della linea germinale ( $G\epsilon$ ) di Immunoglobuline, descritto nella sezione 1.2.5.2, la cui inducibilità trascrizionale è repressa da HMGI(Y) in cellule B di mammifero.  $G\epsilon$  si attiva quando le cellule vengono trattate con IL-4, e quindi si osserva la produzione del suo messaggero. Inoltre la fosforilazione di HMGI(Y) è coinvolta con la trasformazione neoplastica e con la proliferazione, mentre l'IL-4 è capace di inibire la proliferazione di cellule B leucemiche. La combinazione di questi fatti ha portato a pensare che in qualche modo a valle del trattamento dell'IL-4, attraverso una via di trasduzione del segnale, potesse avvenire una modificazione post-traduzionale delle HMGI(Y) che contribuisca alla derepressione del promotore di  $G\epsilon$ . Una linea cellulare B linfoblastoide umana denominata JY è stata trattata con IL-4 e confrontando lo stato della fosforilazione delle proteine nucleari tra cellule trattate e non, si è potuto dimostrare un generale aumento in fosfoproteine, che include l'aumento della fosforilazione delle HMGI(Y) *in vivo*. Attraverso analisi di mappatura di fosfoamminoacidi e fosfopeptidi si è localizzata la nuova fosforilazione in serina a livello della coda acida. Il farmaco immunodepressivo rapamicina inibisce invece parzialmente l'induzione del promotore  $G\epsilon$  attraverso IL-4 in cellule B umane e causa il blocco della fosforilazione della proteina HMGI(Y). Invece la genisteina, che è un inibitore specifico delle tirosinchinasi non inficia la fosforilazione indotta dall'IL-4, suggerendo che la modificazione non avviene attraverso una via di trasduzione che attiva le proteine STAT. Si sa ancora poco sulle *pathway* attivate dall'IL-4, ma il fatto che attivi la PI 3-chinasi associata ad un recettore di membrana, il quale è sensibile al trattamento con

rapamicina ha senso. Un effettore a valle di questa chinasi è la protein-chinasi C, ma nello stesso tempo la modificazione in serina della coda acida richiama sicuramente la Casein chinasi II. Esperimenti di fosforilazione *in vitro* hanno suggerito che la PKC molto probabilmente non c'entra, mentre le HMGI(Y) sono buoni substrati della CK-II. Si è concluso da analisi comparative che la fosforilazione IL-4 dipendente della coda acida di HMGI(Y) *in vivo* è analoga a quella della CK-II *in vitro*; quindi la chinasi che opera a valle dell'IL-4 ha la stessa specificità della CK-II. Esperimenti di ritardo di banda con le HMGI(Y) fosforilate dalla CK-II e non con la sonda derivante dal promotore  $G\epsilon$  dimostrano un abbassamento dell'attività di legame al promotore. (72)

La CK-II è regolata da fattori di crescita come l'insulina, che condivide alcuni meccanismi di trasduzione del segnale con la IL-4, effettua fosforilazioni regolative su fattori di trascrizione e lega direttamente la rotamasi FKBP25, che a sua volta interagisce con la rapamicina. Comunque la CK-II probabilmente non è la chinasi utilizzata *in vivo*, dato che non è stato mai riportato che sia sensibile al trattamento con rapamicina *in vivo* ed il complesso rapamicina-FKBP25 non inibisce la CK-II *in vitro*.

Con l'attivazione del recettore per l'IL-4 vengono reclutate la PI 3-chinasi ed altre molecole come il substrato 1 del recettore dell'insulina (IRS-1) ed il suo omologo 4PS. Questi fattori precoci della *pathway* indotta da IL-4 potrebbero avere un ruolo nella regolazione genistein-indipendente dello stato fosforilato delle HMGI(Y). Molto probabilmente si tratta di una via molto specifica e non della solita via attraverso Raf e Ras e le MAP chinasi, dato che le MAP chinasi non sono sensibili alla rapamicina e l'IL-4 non sembra attivare né Raf né Ras e nemmeno le MAP chinasi. L'unica chinasi che agisce nei passaggi più tardivi della trasduzione del segnale e che nello stesso tempo è una chinasi indotta da fattori di crescita ed è specificamente e potentemente inibita dalla rapamicina è la chinasi pp70 S6. Inoltre in un'altra linea cellulare rispondente all'attivazione da citochine (CTLL) sia IL-4 che IL-2 portano alla fosforilazione rapamicina sensibile di HMGI(Y), suggerendo che la fosforilazione sia provocata da un componente comune tra la famiglia dei recettori delle citochine. Infatti i recettori per l'IL-2 ed IL-4 condividono molte molecole di trasduzione del segnale che sono note per attivare la serin.treonin chinasi pp70/85 S6.

Nonostante queste evidenze anche la chinasi pp70 S6 forse non è il candidato giusto, proprio perché la sua specificità per il substrato non corrisponde a nessun peptide di HMGI(Y). Ulteriori ricerche saranno necessarie per trovare la chinasi responsabile di questa fosforilazione. L'effetto che però questa fosforilazione potrebbe provocare è duplice. Un



meccanismo prevede l'allentamento del legame con il promotore e quindi il possibile reclutamento di TBP e del complesso basale della trascrizione sul sito finallora occupato da HMGI(Y) e quindi l'avvio alla trascrizione. Un'altra possibilità prevede un'alterazione delle interazioni proteina-proteina tra HMGI(Y) fosforilate e altre proteine regolatrici, un'ipotesi avvalorata dal fatto che complessi nucleoproteici (NF-BRE) che si formano tra promotore ed estratti nucleari da cellule trattate con IL-4 escludono l'HMGI(Y) qualora gli estratti siano trattati con una defosfatasi. (72, 73)

La fosforilazione delle proteine HMG dipende sicuramente dallo stato fisiologico della cellula. In cellule di mammifero altamente proliferanti, le proteine HMGI(Y) come anche l'istone H1 sono tra le proteine cromosomiali più fosforilate. Inoltre un aumento nel grado di fosforilazione di 2-3 volte è stato osservato in cellule NIH 3T3 arrestate in mitosi. (74) L'aumento del grado di queste fosforilazioni era dovuto solo a modificazione della Thr-53 operata dalla p34<sup>cdc2</sup>.

Le fosforilazioni della p34<sup>cdc2</sup>, che avvengono a livello dei residui 53 e 78 disturbano il legame della proteina con il DNA *in vitro*. Infatti quattro cariche negative che vengono a porsi ai lati dell'AT-hook, che sfrutta anche le sue cariche positive per legarsi al DNA, indeboliscono tale interazione. (70, 75)

Ricerche analoghe effettuate sulla proteina HMGI di un insetto dittero, *Chironomus tentans*, altamente omologa alla proteina umana, (76) hanno dimostrato che le fosforilazioni sono regolate durante il ciclo cellulare. Infatti più del 30% della proteina è altamente fosforilata *in vivo* in cellule proliferanti mentre con la riduzione della proliferazione e con l'inizio del differenziamento indotto dall'ecdisione le forme fosforilate diminuiscono. I siti di fosforilazione sono stati localizzati a livello della serina 3, 22 e 72 e si è dimostrato che sono substrati della p34<sup>cdc2</sup> (e della chinasi attivata da mitogeni o MAP), della MAP e della proteina chinasi C, rispettivamente. In cellule arrestate in mitosi la serina 3 risulta altamente fosforilata, mentre la 22 non è alterata. In cellule quiescenti si nota una forte riduzione della fosforilazione a livello di entrambe le serine. L'affinità della proteina HMGI di *Chironomus* per il DNA rimane invariata anche quando una delle due serine è fosforilata, mentre nel momento in cui entrambe sono fosforilate si nota un indebolimento di ben 10 volte dell'attività di legame al DNA, ma anche un'alterazione del modo di legarsi. (77)

In linee cellulari di *Chironomus* l'ormone ecdisione stimola il differenziamento, un processo che si accompagna con cospicui cambiamenti nelle attività delle MAP chinasi. Le serine 3 e 22 sembrano essere dei substrati delle MAP chinasi perché corrispondono al loro consensus (S/T\*P). L'utilizzo di una MAP chinasi (Erk-2) *in vitro* dimostra che la proteina HMGI si

fosforila efficientemente e che la Ser-22 è il substrato preferenziale, mentre la Ser-3 viene modificata 3 volte di meno. Usando il 2'-amino-3'-methoxyflavone, un inibitore specifico delle MAP chinasi chinasi (MAPKK), in un esperimento *in vivo*, si è osservata una riduzione sia della specie monofosforilata che della specie doppiamente fosforilata, suggerendo che la fosforilazione delle due serine 3 e 22 è sotto controllo della via di trasduzione del segnale delle MAP chinasi. (77)

Anche la proteina HMGI umana è un substrato delle MAP chinasi. Assieme al fatto che possiede due potenziali siti a livello della Ser-36 e della Thr-53 era già stato osservato mediante *peptide mapping*, che nella proteina fosforilata *in vivo* in cellule proliferanti in fase esponenziale il sito maggiormente fosforilato non corrispondeva al sito modificato dalla p34<sup>cdc2</sup>. (74)

Un ulteriore studio ha evidenziato che la Ser-72 della HMGI di *Chironomus* è substrato della Protein chinasi C (PKC) *in vitro*. È l'unico residuo che corrisponde al consensus di questa chinasi che è (S/T)\*X(K/R) per cui sembra molto probabile che sia substrato delle PKC anche *in vivo*. (77)

La fosforilazione di proteine spesso inizia il processo di degradazione. È stato verificato, almeno per quanto riguarda la fosforilazione in Ser-22, che il *turn-over* di questa proteina non viene inficiata. (77)

Il vero ruolo di queste modificazioni è quindi ancora sconosciuto. Siccome l'HMGI di *Chironomus* e le sue forme fosforilate esibiscono lo stesso basso grado di "turn-over" si può ipotizzare che modificazioni post-traduzionali come la fosforilazione possano servire come meccanismi che governano i cambiamenti della localizzazione cromosomica di questa proteina dipendenti dal ciclo cellulare e da segnali extracellulari. Inoltre la dimostrazione che HMGI di *Chironomus* è substrato per le chinasi p34<sup>cdc2</sup>, MAP e PKC sia *in vitro* che *in vivo* indica che la proteina è funzionalmente molto versatile. (77) I dati ottenuti sulla HMGI in insetto aprono nuove strade per la comprensione delle proteine HMGI in mammifero. Ad esempio una singola fosforilazione, che non altera il legame della proteina col DNA, potrebbe però modulare l'interazione con altre proteine all'interno di una struttura nucleoproteica superiore. Una doppia fosforilazione invece, ai lati di un dominio che lega il DNA, abbassa l'affinità della proteina per il DNA e fa scomparire i complessi distinti in esperimenti di ritardo di banda. Come si sa da studi precedenti il legame più forte col DNA si ottiene con un sito trivalente sul DNA, in cui vi è una corretta spaziatura tra i 3 tratti ricchi in AT, in modo che tutti e tre gli AT-hook della stessa proteina possono prendere contatto e dare un contributo cooperativo intramolecolare altissimo all'affinità di legame al DNA. (49). Con la doppia

fosforilazione si sottrae un dominio di legame, lasciandone solo due, facendo sì che la proteina diventi più mobile e aumenti la possibilità di interazioni divalenti. Dato che la probabilità di trovare 2 tratti di AT correttamente spazati nei cromosomi è sicuramente più alta che trovarne tre il processo di fosforilazione potrebbe facilitare il processo di dissociazione delle proteine HMGI(Y) da un tipo di sito ad altri. (77)

Uno studio sulla proteina HMGI-C murina ha portato ad alcuni chiarimenti molto importanti: innanzitutto la fosforilazione della coda C-terminale mediante la CK-II diminuisce l'affinità di tutta la proteina per il DNA ed aumenta inoltre la stabilità al taglio di proteasi specifiche; invece la modificazione da parte della p34cdc2 che avviene ai due lati del 2° AT-hook inficia alcuni contatti dello stesso con il DNA dell'elemento PRDII del promotore dell'INF $\beta$ . In questo studio si è potuto verificare che HMGI-C murina lega in analogia alla HMGI di *Chironomus* con il primo e secondo motivo AT-hook, mentre la proteina HMGI umana investe i due ultimi motivi più vicini alla coda C-terminale. Ciò probabilmente è dovuto alla distanza relativa tra i centri dei diversi AT-hooks, che è di 20 residui tra il primo e secondo in *Chironomus* HMGI e nella HMGI-C murina, mentre nella HMGI umana sono 32 tra il primo e secondo e 26 tra il secondo e terzo AT-hook. Quindi per il legame con PRDII probabilmente viene scelta la coppia più stretta. (78)

### **Acetilazione**

Recentemente sono stati identificati 3 siti di acetilazione sulle proteine HMGI ed HMGI-C umane nel contesto del *transcriptosoma* del promotore dell'interferone  $\beta$  umano. Infatti come descritto nelle sezioni 1.2.5.1 e 1.2.5.2 la proteina HMGI(Y) ha un ruolo importantissimo nell'assemblaggio e nel mantenimento di questa struttura nucleoproteica superiore. La minima alterazione della proteina può ribaltare completamente la situazione. Infatti è ciò che succede: la proteina HMGI(Y) viene specificamente acetilata dal coattivatore CBP a livello della lisina 65, posta a lato del secondo AT-hook (quello che media l'attacco specifico al DNA) facendo sì che il potere legante della proteina diminuisca e che inizi una globale destabilizzazione di tutto l'*enhanceosoma*. Infatti questa acetilazione provoca un istantaneo abbassamento della cooperatività del complesso HMGI(Y)/NF- $\kappa$ B e la terminazione della trascrizione. (79) Risulta da questo lavoro che i tempi dell'espressione di INF $\beta$  indotta da virus è regolata *in*

non impaccato, e quindi, oltre a liberare frammenti piccoli di DNA dal nucleo, fuoriescono anche le proteine associate a questo DNA. Tipicamente l'istone H1 associato alla cromatina repressa, rimane nell'altra frazione, ossia all'interno dei nuclei. Attraverso l'utilizzo di NAD<sup>+</sup> marcato con 32P si è stabilito che la modificazione osservata sulle HMGI(Y) era ADP-ribosilazione. Inoltre questo fenomeno è stato osservato solo con il Ca<sup>++</sup>, non con il Mg<sup>++</sup> e nemmeno con Na<sup>+</sup> o K<sup>+</sup>, suggerendo un meccanismo Ca<sup>++</sup>-dipendente specifico e non un semplice aumento dell'attività della poly(ADP-ribosio) polimerasi che può essere sensibile a ioni bivalenti in generale. L'ipotesi suggerita in questo caso è che l'aggiunta di Ca<sup>++</sup> nel mezzo sperimentale mimava uno stimolo extracellulare come potrebbe essere quello del TNF- $\alpha$  che induce il rilascio del Ca<sup>++</sup> dai serbatoi intracitoplasmatici. Il Ca<sup>++</sup> libero può propagare nel nucleo dove induce la frammentazione del DNA. Si sa che questo è un fenomeno tipico dell'apoptosi e in seguito ad esso la poly(ADP-ribosio) polimerasi, normalmente inattiva, inizia a poly-ADP-ribosilare molte proteine, come gli istoni, RNA e DNA polimerasi e la topoisomerasi. Le HMGI potrebbero essere incluse in questo gruppo di bersagli della poli(ADP-ribosio) polimerasi e la conseguente modificazione potrebbe essere correlata alle alterazioni morfologiche della cromatina che avvengono durante il collassamento della struttura nucleare nell'apoptosi. (80)

Tabella II.

Schema riassuntivo delle modificazioni post-traduzionali reperite nelle proteine HMGI

Tipo di modificazione	enzima	HMGI	HMGY	HMGI-C
fosforilazione	CK-II	Homo sapiens: Ser-103 Ser-102 Ser-99	Homo sapiens: Ser-92 (103-11) Ser-91 (102-11) Ser-87 (99-11)	Mus muris: Ser-103 Ser-100 Ser-99 Thr-98
		Chironomus tentans: Ser-72		
	P34 <sup>cdc2</sup>	Homo sapiens: Thr-53 Thr-78 Chironomus: Ser-3 Ser-22	Homo sapiens: Thr-42 (53-11) Thr-67 (78-11)	Mus muris: Ser-43 Ser-58
	MAP chinasi	Homo sapiens: Ser-36 Thr-53 Chironomus tentans: Ser-22		
	pp70 S6 (?) (rapamicina)	Homo sapiens: Ser (coda acida)	Homo sapiens: Ser (coda acida)	
	PKC	Chironomus tentans: Ser-72		
acetilazione	CBP	Homo sapiens: Lys-65 Un aa nell'intervallo 6-31		
		P/CAF	Lys-71	
Poly-ADP- ribosilazione	Poly(ADP- ribosio) polimerasi	?	?	?

## 1.2.5 Funzioni Biologiche delle proteine HMGI

Questi fattori proteici presentano un ruolo a livello cellulare molto complesso e ancora lontano dall'essere esaurientemente compreso. Evidenze crescenti li vedono implicati in numerosi processi biologici a livello dell'architettura cromatinica e di regolazione dell'espressione genica sia attraverso meccanismi generali di derepressione che in modo promotore specifico. Le HMGI sembrano particolarmente adatte per queste funzioni per la capacità di legare sia il DNA sia altre proteine, quali specifici fattori di trascrizione.(81)

È essenziale far notare che in generale sono più abbondanti le nozioni sulle proteine HMGI(Y) che sulla proteina HMGI-C, sicuramente per motivi di abbondanza relativa nei modelli cellulari oggetto di studio in ciascun laboratorio. L'omologia di sequenza tra questi fattori indicherebbe però anche omologia funzionale, così da poter trasferire alcune conoscenze sul modo d'agire delle HMGI(Y) anche alla proteina HMGI-C.

### 1.2.5.1 HMGI e cromatina

Originariamente identificate come proteine in grado di interagire con il DNA satellite altamente ripetitivo (50), si è attribuito alle proteine HMGI un ruolo attivo nell'organizzazione del DNA nucleosomale e nella condensazione dell'eterocromatina. (82). In seguito sono state ripetutamente immunolocalizzate a livello delle bande C (costituenti regioni telomeriche e centromeriche) e di zone ricche in AT appartenenti a bande G/Q di cromosomi metafasici, (83) ed anche dei cromosomi nel resto del ciclo cellulare, data la co-immunolocalizzazione con la topoisomerasi II.(84) Questa localizzazione nucleare suggerisce una funzione nelle alterazioni strutturali dinamiche della cromatina durante tutto il ciclo cellulare: a sostegno di questa ipotesi si colloca anche la probabile modulazione reversibile dell'affinità di legame, mediata da p34<sup>cdc2</sup> e quindi coordinata con gli eventi del ciclo cellulare.

Il DNA "four-way junction" (4H) è stato suggerito come modello per il DNA della cromatina nel sito in cui entra nel ed esce dal nucleosoma, (85) oppure come giunzione "Holliday" tipica negli intermedi della ricombinazione omologa *in vivo*. Sperimentalmente questa struttura è data dall'ibridazione di 4 oligonucleotidi contemporaneamente, che formano un DNA cruciforme dinamico (86) nel senso che a seconda della assenza o presenza di ioni Mg<sup>2+</sup> forma

una struttura “aperta”, con tutte e quattro le braccia distese con una massima apertura a livello del punto d’incrocio, oppure chiusa, a formare una X piuttosto che una vera croce, rispettivamente. Studi di cristallografia hanno dimostrato che sia la ricombinasi Cre del fago  $\lambda$  che la proteina batterica RuvA formano dei co-complessi con il DNA 4H solo in forma aperta. In assenza di  $Mg^{2+}$  i punti di diramazione sono chiusi e le braccia della giunzione si dipartono co-assialmente a formare una struttura a X, creando così due filamenti continui e due filamenti di scambio o non continui. Il legame delle HMGI a questo DNA potrebbe riflettere una funzione costitutiva di queste proteine nell’organizzazione della cromatina, contrariamente alla loro funzione specifica nella direzione dei complessi promotoriali di particolari geni. Sia la forma aperta che quella a X hanno il DNA nella conformazione B.

È stato dimostrato che al DNA 4H in forma aperta legano sia l’istone H1, la proteina HMG-1 e HMGI(Y), pur avendo tre tipi di domini di legame al DNA del tutto diversi. Mediante studi di *ritardo di banda* si sono potute calcolare le costanti di dissociazione che sono

$$K_d(\text{HMGI(Y)}) = 6.5 \text{ nM}$$

$$K_d(\text{H1}) = 16 \text{ nM}$$

$$K_d(\text{HMG-1}) = 80 \text{ nM}$$

ed inoltre esperimenti di *competizione* hanno dimostrato che queste tre proteine legano in mutua esclusione tra di loro. (87) Analizzando a livello nucleotidico i punti di contatto tra questo DNA ed le tre diverse proteine mediante *footprinting* con radicali idrossilici, si sono potuti intravedere sia *pattern* di legame unici che comuni tra le tre proteine. Sia HMG-1 che HMGI(Y) contattano specificamente tutta la lunghezza del punto d’incrocio ed interagiscono con punti ben precisi sulle braccia, mentre il dominio globulare dell’istone H1 contatta solo le braccia ma non il quadrivio. Inoltre per HMGI(Y) si è dimostrato che lega sempre nella stessa direzione, utilizzando il secondo e terzo dominio di legame al DNA per contattare contemporaneamente punti opposti al quadrivio. (87) Questi risultati sono interessanti soprattutto alla luce del fatto che HMGI(Y) contatta specificamente il solco minore di tratti ricchi in AT, mentre l’HMG-box ed il dominio globulare dell’istone H1 legano il DNA indipendentemente dalla sequenza. Soprattutto per l’istone è importante non dipendere dalla sequenza visto che deve legarsi al DNA linker, che sembra non mostrare regolarità di sequenza. Si pensa però che tutte e tre le proteine riconoscano specifiche strutture e siano in grado di indurre modificazioni conformazionali una volta legati al DNA e per HMGI(Y) ed HMG-1 è stata dimostrata un’attiva partecipazione in eventi di ricombinazione e di integrazione virale, oltre che ad agire da fattori architetturali nella regolazione trascrizionale.

Il modo preciso di agire della proteina HMG-1 si può dedurre da studi effettuati sul fattore SRY, un fattore umano di determinazione sessuale che contiene una HMG-box. Quando contatta il DNA lega il solco minore del B-DNA lo srotola parzialmente allargando il solco e contemporaneamente induce una curvatura cospicua del substrato. Dato che i punti di diramazione nel DNA 4H assomiglia ad una di queste curvature di DNA a doppio filamento. Coerentemente la maggioranza dei nucleotidi nei punti di diramazione sono protetti dal taglio chimico in presenza della proteina HMG-1. *In vivo* questa situazione risulta sicuramente più complessa, dato che la proteina HMG-1 possiede ulteriori domini indipendenti con cui contattare il DNA.

L'istone H1 sembra utilizzare due superfici di contatto per il DNA posto sullo stesso dominio globulare, in modo tale da interagire contemporaneamente con due doppi filamenti, come un treno su due binari. A livello del DNA 4H si nota una distribuzione simmetrica della proteina sulle braccia ma mai a livello dei nucleotidi del *branch point*.

Molto probabilmente l'unicità e la non completa sovrapposizione dei contatti col DNA 4H delle tre proteine fa supporre diversi ruoli strutturali e/o funzionali durante gli eventi cellulari come la ricombinazione genetica, riarrangiamenti V(D)J delle Immunoglobuline o l'integrazione virale nel genoma dell'ospite. (87)

Proprio riguardo all'integrazione virale è stato asserito che HMGI(Y) ha un ruolo essenziale nell'integrazione del cDNA del virus HIV-1. Il cDNA di HIV-1 si trova in vivo all'interno di complessi nucleoproteici, denominati PICs (Preinitiation Complexes). Questi possono essere isolati sperimentalmente da cellule appena infettate e riescono ad effettuare l'integrazione *in vitro*, se messi a contatto con il DNA bersaglio. (88)

I PICs contengono l'integrasi, codificata dal virus, che media il taglio del bersaglio e l'inserimento del cDNA lineare dell'HIV-1; inoltre contengono anche la matrice virale (MA), il nucleocapside (NC), la trascrittasi inversa (RT) e cosiddette proteine VPR. Il nucleocapside e la trascrittasi inversa sono importanti per la sintesi del DNA mentre la MA e le VPR sono necessari per la localizzazione subcellulare dei PICs. Nessuno di questi componenti media l'integrazione, ma serve un fattore specifico dell'ospite per questo, dato che il trattamento dei PICs con alto sale toglie l'attività integrasica che può essere prontamente ricostituita aggiungendo estratti da cellule non infettate. La purificazione di questa attività ricostituente la capacità integrasica da cellule SupT1 ha portato all'identificazione di HMGI(Y). Come funziona? Innanzitutto HMGI(Y) non è capace da solo di effettuare l'integrazione quindi ha un ruolo come fattore accessorio. In secondo luogo dato che i PICs trattati con alte concentrazioni di sale (0.6 M KCl) e con il cDNA virale già disponibile all'integrazione (cioè



con le estremità 3' tagliate) sono inattivi, significa che HMGI(Y) è importante almeno per il passaggio in cui si ha il trasferimento covalente del filamento di DNA. Inoltre sembra che HMGI(Y) non si leghi al DNA bersaglio, perché pur aggiungendo questa proteina ai PICs in alta concentrazione salina non si ripristina nessuna integrazione. Sembra invece che HMGI(Y) si leghi al cDNA dell'HIV-1. Infatti esistono molti siti potenziali per HMGI(Y) nell'LTR di HIV-1 e si pensa che un eventuale interazione potrebbe alterare la conformazione del cDNA di HIV-1 in modo da promuovere la funzione del complesso cataliticamente attivo. Soprattutto potrebbe portare in stretta vicinanza le 2 estremità del DNA virale, dato che i 2 punti di integrazione nel DNA bersaglio distano sempre soltanto di 5 basi. Infatti c'è un ponte proteico implicato nell'unione le estremità del cDNA dato che il taglio del cDNA con un enzima di restrizione non ostacola la fusione corretta dei DNA virale e genomico. (88)

Simili fattori sono importanti anche nella trasposizione e nella ricombinazione sito specifica ed esistono esempi in batteri, come le proteine IHF e Hu di *Escherichia coli*, che sono importanti per coadiuvare l'integrasi fagica nella ricombinazione del fago  $\lambda$  o Mu, rispettivamente. Anche queste proteine possiedono una capacità intrinseca di curvare il DNA. Il fatto che sia HMGI(Y) il o uno dei fattori prescelti trova alcune spiegazioni: a parte la sua attitudine da fattore architetturale è importante per la distribuzione subcellulare del virus che esso sia concentrato nel nucleo. Così i PICs non utilizzano il proprio DNA come bersaglio eseguendo un'autointegrazione suicida ad esempio già a livello del citoplasma. (88)

L'istone H1 ha un ruolo chiave nella compattazione della cromatina in una fibra trascrizionalmente inattiva attraverso la sua associazione con il DNA *linker* internucleosomiale. Molte osservazioni asseriscono che l'attivazione trascrizionale è associata alla cosiddetta "apertura" della cromatina, ossia uno srotolamento della fibra cromatinica attraverso una diminuzione della componente istonica. Il meccanismo che induce e mantiene tale impoverimento locale di istone H1 è stato studiato a livello di regioni associate allo *scaffold* (SAR) e coinvolge le proteine HMGI(Y). (89)

Le SAR sono sequenze di alcune centinaia di paia di basi con un contenuto maggiore del 70% di Adenine e/o Timine, che legano lo scaffold nucleare *in vitro*; sembrano essere elementi implicati nel delimitare strutturalmente le anse cromatiniche e nell'attivazione trascrizionale. Lo *scaffold* nucleare è una specie di impalcatura nucleare proteica che serve come sostegno strutturale per la cromatina. La specificità delle interazioni tra lo *scaffold* e la cromatina sembra essere mediata da proteine che riconoscono certe strutture di DNA non-B piuttosto che specifiche sequenze. Tale foggia potrebbe essere data dal solco minore di tratti ricchi in AT

incastonati nelle SAR e da curvature intrinseche del DNA. Infatti la distamicina, che lega specificamente il DNA ricco in AT, inibisce *in vitro* le interazioni tra le SAR e le proteine dello *scaffold*.

La topoisomerasi II e l'istone H1 sono due proteine che legano specificamente le SAR. Diverse unità di istone H1 si legano cooperativamente tra loro a formare dei veri centri di nucleazione che provocano in seguito l'assemblaggio di altri istoni H1 posti sulla cromatina fiancheggiante e danno inizio alla condensazione della cromatina. Nei stadi tardivi della compattazione cromatinica subentra la topoisomerasi II. (89)

Le SAR sembrano regioni coinvolte nella decondensazione cromatinica, o *chromatin opening*, in quanto titolandole con agenti come la distamicina si osserva uno spiazzamento degli istoni H1 ed un graduale aumento dell'accessibilità dei nucleosomi, oltre che un globale aspetto meno compattato della cromatina. Diverse evidenze sperimentali avevano implicato le proteine HMGI(Y) nell'apertura cromatinica, dato che sono associate a geni trascrizionalmente attivi e sono capaci di legare le SAR *in vitro*. Ma la scoperta essenziale è quella che le HMGI(Y) sono capaci di dereprimere la trascrizione inibita dall'istone H1, legandosi alle SAR. Ne risulta che l'istone H1, una volta spiazzato, si redistribuisce sulla cromatina ai lati delle SAR, dove prontamente è capace di reprimere la trascrizione di regioni non-SAR.

L'azione delle HMGI(Y) assomiglia molto a quella della distamicina, ma è circa 50 volte più efficace nella competizione con l'istone H1. Anche in questo caso un mutante di HMGI senza la coda COOH-terminale dereprime esattamente con la stessa efficienza. Una dimostrazione *in vivo* della redistribuzione dell'istone H1 in seguito alla presenza delle proteine HMGI è data dal fatto che queste ultime sono arricchite in subfrazioni cromatiniche impoverite di istone H1, dette S1. Inoltre si nota in generale un arricchimento di tutte le proteine della classe HMG in queste subfrazioni. Attraverso esperimenti di recupero delle proteine HMGI dalla frazione S1 è stato possibile calcolare che circa il 5% di tutta la cromatina assieme al 35% del contenuto totale di HMGI(Y) si ripartiscono in questa subfrazione, stabilendo che possiede un arricchimento di 7 volte delle proteine HMGI(Y) rispetto alla distribuzione generale nella cellula. (89)

Domini della cromatina contenenti geni trascrizionalmente attivi sono strutturalmente alterati a diversi livelli e sono più aperti o accessibili rispetto la cosiddetta cromatina "bulk". (18) Le maggiori modificazioni biochimiche della cromatina attiva includono un contenuto sub-stechiometrico di istone H1 e la presenza di lacune della fila di nucleosomi sulle regioni regolative di geni. L'inizio obbligato di queste modificazioni e della conseguente attivazione

trascrizionale sembra essere dato dalla decompattazione della cromatina trascrizionalmente silente attraverso lo spiazzamento degli istoni H1, che porta ad una locale apertura della cromatina. Si osserva infatti che le SAR localizzano generalmente in stretta associazione con elementi regolativi tessuto-specifici ad esempio in *Drosophila* e in molti geni di mammifero. (89)

Eccezioni notevoli sono i geni del “*cluster*” degli istoni delle proteine “heat-shock” che sono ubiquitari. Ma hanno una cosa in comune: sono espressi ad alti livelli durante uno stadio limitato del ciclo cellulare oppure in un numero ristretto di cellule durante lo sviluppo embrionale. Ciò suggerisce in definitiva che le SAR potrebbero servire nell’attivazione rapida nel tempo di questi geni iniziata dalle HMGI(Y) ed effettuata e/o mantenuta poi da fattori tessuto-specifici disponibili. (89)

Le regioni regolative al 5’ di geni costitutivi o “*housekeeping*” come anche di alcuni geni tessuto-specifici sono associate a regioni di DNA ricche in CpG non metilate. (90) La cromatina di queste isole ricche in GC è priva di istoni forse a conseguenza del fatto che H1 ha una bassa affinità per il DNA ricco in GC. Probabilmente in condizioni di quantità di istone H1 sub-stechiometriche la distribuzione degli istoni sarebbe sicuramente a “vantaggio” delle SAR. Quindi al contrario delle SAR le regioni CpG sono sempre aperte e necessitano invece di un meccanismo apposito, come ad esempio l’interazione con specifiche MeCPs (proteine metilanti le citosine a livello delle isole CpG) per reprimere la trascrizione di certi geni. (91)

Interessanti a questo riguardo sono gli studi di mappaggio del clusters genici delle  $\alpha$ - e delle  $\beta$ -globine umane, in quanto nel primo caso su 140 Kbase non vi è nemmeno una SAR, ma in compenso tutti i geni sono associati ad una isola CpG, mentre il cluster delle  $\beta$ -globine possiede ben otto SAR. Coerentemente la cromatina delle  $\alpha$ -globine è sempre aperta come risulta da esperimenti di sensibilità alla DNasi I, sia in cellule eritroidi che non, mentre la cromatina delle  $\beta$ -globine risulta aperta solo in cellule eritroidi. Queste considerazioni suggeriscono quindi che in cellule eucariotiche si sono sviluppate due strategie generali di apertura cromatinica, le SAR e le isole CpG. (89)

Una ricerca recentissima implica le proteine HMGI(Y) nella derepressione del promotore murino del gene per l’INF $\beta$  in seguito ad infezione virale. (92) Il promotore murino è molto simile a quello umano, estesamente descritto nel capitolo XY, tranne che a livello dell’elemento VRE, responsivo all’infezione virale, HMGI(Y) non lega, anche se esiste un sito di legame a bassa affinità. Vi è una regione a monte del VRE, ricca in AT alla quale è stato dimostrato lega l’istone H1, in parte responsabile della repressione trascrizionale del

gene. *In vitro* HMGI è stato capace di spiazzare l'istone H1 dalla regione ricca in AT nel promotore "wild-type", legandosi ad un sito ad alta affinità poche basi a monte del VRE. L'espressione di HMGI(Y) non ha effetto su mutanti del sito ad alta affinità per HMGI(Y), facendo supporre che HMGI(Y) non agisce attivando *Inf-β* direttamente, ma piuttosto dereprimendolo attraverso lo spiazzamento dell'istone H1 legandosi al sito di legame ad alta affinità per HMGI. (92) È importante notare che solo esperimenti di cotrasfezione stabili hanno permesso di stabilire la repressione da H1 e la derepressione di HMGI(Y) sul promotore *wild-type*, dato che nelle trasfezioni transienti il DNA trasfettato si presenta come cromatina aperta. (93)

Oltre a questo ruolo di fattore strutturale associato a strutture di cromatina attiva, in grado di influenzare l'espressione genica mediante interazioni con i nucleosomi e competizione con gli istoni linkers, va definendosi sempre di più l'evidenza di una funzione più specifica a livello del controllo trascrizionale.

### 1.2.5.2 HMGI e Controllo Trascrizionale

Come descritto nella sezione 1.1 recentemente è stata messa in luce l'esistenza di regioni regolative peculiari, la cui organizzazione combinatoriale consta di siti multipli riconosciuti non solo da specifici fattori trascrizionali, ma anche da componenti strutturali, quali HMGI(Y): in questi contesti l'azione delle due isoforme è risultata sempre del tutto analoga. Pur non essendo dotate di capacità transattivante (48) risultano essenziali per l'espressione in quanto mediano l'assemblamento di complessi tridimensionali stereospecifici a loro volta in grado di influenzare positivamente l'attività del complesso di trascrizione basale. Si è ipotizzato che questa complessa architettura, che prevede un preciso orientamento e posizionamento reciproco fra i siti componenti, potrebbe essere tipica di promotori inducibili: infatti fornisce un mezzo di integrazione di multiple vie di trasduzione del segnale innescate da stimoli diversi, visto che l'attivazione risulta solo in seguito all'induzione coordinata di tutti i fattori coinvolti. (94)

In questo contesto l'azione di HMGI(Y) prevede interazioni specifiche sia con sequenze di DNA, che con altri fattori, con un duplice risultato: innanzitutto HMGI(Y) è in grado di facilitare il legame di specifici fattori trascrizionali a livello dei rispettivi siti, e in secondo

luogo è in grado di mediare l'instaurarsi di una serie di interazioni cooperative tra diversi complessi di proteine transattivanti, necessarie alla risultante sinergia trascrizionale. Questo secondo obiettivo viene raggiunto non solo attraverso interazioni specifiche con i fattori coinvolti, ma anche modulando l'architettura intrinseca di queste regioni regolative, creando l'orientamento favorevole all'assemblaggio di uno specifico complesso nucleoproteico di ordine superiore. (59)

Sebbene la schiera di promotori ed enhancers con questo tipo di organizzazione vada progressivamente allargandosi, il sistema meglio caratterizzato riguarda l'enhancer per il gene umano dell'interferone  $\beta$ .

### **Regolazione del gene per IFN $\beta$**

L'IFN $\beta$  appartiene ad un gruppo di polipeptidi secreti che proteggono le cellule da infezioni virali. La sua espressione è tipica dei fibroblasti. Sebbene di norma non sia rintracciabile, il suo trascritto si riscontra in elevate quantità entro 4-8 ore dall'infezione virale o a seguito dell'esposizione a RNA a doppio filamento, tipico di alcune infezioni virali. L'accumulo di questo mRNA è in gran parte dovuto all'aumento del tasso di trascrizione indotto dall'infezione virale stessa, finemente regolata, come descritto qui di seguito:

Le sequenze regolative importanti per la trascrizione si estendono dalla posizione -204 alla +1: oltre all'elemento TATA-box vi sono due regioni che controllano l'inducibilità virale.

La prima regione è costituita da due elementi con azione negativa (NDR1 e NDR2), in quanto mutazioni al loro livello provocano l'espressione costitutiva del gene; da esperimenti di protezione alla metilazione *in vivo* è evidente che vengono occupati solo durante lo stato non indotto da molecole specifiche, che quindi fungono da repressori.

La seconda regione è data dall'enhancer che media l'induzione virale. Ha dimensioni ridotte di appena 50 pb ed è costituito da un insieme di 4 domini a regolazione positiva (PRDI-IV) parzialmente sovrapposti. Questi domini sono inattivi presi singolarmente ed in copia singola, in combinazione di due o più copie formano enhancers inducibili non solo da virus, ma anche dai vari stimoli in grado di attivare il fattore trascrizionale coinvolto. Nella regolazione dell'inducibilità virale dell'IFN $\beta$  umano i quattro elementi funzionano sinergicamente, infatti con la mutazione di uno solo di essi si verifica un calo drammatico dell'induzione virale. (95) Queste osservazioni non solo hanno fornito la base sperimentale per poter semplificare questo sistema allo scopo di caratterizzarlo completamente, ma sottolinea l'importanza

dell'organizzazione strutturale di questo enhancer. (94) Infatti è stata inequivocabilmente dimostrata per l'attività trascrizionale la necessità di precise posizioni relative e di corretta fase sull'elica fra i vari PRDs, a costituire un'unità funzionale in grado di garantire l'induzione altamente selettiva di questo gene, a partire da fattori che individualmente hanno specificità piuttosto bassa.

Fra questi elementi i meno caratterizzati sono PRDI e PRDIII, che presumibilmente interagiscono con un fattore detto IRF-1 (INF regulating factor); al contrario l'interesse concentratosi sugli altri due elementi ha consentito di sviscerarne a fondo i meccanismi d'azione, assegnando ad HMGI(Y) un ruolo centrale nell'attivazione trascrizionale di entrambi e quindi nell'induzione virale.

**PRDII** È un sito composito, riconosciuto dal fattore NF- $\kappa$ B, al cui centro contiene una sequenza legata da HMGI(Y). Un'altra sequenza canonica per NF- $\kappa$ B è quella a livello dell'enhancer del gene per la catena leggera delle immunoglobuline, denominata Ig- $\kappa$ B. Questa però non può essere legata da HMGI(Y).

Il fattore NF- $\kappa$ B è costituito da omodimeri od eterodimeri di membri di una famiglia di fattori trascrizionali, denominata *Rel*. (94) Questo potente attivatore trascrizionale si trova virtualmente in tutti i tipi cellulari in forma inattiva citoplasmatica e viene traslocato, per induzione specifica da parte di stimoli piuttosto eterogenei, a livello nucleare. Nel nucleo controlla una vasta schiera di geni e ciò fornisce indicazioni interessanti su come risposte diverse a differenti segnali possano essere mediate dall'espressione soltanto di uno specifico sottogruppo di geni bersaglio.

HMGI(Y) ed NF- $\kappa$ B sono necessari, anche se non di per sé sufficienti, all'induzione virale. Ciò si è dimostrato attraverso mutagenesi sito specifica e metodologie antisense. Ad esempio mutazioni puntiformi a livello della sequenza PRDII, che *in vitro* abbassano selettivamente l'affinità di HMGI(Y) senza alterare quella di NF- $\kappa$ B, oppure sostituendo con la sequenza per Ig- $\kappa$ B, si determina invariabilmente un drastico calo dell'induzione virale. Con questi esperimenti è stato possibile evidenziare oltre alla sinergia funzionale *in vivo*, anche la capacità di HMGI(Y) di interagire direttamente con NF- $\kappa$ B e di stimolarne il legame su PRDII. A tal proposito è stata fornita una possibile spiegazione meccanicistica della cooperatività del legame a questo livello da parte di NF- $\kappa$ B ed HMGI(Y). (59) Da analisi di fase e studi di permutazione circolare, è emerso che l'elemento PRDII mostra un ripiegamento intrinseco più pronunciato rispetto al sito canonico di Ig- $\kappa$ B. Come descritto nella sezione 1.2.3. HMGI(Y), legandosi nel solco minore, riduce tale ripiegamento per facilitare

l'interazione di NF- $\kappa$ B sul solco opposto. È stato dimostrato inoltre che questo effetto è specifico per gli eterodimeri p50/RelA, che agiscono da attivatori ottimali in questo contesto. Questa forma di facilitazione differenziale mediata da HMGI(Y), che consente di selezionare specifici membri all'interno di famiglie di fattori trascrizionali, sembra essere un meccanismo piuttosto generalizzato.

Malgrado siano conosciuti pochissimi geni bersaglio diretti di HMGI-C è stata dimostrata una omologia funzionale della stessa con la HMGI(Y). Infatti anche HMGI-C è in grado di facilitare l'assemblaggio di un complesso nucleoproteico con NF- $\kappa$ B a livello di PRDII. (96 e sezione 4.1.2.)

**PRDIV** Anche questo è un elemento composito, costituito da un sito riconosciuto da omodimeri della famiglia ATF-2 (Activator Transcription Factor 2) o eterodimeri ATF-2/c-jun, ai lati del quale vi sono due siti per HMGI(Y). Anche a questo livello HMGI(Y) esplica una funzione analoga di fattore accessorio nei confronti di ATF-2, attraverso interazioni specifiche proteina-proteina con la regione bZip di ATF-2, con il duplice risultato di favorirne il legame al rispettivo sito e di stimolarne la dimerizzazione. (97) In questo contesto appare ancora più inequivocabile il ruolo di HMGI(Y) nel selezionare i membri specifici coinvolti nel legame col DNA, dato che è stato dimostrato che ha azione opposta su due isoforme di ATF-2 prodotte per splicing alternativo. Mentre ha un effetto di facilitazione nei confronti dell'isoforma più pesante, ATF-2<sub>(195)</sub>, l'interazione di PRDIV con la forma più leggera, ATF-2<sub>(192)</sub> priva della regione che interagisce non solo con HMGI(Y) ma anche con NF- $\kappa$ B, viene selettivamente contrastata da interferenze steriche con HMGI(Y).

L'effetto di HMGI(Y) sull'attivazione trascrizionale non riguarda soltanto i singoli elementi, ma permette anche ai fattori IRF-1 ed NF- $\kappa$ B di occupare contemporaneamente i rispettivi siti sui corrispondenti PRD, un fatto che non può verificarsi in assenza di HMGI(Y) nemmeno *in vitro*. Il risultato di questi molteplici e complessi effetti mediati da HMGI(Y) permette una fine modulazione dell'architettura dell'enhancer: da una configurazione intrinseca inattiva, viene ad assumere un'orientamento peculiare tale da favorire le interazioni proteina-proteina e proteina-DNA richieste all'assemblamento e stabilizzazione di un complesso multiproteico stereospecifico, denominato *enhanceosoma*, in risposta ad uno stimolo specifico come l'infezione virale. (94) Quindi, l'assemblaggio cooperativo dell'enhanceosoma fornisce un meccanismo di specificità che assicura che il gene per l'INF $\beta$  venga attivato da infezione virale e non da qualsiasi altro induttore capace di stimolare singolarmente NF- $\kappa$ B, IRF-1

oppure ATF-2/c-jun. Il meccanismo con cui i domini di attivazione di alcuni o tutti fattori trascrizionali dell'enhanceosoma promuovono l'inizio della trascrizione da parte della RNA polimerasi II è oggetto di studi recentissimi ed è riportato qui di seguito.

Come descritto nella sezione 1.1. l'attivazione trascrizionale inizia con il reclutamento dell'apparato basale della trascrizione su di un promotore altrimenti silente. Inoltre fattori trascrizionali possono interagire in modo sinergico tra di loro ed il macchinario basale al fine di generare un aumento ingente della trascrizione genica. Sinergia trascrizionale può risultare da contatti multipli e simultanei tra domini di attivazione ed alcuni componenti dell'apparato basale della trascrizione. Tali interazioni possono essere o diretti o indiretti. In quelli indiretti i domini di attivazione affidano i loro effetti stimolatori a proteine intermediarie che funzionano come coattivatori della trascrizione. Nel caso dell'*enhanceosoma* dell'INF $\beta$  i livelli trascrizionali generati da tutti gli attivatori risulta essere maggiore rispetto alla somma dei contributi dei fattori presi individualmente, fenomeno definito appunto come sinergia trascrizionale. Due dei co-attivatori meglio caratterizzati ad oggi, la proteina CBP (*CREB binding protein*) ed il polipeptide strettamente relato p300 (12), vengono efficientemente reclutati dall'enhanceosoma attraverso contatti proteina-proteina multipli, ottenendo attivazione sinergica della trascrizione. Per questo fenomeno è indispensabile che ci sia un'interazione proteina-proteina specifica tra il cosiddetto dominio di interazione di p65 che funziona in modo cooperativo con i domini di attivazione appropriatamente posizionati di p65, IRF-1 ed ATF-2/c-Jun per reclutare il coattivatore CBP/p300 sull'enhanceosoma.

Esperimenti di interazioni proteiche hanno dimostrato che i domini di attivazione di IRF-1, c-Jun e p65 possono interagire individualmente con il frammento amminotermale di CBP. Ma attraverso esperimenti di cotrasfezione usando diversi costrutti di mutazione del promotore dell'INF $\beta$ , in cellule infettate da virus e non, si è dimostrato che CBP viene reclutato a livello del promotore attraverso l'enhanceosoma che funziona come una unità, e non dai singoli attivatori separatamente. Si dimostra quindi che l'*enhanceosoma* correttamente assemblato funziona come una nuova superficie di attivazione e di legame per il complesso CBP/p300, data dall'insieme dei domini di attivazione dei singoli fattori trascrizionali posizionati correttamente nell'*enhanceosoma* e non diversamente. (98) Per ottenere sinergia trascrizionale però non basta aver reclutato CBP/p300 a livello dell'*enhanceosoma*, servono ancora diversi contatti con l'apparato basale della trascrizione. Infatti sono state dimostrate interazioni tra IRF-1 e p65 con diversi bersagli del macchinario basale della trascrizione come TFIID, TFIIA, TFIIB ed il coattivatore USA, (99) ma soprattutto serve CBP/p330, considerato un



componente dell'oloenzima della RNA polimerasi II, il quale, a sua volta, interagisce mediante il suo carbossi-terminale con due subcomplessi già presenti a livello del promotore, ossia l'oloenzima dell'RNA polimerasi II ed il complesso TFIID/A/B. Sembra che il reclutamento stesso di TFIID/A/B venga facilitato da CBP, perché è stato dimostrato che quest'ultimo interagisce anche con TBP (100) o con TFIIB (101), ma soprattutto modifica i nucleosomi acetilando gli istoni (102). Interazioni proteina-proteina tra diversi componenti dei differenti complessi potrebbero stabilizzare le interazioni complessive col DNA e promuovere l'inizio della trascrizione. Sinergia trascrizionale è il risultato netto dell'assemblaggio cooperativo dei due subcomplessi in un *transcriptosoma* (103) che simultaneamente elimina gli effetti negativi della cromatina ed inizia efficientemente la trascrizione. (98) Da tutto ciò risulta evidente l'importanza fondamentale della funzione architetturale di HMGI(Y) che assicura un corretto posizionamento di tutti gli elementi e un'efficiente cooperatività tra gli stessi a livello dell'*enhanceosoma*.



## Altri geni regolati da HMGI(Y)

Il fattore HMGI(Y) è implicato nella regolazione di un numero sempre crescente di geni sia cellulari che virali. È importante sottolineare che oltre all'effetto di modulazione positiva della trascrizione, come descritta per l'enhancer dell'INF $\beta$  e valida per molti altri elementi di controllo trascrizionale (Tabella III), questo stesso fattore è stato implicato anche nella repressione trascrizionale (Tabella IV).

### Controllo positivo:

Tabella III. Geni attivati da HMGI(Y).

Geni	elemento in cis	elemento in trans	interazioni proteina-proteina con HMGI(Y)	referenza
INF $\beta$	PRDII	NF- $\kappa$ B (p50/p65)	NF- $\kappa$ B p50 HMGI aa 50-65	104, 105
	PRDIV	ATF-2/c-Jun	ATF-2 (bZip)	106
Inf $\beta$	VRE			92
Selectina E	PDIII e PDIV	NF- $\kappa$ B (p50/p65)	p50	107
MGSA/GRO		NF- $\kappa$ B (p50/p65)	n.d.	108
IL-2	CD28RE	NF- $\kappa$ B (cRel/p65)	n.d.	109
GM-CSF	CK-1	NF- $\kappa$ B (cRel/p65)	n.d.	109
IL-3	CK-1	NF- $\kappa$ B (cRel/p65)	n.d.	109
IL-2R $\alpha$	PRRI	Elf-1	Elf-1 (ets-domain)	110
IL-2 bovina	(AUAA) <sub>n</sub> 3'UTR			111
Ly-6A/E	Distal element	STAT 1	n.d.	112
HLA-DRA	tratto ricco in AT	Oct-2	n.d.	113
TNF- $\beta$	UAS		n.d.	114
Papovavirus	tratto ricco in AT	Tst-1/Oct-6	Tst-1/Oct-6 POU-domain HMGY aa 46-56	115
Herpes virus I	LAP-2	Sp1	n.d.	116
Herpes virus I	IE-3 (3 elementi)	ICP 4	n.d.	58
iNOS	Oct, NF- $\kappa$ B	NF- $\kappa$ B (p50/p65)	n.d.	117

Del tutto analoga all'enhancer per IFN $\beta$ , appare l'organizzazione del promotore per la selectina E, (107) una molecola transientemente espressa dall'endotelio capillare, per induzione da parte dell'interleuchina 1 (IL-1) e del fattore di necrosi tumorale di tipo  $\alpha$  (TNF  $\alpha$ ), per mediare la diapedesi dei neutrofili e dei linfociti verso i siti di infiammazione acuta o cronica.

Anche in questo caso l'espressione richiede l'azione sinergica di quattro domini regolatori positivi (PDI→PDIV), che interagiscono con NF- $\kappa$ B e ATF-2. In particolare, mentre PDI contiene un sito canonico per NF- $\kappa$ B, le regioni contigue PDIII e PDIV sono elementi compositi che legano sia NF- $\kappa$ B che HMGI(Y) (molto simili quindi al sopra citato PRDII); pur essendo entrambe richieste per l'induzione dell'espressione, in vitro non possono venire contemporaneamente occupate da due molecole di NF- $\kappa$ B. In vivo ciò è reso possibile dalle alterazioni conformazionali indotte in seguito all'interazione con HMGI(Y), che rendono accessibile anche il sito adiacente non occupato, stabilizzando la formazione di un complesso proteico ternario a livello di questi due elementi. Un ulteriore sito per HMGI(Y) si trova presso PDII e, analogamente a quanto descritto per l'elemento PRDIV, l'interazione a questo livello facilita anche il legame di ATF-2: così HMGI(Y) attraverso interazioni specifiche sia con il DNA che con fattori trascrizionali specifici è in grado anche in questo caso di mediare l'assemblamento e la successiva stabilizzazione di complessi nucleoproteici attivi trascrizionalmente.

Sono stati inoltre caratterizzati elementi analoghi ai suddetti PRDII e PDIII e IV, in altri promotori immunologicamente rilevanti.

In primo luogo, un sito variante per NF- $\kappa$ B simile è presente nelle regioni che regolano l'espressione della chemochina MGS/GRO (108) e del gene per l'interleuchina 8 (IL-8), molecole importanti nella chemiotassi dei neutrofilo nel contesto sia di fenomeni infiammatori che tumorigenici, dato che vengono iperespressi nella progressione neoplastica dei melanociti a melanomi maligni. Infatti è stato dimostrato che l'integrità di questo elemento riconosciuto da NF- $\kappa$ B, con sovrapposto un sito per HMGI(Y), è essenziale per l'espressione basale e indotta da citochine; inoltre è stata ipotizzata la cooperazione nella regolazione trascrizionale con un elemento a monte, riconosciuto da Sp1, presumibilmente attraverso la formazione di un complesso ternario attraverso interazioni dirette tra fattori proteici e modificazioni della curvatura del DNA, promosse da HMGI(Y).

Inoltre è stata recentemente evidenziata l'importanza di analoghi siti varianti per NF- $\kappa$ B nell'attivazione dei linfociti T, presenti in geni di diverse citochine, quali l'interleuchina 2 e 3 (IL-2 e IL-3) ed il fattore stimolatore delle colonie dei granulociti e macrofagi (GM-CSF) (Himes Immunology 96), fondamentali nella successiva risposta immune cellulo-mediata. Infatti nei promotori di questi geni, sono stati identificati siti per NF- $\kappa$ B funzionalmente differenti: un sito canonico e un sito variante (denominato CD28-RE nel contesto del promotore dell'IL-2 e CK-1 in quello per GM-CSF e per IL-3) sono coinvolti nell'attivazione

da parte di diversi stimoli. A questo livello HMGI(Y) non si limita ad aumentare l'affinità di legame di NF- $\kappa$ B, ma è stata anche coinvolta nella capacità di rispondere alle differenti forme di induzione, che appare correlata alla formazione di diversi complessi di dimeri fra membri della famiglia Rel: sul sito canonico si lega la forma classica di NF- $\kappa$ B (l'eterodimero p65/p50), mentre sul sito variante l'eterodimero cRel/p65. Infatti è stato dimostrato che HMGI(Y) concorre nella regolazione della dimerizzazione e del legame selettivo delle subunità, in quanto esercita un effetto specifico nell'interazione con il DNA soltanto di cRel, coinvolto soltanto a livello di CD28-RE e CK-1.

Questa forma di regolazione dell'espressione di IL-2 che coinvolge HMGI(Y) può essere legata alla stretta correlazione tra l'espressione di HMGI(Y) e la proliferazione cellulare, visto che è stato dimostrato che questo stesso fattore coopera nell'attivazione trascrizionale del promotore della catena  $\alpha$  del recettore per l'IL-2 (IL-2R $\alpha$ ) (110): ciò indica che HMGI(Y) riveste un ruolo fondamentale nel sostenere la stimolazione autocrina dei linfociti T attivati, via sintesi coordinata di IL-2 e del suo recettore ad alta affinità. A livello del promotore per il gene di IL-2R $\alpha$  sono stati individuati due elementi regolativi positivi (PRRI e PRRII). Mentre PRRI, costituito da motivi CArG per NF- $\kappa$ B, è coinvolto nell'induzione rapida in risposta a stimoli mitogenici, PRRII è essenziale per l'espressione basale cellula specifica di questo gene: contiene siti riconosciuti da HMGI(Y) e da Elf-1, un membro della famiglia Ets, la cui espressione è tipica delle cellule linfoidei. Al contrario di quanto visto a proposito dell'enhancer per IFN $\beta$ , non sembra che HMGI(Y) determini un'accresciuta affinità di legame di Elf-1 per il suo sito, anche se è stata evidenziata l'esistenza di una sinergia funzionale tra questi due fattori. Questo effetto dipende dalla capacità dei due fattori di interagire anche in assenza di DNA: dato che è stato dimostrato che il dominio ets è necessario e sufficiente a quest'interazione, si può ipotizzare che HMGI(Y) legandolo possa mantenerlo in una conformazione trascrizionalmente attiva. Oltre a ciò, anche in questo contesto HMGI(Y) potrebbe contribuire alla coordinazione trascrizionale fra PRRI e PRRII, elementi ampiamente distanziati, alterando la curvatura del DNA per facilitare interazioni proteina-proteina e/o mediandole direttamente, fungendo da "colla molecolare", dato che la capacità di Elf-1 di legare p50 e cRel è accresciuta in presenza di HMGI(Y).

Sempre nel contesto di risposte a stimoli infiammatori si colloca anche il gene iNOS, il cui prodotto, l'ossido nitrico sintasi inducibile, catalizza la formazione dell'ossido nitrico, un gas vasodilatatore importante per la regolazione di sistemi di diversi organi. L'isoforma inducibile dell'ossido nitrico sintasi è trascrizionalmente regolata da stimoli infiammatori, di cui un

mediatore critico è il fattore NF- $\kappa$ B. Diversamente dagli altri promotori il sito per NF- $\kappa$ B non si sovrappone con quello di HMGI(Y), il quale invece si trova a 15 pb a valle del sito per NF- $\kappa$ B, a livello di una sequenza ottamero, sequenza consenso per fattori di trascrizione della famiglia Oct. Tale sequenza, ricca in AT, è risultata essenziale per l'inducibilità di iNOS attraverso stimolazione dell'IL-1 o di LPS in cellule vascolari muscolari lisce, e da LPS in macrofagi. Tale sito viene legato da proteine Oct-2-simili dopo attivazione dei macrofagi con IL-6 oppure da proteine Oct-1-simili in risposta al contatto con LPS. Ma in nessuno di questi studi è stata dimostrata cooperazione con NF- $\kappa$ B. Il ritrovamento che il fattore HMGI(Y) lega tale sito potenziando la transattivazione del gene iNOS da parte di NF- $\kappa$ B, apre un nuovo orizzonte nella comprensione della regolazione del gene iNOS. (117)

Anche questi dati implicano le HMGI(Y) nella regolazione specifica di geni indotti da citochine. Ma proprio nell'esempio del gene iNOS si è visto che in seguito allo stimolo iniziale (IL-1 $\beta$ , LPS) nelle cellule endoteliali si ha un aumento della sintesi delle proteine HMGI(Y) che non è dovuto all'aumento della stabilità della proteina o del messaggero, ma ad un incremento dell'attività del promotore. Infatti, diversamente da ciò che è stato descritto per le HMGI(Y) nel contesto del gene umano *INF $\beta$* , lo stimolo da parte di una citochina influenza innanzitutto la trascrizione di *HMGI(Y)*. (118).

Gli interferoni costituiscono una famiglia di citochine che regolano diverse funzioni cellulari come ad esempio attività antivirali, antiproliferative ed immunoregulative. La maggior parte di queste attività vengono compiute attraverso prodotti di geni trascrizionalmente attivati dagli interferoni. Tali geni spesso contengono 2 tipi di regioni che rispondono a tale stimolazione, denominati ISRE (Interferon stimulated Response Element) e GAS (Interferon- $\gamma$  activation site), a livello dei quali, a seconda del tipo di citochina viene a formarsi un complesso multifattoriale specifico. Spesso la funzione immunoregolutiva degli interferoni porta all'induzione di molecole di superficie cruciali per le interazioni intercellulari, come ad esempio le molecole MHC di istocompatibilità di classe I e/o II a livello di cellule T ed APC (Antigen Presenting Cell). Esistono però tutta una serie di fattori accessori richiesti ad esempio per l'attivazione dei linfociti T, come l'antigene Ly-6A/E, antigene di superficie espresso sulla superficie cellulare in momenti cruciali durante l'ematopoiesi e la risposta immunitaria. L'analisi estesa del promotore di Ly-6A/E ha evidenziato due elementi, uno distale e l'altro prossimale, essenziali per l'inducibilità dall'INF- $\alpha/\beta$  in cellule B. L'elemento distale è inserito nel contesto di un'ampia regione regolativa costituita da un ISRE al 5', richiesto per la risposta ad entrambi i tipi di interferone, un nuovo elemento al 3', importante solo per

l'inducibilità dall'INF- $\gamma$ , e, nel mezzo, un sito di legame per HMGI(Y). Dato che i due elementi richiesti per l'inducibilità da parte dell'INF- $\alpha/\beta$  (distale e prossimale) fisicamente distano più di 2 Kbasi l'uno dall'altro nel promotore di Ly-6A/E, è stata proposta una interazione cooperativa tra le proteine che legano questi elementi mediata da HMGI(Y). Tale ipotesi è avvalorata da esperimenti di trasfezione con l'antisense di HMGI(Y) nei quali viene abolita specificamente l'inducibilità del promotore di Ly-6A/E da parte dell'INF- $\alpha/\beta$ , mentre invece viene preservata l'inducibilità da parte dell'INF- $\gamma$ , escludendo così qualsiasi dubbio su eventuali effetti non specifici dell'antisense nel contesto dell'espressione del gene per Ly-6A/E. (112)

Un altro gruppo di fattori, la cui attività trascrizionale viene influenzata da HMGI(Y), è costituito da membri della famiglia Oct (octamer binding proteins), caratterizzata da proteine in grado di legare il DNA, a livello di specifiche sequenze ottameriche, per mezzo di un dominio peculiare denominato POU domain, e dotate di un dominio transattivante piuttosto debole, solitamente potenziato da coattivatori.

In primo luogo, caratterizzando le regioni regolative del gene HLA-DRA, (molecola componente di MHCII) espressa costitutivamente in linfociti B, in linfociti T attivati e in APC e indotta da IFN $\gamma$  in un'ampia gamma di tipi cellulari, è emerso un elemento del tipo ottamerico aggiuntivo, a monte del promotore prossimale che regola tutti i geni per le componenti dell'MHCII. Complessivamente queste regioni presentano siti multipli per HMGI(Y), che sembra coinvolta nell'assemblamento di due distinti complessi trascrizionali: uno necessario all'espressione costitutiva nei linfociti B, l'altro indotto da IFN $\gamma$ . (113)

Da una parte l'espressione di HLA-DRA indotta da questa citochina richiede l'integrità del promotore prossimale, che contiene 5 siti per HMGI(Y), alcuni dei quali coincidono con elementi in cis riconosciuti da altri fattori trascrizionali o componenti del macchinario di trascrizione basale. L'osservazione che sono necessarie mutazioni a livello di più di uno di questi siti per bloccare la risposta allo stimolo dell'IFN $\gamma$ , suggerisce che il legame di molecole multiple a questo promotore può essere in parte ridondante e perciò ricoprire la funzione generale di facilitare l'assemblamento di complessi trascrizionali attivi.

D'altra parte nell'espressione costitutiva di questo gene nei linfociti B un ruolo critico è ricoperto dall'elemento aggiuntivo sopracitato, che consiste nella sequenza ottamerica, riconosciuta dal fattore Oct-2A (la cui espressione è tipica del tessuto linfoide), con sovrapposto un sito per HMGI(Y). E' stato dimostrato in tale contesto il ruolo essenziale di HMGI(Y), che non solo agisce sinergicamente con Oct-2A *in vivo* e *in vitro* stimolandone il

legame al rispettivo sito, ma è anche in grado di discriminare fra diversi membri di questa famiglia, fungendo da coattivatore specifico per Oct-2A. Infatti si è osservata una facilitazione differenziale da parte di HMGI(Y) nei confronti di Oct-2A rispetto a Oct-1 (la cui espressione è ubiquitaria): anche se il legame con HMGI(Y) coinvolge il POU domain, a determinare il risultato di questa interazione sul legame con la sequenza ottamerica è il dominio C terminale variabile dei fattori Oct. La possibilità di HMGI(Y) di reclutare selettivamente Oct-2A a questo livello fornisce la base dell'attivazione cooperativa tessuto-specifica di questo gene e, in generale, suggerisce un ruolo nel dare origine a risposte specifiche, agendo selettivamente nei confronti di particolari membri all'interno di famiglie di fattori trascrizionali.

Un altro sistema in cui è stata messa in luce la cooperazione tra HMGI(Y) e un membro della famiglia di proteine Oct, consiste nel promotore del Papovavirus JC, (115) un virus che infetta specificamente le cellule gliali del sistema nervoso centrale di individui immunodepressi provocando una grave demielinizzazione nota come leucoencefalopatia multifocale progressiva. Parte del tropismo di questo virus è stata spiegata dalla necessità del fattore Tst-1/Oct-6, la cui espressione è caratteristica di alcuni tipi cellulari fra cui specifiche subpopolazioni neuronali; infatti l'attivazione dei geni virali dipende dal legame di questo fattore su di un sito ricco in AT del promotore virale. Anche in questo caso è stato dimostrato il ruolo sinergico di HMGI(Y) nel facilitarne il legame in vitro e di cooperare nell'attivazione trascrizionale in vivo.

Recentemente è stato evidenziato un altro caso in cui HMGI(Y) risulta implicata nella regolazione dell'espressione virale: uno dei promotori attivi durante la fase di infezione latente dell'Herpes simplex di tipo 1 (HSV1), denominato LAP2, contiene, oltre a due siti per Sp1, un tratto costituito da 23 residui T, al cui livello interagisce una proteina antigenicamente relata a HMGI(Y). (116) Inoltre attraverso tecniche antisenso ne è stato dimostrato l'effettivo coinvolgimento nell'espressione virale. È probabile che anche a questo livello HMGI(Y) possa stimolare il legame di una serie non ancora caratterizzata di fattori trascrizionali ai rispettivi siti e influenzarne l'orientamento reciproco a formare complessi trascrizionalmente attivi; una prima indicazione in questo senso deriva dalla dimostrazione che HMGI(Y) è in grado di stimolare il legame di Sp1 ai siti sopraccitati, proteggendoli dalla metilazione e quindi dalla derivante repressione.

L'attività di un altro promotore del virus HSV1 risulta essere modulata dall'intervento delle proteine HMGI(Y) della cellula ospite infettata: il promotore del gene IE-3, appartenente alla classe dei geni precocissimi risulta represso da ICP4, il regolatore trascrizionale più



importante di HSV1. È una proteina che lega il DNA e agisce come repressore dei geni precocissimi, mentre attiva quelli precoci e tardivi. Dato che sono state dimostrate interazioni tra ICP 4 e le proteine TBP,TFIIB e la TAF 250, si ipotizza che ICP agisca inficiando l'assemblamento del complesso di preinizzazione della trascrizione. (58) ICP4 infatti riconosce e si lega a livello del sito di inizio della trascrizione del gene IE-3, reprimendolo. *In vitro* anche HMGI(Y) lega lo stesso sito e saggi *in vivo* dimostrano che potenzia l'effetto di repressione di ICP4. Ma l'effetto globale delle HMGI(Y) come modulatori della trascrizione di questo gene non sono quelli di repressione. Infatti legano altre due regioni dello stesso promotore, ossia la TATA box a -21/-26 e più a monte un sito che si sovrappone al motivo TAATGARAT a -106/-96 e probabilmente attraverso tali interazioni multiple e cooperative si ottiene un effetto trascrizionale positivo. È probabile che si ottiene un'architettura del promotore tale da reclutare e riunire fattori che legano sequenze regolative a monte.

Studiando l'effetto trasformante dell'infezione a carico di precursori di linfociti B da parte del virus della leucemia di Abelson, HMGI(Y) risulta implicato nella derivante espressione costitutiva del fattore di necrosi tumorale di tipo  $\beta$  (TNF- $\beta$ ), rigidamente controllata in condizioni normali. Infatti il legame di questo fattore a livello di una regione regolativa, denominata UAS (upstream activating sequence) appare necessario, anche se non sufficiente, all'attivazione trascrizionale, probabilmente attraverso meccanismi simili a quelli precedentemente descritti.

### **Controllo negativo:**

Nonostante quanto riportato finora, la proteina HMGI(Y) non è solo in grado di regolare positivamente l'attività trascrizionale dei promotori; si stanno infatti assommando diversi casi (Tabella III) in cui può fungere da repressore attraverso molteplici meccanismi, che vanno da un semplice impedimento sterico all'accesso di fattori trascrizionali chiave ai rispettivi siti, ad una modulazione più fine, che richiede interazioni specifiche fra proteine.

Infatti ricordando quanto detto a proposito della capacità di HMGI(Y) di favorire selettivamente il legame di specifici membri di una data famiglia di fattori trascrizionali, è possibile che la differenza tra attivazione e repressione risultanti possa fondarsi su come viene influenzato il legame di altre proteine a un dato promotore.

Tabella IV: Geni repressi da HMGI(Y).

Geni	elementi in cis	elementi in trans	interazioni proteina-proteina con HMGI(Y)	referenza
IL-4	tratto ricco in AT	NF-AT	n.d.	118, 119
Ige germ line	tratto ricco in AT	NF-BRE	n.d.	121
gp-91 phox	tratto ricco in AT CPI		n.d.	120

Di particolare interesse a questo riguardo è il ruolo dimostrato per HMGI(Y) nel controllo dell'espressione di diversi gruppi di citochine in sottopopolazioni differenti di linfociti T; infatti nei T helper di tipo 1 non partecipa soltanto all'attivazione trascrizionale del gene per IL-2, ma anche alla repressione dell'espressione di IL-4. (118, 119) In particolare è stato caratterizzato il forte enhancer che regola l'espressione cellulo-specifica di IL-4: dei 6 elementi costituenti, 4 sono riconosciuti da NF-AT, fattore specifico dei linfociti T coinvolto anche nell'espressione di IL-2; uno di essi si sovrappone con un sito forte per HMGI(Y) (costituito da un tratto di 12 AT). A questo livello però è stato dimostrato un ruolo di competizione mediato da HMGI(Y). Una possibile base per questo effetto opposto, può risiedere nel fatto che il legame di NF-AT con il DNA si estende parzialmente anche nel solco minore e la competizione potrebbe derivare da interazioni comuni con HMGI(Y) a questo livello.

Inoltre è stato rilevato che HMGI(Y) è in grado di reprimere l'espressione di questo gene con un meccanismo aggiuntivo, senza influenzare il legame di altri fattori: si ipotizza che possa modulare l'architettura di questo promotore, impiegando più di uno dei suoi *AT-hooks* nell'interazione con diversi siti. Un meccanismo analogo potrebbe mediare anche la repressione del gene per IL-5, dato che il suo promotore contiene simili tratti AT lunghi e contigui.

In studi recenti è stato evidenziato un effetto negativo anche sul promotore per l'RNA *germ line* per la catena pesante  $\epsilon$  delle immunoglobuline, la cui trascrizione correla con il tasso di *switching* isotipico in risposta a stimolazione con IL-4 e con lipopolisaccaride (LPS). (121) Questo promotore contiene due elementi importanti, fra cui un sito ricco in AT riconosciuto dal fattore NF-BRE, indotto da IL-4. Paradossalmente mutazioni a questo livello ne

incrementano la trascrizione: ciò può derivare dal coinvolgimento di HMGI(Y), che è stato dimostrato con tecniche antisense.

Un ruolo analogo potrebbe essere implicato anche nella regolazione negativa del gene *gp-91 phox*, componente di un complesso necessario all'attività antimicrobica tipica dei fagociti differenziati. (120) Infatti si è visto che HMGI(Y) è in grado di legare in modo specifico i due CCAAT-box di questo promotore, analogamente al fattore CDP (CCAAT box displacement factor) che ha azione negativa sull'espressione in cellule immature; in seguito al differenziamento il livello di entrambe queste proteine è considerevolmente ridotto e ciò permette l'interazione con un attivatore trascrizionale ubiquitario CP1 (CCAAT binding factor).

### **Interazioni proteina-proteina**

È importante prestare attenzione anche alle interazioni specifiche che le proteine HMGI realizzano con altri fattori di trascrizione. Infatti nella maggior parte dei casi il contributo delle HMGI è quello di facilitare il legame di quest'ultimi al loro sito di riconoscimento sul DNA, non solo aggiustando la conformazione del loro sito di riconoscimento, ma anche contattandoli direttamente.

Un esempio a tale riguardo è quello del fattore SRF o *serum response factor*, fattore di trascrizione appartenente alla famiglia delle proteine con una MADS-box, che regola l'attività di vari promotori, fra cui anche quelli del gene *c-fos* e *SM22 $\alpha$* , specifico della muscolatura liscia, legandosi all'elemento *CarG box*. In esperimenti di co-trasfezione l'aggiunta di HMGI(Y) nel sistema provoca un potenziamento cospicuo (~10 volte) della trascrizione mediata da SRF di entrambi i promotori. La regione di HMGY che interagisce con SRF è stata individuata a livello degli amminoacidi 50-81, che corrispondono al terzo AT-hook; HMGY non sembra essere coinvolta nel legame al DNA. Anche se HMGI(Y) lega il DNA sulla *CarG box* specificamente, questa interazione non è necessaria per l'effetto delle HMGI(Y). (121)

Anche nel caso di Elf-1 la capacità di interagire con l'elemento PRRII non era aumentata dall'aggiunta di HMGI(Y), mentre lo era la capacità di transattivare, dovuto ad un effetto sinergistico dato dall'interazione proteina-proteina con HMGI(Y). (110)

A livello del virus JC interazioni proteina-proteina tra HMGI(Y) e Tst-1/Oct-6 portano all'attivazione sinergica della trascrizione. Sono implicati gli amminoacidi 45-56 di HMGI, comprendenti parte del dominio di legame al DNA. (115)

Esperimenti di transfezione transiente con HMGI(Y) ed il fattore NF-Y sul promotore naturale dell' $\alpha 2(I)$  collagene e su di un costrutto con più elementi CCAAT-*box* specifici per NF-Y hanno dimostrato che HMGI(Y) attiva il fattore NF-Y, attraverso un'interazione proteina-proteina che coinvolge il dominio di legame al DNA della subunità NF-YA ed il 1° oppure 2° AT-hook dell'HMGI(Y). (123) HMGI(Y) non lega il DNA in questo caso ma facilita l'interazione della proteina NF-Y sia ricombinante che nativa con la CCAAT-*box*. (123), indicando che un ruolo di HMGI(Y) *in vivo* potrebbe essere quello di stabilizzare il complesso tra NF-Y ed il DNA a livello della CCAAT-*box*.

Si possono osservare interessanti analogie tra questi fattori ed anche ATF-2<sub>(195)</sub>, in quanto nell'interazione con HMGI(Y) sono sempre coinvolti i loro domini di legame al DNA, cioè bZip (ATF-2), dominio POU (Tst-1/Oct-6), ets domain (Elf-1), DBD (NF-YA) che in parte possono fungere anche da domini di dimerizzazione, che almeno nel caso di ATF-2 è stato dimostrato essere stimolata da parte di HMGI(Y).

L'interazione tra p50 di NF- $\kappa$ B ed HMGI sembra invece piuttosto complessa, perché esperimenti recentissimi di mutagenesi hanno dimostrato che gli stessi amminoacidi implicati nel riconoscimento dell'elemento PRDII da parte di HMGI (secondo AT-hook) sono anche quelli che interagiscono con la subunità p50. (105) Alcune spiegazioni per questo fenomeno potrebbero essere le seguenti: innanzitutto la proteina HMGI in complessi ternari potrebbe adottare un distorcimento tale da interagire con gli stessi amminoacidi con due componenti diversi contemporaneamente; in secondo luogo le mutazioni che allentano il legame di HMGI al DNA fanno sì che nemmeno NF- $\kappa$ B vi si possa legare ed il fatto che senza DNA non c'è cooperatività tra i due fattori implicherebbero che l'interazione HMGI-p50 avvenga indirettamente attraverso il DNA. Sembra più plausibile la seconda ipotesi, dato che per entrambi i polipeptidi sono stati dimostrati contatti con le stesse basi del DNA anche se effettuati da due solchi opposti (p50: lisina 244 fa un ponte H con la T centrale di PRDII o con la G corrispondente in Ig $\kappa$ ; HMGI: contatta le stesse basi con i "core" del 2° e/o 3° AT-hook). Questo fatto invalida ricerche precedenti in cui l'interazione tra le due proteine in assenza di DNA era stata dimostrata *in vitro* (104), ma si sa che con i sistemi del GST pull-down spesso anche solo interazioni deboli possono essere evidenziate. (105)

Attorno ai 9.5 giorni p.c. c'è un'espressione allargata a tutto l'embrione, eccetto che alla regione del cervello in cui l'mRNA di HMGI-C compare solo in una piccola porzione anteriore. Negli stadi successivi di differenziamento si osserva un restringimento dell'area di trascrizione fino poi a completa sparizione. (127)

Tra gli aspetti della distribuzione del trascritto di *Hmgi-c* emergono due peculiarità estremamente interessanti:

- la presenza del relativo mRNA in tutti i centri di condensazione di cellule mesenchimali che prendono parte ai processi di modellamento durante lo sviluppo. Ciò è specialmente evidente a livello scheletrico soprattutto per le coste, le vertebre, gli anelli cartilaginei della trachea e i centri di ossificazione attorno allo sterno ed al palatino;
- un'espressione esclusiva di cellule in attiva proliferazione e con basso grado di differenziamento.

La delezione del gene *Hmgi-c* provoca l'insorgenza di un fenotipo denominato *pygmy*, che è dato da un particolare nanismo nel topo che non ha disfunzioni ormonali, ma ha un difetto della crescita causata o da una risposta mancata ad un fattore di crescita o non modulando eventuali geni bersaglio durante l'embriogenesi. (128) Il topo *pygmy* è molto più piccolo rispetto al normale, con una diminuzione del peso corporeo complessivo all'80% rispetto al topo naturale se la delezione è in eterozigosi, mentre si riduce fino al 40% nel topo con la delezione in omozigosi. In particolare il tessuto adiposo subisce una riduzione fino al 12% rispetto al topo *wild-type*.

Anche nel pollo è stato descritto un nanismo specifico non legato a disfunzioni della via del fattore di crescita I insulino-simile. Si tratta di una mutazione spontanea autosomica recessiva (*adw*) in polli bianchi tipo *Leghorn* che provoca una diminuzione generale del peso corporeo, tranne del capo. Questo fenotipo risulta evidente attorno alle 6-8 settimane di età in quanto la riduzione del peso è del 30%. Questi polli nani sono molto simili al topo *pygmy*, e la mutazione *adw* *-/-* è stata mappata a livello del cromosoma 1 del pollo, in una regione sintenica al cromosoma 10 nel topo e 12 dell'uomo. HMGI-C è un buon candidato come responsabile del fenotipo nano anche nel pollo e in effetti un clone genomico del gene per HMGI-C del pollo e un clone marcatore del locus *adw* coibridizzano in un'analisi FISH effettuata su cromosomi metafasici di pollo. (129)

Tutto ciò suggerisce un possibile ruolo attivo delle HMGI nello sviluppo embrionale, nel differenziamento e/o nel controllo della proliferazione, con uno speciale ruolo tessuto-specifico di HMGI-C nel mesenchima.

Tutti i tessuti o linee cellulari analizzati fin'ora per la presenza di una o più proteine della famiglia HMGI sono elencati nella Tabella V.

Un esempio di un tessuto di origine mesenchimale è il tessuto connettivo del muscolo. Il differenziamento del muscolo scheletrico è accompagnato da una serie di eventi molecolari attraverso i quali i mioblasti in rapida proliferazione si differenziano a formare i miotubi. Tale transizione coinvolge uno *switch* nell'espressione genica che risulta nell'induzione di proteine muscolo-specifici e contrattili e in una concomitante sparizione di molte proteine del citoscheletro dei mioblasti. Si è trovato un sistema adatto per studiare l'espressione delle proteine HMGI sia in cellule in rapida proliferazione che terminalmente differenziate, utilizzando le linee di mioblasti L8 ed L6 ed delle fu-1, derivate dalle L8, ma difettive per la fusione dei miotubi.

L'emivita delle proteine HMGI(Y) in mioblasti e miotubi è di 60h e 53 h, rispettivamente, mentre nelle fu-1 è di 12h soltanto. Ciò potrebbe riflettere delle differenze nel grado di necessità di queste proteine rispetto allo stato metabolico dei diversi tipi cellulari. Un *turn-over* più veloce potrebbe indicare dei rimaneggiamenti cromatinici più rapidi, mentre un'emivita di 60h significa che possono perdurare per almeno due generazioni cellulari. Inoltre i livelli di tutte le HMG si abbassano nel momento in cui si formano i miotubi, rimane da stabilire se le HMGI(Y) diminuiscono perché le cellule, una volta formati i miotubi, non si dividono più o perché si trovano in uno stato più differenziato. (46)

**Tabella V**  
**Espressione dei fattori HMGI in tessuti o linee cellulari**

Tabella V.1								
ORGANISMO	TESSUTO O TIPO CELLULARE	ESPRESSIONE						REF.
		RNA			PROTEINA			
		HMGI	HMGY	HMGI-C	HMGI	HMGY	HMGI-C	
<b>UOMO</b>	HeLa, carcinoma della cervice	+	+	+	+	+	+	33
	K-562, linea cellulare eritroleucemica	+	+	n.d.	n.d.	n.d.	n.d.	
	K-562 + TPA	++	++	n.d.	n.d.	n.d.	n.d.	
	Cellule eritroidi	+	+	n.d.	n.d.	n.d.	n.d.	
	Placenta	+	+	n.d.	n.d.	n.d.	n.d.	33
	Linfociti B	+	+	n.d.	n.d.	n.d.	n.d.	72
	JY (linf. B; linea linfoblastoide	n.d.	n.d.	n.d.	+	+	n.d.	
	Cellule Raji	+	+	n.d.	n.d.	n.d.	n.d.	
	Sangue periferico donatori sani	+	+	-	+	+	-	141
	Placenta	-	-	n.d.	n.d.	n.d.	n.d.	33
	HuT, linfociti T	-	-	n.d.	n.d.	n.d.	n.d.	33
	Cellule staminali CD34+	n.d.	n.d.	+	n.d.	n.d.	n.d.	
	Cellule staminali leucemiche	n.d.	n.d.	+	n.d.	n.d.	n.d.	
	Carcinomi colonrettali (63/64)	+	+	n.d.	+(61/64)	+(61/64)	n.d.	152
	Polipi adenomatosi del colon 2/5	+	+	n.d.	-	-	n.d.	
	Mucosa normale del colon	-	-	n.d.	-	-	n.d.	
	Carcinomi colonrettali 30	+++	+++	n.d.	+++	+++	n.d.	149
	Adenomi del colon 26	++	++	n.d.	+	+	n.d.	
	Mucosa normale 23	-/+	-/+	n.d.	-/+	-/+	n.d.	
	C4-I (Carcinoma cervice differenziato)	+	+	n.d.	++	++	n.d.	151
	CaSKi (Carcinoma metastatico della cervice)	+	+	n.d.	++	++	n.d.	
	ME-180 (Carc. squamoso invasivo cervice)	+	+	n.d.	+	+/-	n.d.	
	Epiteli normale stratificato 28 casi	n.d.	n.d.	n.d.	-/+	-/+	n.d.	
	Neoplasie intraepiteliali cervice CIN 1 25 casi	n.d.	n.d.	n.d.	+	+	n.d.	
	Neoplasie intraepiteliali cervice CIN 2 10 casi	n.d.	n.d.	n.d.	++	++	n.d.	
	Neoplasie intraepiteliali cervice CIN 3 25 casi	n.d.	n.d.	n.d.	+++	+++	n.d.	
	Carcinoma invasivo della cervice 27 casi	n.d.	n.d.	n.d.	++++	++++	n.d.	





**Tabella V.1ter**

ORGANISMO	TESSUTO O TIPO CELLULARE	ESPRESSIONE						REF.
		RNA			PROTEINA			
		HMGI	HMGY	HMGI-C	HMGI	HMGY	HMGI-C	
<b>UOMO</b>	Tumore benigno della prostata (congelato)	-	-	n.d.	n.d.	n.d.	n.d.	144
	Tumore prostatico grado Gleason 1-2 (“)	+	+	n.d.	n.d.	n.d.	n.d.	
	Tumore prostatico grado Gleason 3 (“)	++	++	n.d.	n.d.	n.d.	n.d.	
	Tumore prostatico grado Gleason 4-5 (“)	+++	+++	n.d.	n.d.	n.d.	n.d.	
	Tumore benigno della prostata	-	-	n.d.	n.d.	n.d.	n.d.	
	Tumore prostatico grado Gleason 1-2 (“)	+	+	n.d.	n.d.	n.d.	n.d.	
	Tumore prostatico grado Gleason 3 (“)	++	++	n.d.	n.d.	n.d.	n.d.	
	Tumore prostatico grado Gleason 4-5 (“)	+++	+++	n.d.	n.d.	n.d.	n.d.	
	Metastasi linfonodale	+++	+++	n.d.	n.d.	n.d.	n.d.	

La tabella V contiene tutte le informazioni disponibili sull'espressione o del trascritto o della proteina dei componenti della famiglia HMGI, suddivisa in sezioni a seconda dell'organismo studiato, uomo, roditori ed uccelli. Non si possono confrontare i dati riportati d diversi gruppi, dato che non sono utilizzati gli stessi mezzi di rilevamento. Quindi questa tabella è da considerare piuttosto un riassunto di evidenze sperimentali pubblicate.

n.d. = non documentato; non significa che la specie non c'è, ma semplicemente che non è stata cercata.

(-) = non rilevabile;

(-/+ ) = poca, non significativa;

+ = livello costantemente rilevabile;

++ fino ++++ = quantità crescenti di espressione

**Tabella V.2**

ORGANISMO	TESSUTO O TIPO CELLULARE	ESPRESSIONE						REF.
		RNA			PROTEINA			
		HMGI	HMGY	HMGI-C	HMGI	HMGY	HMGI-C	
<b>TOPO</b>	NIH-3T3, fibroblasti	+	+	+				34
	NIH-3T3 + TPA	++	++	++				
	Cellule Ehrlich di ascite	n.d.	n.d.	n.d.	+	+	n.d.	67
	Swiss 3T3 quiescenti	n.d.	n.d.	n.d.	+	+	n.d.	84
	Swiss 3t3 proliferanti (+siero)	n.d.	n.d.	n.d.	+	+	n.d.	
	Cellule MRL di ascite	+++	+++	n.d.				33
	Cellule Friend	++	++	n.d.				
	LBRM-33 cellule T	+	+	n.d.				
	CTLL (inf. T dipendenti da fattori di crescita)	n.d.	n.d.	n.d.	+	+	n.d.	72
	Cellule F9, teratocarcinoma indifferenziato				++	++		145
	Cellule F9 + acido retinoico (differenziate)				-	-		
	Cellule F9 + dibutiril-cAMP (differenziate)				-	-		
	JB6 P-	1x	1x	n.d.	-	-	n.d.	
	JB6 P+	20x	20x	n.d.	+	+	n.d.	
	JB6 Tx	30x	30x	n.d.	n.d.	n.d.	n.d.	
	JB6 P- (+TPA) induzione dopo	8h	8h	n.d.	++	+	n.d.	
	JB6 P+ (+TPA) induzione	0-24 h	0-24 h	n.d.	++	+++	n.d.	
	Milza	-	-	n.d.	n.d.	n.d.	n.d.	33
	Fegato	-	-	n.d.	n.d.	n.d.	n.d.	
	Cervello	-	-	n.d.	n.d.	n.d.	n.d.	
	Polmone				+	+(Y2)	-	140
	Timo				+	+(Y2)	-	
	LLC (Lewis Lung Carcinoma)				++++	++(Y2e3)+	++++	
	Linfoma NQ35				+	+(Y2)	-	
	Linfoma NQ9				+	+(Y2)	-	
	Linfoma NQ36				+	+(Y2)	-	

**Tabella V.2bis**

ORGANISMO	TESSUTO O TIPO CELLULARE	ESPRESSIONE						REF.
		RNA			PROTEINA			
		HMGI	HMGY	HMGI-C	HMGI	HMGY	HMGI-C	
<b>TOPO</b>	D1 linea derivante da tumore mammario							131
	spontaneo da topi Balb/c							
	+ alcune linee derivanti da D1:							
	CL-S1 preneoplastiche	+	+	n.d.			n.d.	
	-SA non cresce in agar molle (metastatico diff.)	++	++	n.d.	+	+	n.d.	
	+SA cresce in agar molle (metast. non diff.)	+++	+++	n.d.	+++	+++	n.d.	
	PMME (cellule epiteliali primarie mammarie)	+	+	n.d.	n.d.	n.d.	n.d.	

**Tabella V.3**

ORGANISMO	TESSUTO O TIPO CELLULARE	ESPRESSIONE						REF.
		RNA			PROTEINA			
		HMGI	HMGY	HMGI-C	HMGI	HMGY	HMGI-C	
<b>RATTO</b>	FRTL5-C12, cellule epiteliali di tiroide giovane	-	-	-	-	-	-	136
	FRTL5-C12 + KiMSV (v-Ki-ras)	++	++	++	++	++	++	
	FRTL5-C12 + HaMSV (v-Ha-ras)	++	++	++	++	++	++	
	PC C13, cellule epiteliali di tiroide adulta	-	-	-	-	-	-	
	PC C13 + MPSV (v-mos)	++	++	++	++	++	++	
	PC C13 + c-myc trasfettato	-	-	-	-	-	-	
	PC C13 + PyMuLV	-	-	-	-	-	-	
	PC C13 + c-myc + PyMuLV	++	++	++	++	++	++	
	Fibroblasti della tiroide	+	+	n.d.	+	+	-	
	Fibroblasti della tiroide + PyMuLV	++	++	++	++	++	++	
	Fibroblasti della tiroide + HaMSV	++	++	++	++	++	++	
	L8(E63) mioblasti							46
	L6 mioblasti							
	fu-1 derivati dalle L8, difettivi nella fusione							

**Tabella V.3bis**

ORGANISMO	TESSUTO O TIPO CELLULARE	ESPRESSIONE						REF.
		RNA			PROTEINA			
		HMG1	HMGY	HMG1-C	HMG1	HMGY	HMG1-C	
HAMSTER	CHO		++++					44

**Tabella V.4**

ORGANISMO	TESSUTO O TIPO CELLULARE	ESPRESSIONE						REF.
		RNA			PROTEINA			
		HMG1	HMGY	HMG1-C	HMG1	HMGY	HMG1-C	
POLLO	Embrioni normali di 3 gg			+++			++	129
	Embrioni normali di 6 gg			++			+	
	Embrioni normali di 8 gg			+			?	
	Embrioni nani adw di 4 gg			+++			++	
	Embrioni nani adw di 6 gg			++			+	
	Embrioni nani adw di 8 gg			+			?	

In questo senso è interessante notare che la stimolazione di cellule di mammifero in coltura quiescenti con diversi fattori di crescita (ad esempio PDGF, FGF, EGF, esteri del forbolo oppure il siero) porta in poco tempo (1-4 ore) ad un'induzione dell'espressione di numerosi geni "early delayed", tra i quali ci sono anche *HMGI*, *Y* ed *I-C*. (47, 39, 130). Questa induzione è molto specifica, perché ad esempio dei quattro promotori identificati del gene *HMGI(Y)* solo uno risponde al trattamento con gli esteri del forbolo (130) mentre la stimolazione con l'EGF ne attiva due. (45). Ciò indica che ogni singolo promotore/*enhancer* può essere individualmente e specificamente stimolato in risposta a fattori di crescita particolari, e costituisce un fatto di significatività biologica non solo per lo sviluppo embrionale, ma anche per la regolazione dell'espressione dei geni per HMGI in cellule normali somatiche e in cellule trasformate in senso neoplastico.

Gli stimoli che inducono l'espressione delle HMGI sono molteplici e anche di diversa natura, com'è mostrato nella seguente tabella (Tabella VI):

Tabella VI.

## Stimoli che inducono l'espressione di HMGI

STIMOLO	TIPO CELLULARE	SPECIE INDOTTA	OSSERVAZIONI	Re
TGF $\alpha$ =stimolo di regolazione della crescita	Fibroblasti normali della mammella di topo in coltura	mRNA di HMGI(Y)	Tale induzione correla con la densità cellulare, che può cambiare in seguito al trattamento con fattori di crescita e alla perdita di inibizione da contatto	131
PDGF FGF Siero	Fibroblasti 3T3 di topo (tipo A31 da Balb/c 3T3)	mRNA di HMGI(Y) e HMGI-C	Induzione dopo alcune ore, si tratta di geni "early delayed" che seguono dopo l'induzione dei geni precocissimi o "immediate-early"; Serve sintesi proteica probabilmente delle proteine dei geni precocissimi per indurre i geni precoci tardivi. La regolazione delle HMGI è diversa da quella degli istoni H1, la cui induzione avviene più tardi con l'inizio della fase S. Che ruolo hanno le HMGI nella transizione G <sub>0</sub> -S?	47
EGF 17- $\beta$ -estradiolo	Linee cellulari umane MCF7 e Hs578T epiteliali mammarie	Solo in Hs578 mRNA e proteina HMGI(Y)	Nelle MCF7 non metastatiche nessuno dei due stimoli induce espressione delle HMGI(Y), mentre nelle Hs578T, altamente metastatica si.	45
EGF LPS (TPA) AP-1		mRNA HMGI(Y)	Vedi testo.	130
Endotoxin IL-1 $\beta$	Cellule primarie muscolari lisce (RASCs) dell'endotelio di aorta di ratto Sprague-Dawley in coltura e <i>in vivo</i> .	mRNA e proteina di HMGI(Y)	Si osserva un picco d'induzione dopo 48h dallo stimolo ed un aumento di 6.2 volte dell'mRNA e anche un aumento concomitante dei livelli della proteina. Dopo stimolazione delle stesse cellule in vivo con endotossina si ottiene la stessa induzione, e in parallelo si osserva un aumento della iNOS (vedi sezione 1.2.5.2).	118
Ipossia	HUVECs (cellule umane di cellule endoteliali della vena ombelicale)	mRNA e proteina di HMGI(Y)	Sembra che questa induzione serva per indurre il promotore della cicloossigenasi-2 attraverso cooperazione tra HMGI(Y) ed NF- $\kappa$ B.	132
Raf-1 oncogenico	Cellule Pa-4 epiteliali della ghiandola parotide di ratto	mRNA di HMGI-C	In un sistema cellulare inducibile si è potuto dimostrare che l'attivazione della chinasi Raf-1 porta ad un aumento tardivo ma sostenuto di HMGI-C, dipendente dalla sintesi proteica e dall'attività delle MAP chinasi poste a valle della "pathway" di Raf-1.	133

## Continuazione Tabella VI

STIMOLO	TIPO CELLULARE	SPECIE INDOTTA	OSSERVAZIONI	Ref
LPS NDV (Virus della Newcastle Disease)	RAW264.7 linea murina cellulare macrofagica Macrofagi peritoneali di topi C3HeB/FeJ	Proteina HMGI-C	La stimolazione attraverso infezione virale da NDV e/o LPS induce la comparsa di complessi proteici contenenti HMGI-C, che invece è assente in cellule non stimulate. Si ottiene così l'attivazione trascrizionale del gene per la chemochina RANTES, però attraverso due meccanismi distinti, a seconda se si tratta di NDV o lipopolisaccaridi.	134
Proteina E6 (HPV16)	Cellule 10 T ½ E6	mRNA di HMGI(Y)	Attraverso "differential display" di cellule trasfettate con E6 e non si è individuato il messaggero di HMGI(Y). L'induzione sembra diretta (non è inficiata da inibitori della sintesi proteica) e specifica, dato che proteine E6 di altri tipi di papillomavirus umani non inducono HMGI(Y).	135
Gli oncogeni virali elencati qui di seguito o il protooncogene c-myc possono essere considerati come stimoli in quanto mimano le "pathways" indotte da molti stimoli mitogenici differenti.				
v-ras-Ki	FRTL-C12 Cellule di tiroide di ratto giovane	Tutte e tre le HMGI	Le cellule perdono i marcatori di differenziazione e la dipendenza da fattori di crescita e diventano tumorigeniche.	136
v-ras-Ha	c.s.	Tutte e tre le HMGI		
MoMuLV	c.s.	-	Mantiene sia i marcatori di differenziazione che la dipendenza da fattori di crescita. L'infezione con MoMuLV non presentava alcun oncogene.	
v-ras-Ki	PC-C13 Cellule di tiroide di ratto adulto	-	Dopo infezione col KiMSV queste cellule non si sdifferenziano completamente e non sono tumorigeniche.	
v-mos	c.s.	Tutte e tre le HMGI, in special modo HMGI-C	Completamente tumorigeniche e sdifferenziate.	
c-myc	c.s.	-	La trasfezione con c-myc non sembra alterare né il grado di differenziazione né la dipendenza da fattori di crescita.	
middle T	c.s.	-	L'infezione col virus del poyoma porta alla perdita dei marcatori differenziativi ma solo bassa tumorigenicità.	
c-myc/middleT	c.s.	Tutte e tre le HMGI, in special modo HMGI-C	La combinazione dei due trattamenti porta a cellule con fenotipo completamente tumorigenico e sdifferenziato.	
v-ras-Ha	c.s.	Solo HMGI	Le cellule non sono tumorigeniche.	
v-ras-Ha	FRT-Fibro Fibroblasti di ratto (?)	Tutte e tre le HMGI	Tumorigeniche.	
middle T	c.s.	Tutte e tre le HMGI	Tumorigeniche.	

## 1.2.7 HMGI e Trasformazione Neoplastica

Alla luce del fatto che le proteine HMGI sono indispensabili nel processo trascrizionale *in vivo* e nello sviluppo embrionale non sorprende che numerosissime ricerche abbiano scoperto alti livelli di espressione di queste proteine in stretta correlazione con la trasformazione neoplastica di cellule normali e nell'incremento del potenziale metastatico di cellule tumorali. In cellule somatiche normali e differenziate gli RNA messaggeri e le proteine di HMGI sono espressi a livelli bassissimi (33, 61, 137,) o non rilevabili. (36, 42, 136, 138-140) Invece in cellule trasformate in senso neoplastico (136, 139, 141-144) come anche in cellule embrionali che non si sono ancora differenziate (136, 138, 145) i livelli delle tre proteine sono spesso eccezionalmente alti. Tumori spontanei o derivanti dalla carcinogenesi chimica di cellule normali, oppure indotti da radiazioni ionizzanti o UV, oppure originati da trasformazione mediante oncogeni virali (*v-src*, *v-ras*, *v-mos* e *v-myc*) contengono livelli molto più alti di mRNA e proteine HMGI rispetto alla controparte normale.

La proteina HMGI-C ha inoltre un ruolo d'elezione nella tumorigenesi, in quanto la sintesi di questo fattore spesso non è presente in fenotipi poco tumorigenici, ma la sua neosintesi nei tumori è indice di solito di fenotipi sdifferenziati e quindi potenzialmente più tumorigenici.(136)

Dato che i livelli cellulari dei messaggeri delle proteine HMGI variano col grado di proliferazione in cellule normali, essendo bassissimi in cellule quiescenti ed aumentando circa quattro volte durante la crescita esponenziale, è importantissimo segnalare che invece nei tumori gli alti livelli di HMGI sembrano essere presenti indipendentemente dal grado di proliferazione, dato che le corrispettive cellule normali, proliferanti alla stessa stregua, contengono comunque minori quantità di HMGI. (143, 144) Alcune stime riportano che certe linee cellulari maligne contengono costitutivamente 15-50 volte più mRNA rispetto alla controparte cellulare non trasformata. (33, 137, 146)

Considerando il ruolo delle HMGI nell'organizzazione cromatinica e l'aumento delle stesse in un tumore come il carcinoma polmonare di Lewis (LLC) di 4-5 volte rispetto al polmone normale (si passa dal 2% di proteine totali nucleari al 10% in LLC) e assumendo che la massa di 1 molecola di HMGI è metà rispetto all'istone, allora si avranno in ogni nucleo della cellula trasformata 1 molecola di HMGI per 5 molecole di istone H1, anziché per ogni 25. Dovremmo



sicuramente concludere che la struttura cromatinica è modificata in LLC rispetto al tessuto di origine. (140)

La caratteristica di un aumento cospicuo di una o più delle proteine HMGI in seguito ad una transizione cellulare è osservata in una moltitudine di casi studiati. Un'ipotesi sul diverso funzionamento delle proteine HMGI qualora si trovano in bassa concentrazione rispetto alle stesse in alta concentrazione è già stata fatta: si pensa che a basse concentrazioni le HMGI assolvono al compito specifico del controllo trascrizionale di molti geni, legandosi preferenzialmente a siti di alta affinità. Quando invece la loro presenza aumenta fino a raggiungere livelli altissimi allora è probabile che leghino anche siti a minor affinità, saturando letteralmente il DNA ed inducendo così processi non noti. (49)

Il contenuto globale delle diverse specie di HMG in cellule di polmone normale di topo e nella controparte trasformata del Carcinoma polmonare di Lewis (LLC) mostra variazioni più o meno consistenti e diverse tra le varie classi, ma solo le variazioni delle HMGI correlano sempre con la trasformazione fenotipica. Inoltre è stato dimostrato che non vi è ridondanza funzionale tra le tre classi HMG. (140).

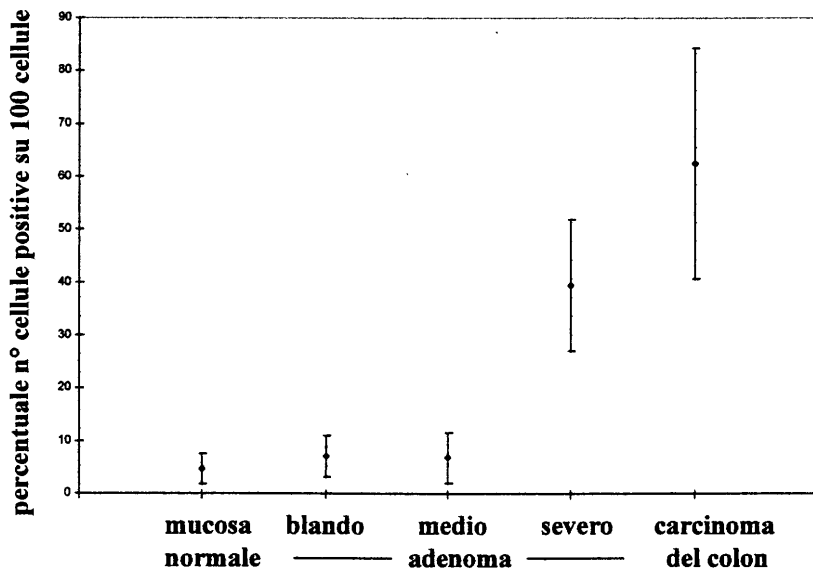
Data la stretta correlazione tra trasformazione neoplastica e livelli costitutivi alti delle proteine HMGI (136, 139, 141) molti ricercatori hanno suggerito di sfruttare questa caratteristica come un aspetto diagnostico non solo del fenotipo neoplastico, ma anche come marcatore di progressione per la metastasi di tumori prostatici nel modello del ratto Dunning (143) in quanto l'entità dell'iperespressione delle HMGI correlava direttamente col grado di aggressività dei tumori piuttosto che con il tasso di crescita. Recentemente tali osservazioni sono state estese a tumori prostatici umani in uno studio retrospettivo dei livelli di RNA messaggero delle HMGI(Y) tramite ibridizzazione *in situ* su materiale in paraffina e materiale congelato ottenuto da pazienti affetti da cancro alla prostata di diverso grado di Gleason. Con chiarezza un aumento dell'espressione delle HMGI(Y) correla con il grado di tumorigenicità (più o meno maligno), e non solo, studi preliminari hanno collegato tale aumento anche con lo stadio del tumore, proponendo le HMGI(Y) come strumento di valutazione della progressione tumorale.(144)

Studi retrospettivi analoghi hanno correlato alti livelli delle proteine HMGI(Y) col fenotipo maligno di neoplasie tiroidee, colon-rettali del seno e della cervice. (147-151) Simili correlazioni tra un incremento dei livelli sia di messaggero che di proteina HMGI(Y) con stadi diversi di progressione tumorale sono stati riportati anche dall'analisi di un sistema di tumore mammario murino ben caratterizzato, stabilendo che l'espressione delle HMGI è un marcatore biochimico affidabile. (131)

Gli sforzi attuali nella diagnosi tumorale sono quelli di riuscire ad ottenere informazioni importanti già a livello della lesione preneoplastica ed inoltre di usufruire di metodi molecolari veloci, precisi e capaci di individuare anche tracce minime di un marcatore specifico. A tale riguardo sono interessanti gli studi effettuati su di una miriade di casi in laboratori differenti, che non giungono sempre alle stesse conclusioni. Un esempio è dato dallo studio delle neoplasie del colon, per i quali un gruppo ha analizzato ben 64 casi già diagnosticati come carcinomi e li ha confrontati con 5 lesioni preneoplastiche (polipi adenomatosi) e con la mucosa normale del colon per la presenza sia di trascritti di *HMGI(Y)* mediante RT-PCR competitiva, che di proteine mediante immunistochemica sulle sezioni biotiche utilizzando un anticorpo anti-*HMGI(Y)*. (152) Risultava che la mucosa era del tutto negativa e gli adenomi saggiati presentavano bassi livelli di RNA, ma le proteine non erano rilevabili, mentre nel 98% di tutti i carcinomi erano alti sia i livelli di messaggero che di proteina. Non avendo trovato però nessuna correlazione tra i livelli di espressione delle *HMGI(Y)* e particolari parametri clinico-patologici tipici del cancro colonrettale (lo stadio e la grandezza del tumore, metastasi linfonodale, grado istologico, invasione linfatica o angiogenetica e lo stato della p53) giungono unicamente alla conclusione che l'espressione di *HMGI(Y)* molto probabilmente inizia in uno stadio precoce della carcinogenesi. (152)

Altri due laboratori concordano pienamente nell'asserire che invece l'aumento dei livelli di RNA e proteina, saggiati in diverse linee cellulari e tumori del colon come anche in lesioni preneoplastiche e la mucosa normale, correlano con il grado di malignità. (149, 153) Non solo, il confronto tra i livelli di espressione di *HMGI(Y)* in lesioni che vanno da pretumorali attraverso i vari adenomi prima differenziati e poi severi fino al carcinoma, possono dare informazioni sulla progressione della lesione.

Un metodo che usano è il confronto delle conte dei nuclei positivi alla reazione immunostochimica in diverse sezioni (% di cellule positive su 100 cellule), mostrato nel grafico qui di seguito:



L'analisi clinico-patologica (profondità dell'invasione tumorale, coinvolgimento linfatico e/o venoso, metastasi linfonodale e stadio di Dukes) effettuata su 30 casi, di cui il 60% (18 casi) mostrava alti livelli, il 30% (9 casi) livelli moderati e il 10% (3 casi) livelli bassi di HMGI(Y) ha mostrato che vi è una differenza significativa nell'espressione delle HMGI(Y) tra i casi con coinvolgimento linfatico e/o venoso e quelli senza. Infatti l'82% (14/17) di quelli che avevano un'alta invasività mostravano alti livelli delle proteine HMGI(Y) mentre 8 su 11 casi non-invasivi avevano livelli di espressione da moderati a bassi. Inoltre una correlazione significativa veniva dimostrata tra gli alti livelli di espressione delle HMGI(Y) e la presenza di metastasi linfonodali ed uno stadio clinico avanzato secondo la classificazione di Dukes, e quindi con una prognosi infausta per i pazienti affetti da cancro al colon. (149)

Divergenze tra i vari gruppi di ricerca possono anche originare dalle metodiche o dagli strumenti su cui ci si basa. Infatti l'anticorpo utilizzato è di primaria importanza nelle analisi immunostochimiche. I gruppi di Abe et al. (149) e Fedele et al. (153) utilizzano l'anticorpo policlonale, ottenuto in coniglio, contro il peptide SSSKSQQPLASKQ, specifico delle

HMGI(Y), mentre Kim et al. ne usano uno commerciale, ottenuto in capra, di cui non viene con precisione indicato l'antigene.

Lo stesso anticorpo più uno specifico contro la proteina HMGI-C sono stati usati per l'analisi di linee cellulari di neuroblastoma, di tessuto delle ghiandole surrenali sia embrionale che adulto ed alcuni neuroblastomi infantili, per vedere se le proteine della famiglia HMGI avessero un ruolo anche in cellule di tipo neurale. Nello studio della differenziazione della cresta neurale era tra l'altro anche stato isolato il gene *HMGI-C* dall'analisi mediante "differential display" di una linea cellulare TC-1S derivata da cellule della cresta neurale e stimulate con EGF. (154) I neuroblastomi sono tumori del sistema nervoso periferico che sembra originarsi da un arresto anomalo nella differenziazione delle cellule multipotenti embrionali della cresta neurale. Se quindi si riesce a far differenziare le cellule di questi tumori, spesso si osserva una regressione spontanea. Un agente differenziante potente *in vitro* risulta essere l'acido retinoico, e vi sono tentativi di terapie con lo stesso anche *in vivo*. Purtroppo non tutti i tipi tumorali sono sensibili all'effetto di inibizione della crescita e alle proprietà differenzianti dell'acido retinoico. Nel tentativo di spiegare una simile differenza al trattamento con l'acido retinoico si è trovato una correlazione tra la refrattarietà intrinseca di cellule e l'espressione di HMGI-C.

L'analisi dell'espressione di HMGI(Y) e HMGI-C dimostra che HMGI(Y) è praticamente espressa in tutti i neuroblastomi analizzati, come anche nelle ghiandole surrenali sia embrionali che adulte. Nuovamente HMGI-C mostra una distribuzione ristretta alla ghiandola surrenale embrionale e solo 5 su 18 neuroblastomi. Valutando la loro espressione in risposta all'acido retinoico somministrato a cinque linee cellulari di neuroblastoma i risultati erano i seguenti:

- trattamento delle linee SK-N-AS, SH-EP, SK-N-SH ha dimostrato la loro insensibilità all'acido retinoico, in quanto anche dopo 6 giorni non si erano accumulate nella fase G1;
- trattamento delle linee SY5Y e KCNR porta all'arresto alla transizione G0/G1 dopo 4 giorni di somministrazione, dimostrando la loro sensibilità.

L'analisi dell'espressione delle HMGI(Y) ha evidenziato che l'acido retinoico provoca una iperespressione di RNA delle HMGI(Y) entro le prime 48h dall'inizio della manipolazione nelle cellule non rispondenti, mentre nelle cellule sensibili vi è un calo sorprendente. Al contrario i livelli di HMGI-C non subiscono nessuna alterazione in nessuna delle cinque linee studiate. Mediante esperimenti di trasfezione in cui si è fatto esprimere o HMGI(Y) o HMGI-C in cellule SY5Y (sensibili all'acido retinoico) trattate in seguito con l'acido retinoico si è

potuto stabilire che l'iperpressione di HMGI(Y) non alterava la sensibilità di SY5Y, mentre HMGI-C, normalmente mai espressa, conferiva la caratteristica di insensibilità all'acido retinoico alle cellule SY5Y. In conclusione, mentre *HMGI(Y)* sembra in qualche modo modulabile da questo agente, HMGI-C non lo è, anzi agisce contro lo stimolo antiproliferativo dell'acido retinoico. (154)

Questo aspetto di HMGI-C di riuscire a mantenere il fenotipo poco differenziato potrebbe basarsi anche sul fatto che è stato dimostrato che HMGI-C può indurre delle modificazioni della composizione di AP-1, che risulterebbe in un'attività che promuove la crescita. (155)

In conclusione HMGI-C è anormalmente espresso in alcuni neuroblastomi *ex vivo*, ed in linee di neuroblastoma resistenti all'acido retinoico e sono responsabili di questa resistenza perché convertono cellule sensibili a cellule resistenti se vi vengono artificialmente espresse. HMGI(Y) è costitutivamente espressa in tutti i tumori e le linee di neuroblastoma, e viene represso solo in cellule acido retinoico rispondenti. Quindi per una efficace e soprattutto non inutile terapia un'analisi preliminare della presenza di HMGI-C e della reattività di HMGI(Y) all'acido retinoico potrebbe anticipare l'effetto di questo agente anti-proliferativo in vivo. (154)

L'idea di usare più di un marcatore tumorale nella diagnosi di tipi specifici di tumori è di grande utilità, dato che spesso le lesioni tumorali non si possono caratterizzare con il 100% di sicurezza dall'analisi istologica e morfologica, e di conseguenza una specifica terapia non può essere stabilita. Un esempio è dato dall'utilizzo dei marcatori HMGI(Y) e Ki-67 nella caratterizzazione di tumori della cervice di diversi gradi di Gleason. (151, vedi sezione 4.2.1).

In questo studio si è preso in considerazione l'ectocervice di donna adulta, formata da un epitelio squamoso stratificato composto da strati di cellule a livello basale, parabasale, intermedio e superficiale. Il differenziamento maturativo dell'epitelio parte dallo strato basale verso gli strati superficiali, con una modificazione progressiva della morfologia cellulare. I processi displastici inducono alterazioni più o meno pesanti sia a livello di epitelio che a livello morfologico.

La gradazione degli epiteli displastici della cervice si basa sulle proporzioni dell'ispessimento dell'epitelio provocato da cellule sdifferenziate dello strato basale o su criteri a livello cellulare come ad esempio atipia nucleare. A questi criteri morfologici l'aggiunta dell'informazione di un marcatore di tumorigenicità può essere utile soprattutto in casi di dubbia istologia. Nello studio effettuato su 60 casi rappresentanti dei gradi CIN 1-3 (vedi tabella V.1 espressioni) si è potuto stabilire una immunoreattività all'anticorpo anti-HMGI(Y) in una percentuale crescente di cellule all'aumentare del grado di displasia. Anche

l'immunoreattività verso l'anticorpo anti-Ki-67 seguiva lo stesso andamento, tuttavia non necessariamente le stesse cellule erano immunoreattive: infatti a livello dello strato basale di una displasia moderata (CIN II) più cellule sono reattive verso l'anticorpo HMGI(Y) che verso l'anti-Ki-67. Questo è importante e anche informativo, dato che i due marcatori sono rappresentanti di due processi biologici diversi, ma strettamente correlati, in quanto Ki-67 è un tipico marcatore di cellule in proliferazione, mentre alti livelli HMGI sono costitutivi di cellule tumorali ed indipendenti dal tasso di crescita. (151)

Nel sistema murino JB6, modello di suscettibilità alla trasformazione neoplastica e di progressione di malignità, è stato dato un primo accenno del fatto che le proteine HMGI possano influenzare l'andamento della cancerogenesi già a livello della suscettibilità, ossia nella caratteristica più precoce di una cellula a diventare tumorale. Infatti, confrontando cellule JB6 resistenti alla trasformazione indotta da agenti cancerogeni come il TPA (12-*O*-tetradecanoil-forbol acetato) con cellule JB6 sensibili a tale trattamento, si notano alcune differenze nell'espressione delle proteine HMGI ed HMGY; l'aumento relativo sia dei messaggeri che delle proteine in seguito al trattamento con TPA è più cospicuo in cellule JB6 resistenti che in quelle sensibili, ma la quantità assoluta di entrambi i prodotti è più alta in cellule sensibili ed i livelli alti sono mantenuti più a lungo rispetto a quelle resistenti. Inoltre si nota per la prima volta una differenza tra i livelli proteici delle due isoforme, infatti la proteina HMGY è la forma proteica preferenzialmente espressa nelle cellule JB6 refrattarie alla trasformazione neoplastica. Si può dedurre da questi fatti che queste proteine nonistoniche cooperano nella regolazione trascrizionale di geni importanti sia per la suscettibilità che per l'induzione e/o mantenimento dello stato neoplastico. Si ipotizza un meccanismo indotto da AP-1 per il quale è stato dimostrato di avere proprietà transattivante in seguito a stimolazione con promotori tumorali solo in cellule JB6 sensibili e non in quelle resistenti alla trasformazione. Un'altra ipotesi da valutare è che le HMGI potrebbero reprimere geni che mediano la resistenza alla cancerogenesi: dove non sono presenti le HMGI allora questi geni vengono espressi, mentre con l'aumento dei livelli di HMGI, come nelle cellule JB6 sensibili, allora potrebbero essere repressi geni che inibiscono la trasformazione, come ad esempio TIS1, TIS2 e una proteina relata a Fra (35).

Inversamente, quando cellule F9 di teratocarcinoma aggressivo altamente sdifferenziate vengono indotte a differenziarsi allora perdono sia i loro alti livelli costitutivi dei prodotti genici di HMGI sia il potenziale tumorigenico. (145)

La dimostrazione più convincente di una relazione causale delle HMGI nella formazione del fenotipo neoplastico proviene da una ricerca recente, in cui l'inibizione della sintesi di

HMGI-C, mediante metodologie antisense, sopprime la capacità di retrovirus oncogeni (recanti *v-mos* o *v-ras-Ki*) di indurre trasformazione neoplastica in cellule epiteliali di tiroide di ratto. (156) Tutte queste evidenze sperimentali coinvolgono le proteine HMGI nei due processi di trasformazione neoplastica e di acquisizione di un aumentato potenziale metastatico.

Pur avendo queste caratteristiche, le proteine HMGI non si comportano come classici oncogeni, inquanto facendoli artificialmente esprimere in cellule normali non portano alla trasformazione (156, 157), suggerendo che in molti casi la loro iperespressione è necessaria, ma non sufficiente, per ottenere il fenotipo neoplastico. È molto probabile che l'attivazione di altri fattori o un'alterazione nelle funzioni architetturali delle proteine HMGI siano richiesti per indurre completamente la trasformazione.

### 1.2.8 HMGI e tumori benigni

Anomalie cromosomiche nei tumori sono già state riconosciute alla fine del secolo scorso, ma il loro significato è stato chiarito solo di recente. Traslocazioni distinte nelle leucemie ed in tumori solidi portano all'attivazione di prodotti di proto-oncogeni, oppure molto comunemente alla creazione di proteine di fusione tumore-specifiche. In entrambe le categorie le proteine coinvolte sono spesso fattori di trascrizione, che una volta alterati inficiano la regolazione della trascrizione che, come si sa, è la causa maggiore dell'eziologia del cancro. Proteine di fusione formate in seguito a traslocazioni cromosomiche sono comuni in tuttuna serie di tipi tumorali e spesso costituiscono antigeni unici e quindi bersagli potenziali per la progettazione di terapie. Esempi eclatanti nelle leucemie sono sicuramente il linfoma di Burkitt, in cui il proto-oncogene *c-myc* viene a porsi sotto il controllo del promotore delle immunoglobuline in linfociti B, oppure la traslocazione PML-RAR $\alpha$  nella leucemia acuta promielocitica e via dicendo. (158)

Studi citogenetici estesi hanno decisamente implicato la regione q13-q15 del cromosoma 12 umano nella genesi di una varietà di tumori solidi maligni e benigni, suggerendo quindi la localizzazione dei geni responsabili associati a tali neoplasie. Tumori maligni con aberrazioni ricorrenti in questa regione sono liposarcomi mixoidi, sarcomi a cellule chiare di tessuti molli

ed un sottogruppo di rhabdomiosarcomi. Il gene responsabile dei liposarcomi mixoidi è il gene *CHOP*, che codifica per un fattore di trascrizione della famiglia C/EBP inducibile dall'arresto del ciclo cellulare e dal danno al DNA. Il gene coinvolto nella genesi dei sarcomi dei tessuti molli codifica per un altro fattore trascrizionale denominato ATF-1.

Tra i tumori benigni un coinvolgimento della stessa regione cromosomica si osserva frequentemente in tumori del tessuto adiposo (lipomi), adenomi pleomorfi delle ghiandole salivari e leiomiomi uterini. Aberrazioni della regione suddetta sono state descritte anche in polipi endometriali, in un sottogruppo di emangiopericitomi, tumori condromatosi ed amartomi polmonari condroidi. Infine casi isolati di tumori solidi benigni con la stessa traslocazione sono stati riportati per tumori della mammella, astrocitomi diffusi ed un tumore dell'osso a cellule giganti (osteoclastoma). (159) (Vedi Tabella VII)

La regione nella quale si localizzano i punti di rottura sul braccio 12q in tuttuna serie di questi tumori benigni è denominata MAR (*Multiple Aberration Region*). Essa si estende per 1.7 Megabasi e contiene le regioni parzialmente sovrapposte di distinti raggruppamenti di punti di rottura, ossia una *breakpoint cluster region*. L'osservazione citogenetica che un particolare segmento della regione MAR è frequentemente inficiato e i raggruppamenti dei punti di rottura si sovrappongono indica che con molta probabilità esiste un comune denominatore nella loro patobiologia. L'analisi a livello molecolare di queste aberrazioni ha portato all'identificazione del gene *HMGI-C* come possibile gene implicato nella patogenesi di questi tumori benigni. (159)

In seguito ad analisi citogenetiche l'approccio sperimentale per identificare le sequenze di DNA colpite da aberrazioni genetiche è la "FISH" (Fluorescence In Situ Hybridization), nella quale si usano grandi frammenti di DNA genomico marcato con fluorocromi come sonde per il riconoscimento di anomalie a livello di cromosomi metafasici. In una ricerca pilota sono stati utilizzati 75 cloni di DNA (YACs) parzialmente sovrapposti, che si estendono su circa 6 Megabasi a livello della regione MAR, per analizzare le aberrazioni cromosomiche in una serie di tumori primari solidi benigni (34 tumori di 8 tipi diversi) e in numerose linee cellulari tumorali (26 linee rappresentanti di 3 tipi tumorali). (vedi Tab VII) Si è potuto restringere la regione contenente i punti di rottura a livello della regione MAR di questi campioni ad una zona di sole 175 Kbasi, posta tra due sequenze genomiche marcatrici riconosciute dalle sonde RM99 ed RM133. (159). A questo punto mediante la tecnica del *3'-terminal exon-trapping e genomic sequence sampling* è stato possibile isolare e sequenziare un frammento di una delle sonde utilizzate, che dava un doppio segnale in FISH dovuto alla rottura del braccio 12q in 3 leiomiomi uterini. Il confronto di questa sequenza contro banche dati di sequenze



nucleotidiche con il programma BLAST ha rivelato un'identità di sequenza con parte della sequenza cDNA della proteina HMGI-C umana precedentemente clonata. (38) In seguito l'intera sequenza genomica di *HMGI-C* è stata determinata e ciò ha permesso di stabilirne l'organizzazione genomica (Figura 1, pannello superiore) e soprattutto l'orientamento della trascrizione, che è diretta in senso telomerico. È importante notare che le dimensioni del gene (170 Kb, vedi Figura 1) e quella dell'intervallo di 175 Kbase della MAR quasi corrispondono. La maggior parte delle aberrazioni nel gene per HMGI-C sono rotture a livello del terzo e quarto introne, che non ci sorprende visto che il terzo introne comprende quasi l'80% di tutta la lunghezza del gene. (159)

I partner di traslocazione o meglio le sequenze ectopiche eventualmente fuse con quelle di *HMGI-C* in seguito all'aberrazione, vengono identificati mediante l'analisi "*CASH*" (*Chromosomal Assignment using Somatic cell Hybrids*), usando un pannello di mappatura di ibridi somatici tra cellule umane e di roditore. Di solito i dati citogenetici corrispondono a quelli dell'analisi CASH, ma ogni tanto, soprattutto quando i riarrangiamenti sono complessi, ci possono essere delle discrepanze. (159)

Le analisi effettuate su moltissime linee cellulari e biopsie di pazienti hanno portato ad evidenziare importanti caratteristiche dei tumori benigni di origine mesenchimale che hanno consentito di formulare delle ipotesi riguardo un comune meccanismo patogenetico.

**I lipomi** sono tumori composti principalmente da cellule adipose mature e sono una delle neoplasie di origine mesenchimale più frequenti del genere umano. Circa il 50% dei lipomi sono caratterizzati da riarrangiamenti citogenetici e l'alterazione predominante sembra essere una traslocazione bilanciata tra 12q14-15 e una grande varietà di partners cromosomici, inclusi i cromosomi 1-7, 10, 11, 13, 15, 17, 21 e X. La variabilità dei partners nelle traslocazioni assieme a duplicazioni, inversioni e delezioni della regione 12q14-15 in questi tumori porta a pensare che il gene posto sul cromosoma 12q14-15 abbia un ruolo primario nella genesi dei lipomi. (160). Esiste una predominanza di un sottogruppo di riarrangiamenti con il locus 12q14-15, data dal locus cromosomico 3q27-28. Qui è posto un gene, denominato LPP, da *lipoma preferred partner*, che come dice il suo nome, costituisce un partner di fusione molto frequente nei lipomi. Esso codifica per una proteina grandissima contenente un dominio ricco in proline, un motivo *leucine-zipper* e tre domini LIM, noti per essere domini di dimerizzazione. La sua omologia con altre proteine LIM fa pensare che LPP possa avere un ruolo nella trasmissione di segnali intercellulari e nella regolazione dello sviluppo. (161) L'analisi mediante RT-PCR ha rivelato due caratteristici tipi di trascritto chimerico di HMGI-

C/LPP derivanti dalla fusione tra i primi tre esoni (contenenti gli AT-hooks) di *HMGI-C* e due o tre domini LIM codificati invece da *LPP*. Il fatto che sia presente sempre questa fusione e non quella reciproca suggerisce una selezione positiva di questa chimera che quindi sembra avere una certa importanza nella patogenesi dei lipomi. (161)

Contrariamente a questo, argomentano molti casi in cui il partner di fusione è sempre diverso o fornisce soltanto qualche amminoacido alla proteina *HMGI-C* troncata, fornendo l'ipotesi che il troncamento stesso dell'*HMGI-C* fosse un evento sufficiente per indurre la tumorigenesi in tessuti di origine mesenchimale. Inoltre, in un recente lavoro la traslocazione t(3;12) (q27;q14-q15), che porta al trascritto *HMGI-C/LPP*, è stata descritta anche in numerosi amartomi polmonari, concludendo che questo prodotto non ha un ruolo specifico nel fenotipo adipocitico. (162)

Altri partners di traslocazione identificati sono degni di nota. Ad esempio l'espressione di trascritti ibridi reciproci tra *HMGI-C* e *FHIT* in un adenoma pleomorfo della ghiandola parotide dà sicuramente qualche problema di interpretazione, dato che entrambi i geni coinvolti sono strettamente associati alla trasformazione neoplastica. I due trascritti reciproci codificano in un caso per i primi tre AT-hooks di *HMGI-C* fusi agli esoni 9 e 10 di *FHIT*, mentre il reciproco possiede 7 esoni di *FHIT* al 5' contenenti la *triade di istidine* fusi agli esoni codificanti per la regione spaziatrice e la coda acida di *HMGI-C*. L'analisi mediante RT-PCR dell'RNA di questa neoplasia ha rivelato inoltre che entrambi gli alleli normali dei due geni erano anche espressi, concordando con l'ipotesi di Tkachenko et al. descritta più avanti. Anche *FHIT* risulta spesso riarrangiato in adenomi pleomorfi e appartiene ad una regione cromosomica definita *fragile*, per cui il suo coinvolgimento potrebbe essere casuale. Si specula anche qui che sono piuttosto rimaneggiamenti al 3' di *HMGI-C* importanti nella patogenesi di più neoplasie benigne, dato che per rimozione del 3' si eliminano otto elementi AUUUA che sono motivi di destabilizzazione dell'mRNA. (163)

Altri partners di fusione potenzialmente importanti nella patogenesi in generale di tumori sono stati identificati "grazie" a diversi riarrangiamenti con il locus 12q14-15. Ad esempio il fattore nucleare NFIB, collocato a livello del locus 13q12 potrebbe rispondere non solo del fenotipo tumorale in lipomi coinvolgenti anche *HMGI-C*, ma in generale in lipomi con una aberrazione a livello di 13q12-22, che è stata trovata nel 7-8% dei lipomi studiati. (164)

*HMGI-C* sembra avere un ruolo chiave nel normale sviluppo del mesenchima primitivo, dato che non solo i lipomi, ma anche leiomiomi uterini (tumori della muscolatura liscia), lipoleiomiomi (costituiti da componenti adiposi e muscolari lisce) ed amartomi condroidi

polmonari (con componenti mesenchimali primitive, muscolari liscia, adipose e cartilaginee) sono tutti proliferazioni clonali caratterizzate da un riarrangiamento del 12q14-15.

**Gli amartomi condroidi polmonari** sono costituiti da proporzioni variabili di mesenchima primitivo, cartilagine matura, tessuto muscolare liscio e tessuto adiposo maturo. Uno studio effettuato su 30 casi (Vedi Tabella VII.3, casi H1-H30) mostra che 15 casi su 30 possiedono aberrazioni clonali nella regione 12q14-15, mentre 2 su 30 possiedono aberrazioni a livello della regione 6p21, dove mappa *HMGI(Y)*, ed infine un caso aveva un riarrangiamento complesso sempre a carico di quest'ultimo locus. Sono rimarchevoli i casi H8 ed H30 con una traslocazione a livello di 12q15 ma con due cromosomi intatti. Studi molecolari degli stessi casi mediante *RACE al 3'* e *FISH* hanno evidenziato che in tutti i 18 casi sopra elencati vi è un riarrangiamento con sequenze ectopiche, ma anche in 6 casi diagnosticati come normali con metodi citogenetici. Comunque un risultato globale indica che il 70% degli amartomi condroidi polmonari mostrano un riarrangiamento di *HMGI-C*.(165)

**I Leiomiomi uterini** sono tumori benigni del tessuto muscolare liscio dell'utero e sono associati con molti disturbi sintomatici e anche con infertilità. Quest'ultimo è il più frequente tumore della pelvi nelle donne, e colpisce circa il 77% di tutte le donne in età riproduttiva. Recentemente si è dimostrato che questi tumori hanno un'origine monoclonale e le aberrazioni citogenetiche più ricorrenti sembrano essere ad oggi le seguenti: t(12;14)(q14;-15;q23-24), delezioni e riarrangiamenti di 7q e 13q, trisomie 12 e riarrangiamenti che coinvolgono i cromosomi 6p o 10q22. Queste aberrazioni consistenti indicano che specifiche alterazioni genetiche sono importanti nella formazione di leiomiomi e facilitano il compito di identificare sequenze patogenetiche critiche. Il sottogruppo t(12;14) rappresenta uno dei più frequenti riarrangiamenti in leiomiomi con anomalie clonali (40%), oltre che in tutt'una serie di tumori benigni di origine mesenchimale (vedi Tabella VII.2).(166, 167)

Sono stati osservati riarrangiamenti cromosomici anche del locus 6p21 in leiomiomi uterini e molti altri tumori benigni, quali angiomiomi, polipi endometriali e lipomi. Il gene *HMGI(Y)* è stato localizzato in questo punto ed è stato dimostrato che esso viene riarrangiato specificamente in un campione di leiomioma con una inv(6)(p21q15). In questo campione risulta un'altissima espressione della proteina *HMGI(Y)* confrontata al miometrio normale circostante. Un'elevata espressione di *HMGI(Y)* si è registrata anche per 8 su 16 campioni che non mostravano specifiche aberrazioni citogenetiche, ed anche in un campione in cui il locus riarrangiato era quello del gene *HMGI-C* (vedi Tabella VII.2). (167). Tutte le analisi su questi

tumori benigni di origine mesenchimale pubblicate indicano che sicuramente meno aberrazioni coinvolgono *HMGI(Y)* rispetto a *HMGI-C*.

Le analisi di tumori benigni di origine mesenchimale fatte sia a livello citogenetico (cariotipo) che molecolare (*FISH*, *RACE* etc.) sono ancora incomplete. Innanzitutto per poter fare delle vere statistiche il numero di campioni analizzati deve sicuramente aumentare. Inoltre vi sono ancora molte discrepanze tra i dati ottenuti con le diverse tecniche e il *RACE* al 3' (mediante il quale si possono identificare sequenze ectopiche di trascritti chimerici), che peccano di alcuni errori di fondo, come ad esempio l'impossibilità di amplificare trascritti lunghi (> di 1Kbase) o trascritti con molte ripetizioni di sequenza. Inoltre la maggior parte di queste analisi si ferma a livello dei trascritti, sia anomali che normali ma aumentati, ma solo in pochi casi viene individuata anche la proteina.

Comunque la scoperta del coinvolgimento dei geni per HMGI in numerose aberrazioni che caratterizzano tumori benigni di origine mesenchimale ha sollevato molte questioni. Una di queste è capire se i geni riarrangiati funzionano come oncogeni dominanti. Infatti nella maggior parte dei casi le alterazioni di *HMGI-C* coinvolgono delle rotture a livello del terzo introne, che è enorme (~140 Kbase), dando origine a trascritti aberranti contenenti i primi tre esoni che codificano per i tre domini di legame al DNA fusi a sequenze ectopiche. Spesso tali sequenze ectopiche danno un contributo di solo alcuni amminoacidi, facendo pensare che non è l'aggiunta di un segmento proteico a creare una nuova forma proteica con proprietà tumorali, ma che piuttosto il semplice troncamento della *HMGI-C* con la perdita del dominio spaziatore e della coda acida sia responsabile della trasformazione cellulare. Esempi di questo tipo sono dati dalle fusioni coi geni *NFIB* e *ALDH1*. (168, 169)

In altri casi però la proteina chimerica è data dalla fusione dei primi tre domini di legame al DNA di *HMGI-C* con domini proteici funzionali codificati dal partner genico del riarrangiamento. Ad esempio nei lipomi il partner di fusione preferenziale è il gene *LPP* che contribuisce alla proteina chimerica con tre domini LIM, che sono in grado di dimerizzare con altri fattori di trascrizione, e ciò rende l'interpretazione di questi riarrangiamenti più complessa. Esprimendo artificialmente la proteina tronca e la stessa in fusione con i domini LIM del gene *LPP* in cellule NIH-3T3, e confrontando i fenotipi risultanti da questi esperimenti con quelli di un costrutto esprimente la proteina normale, si è constatato che entrambe le forme aberrate hanno potere trasformante. (157) Infatti rispetto ai controlli negativi, costituiti dal costrutto con la proteina normale e dal vettore vuoto, che praticamente non mostrano alterazioni nel fenotipo delle cellule trasfettate, le cellule esprimenti le due

proteine aberranti crescono in agar molle ed inducono tumori se iniettati sottocute in topi atimici. Inoltre i rispettivi trasfettanti stabili sono stati saggiati per la capacità proliferativa in terreno normale ed in basso siero, dimostrando che le cellule con le proteine aberranti avevano non solo una densità cellulare di saturazione più alta rispetto ai trasfettanti di controllo, ma anche un tempo di raddoppiamento più breve a basse percentuali di siero. Il confronto con cellule trasfettate con un vettore esprimente Ret-MEN2A, un oncogene noto e altamente tumorigenico, ha permesso di stabilire che le due forme proteiche di HMGI-C, tronca e chimerica, hanno comunque un effetto trasformante simile e modesto sulle cellule NIH-3T3. Questo dimostra quindi che i riarrangiamenti di HMGI-C sono importanti nello sviluppo delle neoplasie e che è sufficiente la perdita dei due ultimi esoni per conferire capacità trasformante. Sembra anche che l'aggiunta di sequenze ectopiche non aumentano tale capacità, almeno per quel che riguarda il sistema delle NIH-3T3 utilizzato in questo studio. In altri tipi cellulari la sequenza ectopica potrebbe avere un'importanza nel modulare il potere trasformante. I dati ottenuti con le cellule NIH-3T3 non possono essere completamente trasferiti in un sistema tumorale *in vivo*, ma il fatto stesso che il fenotipo tumorale indotto dalle proteine tronca e chimerica sia blando è consistente con il comportamento benigno dei tumori associati con le forme di *HMGI-C* riarrangiate. L'analisi di topi transgenici con la proteina tronca o chimerica potrà dare informazioni più precise riguardo agli aspetti istologici dei tumori causati in vivo dai riarrangiamenti di *HMGI-C*. (157).

Un elemento importantissimo nel meccanismo di tumorigenesi di questi riarrangiamenti sembra però essere l'attivazione dell'allele riarrangiato di *HMGI(Y)* o *HMGI-C* nel contesto di una cellula normale e differenziata, che normalmente non esprime proteine della famiglia HMGI, come ad esempio nell'adipocita maturo. Infatti in tumori solidi di origine mesenchimale caratterizzati dal cariotipo 12q13-15 e 6p21 si trovano quasi sempre alleli trascrizionalmente attivi. Quindi le aberrazioni in questi loci attivano alleli di *HMGI* normalmente silenti in cellule normali e l'erronea espressione della relativa proteina nel contesto della cellula mesenchimale differenziata è un passo cruciale nello sviluppo del tumore. In questi sistemi la tumorigenesi risulta da una espressione temporalmente inappropriata nella cellula adulta di un gene normalmente espresso durante lo sviluppo prenatale di una cellula embrionale della stessa linea. (170) Un esempio analogo potrebbe essere l'attivazione di *c-myc* nel linfoma di Burkitt, dove *c-myc* normalmente è espresso in cellule pre-B (171) L'analisi degli alleli non riarrangiati ha dato delle informazioni importanti poter avanzare un'ipotesi di un meccanismo coerente dello sviluppo dei lipomi proposta da Tkachenko et al. (170) Una delezione nella regione 3' non tradotta è probabilmente il minimo

riarrangiamento necessario per la formazione di un tumore. In seguito, in base alla natura del riarrangiamento di *HMGI*, viene selezionata una di due vie alternative di tumorigenesi (vedi Figura 4). Quando il riarrangiamento produce una proteina HMGI senza attività trascrizionale intrinseca, allora lo sviluppo del tumore dipende da una successiva attivazione dell'allele non riarrangiato. Per contro l'obbligata espressione di HMGI nella tumorigenesi è aggirata quando come esito della traslocazione i domini di legame al DNA delle HMGI sono fusi con un dominio di regolazione trascrizionale. Il numero ridotto di eventi coinvolti nella formazione del tumore potrebbe effettivamente spiegare la più frequente traslocazione nei lipomi, cioè la t(3;12) (q29;q15), perché avvicina gli AT-hooks di HMGI-C ai domini LIM che sono coinvolti nella regolazione trascrizionale. (171)

*HMGI-C* copre una regione genomica 17 volte più ampia di *HMGI(Y)* e questo spiega probabilmente perché *HMGI-C* sia coinvolto con una frequenza molto più alta di *HMGI(Y)* nei lipomi. Tumori benigni, a differenza della loro controparte maligna, sono caratterizzati solo da una o poche alterazioni specifiche, per questo l'analisi molecolare di queste neoplasie potrebbe individuare geni di maggiore importanza per la crescita e la proliferazione. Gli studi qui descritti dimostrano sicuramente la significatività della famiglia HMGI nell'adipogenesi e nella progressione tumorale. (171)

Anche nelle leucemie sono state descritte traslocazioni che coinvolgono proteine contenenti un AT-hook, come ad esempio in alcune neoplasie acute linfocitiche (ALL) e acute linfoblastiche, sia pre-B che pre-T. In questi casi il gene *MLL* o *HRX* (Tabella 1, sezione 1.2.3), posto sul locus 11q23, perde il suo Zn-finger, ma fornisce il suo AT-hook alla proteina di fusione che si crea con diversi partners. (158, 172) Riarrangiamenti nel locus 11q23 si sono scoperti anche in un sottogruppo di leucemie mieloidi acute, ma anche nella malattia di Hodgkin o in linfomi non-Hodgkin. Data la variabilità dei partners di fusione, nel caso di *MLL*, come per *HMGI-C*, sembra essere solo la rottura e non la fusione con un altro dominio la vera causa dell'insorgenza della neoplasia. (172)

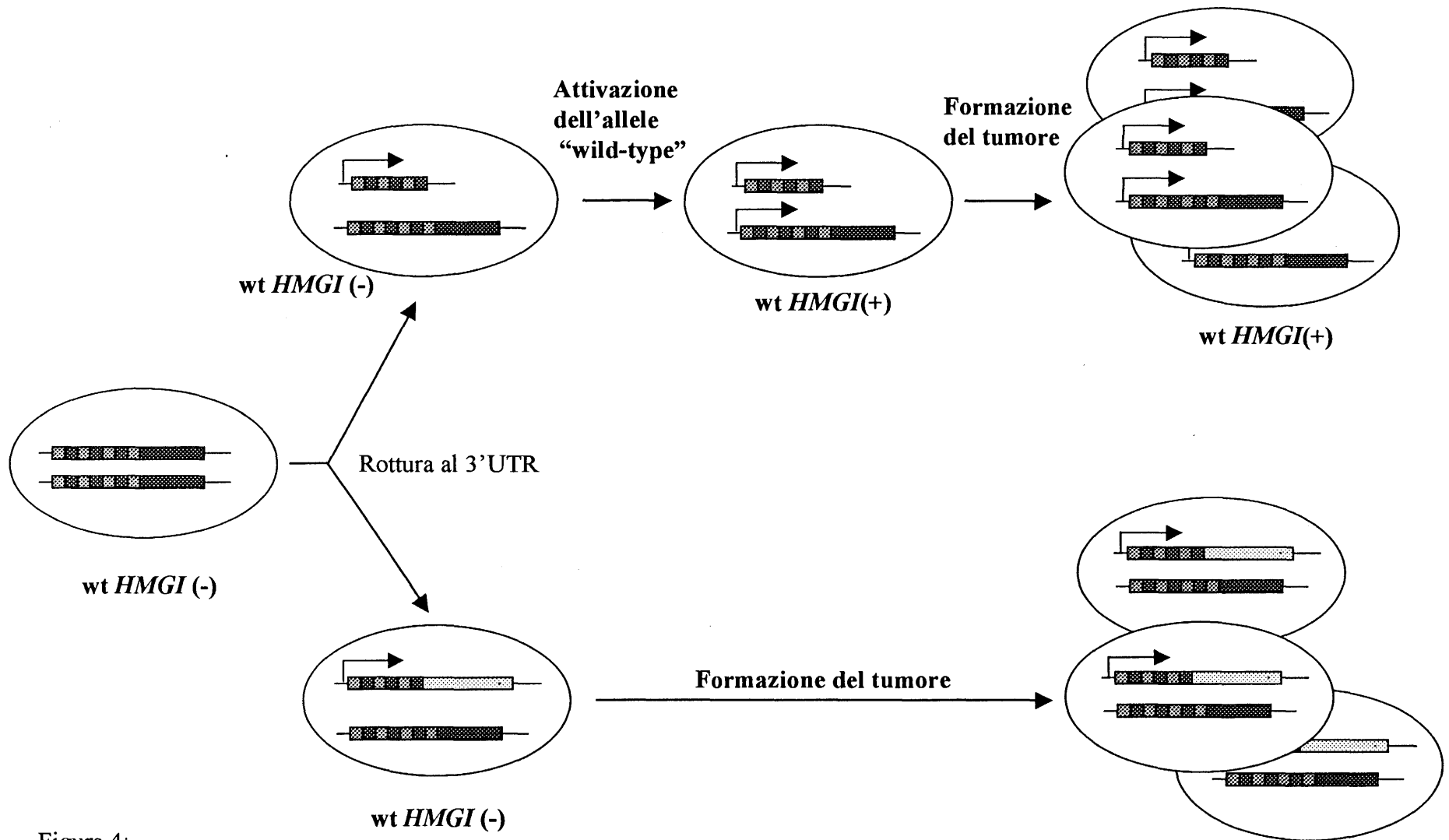


Figura 4:  
 Modello della formazione di un lipoma secondo Tkachenko et. al. Quadrati rosa=AT-hooks, quadrati verdi=regioni spaziatrici, rettangolo blu=coda COOH-terminale, rettangolo giallo=dominio acquisito con attività trascrizionale, freccia piegata=trascrizione dell'allele, wt=allele "wild-type", (-)=non espresso, (+)=espresso.

**Tabella VII**
**Aberrazioni cromosomiche in tumori benigni e solidi di origine mesenchimale**
**Tabella VII.1 Aberrazioni cromosomiche del fenotipo lipomatoso**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
12q14-15	n.d.	n.d.	n.d.	6/6	LIPOMA	6 casi su 6 possiedono un'aberrazione a livello delle 175 Kbase all'interno della regione MAR.	159
12q14-15	t(12;15) (q15;q24)	chimerico	n.d.	1	LIPOMA	Tumore ST90-375; punto di rottura nel gene <i>HMGI-C</i> a livello del 3° introne; traslocazione reciproca produce un trascritto chimerico tra l'RNA codificante per i primi tre AT-hooks e un dominio sconosciuto ad alto contenuto in residui acidi, ricco in serine e treonine, un dominio putativo di attivazione trascrizionale. Il prodotto del reciproco non è documentato.	178
	t(3;12) (q29;q15)	chimerico	n.d.	1	LIPOMA	Tumore ST93-724; punto di rottura nel gene <i>HMGI-C</i> a livello del 3° introne; traslocazione reciproca produce un trascritto chimerico tra l'RNA codificante per i primi tre AT-hooks e 2 domini LIM, tipici di fattori trascrizionali capaci di dimerizzare.	178
	t(12;13) (q14-22;q21-32)	assente	n.d.	1	LIPOMA	Punto di rottura distalmente al <i>HMGI-C</i> , 3' dello stesso risulta deleto.	178
12q13-15	t(3;12) (q28;q13-15)	chimerico	n.d.	1	Linea cellulare di lipoma, Li-14/SV40	Fusione dopo la regione spaziatrice (4° esone) con sequenza ectopica del cromosoma 3.	159
		chimerico	n.d.	1	Li-501/SV40	Fusione dopo il 3° AT-hook con sequenza ectopica del cr. 3.	
		chimerico	n.d.	1	Li-538/SV40	Fusione dopo il 3° AT-hook con sequenza ectopica del cr. 3, ma diversa dalla linea Li-501/SV40.	
12q13-15	n.d.	chimerico	n.d.	1	Li-#294	Fusione dopo il 3° AT-hook con sequenza ectopica del cr. 8q22.	
12q13-15	n.d.	chimerico	n.d.	1	Li-#194	Fusione dopo il 3° AT-hook con sequenza ectopica del cr. 9.	
12q14	t(12;14) (q14;q14)	n.d.	n.d.	1	Li-136/SV40		
	t(12;12) (q14;q23)	n.d.	n.d.	1	Li-166/SV40		



**Tabella VII.1bis Aberrazioni cromosomiche del fenotipo lipomatoso**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
	t(7;13;12) (p15- p21;q14-q21;q13)	n.d	n.d.	1	Li-502/SV40		161
	der(12) t(12;16;?) (q14- q15;q22;?)	n.d.	n.d.	1	Li-501/SV40		
	inv(12) (q12q13-q14)	n.d.	n.d.	1	Li-506/SV40		
12q14-15	t(3;12) (q27;q14)	chimerico HMGI- C/LPP	n.d.	3	Li-501/SV40 Li-538/SV40 Li-14/-SV40	Questo tipo di traslocazione sembra essere presente in ¼ di tutti i lipomi considerati fino al momento, e producono un trascritto chimerico tra HMGI-C (primi tre AT-hooks) e il partner d'elezione LPP.	161
12q14-15	t(3;12) (q27;q14) 6/20 caritipo normale 1/20	LPP		6/20	Lipomi primari	5 casi con la traslocazione mostrano anche il trascritto di fusione con LPP, 1 caso con traslocazione non sembra averlo, mentre un caso con cariotipo normale si.	
	t(12;13) (q14;q12) 1/20 t(1;12) (p34;q14) 1/20 t(2;12) (q35;q14) 1/20 t(2;12) (q37;q14) 1/20 inv(12) (p12q14) 1/20	non hanno visto fusioni con LPP	n.d.	5/20	Lipomi primari	Questi dati risultano sempre da un lavoro in cui si ricercava specificamente LPP come partner di fusione.	
	caritipo normale 4/20	nessuna fusione con LPP	n.d.	4/20	Lipomi primari		
	caritipo complesso 2/20	nessuna fusione con LPP	n.d.	2/20	Lipomi primari		
6p	n.d.	nessuna fusione con LPP	n.d.	2/20	Lipomi primari		
12q14-15	t(12;13) (q14;q12)	chimera con LHFP	n.d.	1	Lipoma primario Li 1321-89	Fusione tra i primi 3 DNA binding domains di HMGI-C con 69 amminoacidi di LHFP (Lipoma HMGI-C fusion partner).	164

**Tabella VII.2 Aberrazioni cromosomiche fenotipo leiomiomi**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
12q14-15	n.d.	n.d.	n.d.	7/8	LEIOMIOMI UTERINI	7 casi su 8 hanno l'aberrazione a livello della MAR nell'intervallo ristretto delle 175 Kbasi.	159
12q14-15	inv(12;12) (q14-15;q24)	chimerico	n.d.	1	LEIOMIOMA UTERINO PRIMARIO	Punto di rottura nel 3° introne di <i>HMGI-C</i> , produce un trascritto di fusione tra i primi 3 AT-hooks di <i>HMGI-C</i> con 10 triplette di sequenza codificante derivanti dal 13° esone del gene per l'aldeide-deidrogenasi mitocondriale.	169
12q14-15	t(12;14) (q15;q24)	n.d.	n.d.	1	Linea cellulare LM-5.1/SV40		159
		chimerico	n.d.	2	LM-30.1/SV40	Due trascritti diversi sono stati evidenziati: Uno risulta dalla fusione dopo il 3° AT-hook con sequenza ectopica del cr. 14, mentre il secondo riporta una fusione dopo la regione spaziatrice.	
		n.d.	n.d.	1	LM-65/SV40		
		n.d.	n.d.	1	LM-608/SV40		
		n.d.	n.d.	1	LM-609/SV40		
		n.d.	n.d.	1	LM100/SV40		
12q14-15	t(12;14) (q14-q15;q24)	n.d.	n.d.	1	LM67/SV40		
12q14	ins(12;11) (q14;q21qter)	n.d.	n.d.	1	LM605/SV40		
12q15	t(X;12) (q22;q15)	chimerico	n.d.	1	LM168.1/SV40	Fusione dopo il 3° AT-hook con sequenza ectopica con il cr. Xq22; dati CASH implicano anche sequenze ectopiche del cr. 14.	
12q14	t(2;3;12) (q35;q21;q14)	n.d.	n.d.	1	LM192.1/SV40		
12q14-15	n.d.	chimerico	n.d.	1	LM-#25	Fusione dopo il 3° AT-hook con sequenza ectopica del cr. 3.	
12q14-15	n.d.	chimerico	n.d.	1	LM-#58	Fusione a livello del 3' non tradotto con sequenza ectopica del cr. 8q22.3qter.	
12q14	t(1;12) (q32;q14) del(7) (q22q32)	n.d.	-	1	LEIOMIOMA Campione 1		167

	del(1) (q42) t(12;14) (q14;q24)	n.d.	-	1	LEIOMIOMA Campione 3		
12q13	t(12;14) (q13;q32)	n.d.	-	1	LEIOMIOMA Campione 4	La stessa paziente presenta all'interno dello stesso leiomioma cellule con cariotipo sia anomalo che normale, è da notare che quelle normali presentano livelli di proteina HMGI(Y) più elevati rispetto al miometrio adiacente. Ciò è molto interessante, dato che l'abberazione colpisce il locus di <i>HMGI-C</i> e non di <i>HMGI(Y)</i> .	167
	normale	n.d.	+ (HMGI(Y))				
12q14-15	t(12;14) (q14-q15;q23-24)	n.d.	-	1	LEIOMIOMA Campione 7		
6p21	t(1;7) (q31;q22),-2,-3,-6.add(11) (q21),18,+mar1	n.d.	-	1	LEIOMIOMA Campione 2	La stessa paziente presenta all'interno dello stesso leiomioma cellule con cariotipo sia anomalo che normale, è da notare che quelle normali presentano livelli di proteina HMGI(Y) più elevati rispetto al miometrio adiacente.	
	normale	n.d.	+ (HMGI(Y))				
6p21	inv(6) (p21q15), inv(9)	n.d.	+++++ HMGI(Y)	1	LEIOMIOMA Campione 5		
	normale	n.d.	++ HMGI(Y)	1	LEIOMIOMA Campione 6		
	del(1) (p22p35),-6,add((10) (q22),19,mar1	n.d.	-	1	LEIOMIOMA Campione 8		
	n.d.	n.d.	+++ HMGI(Y)	1	LEIOMIOMA Campione 10		
	normale	n.d.	+ HMGI(Y)	1	LEIOMIOMA Campione 9	La stessa paziente presenta all'interno dello stesso leiomioma cellule con cariotipo sia anomalo che normale, è da notare che quelle normali presentano livelli di proteina HMGI(Y) più elevati rispetto al miometrio normale adiacente, e quelle col cariotipo anomalo presentano livelli ancora più elevati.	
	normale	n.d.	+ HMGI(Y)	1	Leiomioma		
7 e 14	del(7) (q11) t(7;14) (q22;q32)	n.d.	+++ HMGI(Y)	1	Leiomioma		

**Tabella VII.3 Aberrazioni cromosomiche del fenotipo Amartoma**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
12q14-15	n.d.	n.d.	n.d.	8/9	AMARTOMA CONDROIDE POLMONARE (PCH)	8 casi su 9 hanno il punto di rottura all'interno della MAR proprio nell'intervallo delle 175 Kbasi.	159
	n.d.	n.d.	n.d.	1/1	AMARTOMA DELLA MAMMELLA	1 caso su 1 ha il punto di rottura all'interno della MAR proprio nell'intervallo delle 175 Kbasi.	
12q14-15	t(12;14) (q14-15;q24)	Trascritto normale + alcuni aberranti.	n.d.	3	PCH	In tutti e tre i casi il punto di rottura è localizzato in posizione 5' rispetto ad <i>HMGI-C</i> , suggerendo che l'attivazione trascrizionale può iniziare la crescita tumorale. Inoltre vi sono due cromosomi 12 normali ed un derivativo del cromosoma 14, ma manca il derivativo der(12). I trascritti aberranti non contengono sequenze ectopiche del cr. 14, ma sequenze appartenenti all'introne 3 di <i>HMGI-C</i> stesso; ciò risulta molto probabilmente da siti di "splicing" criptati.	178
12q14-15	Cariotipo normale (46,XX oppure 46, XY)	chimerici	n.d.	4	PCH (Casi H7, 14, 17 e 21)	Pur avendo cariotipo normale si sono evidenziati dei trascritti chimerici dovuti a fusione dopo il 3° AT-hook in due casi, dopo la regione spaziatrice in un caso ed infine all'interno dell'esone 2 nel quarto caso con sequenze ectopiche. Nei due casi H17 e H21 la sequenza "ectopica" è data da regioni introniche o del 3° o 4° introne di <i>HMGI-C</i> .	165
		n.d.	n.d.	6	PCH (Casi H5, 10, 11, 12, 16, 18)	Cariotipo normale, segnali normali col FISH.	
12q14-15	inv(12) (q14q22) t(12;18) (q24.1;q21)	n.d.	n.d.	1	PCH (Caso H1)	Doppio segnale col FISH.	
	t(1;11) (p31;q125) inv(12) (p12q14)	n.d.	n.d.	1	PCH (Caso H2)	Doppio segnale col FISH.	

**Tabella VII.3bis Aberrazioni cromosomiche del fenotipo Amartoma**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
							165
	t(6;12) (q25;q14)	n.d.	n.d.	1	PCH (Caso H3)	Segnale bipartito col FISH.	
	t(3;6) (p25-26;p21.1-21.2) t(9;12) (q34;q14-q15) t(12;16) (q14-q15;q21)	chimerico	n.d.	1	PCH (Caso H6)	Segnale al di sotto del punto di rottura. Fusione con la sequenza ectopica dopo il 3° AT-hook.	
	t(12;14) (q15;q24)	n.d.	n.d.	1	PCH (Caso H8)	Segnale sul cr. 14.	
12q14-15	ins(12;8) (q15;p22pter)	chimerico	n.d.	1	PCH (Caso H13)	Segnale bipartito, fusione tra <i>HMGI-C</i> dopo la regione spaziatrice con sequenze introniche del gene stesso.	
	ins(12;3) (q15;13.2qter)	chimerico	n.d.	1	PCH (Caso H15)	Segnale bipartito col FISH, fusione dopo il 3° AT-hook di <i>HMGI-C</i> con sequenze introniche.	
	t(11;12) (p15;q15) del(12) (q15)	chimerico	n.d.	1	PCH (caso H19)	Segnale bipartito col FISH, fusione dopo il 3° AT-hook di <i>HMGI-C</i> con sequenza ectopica.	
	t(12;14) (q15;q24)	n.d.	n.d.	2	PCH (Casi H9 e H20)	Segnale bipartito col FISH.	
		chimerico	n.d.	1	PCH (Caso H30)	Segnale bipartito col FISH, fusione dopo il 3° AT-hook di <i>HMGI-C</i> con sequenza ectopica.	
	inv(12) (?q15)	n.d.	n.d.	1	PCH (Caso H22)	Doppio segnale col FISH.	
?	inv(9)c	chimerico	n.d.	1	PCH (Caso H23)	Anche se il segnale da FISH con le sonde <i>HMGI-C</i> specifiche risultano normali c'è una fusione di <i>HMGI-C</i> dopo il 3° dominio di legame al DNA con una sequenza ectopica.	
12q14-15	t(1;12;13) (p22qter;q15;q32) del(12) (q15) t(1;13) (p22;q32)	Chimerico, 2 tipi	n.d.	1	PCH (Caso 24)	Segnale bipartito col FISH, ci sono due differenti trascritti, entrambi con fusione dopo il 3° AT-hook di <i>HMGI-C</i> con sequenza intronica.	
	del(12) (q15;q23)	n.d.	n.d.	1	PCH (Caso 27)		
	inv(12) (p11.2;q15)	n.d.	n.d.	1	PCH (Caso 28)		
6p21 <i>HMGI(Y)</i>	t(6;14) (p21.3;q24)			2	PCH (Caso H4)		

**Tabella VII.3ter Aberrazioni cromosomiche del fenotipo Amartoma**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
	t(3;6) (p25-26;p21.1-21.2) t(9;12) (q34;q14-q15) t(12;16) (q14-q15;q21)	n.d. per HMGI(Y)	n.d.	1	PCH (Caso H6)		165
	t(6;8) (p21.3;q12-13) del(14) (q24)	n.d.	n.d.	1	PCH (Caso H29)		
12q14-15	t(3;12) (q27;q14)	Chimerico con LPP	n.d.	3/5	PCH	Dimostrano qui che la traslocazione con 3q27 e concomitante fusione con LPP non è tipico solo dei lipomi, ma anche dei PCH.	162
6p21	t(6;14)p21;q24)	n.d.	n.d.	4	PCH	Le traslocazioni coinvolgenti la banda 6p21 sono specifiche di HMGI(Y), identificato con RT-PCR e FISH.	165
	der(8)t(6;8)(p21.3;q12), der(14)t(6;14)(p21;q24)	n.d.	n.d.	1	PCH	Sono casi selezionati con aberrazioni di 6p21	
	t(6;10)(p21;q22.3)	n.d.	n.d.	1		Sono casi selezionati con aberrazioni di 6p21	
	der(6)t(6;14)(p21;q24), der(14)(6;14)(p21;q24), der(16)t(6;16)(p21;p12qter)	n.d.	n.d.	1		Sono casi selezionati con aberrazioni di 6p21	
	der(1)t(1;14)(p34;q24), der(6)t(6;17)(p21;q23), der(14)t(6;14)(p21;q24) der(17)t(1;17)(p34;q23)	n.d.	n.d.	1		Sono casi selezionati con aberrazioni di 6p21	
	der(6)inv(6)(p21.3q26) t(6;9)(q26;p13), der(9)t(9;14)(p13;q24), der(14)t(6;14)(p21.3;q24)	n.d.	n.d.	1		Sono casi selezionati con aberrazioni di 6p21	

**Tabella VII.4quater Aberrazioni cromosomiche del fenotipo Adenoma**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
	der(4)t(4;6;14)(q35;p21;3p25), der(6)t(4;6)(q35;p11.2), der(14)t(6;14)(p12p21.2;q24)	n.d.	n.d.	1	PCH	Sono casi selezionati con aberrazioni di 6p21	
12q14-15	n.d.	n.d.	n.d.	5/7	ADENOMA PLEOMORFO DELLE GHIANDOLE SALIVARI	7 casi su 7 hanno il punto di rottura a livello della MAR, 5 egli stessi mappano anche all'interno della regione ristretta di 175 Kbasi.	159
12q13-15	t(8;12)(q21.2-21.3;q13-15)	n.d.	n.d.	1	Linea cellulare Ad-211/SV40		159
	ins(12;6)(q15;q16q21)	n.d.	n.d.	1	Ad-248/SV40		
	ins(12)(q15;q24.1)	n.d.	n.d.	1	Ad-263/SV40		
	t(8;12;18) (p12;q14;p11.2)	n.d.	n.d.	1	Ad-295/SV40		
	t(7;12)(q31;q14)	n.d.	n.d.	1	Ad-302/SV40		
	t(1;12)(p22;q15)	chimerico	n.d.	1	Ad-312/SV40	Fusione tra sequenza HMGI-C specifica a livello della regione 3' non tradotta con sequenza ectopica dal cr. 1p22.	
	inv(12)(p13q15)	n.d.	n.d.	1	Ad-366/SV40		
	t(12;14)(q13-15;q13-q15)	n.d.	n.d.	1	Ad-386/SV40		
12q14-15	t(9;12)(p23;q14)	chimera con NFIB	n.d.	2	Adenoma pleomorfo ghiandole salivari	Solo 5 amminoacidi di NFIB vengono aggiunti alla proteina HMGI-C tronca.	168
12q14-15	t(3;12)(p21;q14)	chimera con FHIT	n.d.	1	Adenoma pleomorfo ghiandola parotide		163

**Tabella VII.5 Aberrazioni cromosomiche in altri tumori di origine mesenchimale**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
12q15	r(12q15) "Ring chromosomes" +Giant marker chr. +Double minute chr.		HMGI-C	30	TUMORE ATIPICO LIPOMATOSO (ALT)	Si tratta di una cospicua Amplificazione genica che colpisce <i>HMGI-C</i> nell'85% dei casi; ne risulta un'attivazione a livello trascrizionale che sfocia in alti livelli della proteina evidenziati mediante tecniche immunostochimiche. Questi ALT comprendono liposarcomi differenziati e lipomi atipici, comunque si tratta della controparte maligna dei lipomi.	173
6p	?			2	ALT	<i>HMGI(Y)</i>	
	normale			1	ALT	Non si conosce la causa.	
12q14-15	n.d.	n.d.	n.d.	1	FIBRO- ADENOMA	Fibroadenoma della mammella, l'aberrazione cromosomica risiede all'interno delle 175 Kbasi caratteristica della MAR.	159
12q15	t(4;12) (q27;q15) clonale	n.d.	n.d.	1	FIBRO- ADENOMA	È un fibroadenoma intracanicolare della mammella. Questa aberrazione risulta essere clonale e definisce un sottogruppo di adenomi della mammella.	174
12 2	46, XX, der(2) inv(2) (p25q21) ins(2;12) (p25;q13q14) t(2;12) (q21;q15) der(12) (q13q14) del(12) (q15)	n.d.	n.d.	1	POLIPO ENDO- METRIALE	Riarrangiamento a livello del 3° introne di <i>HMGI-C</i> molto complesso.	175
12q14-15	n.d.	n.d.	n.d.	1/1	ANGIOMIXOMA	1 caso su 1 ha il punto di rottura all'interno della MAR proprio nell'intervallo delle 175 Kbasi.	159



**Tabella XX.5bis Aberrazioni cromosomiche in altri tumori di origine mesenchimale**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
12q14-15	n.d.	n.d.	n.d.	1/1	POLIPO DELLO ENDOMETRIO	1 caso su 1 ha il punto di rottura all'interno della MAR proprio nell'intervallo delle 175 Kbase.	159
MAR 12q13-15	Amplificazione	Si osservano alti fino ad altissimi livelli di espressione e anche trascritti di peso molecolare anomalo.	n.d.	1	ISTIOCITOMA FIBROSO MALIGNO	Da uno studio di ibridazione genomica comparativa effettuato su 122 sarcomi umani di diverso tipo sono emerse evidenze di Amplificazione genica coinvolgenti i geni caratteristici della regione MAR, quali i geni MDM2, CDK4 e SAS. Su alcuni di questi , qui a fianco riportati, sono stati effettuati studi più accurati evidenziando trascritti HMGI-C specifici mediante Northern blot.	176
			n.d.	1	EMANGIO- PERICITOMA		
			n.d.	12	LIPOSARCOMI		
			n.d.	1	SCHWANNOMA MALIGNO		
			n.d.	1	SARCOMA NON CLASSIFICATO		
			n.d.	2	OSTEOSARCOMI		
	Amplificazione + Delezione	chimerico	n.d.	1	LIPOSARCOMA	Questo tumore mostra amplificazione genica di HMGI-C però mancante degli esoni 4° e 5°, fuso a sequenza ectopica (10 triplette codificanti) sconosciuta.	

**Tabella XX.5ter Aberrazioni cromosomiche in altri tumori di origine mesenchimale**

LOCUS	CARIOTIPO (Aberrazione)	Specie espresse		N° CASI	FENOTIPO	OSSERVAZIONI	REF.
		RNA	PROT.				
12q13-15	t(12q)	1100 pb, 600 pb, 350 pb	n.d.	1	LEUCEMIA ACUTA MIELOGENICA (AML)	Si tratta di un riarrangiamento non casuale, i pazienti con questo cariotipo hanno prognosi infausta. Attraverso la tecnica 3'RACE si sono evidenziati 3 tipi di trascritti di 1100 pb, 600 pb e 300 pb che sono tutti troncati dopo il 3° esone e fusi a sequenze ectopiche. Due linee cellulari leucemiche, ML3 e BV173, esprimono l'isoforma più piccola.	177
		350 pb	n.d	3	AML,ML3, BV173		

L'interpretazione della tabella VII deve tener conto del fatto che 1. Spesso i casi sono selezionati proprio per un certo cariotipo; 2. Dove non risulta l'espressione di un certo tipo di HMGI non significa che non c'è, ma semplicemente che l'analisi non ha incluso l'identificazione di quella precisa specie. La tabella VII è suddivisa in più sezioni a seconda del fenotipo tumorale.

n.d. = non documentato

### 1.2.9 HMGI, Mesenchima ed Adipogenesi

Il mesenchima è un tessuto connettivo embrionale, costituito da cellule di forma stellata, immerse in abbondante sostanza fondamentale. Si origina dal foglietto medio dell'embrione (mesoderma) e dalle creste neurali. Le cellule del mesenchima sono in grado di dividersi rapidamente e si distribuiscono in tutto l'embrione, diventando le progenitrici della maggior parte delle cellule fisse dei connettivi. I fibroblasti, i condroblasti, gli osteoblasti e le cellule reticolari sono le cellule del tessuto connettivo, di derivazione mesenchimale, che si sono differenziate per la sintesi e la secrezione delle proteine fibrose e dei proteoglicani. Nei tessuti connettivi si possono inoltre trovare cellule adipose, macrofagi e mastociti.

Il tessuto adiposo è una varietà di tessuto connettivo lasso i cui elementi cellulari sono rappresentati dalle cellule adipose, che accolgono nel loro citoplasma gocce lipidiche. Il tessuto adiposo ha un ruolo importantissimo nell'assicurare un apporto costante di materiale energetico all'organismo; le sue cellule provvedono direttamente alla sintesi dei grassi e alla loro liberazione e assolve inoltre ad una funzione meccanica di sostegno e protezione. Il tessuto adiposo bianco, molto più rappresentato rispetto a quello bruno, è distribuito in molti distretti del corpo, come la tela sottocutanea, nelle regioni ascellari ed inguinali, nonché nei glutei e nelle loggie renali, all'interno dell'addome nel grande omento, nel mesentere e nella cavità retroperitoneale, in associazione con le gonadi e nelle cavità orbitali.

Il gene *HMGI-C* (sia umano che murino) presenta una distribuzione d'espressione ristretta a tessuti embrionali differenziati di origine mesenchimale. Questo fatto, associato alla crescita anormale delle cellule adipose nei lipomi, depone a favore di un presunto ruolo chiave di *HMGI-C* nell'adipogenesi. A supportare questa ipotesi si pongono gli studi sul topo pigmeo descritti nella sezione 1.2.6 (42, 128).

Culture di fibroblasti embrionali da embrioni omozigoti per la mutazione *pygmy* presentano una proliferazione cellulare quattro volte inferiore ai fibroblasti di embrioni normali, suggerendo così un ruolo attivo di *Hmgi-c* nella proliferazione cellulare e probabilmente nello sviluppo peculiare degli adipociti.

Analogamente trasfezioni di cellule pre-adipocitiche 3T3-L1 con costrutti antisenso dell'*HMGI(Y)* (179) hanno indotto un grosso aumento del tasso di crescita cellulare, tanto da supporre che nella proliferazione del tessuto adiposo il processo di sviluppo sia molto più

complesso di quanto supposto, e richieda un bilanciamento corretto tra l'HMGI(Y), che sembra agire da effettore negativo, e HMGI-C, correlata a fenomeni di iperproliferazione.

Dati a supporto derivano da recentissimi studi su topi transgenici esprimenti una forma di HMGI-C priva della regione spaziatrice e della coda C-terminale. (180, 181) Il transgene costituito solo dai primi tre esoni è stato scelto come alternativa ad un costrutto di fusione tipico dei lipomi umani (ad es. LIM domain), proprio perché gli studi sui lipomi ed altre neoplasie benigne e su fibroblasti trasfettati sia con la proteina tronca che chimerica (157) hanno suggerito che fosse la perdita della regione C-terminale piuttosto che la fusione con domini ectopici funzionali a scatenare la formazione delle neoplasie.

Al contrario del topo pigmeo, in questi topi si sviluppa un gigantismo, e pur essendo la proteina tronca espressa in tutti i tessuti, si ha un'abbondanza di tessuto adiposo, con una tendenza all'obesità, ed un'alta incidenza di lipomi. Uno dei due gruppi di ricerca ha stabilito in particolare la presenza di masse addominali adipose abbondantissime, come anche attorno ai reni, nel mediastino e alla base del cuore. Inoltre l'analisi istologica ha evidenziato molte anomalie ed infiammazioni all'apparato urogenitale, probabilmente dovuto a compressione dello stesso per opera dell'aumentato ingombro del tessuto adiposo. (180) Lo stato differenziativo del tessuto adiposo era terminale, come stabilito dall'espressione di diversi marcatori differenziativi, come il gene aP2, PPAR $\gamma$  (recettore  $\gamma$  attivato dal proliferatore del perossisoma), la lipoprotein lipasi (LPL), leptina, e la fosfoenolpiruvico carbossichinasi (PEPCK). (180)

Minime differenze tra i due studi possono essere dovute all'uso di differenti promotori posti artificialmente a monte del gene troncato, in quanto il gruppo di Battista et al. utilizza un promotore virale fortissimo (CMV) mentre Arlotta et al. ne usano uno murino (H-2K<sup>b</sup>), ubiquitario, costitutivamente attivo a livelli moderati nel topo. Infatti l'insorgenza e l'entità del gigantismo, associata a lipomatosi nei topi di Battista et al. sembra dipendere dalla quantità di HMGI-C tronca prodotta.

Le ricerche del secondo gruppo (Arlotta et al) concordano con l'aumento della massa adiposa, ma la differenza di peso tra i transgenici e i topi naturali non è più significativa dopo i tre mesi di crescita. L'analisi istologica dettagliata di 30 topi mostra abbondanza di adipe a livello dell'ilus renale ed alla base del cuore e nella regione mediastinica intorno allo stesso. Un terzo degli animali esibivano anomalie istologiche tali da suggerire un potenziale trasformante della proteina, indipendentemente dal sito di integrazione del transgene, perché i fondatori non erano gli stessi.

Questo studio fornisce l'evidenza diretta che la proteina HMGI-C tronca può funzionare analogamente ad un oncogene tradizionale. Dato che il polipeptide tronco è espresso praticamente in tutti i tessuti dei topi transgenici, è molto significativo il fatto che si sia prodotto un solo tipo di tumore. Infatti osservano la mancanza di tutti gli altri tumori di origine mesenchimale, come quelli delle ghiandole salivari, del miometrio uterino e di tessuti fibrosi. Un fatto del genere contrasta chiaramente con l'espressione di altri tipi di oncogeni, che funzionano in modo più pleiotropico in topi transgenici. (181)

Un'ulteriore scoperta è stata quella di trovare uno stato esteso di steatosi (infiammazione del tessuto adiposo) in quasi metà dei topi transgenici. Tale condizione è virtualmente assente in topi non transgenici ed è dovuta probabilmente ad una reazione autoimmunitaria verso antigeni specifici dell'adipe. (181)

In conclusione è molto probabile che HMGI-C abbia un ruolo importante e specifico nell'adipogenesi normale ed alterata. L'osservazione dell'abbondanza di tessuto adiposo in topi transgenici per la HMGI-C tronca assieme al fenotipo pygmy in topi HMGI-C-deficitari suggerisce inoltre che HMGI-C differisce dalla proteina HMGI(Y) altamente omologa, in quanto dimostra di avere un ruolo biologico molto ristretto ed individuale. Molto probabilmente esiste un gruppo particolare di geni bersaglio della HMGI-C, diversamente dalla HMGI(Y) che appare come facilitatore universale della trascrizione genica.

## Capitolo 2.

### Scopo della tesi

Data la rilevanza dei fattori HMGI nella regolazione della trascrizione, si è voluto analizzare innanzitutto l'interazione di queste proteine con diversi elementi promotoriali. Avendo alta affinità verso sequenze ricche in AT ci è sembrato interessante studiare elementi regolati da fattori trascrizionali contenenti un omeodominio HD, che classicamente riconoscono un motivo centrale 5'-TAAT-3' (124), *in vitro*, mediante la tecnica del *bandshift*. Una volta saggiata l'attività di legame e l'effetto sull'interazione di diverse proteine aventi un HD allo stesso elemento *in vitro*, si potrà valutare, mediante cotrasfezioni in sistemi cellulari, se anche *in vivo* l'effetto osservato si esplica su di un promotore regolato da fattori HD, posto a monte di un gene marcatore.

Analogamente, per poter caratterizzare la funzione biologica di HMGI-C a livello dell'elemento PRDII del promotore dell'Interferone  $\beta$  umano e paragonarla con quella evidenziata per HMGI(Y) (48), si è voluto valutare la reale capacità di interazione specifica *in vitro* fra i suddetti fattori e questa sequenza regolativa attraverso una serie di esperimenti di *bandshift*. A questo proposito dovranno essere prodotte le relative proteine in sistemi procariotici mediante vettori di espressione opportunamente preparati con tecnologie di DNA ricombinante. Sono inoltre in programma saggi di attività per valutare le potenzialità di HMGI-C anche *in vivo*, usando promotori artificiali costituiti da elementi PRDII.

Alti livelli di espressione delle proteine HMGI sono una caratteristica del fenotipo tumorale.(136) Molti gruppi di ricerca hanno valutato la presenza dei trascritti di queste proteine in diversi tipi di neoplasie umane, correlando i diversi gradi di displasie con l'incremento dell'espressione delle stesse. È stato quindi prodotto un anticorpo contro l'N-terminale della proteina HMGI(Y) che si è rivelato un ottimo anticorpo in esperimenti di *Western blot*, ma in seguito anche nell'analisi immunostochimica di diversi carcinomi umani, come quelli del colon e della tiroide.(147, 148) È stato di nostro interesse saggiare la bontà di questo anticorpo su carcinomi della cervice uterina, per confrontare i dati ottenuti con un altro anticorpo, contro un marcatore di proliferazione Ki-67, correntemente usato per

la diagnosi ospedaliera. Inoltre si è voluto vedere se era possibile stabilire una correlazione tra la gravità della displasia diagnosticata morfologicamente con l'aumentare dell'immunoreattività dei tessuti saggiati.

Infine, dato che la riespressione di HMGI-C durante la trasformazione neoplastica è dovuta ad un'attivazione a livello trascrizionale (156) un'ampio studio sulla regolazione dell'espressione di *Hmgi-c* è stata iniziata cercando di isolare innanzitutto il rispettivo promotore per poterlo saggiare in modelli cellulari di confronto tra cellule normali e trasformate. Lo scopo finale sarebbe quello di scoprire i meccanismi molecolari con cui questo gene è regolato, soprattutto nel contesto della trasformazione neoplastica.

## Capitolo 3.

### Materiali e Metodi

#### 3.1 Analisi del DNA

##### 3.1.1 Estrazione di DNA plasmidico

Fra i numerosi metodi di estrazione del DNA plasmidico esistenti, sono stati scelti protocolli in uso corrente, descritti nel manuale di *Molecular Cloning* (182), in base al grado di purezza e alle quantità richieste dalle successive applicazioni.

Sistemi commerciali (GENOMED, Quiagen) sono stati usati per preparazioni che permettono di ricavare DNA plasmidico di elevata purezza, utilizzabile per esperimenti di trasfezione.

##### 3.1.2 Elettroforesi in gel d'agarosio

Con questo metodo è possibile separare frammenti diversi di DNA a seconda del loro peso molecolare. Infatti l'agarosio solidificando forma una specie di reticolo, le cui maglie agiscono da setaccio molecolare, ritardando tanto più la migrazione dei frammenti quanto più grandi sono le loro dimensioni. In funzione delle dimensioni dei frammenti in analisi si stabilisce la percentuale di agarosio (in un intervallo tra 0,3% e 3%) e quindi la porosità del gel: un gel al 3% consente di separare frammenti al di sotto di 0,1 Kb, mentre con gels allo 0,3% si possono risolvere anche molecole di 50 Kb.

Questo metodo di separazione ha trovato largo impiego, sia a scopi analitici che preparativi. E' possibile infatti sia ricavare informazioni sulle dimensioni di un frammento di DNA in analisi che stimarne la concentrazione, mediante confronto con adeguati riferimenti a peso molecolare e a concentrazione noti.

##### Preparazione del gel

L'agarosio viene sciolto ad una temperatura superiore a 90°C in tampone TAE 1X, addizionato con bromuro di etidio alla concentrazione finale di 0,5 µg/ml: questa molecola fluorescente è un intercalante del DNA che, legandosi ad esso, subisce un effetto ipercromico nel campo degli ultravioletti e in questo modo consente di visualizzare la catena nucleotidica con un apposito transilluminatore.

Questa soluzione viene versata in una cassetta elettroforetica orizzontale, dotata di spaziatori e pettine, e lasciata solidificare. Rimossi gli spaziatori e il pettine, la cassetta viene riempita di tampone di corsa TAE 1X.



## **Preparazione dei campioni e separazione elettroforetica**

Al campione in analisi si aggiunge, in rapporto 1:10 in volume, il tampone di caricamento 10X e si deposita nel pozzetto la miscela con una micropipetta.

Si effettua una corsa anodica a 60-80 V e al termine si osserva il gel su di un transilluminatore a luce ultravioletta.

## **Estrazione di DNA da gel**

E' possibile separare frammenti di DNA di diverso peso molecolare mediante elettroforesi in agarosio e successivamente recuperare la banda contenente il DNA d'interesse, ritagliandola dal gel, grazie al fatto che può essere evidenziata con un transilluminatore ad UV. A seconda delle dimensioni del frammento da recuperare si utilizzano due metodi alternativi: per molecole di grandi dimensioni si può utilizzare gel d'agarosio 1%, ma per frammenti più piccoli al di sotto di 300 bp invece si impiega gel d'agarosio a basso punto di fusione. Di conseguenza si distinguono anche le tecniche per l'estrazione dalla banda del gel.

### **Estrazione da gel d'agarosio 1%**

E' possibile sciogliere l'agarosio anche a temperature relativamente basse, in modo da non danneggiare il DNA,

- ritagliata la banda d'interesse da gel, si aggiungono 3 volumi di una soluzione di sodio perclorato (un sale caotropo), che favorisce la disgregazione della matrice del gel;
- si incuba 10 minuti a 45-55°C, accertandosi che tutto l'agarosio venga ben sciolto;
- per estrarre il frammento maggiore di 200pb, si sfrutta una specifica matrice (*BIO-RAD*) che è in grado di interagire con il DNA, lasciando il campione in incubazione per almeno 30 minuti a temperatura ambiente;
- si centrifuga a 15000g per 5 minuti e, scartato il surnatante, si risospende in altri 0.5 ml di sodio perclorato e si attua un'ulteriore centrifugazione;
- si effettuano tre lavaggi con 0,5 ml di soluzione di lavaggio.
- Eliminato il surnatante, si aggiungono 30 µl d'acqua e si eluisce per 10 minuti a 45-55°C;
- si centrifuga brevemente e si recupera il surnatante;
- il fondello di matrice viene ulteriormente risospeso in 15 µl d'acqua e, ripetute l'eluizione e la centrifugazione, si unisce il surnatante al precedente.

### **Estrazione da gel a basso punto di fusione**

Si usa un gel a basso punto di fusione, solitamente per frammenti piccoli che non si estraggono con le palline di vetro c.s.

- La banda d'interesse, visualizzabile agli UV, viene ritagliata con una lametta e posta in un eppendorf con due volumi di acqua e NaCl 0.2 M;
- l'agarosio viene sciolto incubandolo per alcuni minuti a 68 °C;
- infine, si purifica il DNA con un'estrazione fenolo/cloroformio.

### 3.1.3 Manipolazioni del DNA

#### Reazione a catena della polimerasi (PCR)

Questa tecnica messa a punto alla metà degli anni '80, permette la rapida amplificazione di frammenti di DNA lunghi fino a 1000 pb. Secondo il protocollo base, l'attuazione di questa metodica prevede la sintesi di due oligonucleotidi, i quali presentano tipicamente una composizione nucleotidica complementare a sequenze che:

- giacciono sui filamenti opposti del DNA stampo;
- fiancheggiano il segmento di DNA da amplificare.

La reazione si basa su cicli ripetuti di :

- denaturazione ad alta temperatura, che permette l'apertura del DNA stampo;
- rinaturazione, necessaria per avere l'ibridazione degli oligonucleotidi con le sequenze target;
- estensione degli inneschi, catalizzata dall'azione della DNA polimerasi di *Thermus aquaticus*.

Quest'enzima è termostabile: può, infatti, sopportare prolungate incubazioni a 95 °C, senza subire inattivazione durante il passaggio di denaturazione del DNA, perciò non necessita di essere rimpiazzato ad ogni ciclo d'amplificazione. Ne esiste anche una forma manipolata geneticamente e sintetizzata in *E. coli* (AmpliTaq™), che mantiene la caratteristica attività polimerasica.

A partire dal terzo ciclo in poi l'amplificazione segue un andamento esponenziale e, dopo 20-30 cicli, il segmento di DNA definito dalle terminazioni 5' degli inneschi, e la cui lunghezza è pari alla distanza tra i due *primers*, rappresenta il 100% del DNA totale (il DNA stampo originale ne costituisce un'infima parte).

La reazione viene condotta in un termostato *MiniCycler* (*MJ Research*) che può essere programmato impostando:

- le condizioni di temperatura,
- e la durata delle varie fasi della reazione.

Tali parametri vengono adattati in funzione della quantità di DNA di partenza disponibile, della lunghezza del tratto da amplificare e del valore della  $T_m$  degli inneschi oligonucleotidici.

#### Processo di amplificazione

La metodica utilizzata prevede che al campione denaturato per bollitura vengano poi addizionati:

- 50 pmoli di ciascuno dei due inneschi;
- 2 µl di una soluzione di dNTP 10 mM;
- 5 µl del tampone 10× per PCR a pH 8.8;
- 3 µl di MgCl<sub>2</sub> 25 mM (1.5 mM finale);
- 0.5 µl di Taq DNA polimerasi (*Promega*);

H<sub>2</sub>O a 50 µl finali.

Il profilo della reazione di amplificazione prevede i seguenti passaggi:

- 4 minuti di denaturazione a 94°C seguita da 25 cicli di:
- 1 minuto di denaturazione a 94°C;

- 1 minuto di ibridazione a 50°C;
  - 1 minuto e trenta secondi di allungamento a 75°C;
- l'ultimo passaggio prevede 5 minuti di ulteriore allungamento a 75°C.
- Prima di far procedere la reazione viene depositata qualche goccia di olio minerale sopra la miscela in modo tale da impedirne l'evaporazione.
- I prodotti vengono poi caricati su gel d'agarosio confrontandoli con riferimenti di peso molecolare noti.

### **Subclonaggio di prodotti da PCR**

La PCR ha il vantaggio di fornire sequenze di DNA che possono anche essere subclonate in vettori d'amplificazione e/o espressione, ciò prevede però alcuni passaggi per la preparazione del frammento.

### **Trattamento con l'enzima Klenow Fragment**

Questa reazione è necessaria per riempire potenziali discontinuità lasciate dalla Taq DNA polimerasi.

La reazione viene allestita come segue con 45 µl della reazione da PCR a cui vengono addizionati:

- 1 µl della soluzione di dNTP 10 mM;
- 0.5 µl del frammento Klenow della DNA polimerasi (1U/µl);
- MgCl<sub>2</sub> 5 mM finale.

La reazione procede per 30' a temperatura ambiente, con successiva denaturazione dell'enzima a 68°C per 3'.

- Si realizza poi un'estrazione fenolo/cloroformio, seguita da una precipitazione con etanolo.

### **Trattamento con la polynucleotidilchinasi (PNK)**

I prodotti da PCR presentano le estremità 5' prive di gruppo fosfato; elemento essenziale perchè la DNA ligasi possa attuare la reazione di ligazione tra vettore ed inserto. Per ovviare a ciò si tratta il campione con la polinucleotidilchinasi (PNK) che catalizza il trasferimento di un gruppo fosfato dall'ATP al 5'-OH del DNA.

Il protocollo prevede che a 0.1 µg di amplificato vengano aggiunti:

- 5 unità di PNK (10U/µl);
- ATP 10 µM finale;
- 3 µl di OPA (One Phor All) 10×

H<sub>2</sub>O fino a 15 µl.

segue un'incubazione a 37°C per 1 h e quindi l'inattivazione enzimatica a 80°C per 10'.

### **Digestione del DNA con enzimi di restrizione**

Un'esonucleasi è un enzima che digerisce DNA o RNA a partire da una delle terminazioni (5' e/o 3') della molecola. Un'endonucleasi, invece, digerisce a livello di siti interni:

- le endonucleasi di restrizione di tipo I operano a distanze indeterminate dal sito *target*,

- mentre quelle di tipo II digeriscono il DNA a livello di una specifica coppia di basi nel contorno definito di una sequenza *consensus*.

I numerosi enzimi di restrizione disponibili in commercio riconoscono infatti specifiche sequenze consensus, costituite da 4 a 8 b, al cui livello operano un taglio che può originare estremi coesivi o tronchi, a seconda dell'enzima stesso. Oltre a questo l'attività di ogni enzima è caratterizzata da un diverso *optimum* per quanto riguarda pH, temperatura, forza ionica e tipo di sali presenti; in ogni caso solitamente insieme all'enzima viene fornito anche l'appropriato tampone di reazione 10X.

La reazione di digestione generalmente viene allestita mantenendo i seguenti rapporti:

- 1 µg di DNA plasmidico
- 1 unità di enzima
- 1 µl di tampone specifico 10X

in un volume totale di 10 µl.

Per la maggior parte degli enzimi la temperatura ottimale è di 37°C.

Se lo scopo della digestione è analitico bastano 45 minuti di reazione, se invece si vuole digerire completamente una quantità maggiore di DNA, a scopi preparativi, è necessaria almeno un'ora e mezza.

### **Digestione con nucleasi**

Fra i diversi tipi di nucleasi di restrizione hanno assunto un particolare impiego in biologia molecolare la Bal 31 ottenuta da *Alteromonas espejiana* e la nucleasi S1 ottenuta da *Aspergillus oryzae*.

La Bal 31 rimuove progressivamente nucleotidi dall'estremità di frammenti di DNA a doppia elica, senza specificità direzionale. Date le sue caratteristiche viene ampiamente utilizzata nella formazione di delezioni progressive di regioni di DNA.

Le condizioni di reazione dei due enzimi sono le seguenti:

### **Digestione con la Bal 31**

- ◆ 5 µg DNA
- ◆ 20 µl tampone specifico 5X
- ◆ 2,5µl BSA 0,1 µg/µl
- ◆ Bal 31 (Amersham) 2,5 U/µg DNA

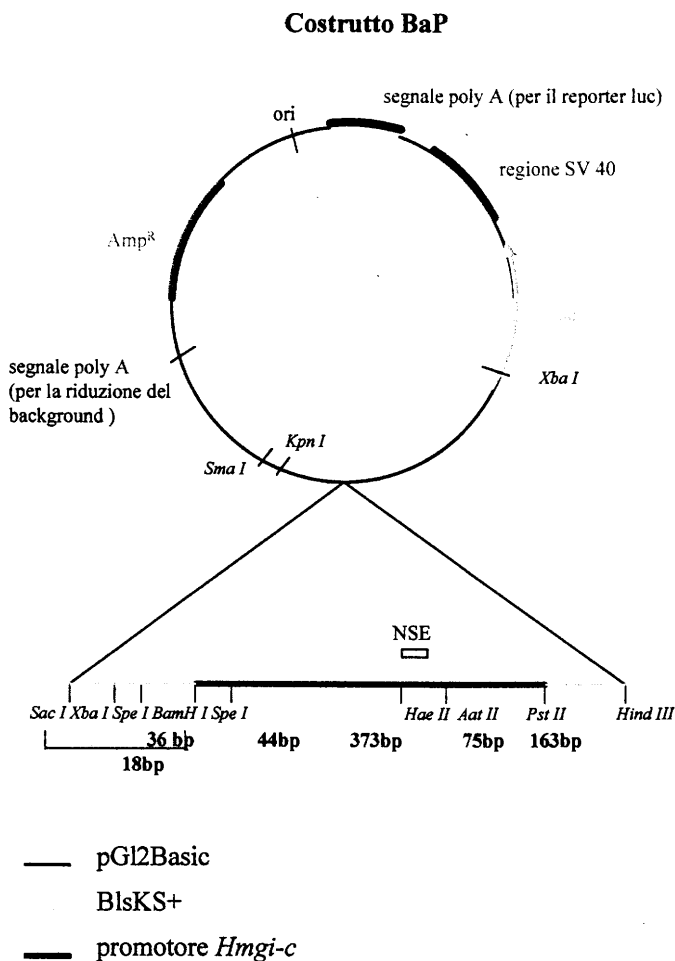
la reazione avviene in un volume di 100 µl alla temperatura di 30°C per tempi variabili fra 30 e 90 minuti. L'inattivazione dell'enzima avviene con l'aggiunta di 10 µl EDTA 0,5 M pH 8 e l'incubazione alla temperatura di 75°C per 10 minuti.

### **Digestione con la nucleasi S1**

- ◆ 6 µg DNA
- ◆ nucleasi S1 (Boehringer) 5U/µg DNA
- ◆ 10 µl tampone d'incubazione 10X

il tutto in un volume finale di 100 µl, e mantenuto a 42°C per 25 minuti per la linearizzazione del plasmide, mentre per ottenere un plasmide nella forma rilassata il tempo di reazione viene ridotto a 5 minuti con 1U di enzima per µg DNA.

Qui di seguito è visualizzato il costrutto BaP di *Hmgi-c*, usato negli esperimenti di sensibilità alla S1 nucleasi:



### Trattamento con la T4 DNA polimerasi

Il taglio condotto da alcuni tipi di enzimi di restrizione o dalle nucleasi può lasciare le estremità del DNA coesive in modo che il duplex sia presente con estremità 5' oppure 3' sporgenti. Qualora si voglia riunire estremità di DNA di questo tipo, c'è la necessità di renderle precedentemente piatte, in modo da permettere la ligazione. L'operazione può basarsi sul riempimento del sito di restrizione mediante una sola reazione di polimerizzazione, quando il filamento 5' è sporgente, oppure combinato con un precedente trattamento di eliminazione dei nucleotidi qualora si tratti di un'estremità 3' sporgente.

Gli enzimi utilizzabili a questi scopi sono diversi e commercialmente disponibili. Nel caso della T4 DNA polimerasi, questo enzima è dotato di una forte attività esonucleasica 3'-5' condotta in assenza di nucleotidi, ed anche di una discreta attività polimerasica 5'-3', la quale richiede sia uno stampo, sia del DNA che funga da innesco come ovviamente i nucleotidi.

Le condizioni di reazioni, generalmente consigliate dalla casa produttrice, prevedono l'utilizzo:

- 3-5 µg di DNA (con estremità coesive)
- 5µl di tampone specifico 10X (*NEB*) (oppure adattando le condizioni del tampone di reazione precedente a quello attuale)
- 1,5 µl di T4 DNA polimerasi 3U/µl (*NEB*)
- 2,5 µl di BSA 1µg/µl
- portati a 50 µl finali con acqua.

La reazione è condotta a 12°C per 15 minuti, per permettere all'enzima di rimuovere i nucleotidi in direzione 3'-5'. Successivamente segue una fase di polimerizzazione con lo scopo di rendere le estremità completamente piatte, qualora la reazione precedente si sia protratta oltre l'estremità sporgente. Perciò alla miscela di reazione vengono aggiunti

- 2 µl di una miscela dei quattro deossinucleotidi (dNTP 10 mM).

La reazione viene incubata a 37°C per 15 minuti.

### Defosfatazione del vettore

Per preparare un vettore plasmidico al fine di inserirvi un frammento di DNA a seguito di un processo di restrizione è necessario rimuovere i gruppi fosfato alle sue estremità 5', in modo tale da impedirne la ricircularizzazione se privo del segmento di DNA d'interesse. La defosfatazione viene operata anche nel caso di vettori digeriti con due enzimi di restrizione che generano estremi non compatibili perchè, in caso di digestione parziale, c'è comunque la possibilità che il vettore si richiuda su se stesso.

Quindi, a seguito della restrizione, si porta il pH del tampone a 8 con l'aggiunta di Tris/HCl 1 M (pH 8) e si incuba a 37°C per un'ora con 1 µl di fosfatasi alcalina (1,7 U/µl).

### Ligazione vettore-inserito

La DNA ligasi del batteriofago T4 è un enzima chiave per la costruzione di molecole ricombinanti. In particolare, quest'enzima ha l'abilità di legare, in presenza di Mg<sup>2+</sup> e ATP, non solo frammenti di DNA con estremità coesive e complementari, ma anche quelli con terminazione tronche.

La miscela di reazione è composta da:

- 20-40 ng di vettore, defosforilato con fosfatasi alcalina;
- frammento in concentrazione molare doppia rispetto al vettore;
- 1 µl di tampone specifico 10X;
- 1 µl di T4 DNA ligasi (2 U/µl);

H<sub>2</sub>O fino a 10 µl.

Si incuba per 12 ore a 16°C. Le reazioni di ligazione vengono inseguite inserite in batterio mediante trasformazione per produrre il plasmide ottenuto dalla ligazione.

### Costruzione dei vettori reporter di *Hmgi-c*/luciferasi

Un frammento BamHI-PstI derivato dal clone genomico λ I-C.5 (37), contenente il sito di inizio della trascrizione è stato inserito nel sito di clonaggio multiplo del plasmide pBlsKS+ (Stratagene). Lo stesso frammento è stato escisso con SacI ed HindIII e clonato nel sito multiplo di clonaggio del vettore reporter con la luciferasi pGL2-Basic (Promega). Questo costrutto è chiamato BaP. Il costrutto più lungo di 3.2 Kbasi (SaP)

stato ottenuto subclonando un frammento genomico SacI- SpeI in BaP, sfruttando SacI del vettore al 5' e SpeI interno al frammento BamHI-PstI. A partire da SaP sono stati preparati dei mutanti per delezione al 5' usando i siti di restrizione PvuII, SacII, HaeII ed AatII interni alla sequenza di Hmgi-c, ottenendo i costrutti PvuP, SsaP, HaP ed AaP. Il costrutto 4P è stato ottenuto per ligazione di un frammento di PCR ( da -282 a +187 rispetto al sito di inizio della trascrizione), tagliato con PstI ed inserito nel vettore BaP tagliato con SmaI e PstI. 4P-HA è un mutante di delezione di 4P, ottenuto tagliandolo con HaeII e AatII, rendendo le estremità piatte e rilegandole. Delezioni al 3' sono state ottenute digerendo SaP con HindIII, che taglia a -1325 nel promotore e nel sito multiplo di clonaggio, e rilegandolo. SaB è stato costruito subclonando un frammento genomico SacI-BamHI in pGL2-Basic.

NF1mut: per preparare questo costrutto è stata fatta una PCR con l'oligo mD senso (vedi sezione bandshift) ed l'oligonucleotide pGL2 antisense specifico del vettore reporter pGL2 Basic (Promega). L'amplificato è stato digerito al 5' con AatII e al 3' con HindIII e subclonato nel vettore HaP previamente tagliato con gli stessi enzimi.

### 3.1.4 Trasferimento di DNA su filtro mediante “Southern blotting”

Al fine di utilizzare sonde radioattive per l'analisi di un campione di DNA, è necessario prima di tutto separarlo mediante elettroforesi in gel d'agarosio per poi trasferirlo per capillarità su di una membrana di nylon (Genescreen NEN/ Du Pont): questa tecnica messa a punto da E. M. Southern consiste infatti nel blotting, ossia nell'assorbimento del DNA su di un filtro.

Il gel viene innanzitutto immerso per 10 minuti in una soluzione NaCl 1,5 M con NaOH 0,5 M, cosicché l'ambiente alcalino denaturi in situ il DNA che in seguito potrà legarsi al filtro come singolo filamento.

Quindi si imbeve una spugnetta con questa soluzione denaturante, vi si sovrappongono tre fogli di carta assorbente (Wathman 3MM) in modo che anch'essi siano imbevuti della stessa soluzione senza formare bolle d'aria, e vi si pone il gel; su di esso viene collocata una membrana delle stesse dimensioni, precedentemente equilibrata in acqua distillata per 10 minuti. Sulla membrana vengono sistemati altri tre fogli di carta 3MM e infine strati successivi di carta assorbente per uno spessore di almeno 10 cm. Completato l'assemblaggio, si applica una pressione costante con un peso di circa 500 x g. In questo modo la soluzione denaturante sale per capillarità dalla spugnetta alla carta sovrastante passando attraverso il gel e determinando il trasferimento del DNA sulla superficie della membrana: questo processo si completa in un intervallo di tempo che varia da 2 a 12 ore a seconda dello spessore del gel.

Al termine il filtro viene lavato brevemente in 2X SSC per rimuovere eventuali residui di agarosio, asciugato ed esposto per 1 minuto ai raggi UV del transilluminatore al fine di formare legami covalenti tra il DNA e la membrana.

### 3.1.5 Marcatura del DNA mediante “Random Primed Radiolabelling”

Questo metodo viene usato allo scopo di ottenere una sonda di DNA radioattiva. Il frammento di DNA da marcare viene denaturato mediante bollitura e poi fatto ibridare con una miscela di esanucleotidi di sequenza casuale, i quali fungono da inneschi della reazione di polimerizzazione ad opera del Klenow Fragment della DNA polimerasi I. Il frammento da marcare (~100ng) viene fatto bollire 3' in presenza di 30 ng della miscela

degli esanucleotidi in un volume di 12  $\mu\text{l}$ . Dopo aver raffreddato la miscela in ghiaccio vengono aggiunti 2  $\mu\text{l}$  del tampone di marcatura 10X, 3  $\mu\text{l}$  di  $^{32}\text{P}$ -dCTP (3000mCi/mmol) ed 1  $\mu\text{l}$  di DNA polimerasi I (3U) in un volume finale di 20  $\mu\text{l}$ . La reazione viene fatta avvenire a  $37^\circ\text{C}$  per 2-3 ore. La sonda viene separata dai nucleotidi non incorporati mediante cromatografia su colonna per gel filtrazione. La colonna viene allestita impaccando la resina G-50, equilibrata nel tampone TNES, in una siringa di plastica da insulina contenente sul fondo della alna di vetro. Per equilibrare la colonna si depositano sulla sua sommità 0.1 ml di tampone TNES e quindi si lascia eluire; questo passaggio viene ripetuto 3 volte. Dopo aver portato il volume della reazione a 100  $\mu\text{l}$  con TNES la sonda viene caricata sulla colonna ed eluita con aggiunte successive di 0.1 ml di TNES, aiutandosi col contatore Geiger per individuare la/le frazione/i eluita/e radioattiva/e. L'attività specifica della sonda così ottenuta è in media di  $5 \times 10^8$  cpm/ $\mu\text{g}$  di DNA.

### 3.1.8 Ibridazione di acidi nucleici su filtro di nylon

Dopo aver effettuato il trasferimento di un acido nucleico su filtro di nylon tramite Southern o Northern blotting o direttamente da colonia, è possibile evidenziare una specifica sequenza mediante ibridazione con una sonda specifica marcata radioattivamente.

E' importante stabilire le condizioni sperimentali in modo da evitare l'interazione aspecifica della sonda con altre sequenze di DNA: vanno quindi fissati temperatura e forza ionica della soluzione di ibridazione. Questi parametri influenzano le condizioni di "stringenza" (alta temperatura e bassa forza ionica destabilizzano la doppia elica), che variano a seconda del grado di omologia della sonda con la sequenza nucleotidica da rilevare.

In particolare quando si sfruttano sonde oligonucleotidiche, è fondamentale stimare la temperatura di semidenaturazione della doppia elica considerata ( $T_m$ ) attraverso una formula empirica:

$$T_m = 16,6 \log(M) + 0,41 (\%G/C) + 81,5 - 820/L$$

con M : concentrazione molare dei sali

L : numero di basi della sonda

E' opportuno operare almeno  $10^\circ\text{C}$  al di sotto del valore così ottenuto sia per l'ibridazione che per i successivi lavaggi.

L'ibridazione avviene in bottiglie di vetro Pyrex poste su di un supporto rotante all'interno di un'incubatore a temperatura variabile.

Il filtro viene equilibrato con 25 ml di una miscela di preibridazione (NaCl 1 M, SDS 1%) per due ore alla temperatura prescelta e in seguito incubato per 12/18 ore alla stessa temperatura nella miscela di ibridazione che contiene anche:

- $0,5 \times 10^6$  cpm/ml di sonda radioattiva, denaturata mediante bollitura;
- mg/ml di DNA competitore di sperma di salmone (la cui funzione è di saturare i siti di legame aspecifico del DNA immobilizzato su filtro);
- soluzione Denhardt a concentrazione finale 1X (che satura il filtro stesso bloccando interazioni aspecifiche con la sonda).

Dopo l'ibridazione, la membrana viene sottoposta a una serie di lavaggi:



- volte con la soluzione 2X SSC per 2 minuti a temperatura ambiente, per eliminare la sonda in eccesso che non ha interagito;
- volte con 2X SSC e SDS 1 % per 15 minuti a temperatura di 10°C al di sotto della corrispondente  $T_m$ ;
- volta con 0,2X SSC e SDS 0,1% per 15 minuti, 10°C al di sotto della corrispondente  $T_m$ ;

in questo modo si aumenta progressivamente la stringenza così da destabilizzare interazioni aspecifiche della sonda, la cui rimozione viene seguita con un contatore Geiger.

Il legame della sonda può essere rilevato in quanto l'emissione di particelle  $\beta$  da parte del  $^{32}\text{P}$  è in grado di impressionare una lastra fotosensibile, che viene sovrapposta al filtro avvolto con una pellicola di polietilene. L'esposizione ha luogo in un'apposita cassetta autoradiografica dotata di schermi intensificatori a -80°C per un periodo variabile da 30 minuti a 2 giorni. Lo sviluppo è condotto in camera oscura mediante le apposite soluzioni di sviluppo e fissaggio (*Kodak*).

### 3.1.9 Trasformazione batterica

E' un processo spontaneo attraverso il quale alcuni ceppi batterici possono captare dal mezzo molecole di DNA esogeno: in natura però questo fenomeno si verifica con una frequenza molto bassa.

La trasformazione consente di inserire dei costrutti plasmidici, anche ricombinanti, all'interno di cellule batteriche e viene utilizzata per amplificarli ed ottenerne grandi quantità, potenzialmente utilizzabili per ulteriori applicazioni.

Sono stati messi a punto diversi metodi per aumentare l'efficienza di questo processo, fra cui l'uso di sostanze chimiche in grado di permeabilizzare la parete batterica.

Le cellule batteriche utilizzate appartengono ai seguenti ceppi:

- XL-1
- DIH 101
- BL21 (DE3): portano integrato nel loro genoma il gene per la T7 RNA polimerasi sotto il promotore inducibile *Lac UV5*; presentano un basso contenuto di proteasi e assenza di resistenza agli antibiotici.

La prima parte della tecnica permette di preparare le cellule batteriche a ricevere DNA esogeno (i batteri vengono resi "competenti") e ha il vantaggio di poter essere eseguita con grandi quantità di batteri poiché quelli che non saranno subito utilizzati potranno essere congelati in  $\text{N}_2$  liquido e conservati a -80°C in aliquote da 100 e

200  $\mu\text{l}$ . Queste, dopo lento scongelamento, possono essere usate nell'arco di diversi mesi senza significative conseguenze sull'efficienza della trasformazione.

#### Preparazione di cellule competenti

- Si diluiscono alcune gocce della sospensione di batteri in glicerolo 50% conservata a -20°C in 1 ml di terreno di coltura SOB lasciato in agitazione a 37°C;
- al raggiungimento della fase logaritmica di crescita ( $\text{OD}_{600} = 0.5$ ) si diluisce ulteriormente in 250 ml dello stesso terreno e si lascia di nuovo crescere (sempre a 37°C) fino alla corretta densità ottica;

- si lascia per 10 minuti in ghiaccio la coltura, e poi si centrifuga a  $3000\times G$  per 15 minuti a  $4^{\circ}C$  e il fondello accuratamente ripulito dal terreno di coltura;
- si risospendono i batteri in 1/3 del loro volume con il tampone di trasformazione (CCMB 80);
- si incuba per 20 minuti in ghiaccio, si ripete l'ultima centrifugazione nelle stesse condizioni e si risospendono i batteri nuovamente nel tampone CCMB 80 ma soltanto in 1/12 del loro volume.

A questo punto le cellule possono essere aliquotate e conservate a  $-80^{\circ}C$  dopo averle congelate per immersione in  $N_2$  liquido.

### Trasformazione

- Quando richiesto le cellule, scongelate lentamente in ghiaccio, vengono mescolate con la soluzione di ligazione nelle giuste proporzioni (50  $\mu l$  di batteri e 5  $\mu l$  di ligazione);
- si incuba per 30 minuti in ghiaccio i batteri;
- si procede a uno shock termico ( $42^{\circ}C$  per 90 secondi) che aumenta la permeabilità della membrana favorendo la penetrazione del plasmide. La reazione viene subito bloccata lasciando i batteri 5 minuti in ghiaccio e quindi si lasciano crescere, in terreno SOC, per 1 ora a  $37^{\circ}C$  in agitazione.
- La sospensione viene versata e fatta assorbire, con un'ansa sterile, su una capsula Petri contenente LB-agar con l'anibiotico adatto.

Le colonie sono visibili dopo 12 ore di incubazione a  $37^{\circ}C$ .

Una buona efficienza di trasformazione, anche se dipendente dal ceppo batterico, si aggira attorno a  $2 \times 10^6$  per  $\mu g$  di DNA.

## 3.1.10 Sequenziamento del DNA

### Metodo dei dideossinucleotidi (Sanger)

Questa tecnica di sequenziamento del DNA si basa sull'utilizzo di nucleotidi modificati (2',3' dideossinucleotiditriposfati) mancanti del gruppo ossidrilico al 3' del ribosio. Il DNA stampo (quello di cui si vuole conoscere la sequenza) è bersaglio della DNA polimerasi del fago T7, che catalizza la sintesi *in vitro* di un filamento complementare a partire da un innesco costituito da un oligonucleotide sintetico. L'allungamento procede finché non viene incorporato nell'elica nascente uno dei 2',3' dideossinucleotiditriposfati presenti nella miscela di reazione. Infatti, dal momento che queste molecole hanno un normale gruppo fosfato in 5', la DNA polimerasi le incorpora normalmente nel DNA; tuttavia una volta incorporato nell'elica, un ddNTP non può formare un legame fosfodiesterico con il successivo dNTP perché privo dell'ossidrilico in posizione 3' e funziona quindi da terminatore specifico dell'allungamento della catena del DNA.

Per questo vengono preparate quattro reazioni separate, ciascuna contenente una quota di un solo ddNTP (in un rapporto ottimale bilanciato con quello del rispettivo dNTP); dato che la polimerasi incorpora casualmente i ddNTP, genera una serie di frammenti di lunghezza diversa, che probabilisticamente rappresentano tutti quelli terminanti con quel ddNTP originabili dallo stampo.

La sintesi avviene in presenza di un dNTP radioattivo ( $\alpha^{32}\text{P}$  dCTP), cosicchè i prodotti delle reazioni possano essere rilevati, in seguito alla separazione elettroforetica, su lastra autoradiografica.

### Reazione

La reazione ha come substrato DNA a doppio filamento e comprende cinque fasi distinte:

**Denaturazione:** perchè il DNA possa fungere da stampo, va innanzitutto denaturato attraverso un ambiente basico e per mezzo del calore.

In 10  $\mu\text{l}$  totali si aggiungono:

- 1  $\mu\text{g}$  di DNA stampo
  - 100 ng di oligonucleotide innesco
  - 2  $\mu\text{l}$  di NaOH 0,5 M
- si incuba a 68°C per 10 minuti.

**Ibridazione:** in modo che l'oligonucleotide innesco possa appaiarsi correttamente allo stampo, è necessario neutralizzare la miscela e stabilire le corrette condizioni di reazione, con l'aggiunta di:

- 2  $\mu\text{l}$  di HCl 0,5 M
  - 2  $\mu\text{l}$  di tampone di ibridazione 10X
- e un'incubazione a temperatura ambiente per 10 minuti.

**Elongazione:** per innescare la reazione di polimerizzazione, si addiziona:

- 3  $\mu\text{l}$  di tampone di elongazione
- 0,5  $\mu\text{l}$  di  $\alpha^{32}\text{P}$  dCTP
- 3 unità di DNA polimerasi del fago T7

Si fa iniziare la sintesi lasciando 5 minuti a temperatura ambiente.

**Terminazione:** la miscela viene suddivisa in 4 aliquote, a ciascuna delle quali vengono addizionati 2,5  $\mu\text{l}$  per ciascuna soluzione di terminazione e si lascia procedere la reazione per altri 10 minuti a 37°C.

**Stop:** ad ogni aliquota vengono aggiunti 5  $\mu\text{l}$  di soluzione di stop.

### Separazione elettroforetica

I prodotti delle reazioni di sequenza si separano mediante migrazione elettroforetica su gel di poliacrilammide denaturante (7 M urea), a una percentuale in grado di risolvere frammenti che differiscono anche per una sola base.

L'elettroforesi avviene in celle verticali, costituite da lastre di vetro (20x40 cm).

I due vetri tra i quali viene colato il gel sono trattati diversamente in modo che, terminata la corsa, il gel aderisca completamente a uno solo di essi: ciò facilita le successive manipolazioni, altrimenti complicate dal ridotto spessore del gel. A tal fine il vetro a cui il gel deve aderire, dopo essere stato pulito accuratamente e asciugato con etanolo, viene trattato con 5 ml di soluzione legante attivata, mentre l'altro con 5 ml di soluzione non legante, che facilita il distacco del gel. Dopo aver rimosso l'eccesso delle soluzioni pulendo nuovamente i vetri con etanolo, si procede all'assemblamento della cella mediante spaziatori laterali di spessore pari a 0,2 mm. Quindi 30 ml della soluzione di acrilamide al 7%, filtrata, degassata e attivata con i catalizzatori (APS e TEMED),

vengono rapidamente versati tra i due vetri e viene inserito un pettine in modo da creare i pozzetti.

Dopo almeno un'ora necessaria affinché la polimerizzazione possa completarsi, la cella viene montata sull'apposito supporto verticale, le cui vaschette anodica e catodica sono riempite di tampone TBE 1X.

Le quattro miscele che costituiscono la sequenza vengono denaturate a 90°C per 3 minuti (in modo da eliminare eventuali strutture secondarie che altererebbero la mobilità dei rispettivi frammenti) e caricate in pozzetti adiacenti (3 µl per ciascuna aliquota).

La corsa viene effettuata al voltaggio costante di 1800 V; visto che la resistenza del gel è considerevole, per effetto Joule nel gel si sviluppa del calore che mantiene lo stato denaturato dei frammenti di DNA e che, per questo, dev'essere distribuito uniformemente sulla superficie, sovrapponendo alla cella una lastra di metallo.

Il procedere dell'elettroforesi viene seguito osservando la migrazione dei due traccianti colorati, presenti nella soluzione di stop: il blu di bromofenolo (la cui mobilità corrisponde a quella di frammenti di 26 basi in un gel al 6%) e lo xilene cianolo ( che migra analogamente a catene di 106 basi).

La durata della corsa è variabile: per risolvere sequenze fino a 150 b servono 90 minuti, mentre per leggere fino a 500 b sono necessarie altre 3 ore.

Al termine della corsa, la cella viene disassemblata; il gel, che aderisce a un vetro soltanto, può essere così agevolmente fissato (in acido acetico 10% per 10 minuti) e lavato sotto acqua corrente, essiccato in forno a microonde ed esposto a una lastra autoradiografica in un'apposita cassetta per 12 ore a temperatura ambiente.

### Sequenziamento per degradazione chimica (Maxam & Gilbert)

A differenza della tecnica di terminazione della catena, che prevede un processo di sintesi enzimatica, il metodo Maxam e Gilbert si basa sulla degradazione chimica del DNA originale.

Base	Modificazione specifica
G	Metilazione dell'N <sub>7</sub> con dimetil solfato a pH 8.0 in modo tale da rendere specificatamente il legame C <sub>8</sub> -C <sub>9</sub> più suscettibile al taglio.
A + G	Trattamento con piperidina a pH 2.0, con indebolimento del legame glicosidico dell'adenina e guanina grazie a protonazione degli atomi di N degli anelli purinici, che porta a depurinazione.
C + T	Trattamento con idrazina: apre gli anelli pirimidinici, che poi riciclizzano in una forma più suscettibile al taglio.
C	Trattamento con idrazina + NaCl 1.5 M.
A > C	Trattamento con NaOH 1.2N a 90°C (colpisce le A piuttosto che le C).

In questa procedura, un frammento di DNA radiomarcato ad un'estremità viene parzialmente tagliato in cinque reazioni chimiche separate, ognuna delle quali è

specifica per una particolare base o tipo di base. Questo genera cinque popolazioni di molecole radiomarcate che si estendono da un punto comune (la terminazione "calda") al sito di taglio chimico, con lunghezza determinata dalla localizzazione di una particolare base lungo il DNA originale.

Le varie popolazioni sono quindi risolte da elettroforesi su gel di poliacrilammide, e le molecole radiomarcate vengono evidenziate tramite autoradiografia.

Il successo di questa metodica dipende interamente dalla specificità delle reazioni di taglio, che si avvalgono di due stadi:

- modificazione chimica di una specifica base, o tipo di base;
- rimozione della base modificata, e taglio dei legami fosfodiesterici 5' e 3' a livello della base modificata.

In ogni caso, queste reazioni vengono portate avanti in modo da assicurare che una sola delle basi target, in ogni molecola di DNA, venga modificata, permettendo la lettura di al massimo 250 basi dall'estremità radiomarcata.

Il metodo Maxam e Gilbert, anche se ormai sorpassato, presenta il vantaggio che la sequenza viene ricavata direttamente dalla molecola originale di DNA e non da una copia enzimatica; ciò fa sì che venga utilizzato in vari contesti quali:

- sequenziamento di oligonucleotidi sintetici;
- analisi di modificazioni del DNA, quali la metilazione;
- studi sulle strutture secondarie di DNA;
- studi di interazione proteina/DNA, grazie a processi di protezione (*footprinting*).

## 3.2 Analisi dell'RNA

Tutte le manipolazioni di RNA richiedono particolare cura per evitare contaminazioni da RNAsi. I materiali devono essere sterilizzati ed l'operatore deve portare i guanti.

### 3.2.1 Estrazione dell'RNA

Per estrarre l'RNA dai tessuti o da cellule provenienti da colture cellulari (183)

- si aggiunge a questi la soluzione D in un rapporto di 10ml di soluzione per grammo di tessuto, oppure 4ml per piastra di coltura con diametro di 10cm;
- immediatamente si frantuma mediante un Polytron per 30 sec. a velocità media, così il guanidinio tiocianato della soluzione D denatura immediatamente tutte le RNAsi che si liberano, impedendo così la denaturazione;
- successivamente si aggiungono all'omogenato 1/10 di volume di tampone acetato 2M a pH 4.5;
- poi si aggiunge 1 volume (=al volume della soluzione D iniziale) di fenolo equilibrato in acqua e 2/10 del volume di cloroformio e si agita energicamente;
- in ghiaccio 15 min.;
- si procede alla separazione della fase acquosa da quella organica mediante centrifugazione a 10000 x g per 15 min.;
- l'RNA disciolto nella fase acquosa viene recuperato;

- si precipita l'RNA selettivamente aggiungendo un ugual volume di isopropanolo e lasciando per 4 ore a -20°C;
- si raccoglie il precipitato mediante centrifugazione a 10000xg per 15 min.;
- si risospende il fondello in 1/20 del volume iniziale di soluzione D e
- si riprecipita aggiungendo 2.5 volumi (risp. al passaggio precedente) di Etanolo al 96% e lasciando a -20°C per un'ora;
- si centrifuga nuovamente per 15 min. a 10000 x g;
- si lava il fondello con Etanolo 70%;
- si risospende in 1/40 del volume iniziale in acqua autoclavata.

### 3.2.2 Selezione dell'RNA poly(A)+

RNA totale è stato preparato da cellule ottenuta dal carcinoma polmonare di Lewis (LLC) e da fegato normale di topo col metodo del guanidinio isotiocianato.(Chomczynsky) Per la preparazione di RNA poly(A)<sup>+</sup> a partire da RNA totale si è effettuata una cromatografia di affinità usando il “*PolyAtract mRNA Isolation System*” (Promega). Per l'analisi di *Northern blot* uguali quantità di poly(A)<sup>+</sup> RNA sono stati caricati in un gel all'1% di agarosio denaturante, separati mediante elettroforesi e trasferiti su membrana GeneScreen (Du Pont). I Northern blot sono stati in seguito ibridizzati con diverse sonde derivate dal cDNA murino di *Hmgi-c*. (Manfio 95) e sono elencati qui di seguito (la numerazione si riferisce al sito di inizio della trascrizione identificato (Rustighi 99):

sonda “ORF”: frammento *PstI-PstI* di 437 pb, da +157 a +593, contenenti la sequenza codificante ;

sonda “A”: frammento *AatII-PstI* di 164 pb, da -8 a +157;

sonda “B”: frammento *EcoRI-AatII* di 261, da -270 a -9

### 3.2.3 Elettroforesi di RNA su gel d'agarosio

La separazione elettroforetica dell'RNA viene effettuata in gel di agarosio in condizioni denaturanti. Si prepara un gel all'1% di agarosio sciogliendolo prima in acqua in un volume minimo che permetta poi di andare al volume prescelto aggiungendo 1/10 del tampone MOPS 10X e formaldeide 6.2% finali, una volta che la temperatura della soluzione dell'agarosio nell'acqua sia scesa sotto i 60°C. La formaldeide serve a mantenere gli RNA denaturati durante la separazione elettroforetica.

La soluzione viene quindi versata in una cella orizzontale e lasciata gelificare. Prima dell'uso cella e pettine vengono lavati con acqua ossigenata e poi sciacquati con acqua. I campioni vengono diluiti tre volte nel tampone di caricamento, denaturati a 68°C per 5 min. e quindi caricati sul gel. Il tampone di caricamento contiene formaldeide e formammide per mantenere denaturati gli RNA. Anche l'etidio bromuro è contenuto nel tampone di caricamento, ma non nella matrice del gel, in quanto si intercala con minor efficienza nell'RNA rispetto al DNA.

La separazione viene condotta a 60 V costanti.

### 3.2.4 Trasferimento su filtro di RNA mediante *Northern blotting*

Una volta separato mediante elettroforesi, l'RNA viene trasferito da gel di agarosio su una membrana di nylon, in modo da venir analizzato attraverso l'ibridazione con sonde radioattive. La procedura è la seguente:

si incuba il gel con NaOH 50mM per 15 min. in agitazione per denaturare l'RNA (facoltativo);

si neutralizza il gel in tampone Tris/HCl 1M pH 7.5 per 15min.;

si allestisce il sistema di trasferimento adagiando il gel su alcuni fogli di carta da filtro (3MM Whatman) poggiati su di una spugna inbevuta di una soluzione di trasferimento data dal tampone fosfato  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$  25 mM pH 6.5. I quattro bordi del gel vengono ricoperti da altrettante strisce di parafilm. Sul gel viene posta la membrana e sopra quest'ultima vengono aggiunti alcuni fogli di carta da filtro ed un blocchetto di carta assorbente dello spessore di ca. 10 cm. Il tutto viene mantenuto sotto una pressione costante, dovuta ad un peso pari a 0.5 Kg.

Per capillarità il tampone passa dalla spugna attraverso il gel portandosi dietro l'RNA, determinando così il trasferimento dello stesso dal gel al filtro, al quale si lega. Il trasferimento richiede ca. 18 h.

Dopo questo tempo il filtro viene recuperato, lavato per 5 min. in tampone fosfato per rimuovere eventuali tracce di gel, quindi asciugato ad una temperatura di 80°C per 2 h. Infine si espone il filtro per 1 min. alla radiazione UV per rendere il legame delle molecole di RNA con la membrana irreversibile.

### 3.2.5 Caratterizzazione dell'estremità 5' del filamento di RNA

#### Allungamento dell'innesco ("*Primer elongation*")

Per identificare sequenze all'estremità sia 3' che 5' dell'RNA, che corrispondono rispettivamente all'inizio e la fine di un gene, si sfrutta la capacità della trascrittasi inversa che copia l'RNA a partire da un innesco radiomarcato finché c'è RNA che funge da substrato. Quando l'enzima ha retrotrascritto l'ultimo nucleotide all'estremità 5' allora cessa la polimerizzazione e si stacca. Il prodotto di questa reazione sarà lungo quanto la distanza sull'RNA tra il nucleotide complementare al primo nucleotide dell'oligo antisense e il primo nucleotide trascritto.

Per questa analisi 1 ng dell'oligonucleotide AR16, marcato con  $^{32}\text{P}$  ( $2.5 \times 10^5$  cpm) mediante la PNK, è stato ibridizzato per 1 ora a 50°C con 10  $\mu\text{g}$  di RNA totale di LLC o di fegato. Per l'allungamento di AR16 si sono utilizzate 200 U di trascrittasi inversa del virus della leucemia murina di Moloney (RTase M-MuLV, Gibco-BRL) per 90 min. a 42 °C. Il prodotto sintetizzato è stato analizzato per la sua lunghezza in un gel di poliacrilammide denaturante al 7% facendolo migrare assieme alla sequenza del DNA genomico effettuata con lo stesso oligo AR16 come pesi molecolari. Il gel è stato trasferito su carta da filtro (3MM, Whatman) essiccato sotto vuoto ed esposto ad una lastra autoradiografica (Kodak XAR-5) tra due schermi intensificatori per circa 3 giorni a -80°C.

### **RACE al 5' (Rapid Amplification of cDNA ends)(184)**

Per identificare sequenze all'estremità sia 3' che 5' dell'RNA, che corrispondono rispettivamente all'inizio e la fine di un gene, si sfrutta la capacità della trascrittasi inversa che copia l'RNA a partire da un innesco. In questo caso, un oligonucleotide antisenso AR3 (5'-CCGCTACCACTGCCTCTA -3', da +187 a +170)(Rustighi 99) è stato usato per retrotrascrivere 1 µg di RNA poly(A)+ di LLC con 200 U di trascrittasi inversa del virus della leucemia murina di Moloney (RTase M-MuLV, Gibco-BRL). Al cDNA ottenuto è stata aggiunta una coda di polyG attraverso un trattamento con la Terminal deossinucleotidil transferasi per poterlo poi amplificare utilizzando due oligonucleotidi, di cui quello al 5' (AR15, 5'-CTCGGCCTCCGCCGCCCGAAT-3', da +148 a +127) corrisponde ad una sequenza di *Hmgi-c* più monte rispetto AR3, mentre l'oligo antisenso 15CEco (5'-GATGGAATCCCCCCCCCCCCCCC-3'), complementare alla coda polyG e in più possiede il sito di riconoscimento per l'enzima di restrizione *EcoRI*, utile per un successivo subclonaggio. Le condizioni di amplificazione sono state ottimizzate ad 1 min. a 94°C, 1' a 55°C per 1 min. 30 sec. a 72°C per 40 cicli. I prodotti di PCR sono stati analizzati su un gel di agarosio all'1.5% e mediante Southern blot usando come sonda un altro oligonucleotide interno AR16 (5'-TTCGGGAGACGGGATGTA-3', da +113 a +96). Una banda AR16-positiva di circa 160 pb è stata identificata e subclonata nel vettore di clonaggio plasmidico pBluescriptKS+ (Stratagene). I cloni risultanti sono stati sequenziati col metodo Sanger.

## **3.3 Analisi delle proteine**

### **3.3.1 Espressione di proteine ricombinanti**

Il sistema d'espressione usato è di tipo procariotico, ed è realizzato in cellule BL21(DE3) di *E. coli*, in quanto si tratta di un ceppo che

- ha una basso tasso di proteasi endogene;
- porta integrato nel proprio genoma il gene per la T7 RNA polimerasi sotto il controllo inducibile del promotore *lac UV5*.

Una colonia da piatto di cellule BL21(DE3), trasformate con il plasmide d'interesse, viene inoculata in 10 ml di LB-broth con ampicillina (50mg/ml) e fatta crescere per 12 ore a 37°C in agitazione. 3 ml della sospensione così ottenuta vengono versati in una beuta contenente 250 ml di terreno fresco e i batteri lasciati crescere a 37 °C, in agitazione, per circa 3 ore fino a che la densità ottica della sospensione, misurata a 600 nm, non è compresa tra 0.6 e 0.8 (unità ottiche di assorbanza).

L'espressione viene indotta mediante aggiunta di IPTG (isopriltio-β-D-galattoside) 0.5 mM finale. L'IPTG è un analogo dei β-galattosidi in grado di indurre l'attivazione dei promotori *lac*. In assenza di β-galattosidi al promotore è legato un repressore che impedisce la trascrizione dei geni posti a valle; quando nel mezzo di coltura viene aggiunto l'IPTG esso si lega al repressore modificandolo conformazionalmente e provocando il suo distacco dal promotore. In questo caso è stato usato un ceppo batterico particolare, in cui l'aggiunta di IPTG induce la produzione della T7 RNA polimerasi (controllata da un promotore *lac*) la quale, a sua volta, riconosce il promotore T7 sul vettore d'espressione e permette così la trascrizione dell'inserito clonato a valle.



Dopo l'induzione i batteri sono lasciati in agitazione a 37°C per 12 ore e quindi centrifugati 30 minuti a 2500g in una centrifuga termostata a 4°C. si elimina il surnatante e il pellet viene lavato con 25 ml di PBS e congelato a -80°C fino al momento dell'estrazione.

### 3.3.2 Estrazione di proteine con acido perclorico

L'intera operazione viene realizzata in ghiaccio in modo da ridurre i rischi di degradazione delle proteine.

I pellet batterici, precedentemente congelati in modo da favorire la lisi della cellula batterica, vengono risospesi in una soluzione acquosa di acido perclorico al 5% (v/v) in rapporto 2 volumi acido perclorico : 1 volume di batteri; alla sospensione vengono aggiunti anche 10 µl di tre inibitori delle proteasi, quali: Leupeptina (soluzione acquosa 0.5 mg/ml), Aprotinina (soluzione acquosa 1 mg/ml) e PMSF (fenil-metil-sulfonil-fluoruro, soluzione satura in isopropanolo) (*Sigma*).

I batteri vengono frullati con un omogenizzatore per 30 secondi e pestellati con una bacchetta di vetro; l'operazione viene ripetuta per tre volte ed è seguita da una centrifugazione della sospensione a 2500g per 25 minuti a 4°C. si recupera il surnatante (anche se non ancora limpido) e sul fondello si ripete una seconda estrazione con 1 volume di acido perclorico al 5% (v/v) addizionata di inibitori. Si ripete la centrifugazione a 2500g per 15 minuti e si aggiunge il surnatante così ottenuto al precedente, il tutto viene infine ricentrifugato un'ultima volta per 30 minuti per eliminare eventuali particelle in sospensione ancora presenti. A questo punto al surnatante vengono aggiunti 6 volumi di acetone, 50 µl di HCl al 37% e il tutto è posto a -20°C per 12 ore. In questo modo si ottiene la precipitazione selettiva delle proteine HMG; il precipitato viene recuperato mediante centrifugazione a 2000g per 20 minuti a 4°C, e risospeso, dopo aver eliminato l'acetone e asciugato brevemente il *pellet*, in un opportuno volume di H<sub>2</sub>O mQ. I campioni sono conservati a -20°C fino al momento della purificazione mediante cromatografia in HPLC

### 3.3.3 Analisi elettroforetiche

#### Elettroforesi in acido acetico/urea

Il gel ha dimensioni di 14 cm in larghezza, 24 cm in lunghezza e 0,075 cm in spessore. È costituito da due parti: un gel di caricamento (*stacking gel*), che serve a concentrare i campioni, ed un gel di separazione (*running gel*).

Prima di caricare i campioni, si effettua una precorsa con una soluzione di cisteammina (cisteammina 1M, 15% di saccarosio, in solvente running) a 400V per circa 4 ore, fino a che la corrente scende da 60-70 mA iniziali a 5-7 mA, ad una differenza di potenziale di 200V. La precorsa ha lo scopo di ridurre tramite la cisteammina eventuali ossidanti presenti nel gel. Di tale soluzione si caricano 30 µl in ciascun pozzetto del gel.

Terminata la precorsa, si possono caricare i campioni. Se sono già stati risospesi nel solvente campioni (1x) sono già pronti. Se invece sono stati risospesi in H<sub>2</sub>O, verranno sciolti nel solvente campioni 2x, in rapporto volumetrico di 1:1.

La corsa elettroforetica consiste di una fase di accumulo, effettuata per non meno di un'ora a 50V, e di una di separazione a 200V per 16 ore. Sia la precorsa che la corsa sono di tipo catodico.

## Elettroforesi in SDS/Tris-Tricina con sistema discontinuo

Il sistema è detto discontinuo perché è costituito da due parti di diversa composizione, sia per la concentrazione di acrilamide, che per il tampone.

Le dimensioni del gel sono le stesse viste per l'acido acetico/urea. Anche in questo caso, sono presenti uno stacking gel ed un running gel. Quest'ultimo può avere diverse percentuali (p/v) di acrilamide.

La corsa elettroforetica prevede una fase di accumulazione della durata di 30-60 minuti, a 50V, ed una di separazione a 200V, o 150V per i gel al 15% ed al 7,5% di acrilamide, per 17 ore. La corsa è di tipo anodico.

I campioni da caricare devono essere sciolti nell'opportuno solvente.

## Colorazione dei gel

Al termine delle elettroforesi, i gel vengono colorati con una soluzione di Blue Coomassie allo 0,05% p/v in metanolo, acqua ed acido acetico, in rapporto 5:4:1. La presenza del metanolo permette di fissare le proteine al gel. La decolorazione viene effettuata dopo un giorno con una soluzione composta da tre parti di H<sub>2</sub>O ed una di metanolo.

### 3.3.4 Purificazione proteica mediante HPLC

Viene usato un sistema di separazione cromatografica ad alta pressione: HPLC (High Performance Liquid Chromatography). Il sistema comprende sei componenti:

- due pompe *Waters* modello 510, che permettono di ottenere pressioni molto elevate (360 atm), grazie alle quali è possibile avere una buona risoluzione dei picchi cromatografici;
- *Waters automated gradient controller* modello 640, che agisce sulle due pompe in modo da creare una miscela dei solventi di eluizione nelle proporzioni opportune;
- *Waters injector* modello U6K, che serve per iniettare il campione in colonna senza causare alcune variazioni della pressione e/o del flusso;
- colonna cromatografica a fase inversa *Biorad RP 304*, del tipo C4, di dimensioni 250 mm × 4.6 mm;
- *Waters absorbance detector* modello 441, che rileva la formazione dei picchi ad una data lunghezza d'onda (nello specifico a 229 nm);
- *Waters data module* modello 740, che registra i picchi precedentemente rilevati.

La capacità della colonna è di 1-3 mg di proteina e può lavorare in un intervallo di pH tra 2 e 7; la matrice è costituita da particelle di silice (di 5 µm di diametro formanti pori di 3 µm) alle quali sono legati dei pendenti idrofobici costituiti da catene idrocarburiche a 4 atomi di carbonio. In questo modo si ha una separazione in base all'idrofobicità delle molecole e quindi alla loro affinità per i gruppi funzionali legati alla matrice. Molecole contenenti gruppi polari o debolmente idrofobici interagiscono poco con la fase stazionaria della colonna e tendono ad eluire velocemente; al contrario molecole molto

idrofobiche si legano fortemente con la fase stazionaria e vengono quindi eluite a tempi di ritenzione ( $R_t$ ) più elevati.

Il solvente di eluizione è costituito da due soluzioni che possono essere mescolate in rapporti diversi nel tempo in modo da creare un gradiente di eluizione adattabile ai vari campioni da purificare. Questo permette di cambiare la forza dell'eluente che entra in colonna con il procedere della cromatografia e di modulare il distacco delle proteine legate alla matrice, al fine di migliorare la risoluzione del sistema. L'eluente deve essere estremamente puro e i solventi usati vanno degassati per evitare che eventuali bolle possano portare ad una variazione del flusso.

I due solventi impiegati sono:

- solvente A, costituito da una soluzione di acido trifluoroacetico allo 0.1% in  $H_2O$ ;
- solvente B, che contiene acetonitrile per il 95%,  $H_2O$  per il 5% e acido trifluoroacetico per lo 0.1%.

La colonna viene mantenuta in una soluzione di acetonitrile, equilibrandola poi, prima di procedere con la cromatografia, con le stesse percentuali dei solventi che si hanno all'inizio del gradiente, facendo passare 10 volumi di questi ultimi in colonna fino a che l'eluuto non dà un valore di assorbanza costante.

Per la purificazione della proteina HMGI-C 109 e relative forme mutate e chimeriche, è stato usato un gradiente non lineare che passa dal 92% di solvente A e 8% di solvente B al 100% di solvente B.

I campioni, conservati a  $-20^\circ C$  in soluzione acquosa vengono iniettati in colonna e le frazioni raccolte sono raccolte, evaporate e conservate nuovamente a  $-20^\circ C$ .

### 3.3.5 Analisi delle proteine mediante spettrometria di massa

Lo spettrometro di massa è uno strumento in grado di determinare il peso molecolare di una proteina (o di composti di altro genere) con un errore percentuale dello 0.01-0.02. lo strumento è costituito essenzialmente da tre componenti:

- la sorgente di ioni;
- l'analizzatore di massa;
- il detector.

Il campione viene ionizzato e l'analizzatore di massa separa gli ioni in base al rapporto massa/carica ( $m/z$ ); ogni molecola ionizzata viene poi rilevata dal detector dando un segnale proporzionale alla propria quantità relativa.

Lo spettro di massa della molecola si ottiene riportando le quantità relative di ciascuno ione in funzione del rapporto  $m/z$ .

La grande innovazione degli odierni strumenti è stata l'utilizzo della tecnica "ion spray" che permette di far passare il campione in fase gassosa senza ricorrere alla vaporizzazione e quindi ad un aumento di temperatura che sarebbe deleterio per campioni termolabili quali le proteine. Con questo sistema gli ioni preformati in soluzione vengono emessi in fase gassosa: il liquido contenente il campione da analizzare è spinto attraverso un vaporizzatore mantenuto ad alto voltaggio in modo che si formino delle goccioline cariche. Quando queste evaporano gli ioni passano alla fase gassosa attraverso un processo a bassa energia che non altera il campione. In tal modo i composti contenenti una o più cariche sono analizzati quasi nella stessa situazione in cui si trovano nella soluzione originale.

Lo strumento usato è un Perkin Elmer Sciex API I.

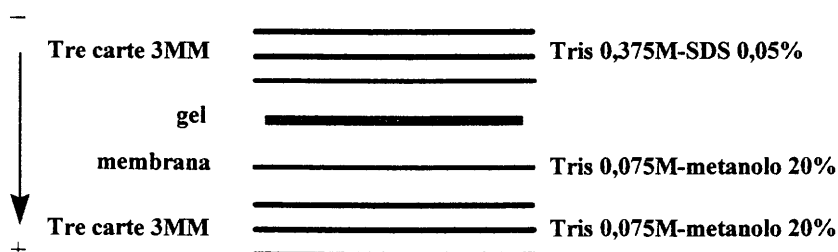
I campioni proteici purificati assolutamente privi di detergenti sono risospesi in una soluzione H<sub>2</sub>O/acetonitrile in rapporto 1:1 (v/v) addizionata di acido formico allo 0.1%. A questo punto la miscela viene iniettata nello spettrometro con una velocità di flusso di 2 µl/minuto. Lo spettro è stato analizzato in un intervallo da 600 m/z a 1800 m/z. La scala di calibrazione di massa è stata messa a punto con il polipropilene glicole e saggiata mediante misura di massa della mioglobina (16950 Da).

### 3.3.6 Trasferimento delle proteine su membrana (*Western blotting*)

Le proteine separate su gel in SDS, vengono trasferite su membrana di fluoruro di polivinilidene (PVDF) con pori di dimensioni di 0,2µm (Immobilon<sup>TM</sup>-P<sup>SQ</sup>, Millipore), usando il metodo del tipo "semidry". A tale scopo utilizza il Mini Trans-Blot Cell (Bio-Rad).

#### Trasferimento su membrana di PVDF

- Al termine della corsa elettroforetica, si taglia il gel delle dimensioni opportune, in base alla parte del tracciato che si desidera trasferire su membrana;
- di norma si utilizzano due membrane, per evitare di perdere le proteine che eventualmente riescano ad oltrepassare la prima. Il PVDF è un materiale idrofobico: prima di idratarlo deve essere immerso in metanolo, quindi in H<sub>2</sub>O, e poi nella soluzione di trasferimento 1.
- Le dimensioni delle membrane sono di 1-1,5 cm superiori a quelle del gel; si dispongono le membrane a contatto con il gel, e su queste tre fogli di carta Watmann 3MM, con dimensioni uguali alle membrane ed imbevute nella stessa soluzione;
- si mettono tre fogli di carta 3MM anche sull'altro lato del gel, imbevute nella soluzione di trasferimento 2;
- il "sandwich" così costituito è ora pronto per essere posto all'interno del Trans-Blot, in modo tale che la differenza di potenziale applicata faccia trasferire le proteine dal gel (posto verso il catodo) alla membrana (posta verso l'anodo);
- il trasferimento dura per 45-60 minuti, ad un voltaggio tale che la corrente non superi i 300mA. Ogni 15 minuti il Trans-Blot viene aperto e si passa una bacchetta sopra il "sandwich" per eliminare eventuali bolle che potrebbero interferire con il passaggio della corrente.



- Al termine del trasferimento, si colora il gel con la soluzione di Blue Coomassie e le membrane con una soluzione di rosso Ponceau S, per 10 minuti (per verificare che le proteine si siano trasferite);
- le membrane vengono decolorate e poste in una soluzione di saturazione per 1-2 ore, sotto agitazione.

Per avere informazioni circa la sequenza di una particolare proteina (si consulti a questo proposito il capitolo “RISULTATI E DISCUSSIONE”) la membrana è stata colorata con la soluzione di Blue Coomassie utilizzata per la colorazione dei gel. Si è proceduto come segue:

- si colora la membrana per 1 minuto;
- si decolora per qualche minuto con la stessa soluzione, priva del colorante;
- si effettuano due lavaggi di 15-20 minuti con una soluzione di H<sub>2</sub>O e etanolo in rapporto di 3:1;
- si lascia asciugare la membrana.

### 3.3.6 Riconoscimento con anticorpi

Sulle membrane, dopo il trasferimento e la successiva saturazione, sono stati effettuati vari riconoscimenti con anticorpi, per rilevare la presenza o meno di specifiche proteine. Il procedimento consiste in:

- un fase della durata di 1 ora, nella quale la membrana viene incubata con l'anticorpo primario (cioè quello diretto contro una specifica proteina), diluito opportunamente nella soluzione di saturazione;
- una serie di tre lavaggi di 5 minuti ciascuno con la soluzione di saturazione, sotto agitazione;
- un'incubazione di 1 ora con l'anticorpo secondario (diretto contro il primario) coniugato con la perossidasi di rafano, in concentrazione opportuna nella soluzione di saturazione;
- tre lavaggi di 5 minuti ciascuno con la soluzione di saturazione, sotto agitazione;
- un lavaggio di 20 minuti con PBS freddo;
- un trattamento della membrana per 5 minuti con il substrato della perossidasi (ECL+Plus/Amersham Pharmacia Biotech);
- eliminazione dell'eccesso di ECL+Plus; sulla membrana si pone quindi una lastra fotografica (Hyperfilm<sup>TM</sup> ECL), che verrà impressionata in corrispondenza del legame degli anticorpi; i tempi di esposizione variano tra i 30 secondi e i 10 minuti, a seconda del caso.

La membrana sulla quale è già stato effettuato un riconoscimento, può essere successivamente sottoposta ad un riconoscimento con un anticorpo diverso. Per evitare interferenze del primo, si incuba la membrana per un'ora con una soluzione, detta *stripping buffer*, a 37°C, in agitazione. Questa permette di staccare gli anticorpi (primario e secondario) precedentemente legatisi. Si eseguono quindi due lavaggi di 20 minuti con PBS, e al termine si può procedere con un nuovo riconoscimento.

Gli anticorpi utilizzati sono:

- **anticorpi primari:**
  - anticorpo policlonale ottenuto in coniglio contro un peptide sintetico la cui sequenza è contenuta nella parte N-terminale di entrambe le proteine HMGI ed HMGY, diluizione 1:1000;
  - anticorpo policlonale ottenuto in coniglio contro l'intera proteina ricombinante HMGI-C, diluizione 1:1000;
- **anticorpi secondari:**

anticorpo ottenuto in capra diretto contro le IgG (H+L) di coniglio, coniugato con la perossidasi di rafano (PIERCE), diluizione 1:5000;

### 3.3.6 Interazioni proteina-DNA

#### Analisi di ritardo di banda (“*Bandshift* o *EMSA*”)

Uno fra i metodi più diffusi per studiare in vitro l'interazione specifica tra regioni regolative di DNA e fattori coinvolti nella regolazione della trascrizione è il band-shift o EMSA (Electrophoretic Mobility Shift Assay). Questa tecnica consente non solo di verificare l'effettiva associazione fra sequenze di DNA con significato regolativo e proteine purificate note, ma anche di mettere in luce, in estratti di particolari linee cellulari, la presenza di complessi formati da uno o più fattori, in grado di legare selettivamente le regioni di DNA analizzate.

La base del cosiddetto "ritardo di banda" è la riduzione della mobilità elettroforetica, in un gel non denaturante, delle molecole di DNA, quando sono associate a una o più proteine. Infatti la migrazione di tali complessi specifici DNA-proteina è determinata principalmente dalle dimensioni e dalla conformazione del fattore proteico, anche se in alcuni casi viene condizionata anche dal grado di curvatura del DNA.

Il frammento di DNA in esame viene dapprima marcato radioattivamente e in seguito incubato con le componenti proteiche; la presenza di eventuali complessi viene successivamente evidenziata, previa separazione elettroforetica, mediante autoradiografia.

#### Preparazione di estratti nucleari proteici

Spesso negli esperimenti di *ritardo di banda* si usano estratti proteici, totali o nucleari, derivanti da cellule in coltura, anziché proteine pure. Dato che nella nostra ricerca si studiano fenomeni che avvengono principalmente nel nucleo abbiamo fatto uso di estratti nucleari.

Non avendo cellule sincronizzate gli estratti nucleari vengono preparati dalle linee cellulari in coltura qualora si trovino in uno stadio di confluenza.

Tutta la procedura di estrazione viene effettuata alla temperatura di 4°C.

Le cellule vengono lavate due volte con PBS in modo da eliminare qualsiasi traccia di terreno di coltura, poi con l'aiuto di una spatolina vengono raccolte in provetta e centrifugate a 3500 RPM per 5 minuti. Successivamente il fondello di cellule formatosi viene risospeso in 50 µl della soluzione I e passato 10 volte attraverso una siringa da 1 ml (25 Gauge) per permettere la rottura meccanica delle cellule. La sospensione viene centrifugata a 2000 RPM per 10 minuti, ed il fondello costituito da nuclei viene lavato ulteriormente con l'aggiunta di 50µl di soluzione I. Dopo una centrifugazione a 2000 PRM per 10 minuti vengono aggiunti 75 µl della soluzione II per permettere la risospensione dei nuclei attraverso delicate rotazioni per circa 30 minuti, durante i quali le proteine nucleari vengono rilasciate da questo compartimento cellulare. A questo punto dopo una centrifugazione a 20.000 RPM (rotore A8.24 *Kontron*) per 30 minuti viene raccolto il surnatante contenente l'estratto proteico nucleare, aliquotato e rapidamente congelato a -80°C fino al suo utilizzo.

## Quantizzazione dell'estratto proteico (saggio Bradford)

Per poter usare delle quantità comparabili di estratto nucleare è necessaria una fine quantizzazione. Il metodo di misurazione utilizzato è il cosiddetto saggio Bradford (*Biorad*) e viene effettuata su microscala. Questa analisi spettrofotometrica impiega il colorante Blu di Coomassi come reagente in grado di legare, secondo una certa stechiometria, le proteine. Grazie al legame con le proteine il colorante cambia il suo massimo di assorbimento da 465nm a 595nm. L'assorbimento a questa lunghezza d'onda è direttamente proporzionale alla concentrazione delle proteine.

Inizialmente viene allestita una curva di taratura impiegando 5 standard a concentrazione nota, dalla quale si ricava successivamente il valore di concentrazione dei vari estratti nucleari.

## Marcatura radioattiva della sonda

Viene sfruttata l'attività enzimatica della polinucleotidechinasi (PNK) che consente di marcare l'estremità 5' di oligonucleotidi sintetici utilizzando  $\gamma^{32}\text{P}$ -ATP.

In un volume finale di 50 ml si fa reagire per 30 minuti a 37°C la seguente miscela:

- 10 pmol di oligonucleotide
- 5 ml di tampone specifico 10x
- 6 unità di PNK
- 4  $\mu\text{l}$  di  $\gamma^{32}\text{P}$ -ATP (2,5 mCi/ml)

In seguito per denaturare l'enzima si incuba 2 minuti a 95°C.

Quindi, dopo aver reso la miscela NaCl 300 mM, si aggiungono 20 pmol dell'oligonucleotide complementare e si incuba a 100°C per 10 minuti; infine si lascia raffreddare gradualmente fino a raggiungere temperatura ambiente. In questo modo gli oligonucleotidi complementari possono appaiarsi correttamente e, considerato l'eccesso molare del complementare freddo, la maggior parte delle molecole marcate costituisce un duplex.

A questo punto la sonda deve venire separata dall'ATP radioattivo che non ha reagito, sfruttando la cromatografia per gel permeazione. Si utilizzano in questo caso le cosiddette "spin columns", il cui vantaggio principale è la rapidità: infatti, sfruttando la forza centrifuga, è possibile accelerare considerevolmente il flusso attraverso le colonne, riducendo notevolmente la durata del processo di purificazione.

La resina utilizzata è Sephadex G-50 (*Pharmacia*), equilibrata con tampone TEN e impaccata in una siringa da 1 ml con il fondo bloccato con della lana di vetro. L'equilibratura della colonna viene effettuata tramite successive aggiunte di tampone TEN, seguite da centrifugazione di 2 minuti a 13000xG, fino a che il volume dell'eluato non è pari a quello caricato in colonna.

A questo punto si porta la miscela di reazione a un volume finale di 100 ml con l'aggiunta di tampone TEN e si applica la soluzione ottenuta sulla sommità della colonna; l'eluizione della sonda avviene mediante centrifugazione a 13000xG per 2 minuti.

Infine si misura l'attività specifica della sonda con un contatore a scintillazione liquida (*Betamatic V, Kontron*), il valore ottenuto si colloca in media in un intervallo tra 20.000 e 40.000 cpm/ $\mu\text{l}$ .

Sono qui di seguito elencati oligonucleotidi usati negli esperimenti di bandshift (solo il filamento senso è mostrato, in grassetto sono indicate le mutazioni):

- A 5'-CAGCGCCTCCTCCTTTCCTCCTCCTC-3'  
 B 5'-TCCTCCTCTCCTCCTCCTCTCT-3'  
 C 5'-TCTCTTCCTCCTCCCCTCTCTTTTT-3'  
 D 5'-TCTTTTTGGCAGCCGCTGGACGTCCGGTGTTGA-3'  
 mD 5'-TCTTTTTTGCAGCCGCTGGACGTCCGGTGTTGA-3'  
 E 5'-GACGTCCGGTGTTGATGGTGGCCAG-3'  
 F 5'-GCAGCGGCCGCGCAGCCTAAGCAGCA-3'  
 G 5'-AGCAGCAGTAGTCCCCGGCGCACT-3'  
 H 5'-CACTCGCCAGCTCGCCTGCTCTCG  
 Sp1 5'-ATTCGATCGGGGCGGGGCGAGC-3'  
 AP1 5'-CGCTTGATGAGTCAGCCGGAA-3'  
 NF1 5'CTAGCTATTTTGGATTGAAGCCAATATG-3'

mD-H è stato generato con la PCR usando l'oligonucleotide mD senso e H antisenso su di uno stampo derivato dal costruito HaP (sezione 3.1.3)

### La reazione

Idealmente sarebbe auspicabile poter riprodurre le condizioni del nucleo cellulare nello studio dell'interazione di un fattore proteico con il DNA; siccome ciò è irrealizzabile è stato ottimizzato un tampone di legame (BB 10X) in grado di limitare le differenze con l'ambiente di reazione in vivo.

Quindi per ogni reazione si allestisce la seguente miscela:

- 2 µl di BB 10X
- 1 µl di BSA (1 mg/ml), che funge da competitore per i legami aspecifici con la sonda
- 0,16 µl di ditiotreitolo (1 M), riducente che limita la formazione di aggregati proteici
- 0,25 µl di blu di bromofenolo (sciolto in acqua con poco tampone TBE)
- 3 µl di glicerolo 87%

e acqua a 20 µl totali, tenuto conto dei volumi di proteina e DNA da aggiungere. Per i band-shift con Sp1 si aggiunge lo ione  $Zn^{++}$  20 nM e per quelli con CTF/NF-1 bisogna raggiungere una concentrazione di NaCl di circa 250 mM finale nella miscela intera.

Per poter evidenziare la specificità del legame, si può includere nella miscela di reazione un eccesso molare di DNA competitore aspecifico, come ad esempio un oligonucleotide costituito da poly (dG•dC) (5 ng/µl) per gli esperimenti con le proteine HMGI oppure poly (dI•dC) per tutte le altre proteine.

Si stabilisce quanta sonda far reagire a seconda dell'attività specifica misurata: se è elevata, si cerca di limitarne la quantità in modo da ottenere bande nette. In ogni caso bisogna tener conto che vi è una soglia limite nella concentrazione del duplex, al di sotto della quale la formazione del complesso non è più favorita.

Le reazioni vengono allestite in una piastra da microtiter, unendo sequenzialmente:

- miscela di reazione;
- DNA competitore (se previsto);
- soluzione proteica; da qui in poi è sconsigliato miscelare "spipettando";
- DNA sonda

Per ogni serie di reazioni, se ne include una contenente solo il DNA sonda e la miscela (portata a 20 µl finali con la soluzione di blu di bromofenolo): in questo modo si ricava un'indicazione della mobilità del duplex libero.



### Esperimenti di supershift

Tutto c.s. tranne che per l'aggiunta di anticorpo specifico 20 minuti prima di aggiungere il DNA nella miscela di incubazione. Anticorpi specifici  $\alpha$ -Sp1 ed  $\alpha$ -Sp3, come anche  $\alpha$ -Oct2 (Santacruz) sono stati aggiunti ad una concentrazione di 5-10 ng/ $\mu$ l.

### Separazione elettroforetica

La formazione dei complessi proteina-DNA viene evidenziata attraverso migrazione elettroforetica in un gel di poliacrilammide in condizioni non denaturanti: si è fissata la percentuale del gel al 7% in considerazione della lunghezza della sonda (24-35 basi) e delle dimensioni delle proteine (in media 10 KDa).

Fra le condizioni da stabilire, ha particolare importanza il tampone di corsa, la cui forza ionica non dev'essere troppo elevata, sia perchè altrimenti il complesso DNA-proteina conduce una frazione troppo piccola della corrente (e di conseguenza non migra a sufficienza), sia perchè un'alta concentrazione di sali può anche incrementare il calore che si sviluppa per effetto Joule durante l'elettroforesi. Inoltre non bisogna trascurare, visto che si utilizza tampone TBE, il fatto che lo ione borato può interagire con le proteine provocando artefatti. D'altra parte se la concentrazione è troppo bassa, si può verificare una sensibile variazione del pH (a causa dell'elettrolisi dell'acqua che si verifica agli elettrodi), il che influisce pesantemente sulla stabilità del complesso.

L'elettroforesi è condotta in celle verticali, costituite da lastre di vetro (22x32,5 cm) che vengono assemblate utilizzando spaziatori laterali di spessore pari a 1,5 mm.

120 ml della soluzione di acrilammide al 7%, attivata con i catalizzatori (APS e TEMED), vengono rapidamente versati tra i due vetri e viene inserito un pettine in modo da formare i pozzetti.

Perchè si eviti la polimerizzazione è necessaria almeno un'ora; dopo di che il gel viene montato sull'apposito supporto verticale, le cui vaschette anodica e catodica sono riempite con tampone TBE 0,5X preraffreddato. Il tutto viene mantenuto a 4°C: è essenziale, infatti, che la corsa avvenga in tali condizioni, in modo da disperdere la maggiore quantità possibile di calore (generato per effetto Joule), che ha conseguenze deleterie sulla stabilità del complesso.

Caricate le reazioni dopo una precorsa di circa 15 minuti a 300 V, la corsa viene effettuata allo stesso voltaggio costante; la separazione viene seguita per mezzo del tracciante utilizzato, il blu di bromofenolo, che viene fatto migrare a una distanza prefissata in relazione alle dimensioni della sonda oligonucleotidica.

Al termine della corsa, la cella viene disassemblata e il gel viene fissato in acido acetico 10%, lavato con acqua distillata ed essiccato su carta 3MM; infine viene esposto a una lastra autoradiografica in un'apposita cassetta dotata di schermi intensificatori. La durata dell'esposizione dipende dall'attività specifica della sonda e può variare da 30 minuti a temperatura ambiente a 12 ore a -80°C.

### Analisi di protezione da DNasi I

Il saggio di protezione da DNasi I (o *footprinting*) fornisce dirette ed accurate informazioni sulle sequenze di DNA legate da specifiche proteine. Inizialmente nata come tecnica meramente qualitativa, volta a localizzare siti di legame proteici sul DNA, si è poi evoluta in modo tale da fornire informazioni anche sulla forza e sull'affinità del

legame proteico; comunque non risulta essere sufficiente per determinare precisamente la natura dei siti d'interazione o il contributo dei singoli nucleotidi.

La metodica si basa su un processo di protezione dello scheletro fosfodiesterico del DNA dall'idrolisi catalizzata dalla DNasi I, attuata da fattori proteici derivanti sia da preparazioni pure di proteine, che da estratti cellulari.

In pratica, il DNA previamente radiomarcato ad un'estremità, viene incubato con il/i fattore/i proteico/i in esame; i risultanti complessi vengono incubati con DNasi I sotto condizioni che ne favoriscono la parziale digestione. Questo assicura la produzione di frammenti radiomarcati a livello delle zone non protette, al contrario delle regioni in cui si lega un elemento proteico, che rimangono non digerite.

Le reazioni vengono quindi separate su gel di poliacrilammide ad alta risoluzione (devono essere separabili frammenti che differiscono per una base) e confrontate con una digestione approntata in assenza di qualunque proteina legante il DNA.

Dopo autoradiografia del gel, le "tracce" (*footprint*) lasciate dalle proteine sul DNA si evidenziano come zone prive di frammenti marcati, grazie alla protezione proteica.

### **Marcatura e purificazione del frammento di DNA**

Si digerisce una quantità di plasmide variabile tra le 5 e le 15 pmoli con appropriati enzimi di restrizione; effettuata una defosfatizzazione con la fosfatasi alcalina (vedi sez. 2.1.7.) si procede alla marcatura con la T<sub>4</sub> polinucleotide chinasi (PNK). Una reazione tipo prevede:

- 5 pmoli di estremità;
- 1 µl di PNK (5U/µl, Epicentre Technologies);
- 2 µl di tampone specifico 10×;
- 5 µl di  $\gamma^{32}\text{P}$ -ATP;

H<sub>2</sub>O a 20 µl finali.

La purificazione del plasmide radiomarcato dal  $\gamma^{32}\text{P}$ -ATP libero viene effettuata con colonne cromatografiche ad esclusione preimpaccate della BIORAD (*Micro Bio-Spin 30*, con limite di esclusione di 20 pb), calcolando poi la percentuale di incorporazione.

Il passaggio successivo prevede una seconda restrizione con endonucleasi specifiche, che permette di ottenere un frammento di DNA da usare nel *footprinting*, che ha un'unica estremità radiomarcata su un solo filamento (è ovvio che per analizzare entrambi i filamenti debbano essere approntate due diverse reazioni).

La purificazione della sonda viene effettuata sfruttando un gel preparativo di poliacrilammide non denaturante al 6% sul quale vengono separati i frammenti radiomarcati; si recupera la sonda secondo il seguente protocollo:

- si ritaglia e si trasferisce la banda d'interesse in un'ependorf, disgregando in parte la matrice polisaccaridica, meccanicamente;
- calcolato approssimativamente il volume della banda si aggiungono 1-2 volumi di tampone di eluizione e si incuba a 37°C in agitazione. Frammenti < 500 pb sono eluiti in 3-4 h, mentre frammenti più lunghi richiedono 12-16 h.
- si centrifuga il campione a 12000g per 1 minuto a 4°C, recuperando il surnatante,
- si aggiungono ulteriormente 0.5 volumi di tampone di eluizione, si agita vorticosamente e si ricentrifuga, riunendo i due surnatanti.
- Rimossi i rimanenti frammenti di poliacrilammide tramite filtrazione, si aggiungono 2 volumi di etanolo al 96% e si attua un'incubazione in ghiaccio per 30 minuti;

- si recupera il DNA con una centrifugazione a 12000g per 10 minuti a 4°C, e si risospende in 200 µl di TE (pH 7.6), con addizione di 25 µl di sodio acetato 3 M (pH 5.2); si riprecipita il DNA con due volumi di etanolo;
- dopo un lavaggio in etanolo al 70% si risospende in 10 µl di TE (pH 7.6).

### Reazione di legame DNA-proteina

La reazione di legame prevede l'addizione di una quantità appropriata di proteina, o di estratto cellulare proteico anche non purificato, alla sonda previamente preparata, in condizioni di reazione il più simili possibili all'ambiente nucleare cellulare.

E' possibile allestire un numero variabile di reazioni con quantità crescenti di ogni proteina (od estratto) in modo tale da valutare su più punti l'interazione con il DNA.

La reazione standard utilizzata per le HMGI:

- 5 µl di FBB 10×(Footprinting Binding Buffer);
- 10.000 cpm di sonda;
- 1 µg di poly dG/dC, usato come competitore aspecifico;
- 10 µg di BSA (Bovine Serum Albumine), che agisce da proteina competitorice nel caso in cui si utilizzino proteine pure;
- quantità variabili di proteina (od estratto);

H<sub>2</sub>O a 50 µl finali.

Le condizioni usate per Sp1 o CTF/NF-1

Segue un'incubazione di 30 minuti a temperatura ambiente, condizione ottimale per le interazioni DNA-HMGI già testata in analisi EMSA.

### Trattamento della miscela DNA-proteina con DNasi I

La Dnasi I, prima di essere usata in esperimenti di *footprinting* su complessi DNA-proteina/e, viene testata esclusivamente sulla sonda per valutare la quantità di enzima e il tempo di reazione più appropriati per il proprio sistema.

Per ogni reazione di legame DNA-proteina approntata precedentemente,

- si addiziona MgCl<sub>2</sub> 4 mM finale (essenziale per un'ottima funzionalità dell'enzima);
- e 0.2 U di Dnasi I, incubata per 12 minuti esatti.

L'attività enzimatica viene bloccata con 50 µl di soluzione di stop 2×.

- Segue un'estrazione fenolo/cloroformio e precipitazione con etanolo per eliminare la Dnasi I.
- I frammenti di DNA vengono poi risospesi in H<sub>2</sub>O e tampone di caricamento.

### Separazione elettroforetica

Le reazioni di digestione si separano mediante migrazione elettroforetica su gel denaturante (8 M urea) di poliacrilammide al 7%, in modo tale da risolvere differenze di una sola base.

L'elettroforesi viene effettuata in celle verticali, con dimensioni 20×40 cm e spessore di 6 mm.

Durante la precorsa del gel (20 minuti a 1800 Volts), si incubano i campioni a 90°C per 3 minuti, in modo tale da eliminare possibili strutture secondarie che potrebbero interferire con la corretta migrazione del campione.

La corsa viene effettuata come descritto nella sezione 3.1.4 per i gel di sequenza.

Al termine della corsa, la cella viene disassemblata; il gel viene fissato (in acido acetico 10% per 10 minuti) e lavato con acqua distillata e poi essiccato su carta 3MM.

L'autoradiografia viene effettuata in apposite cassette dotate di schermi intensificatori per un tempo variabile dipendente dall'attività della sonda (da un minimo di 12 h ad oltre 48 h) a -80°C.

La lettura della lastra risultante è resa possibile dall'uso di *standards* costituiti da prodotti di sequenza, approntati con la tecnica di Maxam e Gilbert, della stessa sonda di DNA radiomarcata usata per il saggio di protezione.

## 3.4 Linee cellulari e condizioni di coltura

### 3.4.1 Linee cellulari

**NIH-3T3** : fibroblasti di topo, crescono in adesione.

**FRTL C12** : sono cellule di tiroide di ratto giovane che conservano inalterati i markers del differenziamento. Crescono in adesione in modo monostratificato.

**FRTL KiMol** : FRTL 5 C12 trasformate con KiMSV, crescono come le cellule normali, ma non necessitano della miscela di ormoni.

**PC C13** : sono cellule di tiroide di ratto adulto che conservano inalterati i markers del differenziamento. Crescono in adesione in modo monostratificato.

**PC 663 mos** : PC C13 trasformate con l'oncogene v-mos del MPSV.

**PC E1a/raf** : PC C13 trasformate con v-raf e l'oncoproteina E1a.

**FRL** : Cellule epatiche di ratto Fischer, crescono adese alla piastra.

**FRL-ras-V** : epatoma derivato da cellule FRL trasformate con l'oncogene v-ras-Ki.

**C5N** : Cheratinociti di topo crescono formando un monostrato in adesione.

**A5** : Carcinoma della pelle di topo, crescono in adesione.

**C4-I** : Linea cellulare di carcinoma squamoso della cervice. Mostra marcatori di differenziazione e cresce in adesione.

**CaSKi** : Linea cellulare umana di carcinoma della cervice, crescono in adesione.

**ME-180** : Linea cellulare derivata da metastasi di carcinoma squamoso della cervice. Crescono in adesione.

**QT6** : fibrosarcoma di quaglia, crescono in adesione.

**SL2** : Cellule Schneider di *Drosophila*, crescono in semi-adesione.

**KG1** : Linea cellulare di leucemia mieloide acuta.

**Nalm6** : Linea cellulare di linfociti pre-B

**Ri-1** : linea di cellule linfocitarie pre-B.

**Jurkat** : Linea cellulare leucemia a cellule T.

Le ultime quattro linee crescono in sospensione.

### 3.4.2 Condizioni di coltura

Le cellule NIH-3T3, QT6, FRL, FRL-ras-V, C5N, A5, ME-180, CaSKi e C4-I crescono in terreno DMEM (Gibco BRL) supplementato con siero fetale bovino al 10%, con una miscela dei antibiotici streptomina/penicillina (100mg/ml e 100U/l rispettivamente) e glutammina 2mM.

Le cellule KG1, Nalm6, Ri-1 e Jurkat sono state coltivate in terreno RPMI (Gibco BRL) supplementato con siero fetale bovino al 10%, con una miscela dei antibiotici streptomicina/penicillina (100mg/ml e 100U/l rispettivamente) e glutammina 2mM.

Le cellule epiteliali di tiroide di ratto (185) FRTL5 Cl2 e PC Cl3 e le linee trasformate derivanti sono state coltivate in terreno F12 Coon's modificato (*Sigma*) con 5% FCS (Fetal Calf Serum), anch'esso supplementato con glutammina 2 mM, penicillina 100 U/l e streptomicina 100 mg/ml.

Inoltre necessitano dell'aggiunta di una miscela di ormoni (6H), costituita da:

- insulina 10 µg/ml
- tiotropina 10 mU/ml
- transferrina 5 mg/ml
- idrocortisone 10 nM
- somatostatina 10 ng/ml
- glicil-L-istidil-L-lisina 10 ng/ml

Le stesse cellule trasformate sono state coltivate nello stesso terreno ma senza ormoni.

Le cellule SL2 di *Drosophila* sono state mantenute in Schneider's Medium (Life Technologies Inc.) supplementato con 10% siero fetale bovino e la miscela di antibiotici c.s. Non serve l'aggiunta di Glutammina, dato che è già presente, in forma stabile, nel terreno acquistato.

Tutte le cellule sono state cresciute in un incubatore a 37°C con CO<sub>2</sub> al 5% in modo da mantenere costante il pH tranne le cellule di *Drosophila* a 25°C senza CO<sub>2</sub>.

### 3.4.3 Trasfezione di cellule eucariotiche in adesione

Per studiare la funzione in vivo di una presunta sequenza regolativa o di un determinato prodotto genico è necessario introdurre DNA nel nucleo cellulare.

Le cellule degli eucarioti superiori però non captano naturalmente il DNA: in realtà hanno messo in atto durante l'evoluzione tutti i mezzi per proteggersi dal materiale genetico invasore.

Quindi, al fine di aggirare le barriere naturali poste al trasferimento genico, è stato impiegato un certo numero di espedienti tecnici: la trasfezione mediata dal calcio fosfato è uno fra i metodi più diffusi, in quanto la sua efficienza è tale da garantire l'espressione transiente di DNA esogeno in un elevato numero di cellule.

Per stabilire che non vi siano interferenze dovute ad integrazione cromosomica si possono effettuare anche trasfezioni stabili. Il confronto tra diversi cloni di trasfettanti stabili può dare queste informazioni.

#### Trasfezione transiente di coprecipitati di Calcio fosfato e DNA

Questo metodo si basa sulla tecnica elaborata da Graham e van der Eb (1973), che sfrutta il fatto che l'assunzione di DNA esogeno da parte di cellule in coltura è notevolmente incrementata quando l'acido nucleico viene presentato come fine precipitato di calcio fosfato-DNA (che nel suo scheletro contiene gruppi fosfato). Anche se il meccanismo rimane oscuro, si pensa che il DNA trasfettato entri nel citoplasma delle cellule mediante endocitosi e poi sia trasferito nel nucleo.

L'efficienza di questo metodo è generalmente piuttosto alta, ma dipende strettamente dal tipo cellulare usato: in alcuni casi quindi può essere incrementata mediante operazioni aggiuntive, fra cui lo shock da glicerolo.

### Condizioni di trasfezione

Le cellule FRTL5 Cl2 e PC Cl3 (e i corrispettivi trasformati) vengono trasferite 48 ore prima dell'esperimento in piatti del diametro di 6 cm a una densità di  $1,8 \times 10^4/\text{cm}^2$ . Tre ore prima della trasfezione il mezzo di coltura viene sostituito con terreno DMEM con 5% FCS, supplementato con:

- glutammina 2 mM
- penicillina 100 U/ml
- streptomicina 100 mg/ml
- HEPES 10 mM (pH 7,35)
- miscela di ormoni (6H).

Per ogni piatto da 6 cm di diametro si preparano i precipitati:

A 135,5  $\mu\text{l}$  della miscela dei DNA plasmidici che si vogliono trasfettare (la cui quantità totale dev'essere almeno di 5  $\mu\text{g}$  perchè l'efficienza sia soddisfacente), si aggiungono 14,5  $\mu\text{l}$  di  $\text{CaCl}_2$  (2,5 M). Se l'esperimento prevede quantità minori di DNA, è necessario aggiungere DNA carrier fino a 5  $\mu\text{g}$ .

Questa soluzione di DNA e  $\text{CaCl}_2$  viene aggiunta a goccia a goccia a 133  $\mu\text{l}$  di HBS 2X, agitando delicatamente ad ogni aggiunta. Durante un'incubazione a temperatura ambiente di 20-30 minuti, si forma un fine precipitato: minori sono le dimensioni dei microcristalli di DNA e fosfato, migliore è l'efficienza di trasfezione.

A questo punto si risospende questa miscela e si stratifica uniformemente la sospensione di calcio fosfato e DNA sulla superficie del terreno al di sopra del monostrato di cellule. Se non sono richieste ulteriori operazioni, le cellule vengono incubate per 16 ore. Quindi si rimuove il terreno insieme ai precipitati e, dopo due lavaggi con PBS, si aggiungono 5  $\mu\text{l}$  di terreno preriscaldato.

Nel caso delle linee di tiroide utilizzate, è necessario un breve trattamento aggiuntivo delle cellule trasfettate con glicerolo: in questo modo si è visto che aumenta l'efficienza di trasfezione e l'espressione transiente del DNA esogeno. Dopo un'ora dalla trasfezione, si sostituisce il terreno con 1,5 ml di soluzione per shock da glicerolo e si incuba per 2 minuti a 37°C; quindi, aspirata la soluzione di glicerolo, si lava due volte con PBS per rimuoverne tracce residue (visti i suoi effetti tossici per le cellule) e si aggiungono 5 ml di terreno completo.

Lo stesso metodo è stato usato per trasfettare le cellule NIH-3T3 e di *Drosophila*, avendo però  $0,5 \times 10^6$  o  $2 \times 10^6$  cellule per piastra e trasfettando con 5 o 10  $\mu\text{g}$  totali di DNA, rispettivamente. Inoltre la raccolta avviene 48 h dopo la trasfezione avendo fatto 1 lavaggio 24 h dopo la trasfezione in NIH-3T3 e nessun trattamento per le SL2, nel mentre.

### Trasfezioni stabili

$1,5 \times 10^6$  cellule NIH-3T3 sono state trasfettate col metodo del calcio fosfato c.s. con 18  $\mu\text{g}$  di costrutto *reporter* e 2  $\mu\text{g}$  del plasmide pWL neo (Stratagene). 48 ore dopo la trasfezione le cellule sono state selezionate per la resistenza a G418 aggiungendo al

mezzo di coltura 0.5 mg/ml di G418 (GIBCO). Dopo 30 giorni si sono contate le colonie e sono state piastrate assieme (tra 50 e 100 colonie per ogni costrutto trasfettato). Lo stesso numero di cellule è stato fatto crescere per 48 ore in 10% siero fetale bovino prima di saggiarne l'attività luciferasica.

### Saggi di trasfezione

Per poter evidenziare l'effetto di particolari fattori sull'attività trascrizionale di specifiche regioni di DNA vengono utilizzati numerosi geni reporter, per lo più (anche se non esclusivamente) di origine procariotica. In generale nella scelta di un gene reporter sono importanti le seguenti proprietà dell'enzima codificato:

- non deve interferire o competere con altre attività enzimatiche delle cellule di mammifero;
- la sua attività dev'essere facilmente e univocamente distinguibile da altre attività analoghe presenti nelle cellule prima della trasfezione;
- i saggi per rilevarne l'attività devono essere sensibili e possibilmente rapidi.

### Saggio di attività del marcatore (reporter)

al 1986 è stato introdotto il gene per la luciferasi di lucciola (*Photinus pyralis*) come enzima reporter in studi di regolazione e di espressione genica.

La proteina non richiede modificazioni post-traduzionali per essere attiva e perciò funziona subito dopo essere stata sintetizzata, catalizzando l'ossidazione della luciferina di coleottero in una reazione che richiede ATP,  $Mg^{2+}$  e  $O_2$  ed è accompagnata dall'emissione di un lampo di luce. I reagenti per il saggio contengono inoltre coenzima A che migliora la cinetica di questa reazione e consente di ottenere segnali luminosi la cui intensità è notevolmente incrementata rispetto a saggi convenzionali e viene rapidamente misurata per mezzo di un luminometro.

Questo sistema oltre alla rapidità, garantisce una notevole sensibilità visto che consente di rilevare quantità di enzima al di sotto di 1 fg e inoltre ha un intervallo di linearità di 8 ordini di grandezza.

Per le diverse linee cellulari utilizzate, sono stati usati due differenti sistemi di reagenti e metodi di misurazione:

**NIH 3T3** : Si uniscono 5  $\mu$ l del lisato a 50  $\mu$ l di substrato (*Luciferase Assay Sistem - Promega-*); il luminometro viene programmato per 5 misure successive di 5 secondi ciascuna e dai valori ottenuti si ricava la media.

**FRTL-5**: Si utilizzano 10  $\mu$ l del lisato per 50  $\mu$ l di substrato (*Dual Luciferase Assay Sistem -Promega-*) e il luminometro effettua una sola misurazione della durata di 30 secondi.

In entrambi i casi va considerata l'autoluminescenza del substrato nella successiva elaborazione dei dati.

### Saggi di efficienza di trasfezione

Per poter confrontare i risultati forniti dal reporter, che riflettono l'effetto di specifiche condizioni sperimentali sul livello dell'espressione genica, è necessario includere un controllo interno attraverso il quale è possibile normalizzare i valori del reporter "sperimentale" ottenuti da ciascun gruppo di trasfezioni. In questo modo si riesce ad

eliminare variabili intrinseche come differenze nel numero e nella vitalità delle cellule da piatto a piatto o anche nell'efficienza di trasfezione e di lisi.

Sono stati sfruttati a questo scopo due enzimi e di conseguenza diversi sistemi di saggi che offrono un differente grado di sensibilità.

### Saggi di attività della $\beta$ -Galattosidasi

Il gene *lacZ* di *E. coli* è stato sfruttato come controllo interno per le trasfezioni in NIH 3T3, QT6 e SL2: è stato utilizzato un costrutto (*SV40*  $\beta$ -gal) che reca questo gene sotto il controllo dell'intera regione promotrice di *SV40*, oppure *pPac* $\beta$ -gal per le cellule *SL2*, dove il promotore è specifico per *Drosophila*. (Hagen & Suske)

Gli estratti usati per i saggi di attività luciferasica possono essere usati anche per misure dell'attività della  $\beta$ -galattosidasi; sono stati utilizzati due diversi sistemi di saggi che sono caratterizzati da diversa sensibilità:

### Saggio colorimetrico

Sfrutta la capacità della  $\beta$ -galattosidasi di convertire un substrato artificiale, l'0-nitrofenil- $\beta$ -galattoside (ONPG -*Sigma*-), in un prodotto di colore giallo; l'attività enzimatica può essere quindi stimata attraverso misurazione spettrofotometrica alla lunghezza d'onda di 410 nm; questo sistema non è molto sensibile dato che il suo limite di rilevazione non va al di sotto di 1 ng circa di enzima.

Si procede preparando la seguente miscela per ogni campione:

- 66  $\mu$ l di ONPG (0,88 mg/ml finali)
- 216  $\mu$ l di tampone P
- 3  $\mu$ l di Mg 100x
- 10  $\mu$ l di lisato cellulare

Inoltre è essenziale includere un bianco per stimare l'assorbimento dovuto al solo substrato, sostituendo all'aliquota di estratto cellulare un egual volume di tampone di lisi.

Si incuba a 37°C per un'ora o fino a che non si sviluppa un colore giallo; a questo punto la reazione viene bloccata aggiungendo 500  $\mu$ l di Na<sub>2</sub>CO<sub>3</sub> e si effettua la misurazione della densità ottica (l'intervallo di linearità precedentemente stimato è compreso tra 0,05 e 0,5 OD nelle condizioni sperimentali adottate).

### Saggio fluorimetrico

L'utilizzo del metilumbrelliferil- $\beta$ -D-galactopiranoside (MUG -*Sigma*-) per valutare l'attività della  $\beta$ -galattosidasi ha il vantaggio di essere molto più sensibile del precedente, dato che, sfruttando l'emissione del composto fluorescente che viene prodotto, è in grado di rilevare fino a 1 pg di enzima.

La miscela di reazione è composta da:

- 240  $\mu$ l di tampone P + MgCl<sub>2</sub>
- 100  $\mu$ l di MUG 1x
- 10  $\mu$ l di estratto cellulare.



Anche in questo caso si prepara un bianco per stimare la fluorescenza dovuta al solo substrato.

Dopo un'incubazione di un'ora a 37°C, la reazione viene bloccata con 200 ml dell'apposita soluzione di stop e si procede alla misurazione spettrofluorimetrica con  $\lambda(\text{eccitazione}) = 390 \text{ nm}$  e  $\lambda(\text{emissione}) = 460 \text{ nm}$ .

### Saggio di chemiluminescenza

Nelle trasfezioni delle cellule FRTL-5 e Ki Mol è stato utilizzato come controllo interno, in combinazione al gene reporter per la luciferasi di lucciola, il gene per la luciferasi di *Renilla Reniformis* sotto controllo del promotore della timidina chinasi (TK) dell'*Herpes Simplex* (pRL-TK Vector -*Promega*-).

Questo sistema (*Dual-Luciferase Reporter Assay Sistem -Promega-*) permette di poter misurare sequenzialmente l'attività delle due luciferasi in estratti cellulari all'interno della stessa provetta, eliminando così anche la variabilità introdotta durante il trasferimento dei campioni da saggiare. Inoltre un vantaggio fondamentale è rappresentato dalla sensibilità, molto maggiore rispetto ai metodi precedenti, dato che consente di rilevare fino a meno di 10 fg dell'enzima, con un intervallo di linearità di 7 ordini di grandezza.

La base del metodo consiste nel fatto che è possibile discriminare selettivamente fra le rispettive reazioni bioluminescenti, in quanto i due enzimi hanno non solo diversa origine evolutiva e struttura, ma anche necessitano di substrati differenti. Così la luminescenza prodotta dalla reazione catalizzata dall'enzima reporter può venire smorzata, attivando contemporaneamente l'emissione di luce prodotta per azione della luciferasi di *Renilla*.

Questo enzima in vivo è una glicoproteina, ma si è visto che la glicosilazione non è necessaria per la sua attività catalitica e quindi, subito dopo la traduzione, è in grado di ossidare il suo substrato, la luciferina di celenterato (celenterazina), generando un segnale luminoso che si estingue molto lentamente.

La misurazione della luminescenza prodotta dal controllo interno si effettua al luminometro subito dopo la lettura dell'attività dell'enzima reporter: allo stesso campione infatti si aggiungono immediatamente 50 ml di soluzione di "Stop & Glow" e si miscela con il vortex per qualche secondo; in questo modo viene smorzata l'emissione luminosa della luciferasi di lucciola (di un fattore superiore a  $10^5$ ) e contemporaneamente viene attivata la luciferasi di *Renilla*, con l'effetto aggiuntivo di ridurre l'autoluminescenza della celenterazina. A questo punto si misura il segnale prodotto per 30 secondi.

## 3.5 Analisi dei topi transgenici

### 3.5.1 Estrazione di DNA da pezzi di coda o dita

Questa estrazione serve per analizzare la presenza del transgene nei topi della generazione filiale F1, proveniente dall'accoppiamento tra un fondatore transgenico con un individuo dell'altro sesso negativo, dato che per la legge di Mendel solo il 25% della F1 dovrebbe contenerlo, se non vi sono effetti posizionali. Con questo protocollo si riesce ad avere sufficienti quantità per analisi mediante PCR e nello stesso tempo

l'individuo da cui proviene il DNA resta in vita per successive analisi. Si procede nel modo seguente:

- tagliare pezzo di coda o dito di circa 2mm;
- immergere in 50  $\mu$ l di tampone per PCR 1 X + proteinasi K (100  $\mu$ g/ml finale);
- tenere in agitazione a 60°C per 1-2 ore;
- sminuzzare il resto del tessuto con un puntale pulito, con attenzione particolare a non contaminare una provetta con l'altra, e agitare ancora per 30 min. - 1 ora a 60°C;
- mettere 7 minuti a 95°C per denaturare la proteasi;
- centrifugare brevemente;
- diluire 10  $\mu$ l in 100  $\mu$ l finali di acqua e usare 5  $\mu$ l di questa soluzione per la PCR.

### 3.5.2 Estrazione del DNA genomico da sacco vitellino (Yolk sac)

Essendo una produzione embrionale il sacco vitellino è adatto per essere sottoposto ad una verifica della presenza o meno del transgene, senza rovinare l'embrione che in seguito dovrà essere manipolato e rimanere intatto fino all'analisi istologica.

- Tutto il sacco vitellino viene messo in 100  $\mu$ l di "tail buffer" + proteinasi K (0.5 mg/ml);
- si incuba per 1-2 ore a 37°C;
- poi si effettua una separazione mediante 1 volume fenolo equilibrato in Tris/HCl 0.1 M pH 7.5;
- una mediante 1 volume fenolo/cloroformio 50:50
- ed una con 1 volume di cloroformio;
- la fase acquosa recuperata viene precipitata in 0.3 M NaCl più 2.5 volumi di Etanolo 96% a -80°C;
- dopo un lavaggio in etanolo al 70%
- si risospende in 100  $\mu$ l di T.E. senza usare la pipetta. Per facilitare la risospensione del DNA genomico si mette a 45°C per alcune ore.

La resa è di solito di circa 10  $\mu$ g totali che corrispondono in tutto a  $3 \times 10^6$  genomi, considerando che un genoma murino possiede all'incirca  $4.7 \times 10^9$  coppie di basi.

### 3.5.3 Southern blot genomico

Per analizzare i campioni per la presenza del transgene, oltre alla PCR esiste anche questo metodo. Si sceglie di usare questo metodo oltre a tutto anche per capire quante copie del transgene sono contenuti nel genoma degli animali transgenici, in modo da poter rapportare anche l'intensità della colorazione blu dovuta alla  $\beta$ -galattosidasi espressa con il numero di transgeni presenti. Questo metodo inoltre non dà problemi di contaminazione com'è possibile invece con la PCR. Di solito basta un'analisi su alcuni embrioni della stessa generazione F1, per controllare che la colorazione corrisponda effettivamente alla presenza del gene espresso e che non ci sia una colorazione aspecifica. Succede in effetti che pur avendo ottenuto individui transgenici (i fondatori) che sono positivi per Southern, non trasmettono il transgene alla progenie o trasmettano un gene inattivato (ad es. per metilazione). Questo significa che si trovano individui nella F1 il cui sacco vitellino presenta il transgene, ma l'embrione non si colora. In casi di questo genere si termina l'analisi su quella linea. Invece se si trova sia il transgene che l'espressione allora basta aver analizzato una F1.

## Trattamento dei campioni di DNA provenienti dal sacco vitellino

I 10 µg di DNA vengono digeriti con un enzima di restrizione appropriato (2-5 U/µg di DNA) in un volume abbondante (più di 200 µl) per molte ore. Poi si tratta con fenolo/cloroformio e si precipita come al solito per rendere il DNA più pulito possibile. Si lava il pellet con 70% di etanolo e si risospende in un volume appropriato da poter caricare un campione per pozzetto in un gel di agarosio. Per migliorare l'entrata nel gel e la migrazione di frammenti genomici ad alto peso molecolare, si è allestito un gel d'agarosio discontinuo, in cui l'entrata nel gel ed il primo centimetro di migrazione avvengono in un gel allo 0.8% a 40 V a 4°C, mentre la vera separazione si verifica in agarosio all'1% sempre a 40 V e a 4°C. Il gel viene preparato senza etidi bromuro, dato che si prevedono corse lunghe durante le quali l'etidio bromuro diffonderebbe dallo stesso. Dopo circa 24 ore di migrazione, in modo da aver separato anche frammenti di circa 3 e 4 Kbasi, si allestisce il trasferimento del DNA digerito su filtro di nylon mediante un protocollo di "Southern blotting" ottimizzato come segue:

per primo si colora il gel in TAE 1X + etidio bromuro per circa 30' in agitazione; poi si fa una foto con i riferimenti appropriati (es. righello accanto al gel), dopodiché si inizia il trattamento del gel con:

- depurinazione per 19 min. in 0.25 N HCl;
  - denaturazione in 0.5 M NaOH/1.5 M NaCl per 20-30 min. per due volte;
  - neutralizzazione in 1 M Tris/HCl pH 7.5/1.5 M NaCl per 20-30 min. per due volte;
- e si assemblano i vari strati del trasferimento come già descritto nella sezione 3.1.6, con l'unica variante che il tampone in questo caso è il SSPE 20X, e si lascia almeno per 12 ore.

Si disassembla il blot prestando attenzione di segnare la posizione dei pozzetti del gel sulla membrana (con una matita) prima di togliere la membrana. Si lava brevemente la membrana in 2X SSPE e si lascia asciugare bene tra 2 filtri di carta (3MM Whatman). In seguito si fissa irreversibilmente il DNA alla membrana mediante esposizione ai raggi UV per 1 min.

## Ibridizzazione con la sonda radioattiva

Tutto come descritto inella sezione 3.1.8 tranne che per i tamponi, in quanto si usa il tampone SPE per far aderire bene la membrana alla parete della bottiglia del fornello ibridizzatore, poi si usa il "Pre-Wet-buffer" per idratare bene la membrana per alcuni minuti. Si sostituisce con il tampone CG di ibridazione e si lascia pre-ibridare per alcune ore, finché non viene aggiunta la sonda radioattiva.

I lavaggi vengono eseguiti monitorando il distaccamento della sonda dal filtro ed aumentando la stringenza utilizzando il tampone CG di lavaggio.

## 3.5.4 Colorazione e disidratazione degli embrioni

Gli embrioni recanti il transgene vengono analizzati per l'espressione dello stesso nei giorni 10, 12, 14 e 16 p.c. attraverso una colorazione specifica per la β-galattosidasi prodotta nei tessuti dove il promotore dell'*Hmgi-c* inserito risulta attivato. Dopo il prelievo dell'embrione dall'utero della femmina sacrificata si procede così:

- breve lavaggio in PBS;

- 3 lavaggi in PBS + Nonidet P-40 allo 0.02% usando poco volume per i primi lavaggi ed aumentando la quantità negli ultimi;
- tagliare sagittalmente gli embrioni se necessario (se sono maggiori di circa 1cm di lunghezza);
- fissarli in paraformaldeide al 4% in PBS + NP-40 allo 0.02% per circa 15 min. a 4°C;
- lavare bene il fissativo con PBS + NP-40 allo 0.02% ( i lavaggi che seguono l'incubazione con la paraformaldeide vanno eseguiti con molta cura e a tempi più lunghi;
- incubare gli embrioni con la soluzione colorante durante tutta la notte a 30°C. La soluzione colorante è costituita da:
  - 0.4 mg/ml X-gal (5-Bromo-4-cloro-3-indolil  $\beta$ -D-galactopiranoside)
  - 2 mM Ferrocianuro di potassio
  - 2 mM Ferricianuro di potassio
  - 2 mM MgCl<sub>2</sub>
  - 0.02% NP-40
  - in PBS.

Dopo la colorazione si lavano gli embrioni brevemente in PBS freddo e si passa alla conservazione degli stessi per disidratazione alcoolica (sempre in ghiaccio o a 4°C):

- un passaggio in etanolo al 30% per 30 min.;
- un passaggio in etanolo al 50% per 30 min.;
- un passaggio in etanolo al 70% per 30 min., poi si lascia in etanolo al 70% fino all'inclusione in paraffina. (Durante ogni incubazione a metà passaggio fare 1 cambio del solvente). In questa forma si conservano anche per mesi.

## 3.6 Composizione delle soluzioni

### 3.6.1 Estrazioni ed analisi del DNA

**TAMPONE TAE 100X** 484 g/l Tris base; 114 g/l acido acetico; 200 ml/l EDTA 0,5 M (pH 8)

**TAMPONE DI CARICAMENTO 10X** 0,1 ml tampone TAE 100X; 0,2 ml glicerolo; 0,04% blu di bromofenolo, acqua distillata fino a 1 ml totale.

**SOLUZIONE SODIO PERCLORATO** 6 M NaClO<sub>4</sub>; 50 mM Tris (pH 8); 10 mM EDTA (pH 8)

**SOLUZIONE DI LAVAGGIO** 40 mM Tris/HCl (pH 7,4); 0,8 M NaCl; 4 mM EDTA (pH 8); aggiungere un volume di etanolo.

**TAMPONE TAQ 10× PER PCR** 500 mM KCl; 100 mM Tris/HCl 100 mM NaCl; 1% Triton X-100.

**TAMPONE OPA (One Phor All) PER PNK**

100 mM Tris/acetato; 100 mM acetato di magnesio; 500 mM acetato di potassio.

**TAMPONE OL10X**

0.7 nM esanucleotidi, 0.8M TRIS/HCl pH 7.5, 1 mM Ditiotreitolo, 0.6mM dATP, dTTP, dGTP

**TAMPONE TNES**

10 mM TRIS/HCl pH 7.5, 0.1M NaCl, 1 mM EDTA, 0.1% SDS

**TAMPONE DI DILUIZIONE PER LA NUCLEASI S1**

20mM Tris pH 7.5, 50mM NaCl, 0.1mM ZnCl<sub>2</sub>, 50% glicerolo v/v

**TAMPONE D'INCUBAZIONE PER LA NUCLEASI S1**

33mM CH<sub>3</sub>COONa pH 4.5, 50 mM NaCl, 0.03 mM ZnSO<sub>4</sub>

**3.6.2 Soluzioni per colture batteriche****TERRENO LIQUIDO DI LURIA-BERTANI (LB BROTH)**

10 g/l NaCl, 10 g/l bacto-triptone, 5 g/l bacto-yeast

**TERRENO SELETTIVO CON ANTIBIOTICI**

LB broth + 25 mg/l kanamicina e/o 50 mg/l ampicillina

**TERRENO SOLIDO DI LURIA-BERTANI (LB AGAR)**

LB broth + 14 g/l bacto-agar

**SOB**

20 g/l bacto-tryptone, 5 g/l bacto-yeast extract, 0.5 g/l NaCl

**SOC**

SOB + 20 mM glucosio e 2 M MgCl<sub>2</sub>

**CCMB**

80 mM CaCl<sub>2</sub>

20mM MnCl<sub>2</sub>

9 mM MgCl<sub>2</sub>

0.1 M glicerolo

**3.6.3 Soluzioni per il sequenziamento del DNA****TAMPONE TBE 10X**

108 g/l Tris base, 27,5 g/l acido borico, 40 ml/l EDTA 0,5 M (pH 8)

**SOLUZIONE DI ACRILAMMIDE AL 7%**

7 ml acrilammide/N,N' metilen-bis-acrilammide (29:1 g/100 ml), 3 ml tampone TBE 10X, 15 g urea e acqua distillata a 30 ml totali, filtrare, degassare e aggiungere:

60 µl ammonio persolfato (APS) al 25%

30 µl N,N,N',N' tetrametilen-diammina (TEMED)

**SOLUZIONE NON LEGANTE**

50% dimetil-cloro-silano 20 g/l in tricloroetano (silano repel), 50% etanolo

**SOLUZIONE LEGANTE**

0,3% silano legante in etanolo 95%: a 5 ml della soluzione aggiungere 75 µl di acido acetico 10%

**TAMPONE DI IBRIDAZIONE**

280 mM Tris/HCl (pH 7,4), 100 mM MgCl<sub>2</sub>, 350 mM NaCl, 150 mM ditiotreitolo (DTT)

**SOLUZIONE DI ELONGAZIONE**

7,5 mM dGTP, dATP, dTTP

**SOLUZIONI DI TERMINAZIONE**

80 mM dGTP, dATP, dTTP, dCTP +

- 8 mM ddGTP : G
- 8 mM ddATP : A
- 8 mM ddTTP : T
- 8 mM ddCTP : C

**SOLUZIONE DI STOP**

0,05% blu di bromofenolo, 0,05% xilene cianolo, 20 mM EDTA, 95% formammide deionizzata

**3.6.4 Soluzioni per l'analisi dell'RNA****SOLUZIONE D**

4M guanidinio tiocianato, 25 mM sodio citrato pH 7, 100 mM β-mercaptoetanol, 0.5% laurilsarcosina, 0.1% antifoam

**TAMPONE MOPS 10X**

0.2 M MOPS, 0.01M EDTA, 0.5M CH<sub>3</sub>COONa, pH 7.0

**TAMPONE DI CARICAMENTO PER RNA**

50% formammide deionizzata, MOPS 1.5 X, 6.2 % formaldeide + etidio bromuro (0.05 mg/ml) + 0.04% blu di bromofenolo

**TAMPONE DI CORSA PER RNA**

1 X MOPS, 3.7 % formaldeide

### 3.6.5 Soluzioni per l'analisi di proteine

#### GEL IN ACETICO/UREA

##### STACKING GEL:

20mg di N,N' metilen-bis-acrilamide (BIS), 20mg di persolfato di ammonio, 0,75g di acrilamide (7,5% p/v) sciolti in 10ml di solvente running, a cui si aggiungono 100 $\mu$ l di TEMED (N,N,N',N'-tetrametilendiammina), come catalizzatore della velocità di polimerizzazione.

##### RUNNING GEL:

60mg di BIS, 100mg di persolfato di ammonio, 4.5g (15% p/v) di acrilamide, sciolti in 30ml di solvente running, a cui si aggiungono 200 $\mu$ l di TEMED.

##### SOLVENTE RUNNING:

acido acetico 0,9N, urea 4M.

##### SOLVENTE BAGNO:

per la corsa elettroforetica, acido acetico 0,9N.

##### SOLVENTE CAMPIONI 1X:

solvente running al 15% di saccarosio (p/v) e al 5% di  $\beta$ -mercaptoetanolo.

##### SOLVENTE CAMPIONI 2X:

urea 8M, acido acetico 2M, saccarosio 20% (p/v),  $\beta$ -mercaptometanolo 5%.

#### GEL IN SDS

##### STACKING GEL:

20 mg di BIS, 7 mg di persolfato di ammonio, 0.75 g (7.5% p/v) di acrilamide sciolti in 9 ml di H<sub>2</sub>O mQ, 1 ml di solvente running 1, 100  $\mu$ l di SDS al 10%. Si aggiungono 7 $\mu$ l di TEMED.

##### RUNNING GEL:

63 mg di BIS, 60mg di persolfato di ammonio, 4,7g (15% p/v) oppure 3,1g (10% p/v) oppure 2,3g (7,5% p/v) di acrilamide; nell'ultimo caso ovviamente non esiste lo stacking gel; il tutto viene sciolto in 26ml di solvente running 1 e in 5ml di solvente running 2; si aggiungono 60  $\mu$ l di TEMED.

##### SOLVENTE BAGNO:

per la corsa elettroforetica: SDS 0.1%, Tricina 0.025M, portato a pH 8.1 con Tris base.

##### SOLVENTE CAMPIONI 1X:

SDS 1%, saccarosio 15%,  $\beta$ -mercaptoetanolo 4%, Tris-HCl 0.075M, pH 8.8

### 3.6.4 Soluzioni per Western blotting

#### SOLUZIONE DI TRASFERIMENTO 1

Tris base 0,075M al 20% di metanolo (facilita il legame delle proteine alla membrana)

#### SOLUZIONE DI TRASFERIMENTO 2:

Tris base 0.375M pH 8.8, SDS 0.05% (facilita il trasferimento)

#### SOLUZIONE ROSSO PONCEAU S:

Ponceau S 2%, acido tricloroacetico 30%, acido solfosalicilico 30%, da diluire 1:10

#### SOLUZIONE DI SATURAZIONE:

Tris 0.05M, pH 7.5, NaCl 0.5M, Tween 20 0.01%, latte magro in polvere 5%, albumina sierica bovina (BSA) 0.2%

#### STRIPPING BUFFER:

$\beta$ -mercaptoetanololo 100mM, SDS 2%, Tris-HCl 62.5mM pH 6,7

### 3.6.5 Soluzioni per bandshift

#### TAMPONE TBE 5X

0.45 M Tris base, 0.44 M acido borico, 0.01 M EDTA (pH 8)

#### TAMPONE TEN

10 mM Tris/HCl (pH 7.5), 1 mM EDTA (pH 8), 100 mM NaCl

#### TAMPONE DI LEGAME (BB10X)

200 mM Tris/HCl (pH 7.5), 750 mM KCl, 10 mM ditiotreitolo (DTT), 50 mg/ml albumina serica bovina (BSA)

#### SOLUZIONE DI ACRILAMMIDE AL 7%

28 ml acrilammide/N,N' metilen-bis-acrilammide (29:1 g/100 ml), 12 ml tampone TBE 5X, acqua distillata a 120 ml totali. Aggiungere: 400 ml ammonio persolfato (APS) al 25%, 100 ml N,N,N',N' tetrametilen-diammina (TEMED)

### Soluzioni per gli estratti nucleari

#### SOLUZIONE I

10 mM HEPES pH 7,9, 10 mM KCl, 1.5 mM Mg Cl<sub>2</sub>, 0.1 Mm EGTA  
0.1 Mm DTT, 0.5 mM PMSF

#### SOLUZIONE II

soluzione I, 5% glicerolo, 400mM NaCl

### 3.6.6 Soluzioni per il footprinting

#### TAMPONE DI ELUIZIONE

0.5 M ammonio acetato, 10 mM magnesio acetato, 1mM EDTA (pH 8.0), 0.1% SDS



**FBB 10× (HMGI)**

250 mM Hepes pH 7.6, 500 mM NaCl, 10 mM DTT, 10 mM EDTA, 0.1% NP40

**FBB 10X (Sp1 e CTF/NF-1)**

25 mM Tris/HCl pH 7.9, 0.5 mM EDTA, 0.6mM DTT, 0.65mM MgCl<sub>2</sub>, 50 mM KCl, 0.2 μM ZnCl<sub>2</sub> solo per Sp1, 10% glicerolo

**SOLUZIONE DI STOP**

0.2 % SDS, 20 mM EDTA, 20 ng/μl DNA di sperma di salmone

**3.6.7 Soluzioni per trasfezione****PHOSFATE BUFFER SALINE (PBS)**

137 mM NaCl, 27 mM KCl, 4.3 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.4 mM KH<sub>2</sub>PO<sub>4</sub>

**HBS 2X**

50 mM Hepes, 1.5 mM Na<sub>2</sub>HPO<sub>4</sub> portare a pH 7,12 con NaOH

**SOLUZIONE PER SHOCK DA GLICEROLO**

15 % glicerolo in HBS 1X

**SOLUZIONE PER IL SAGGIO PER LA β-GALATTOSIDASI**

1 mM MgCl<sub>2</sub>, 45 mM β-mercaptoetanolo, 0.8 mg/ml ONPG (o-nitrofenil β-D-galattopiranoside) in tampone P (=tampone fosfato 0.1 M , con Na<sub>2</sub>HPO<sub>4</sub> 60 mM e 40 mM NaH<sub>2</sub>PO<sub>4</sub>) oppure 1 X MUG (metil-umbrelliferil β-D-galactopiranoside) in tampone P + 1mM MgCl<sub>2</sub>

**SOLUZIONE DI STOP PER ONPG**

1 M Na<sub>2</sub>CO<sub>3</sub>

**SOLUZIONE DI STOP PER MUG**

300 mM GLICINA, 15 mM EDTA; portare a pH 11.5 con NaOH

**3.6.8 Soluzioni per l'analisi di topi transgenici****“TAIL BUFFER”**

50 mM TRIS/ HCl pH 8, 100 mM EDTA, 100 mM NaCl, 1% SDS

**20 X SSPE (Standard-Saline-Phosphate-EDTA)**

3 M NaCl, 200 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7.4, 20 mM EDTA

**TAMPONE SPE (Sodium Phosphate-EDTA)**

25 mM NaH<sub>2</sub>PO<sub>4</sub> pH 6.8, 1 mM EDTA

**TAMPONE DI IDRATAZIONE (Pre-Wet-Buffer)**

= tampone GC senza albumina sierica bovina

**TAMPONE CG DI IBRIDAZIONE**

500 mM  $\text{NaH}_2\text{PO}_4$ , 7% SDS, 1% BSA, 1 mM EDTA

**TAMPONE GC DI LAVAGGIO**

40 mM  $\text{NaH}_2\text{PO}_4$  pH 7.2, 1% SDS, 1 mM EDTA

## Capitolo 4.

### Risultati e discussione

#### 4.1 Le proteine HMGI in associazione con il DNA

##### 4.1.1 Le proteine HMGI interferiscono con gli omeodomini legandosi al DNA (125, Allegato n°1))

Gli *Omeodomini* (HD) sono sequenze di 61 residui amminoacidici capaci di interagire con sequenze specifiche del DNA. Rappresentano il dominio di legame al DNA di un grande numero di fattori di trascrizione e controllano momenti decisivi dello sviluppo in molteplici organismi ad iniziare dal lievito e dagli insetti fino ai vertebrati superiori. La struttura degli omeodomini ed il loro modo di interagire col DNA risultano conservati e, con pochissime eccezioni, l'elemento centrale riconosciuto è 5'-TAAT-3'. Nonostante riconoscano motivi simili di sequenze di DNA, fattori aventi un HD ed espressi nello stesso contesto temporale e spaziale, possono mostrare diverse attività biologiche. Quindi la specificità per il DNA di queste proteine non può essere l'unica caratteristica molecolare che determina il loro effetto biologico. Sono stati scoperti infatti alcuni meccanismi molecolari, oltre alla specificità di legame col DNA, che permettono a distinti fattori con un *omeodominio* di selezionare i geni bersaglio.

Uno di questi meccanismi si basa sull'interazione differenziale dei fattori aventi un HD con fattori accessori che mediano i contatti con l'apparato basale della trascrizione. Ad esempio i fattori trascrizionali Oct-1 e Oct-2, contenenti entrambi un dominio HD, riconoscono la stessa sequenza di DNA, ma solo Oct-1 è capace di interagire con l'attivatore trascrizionale VP16. Un altro esempio di questo fenomeno è l'interazione tra le proteine Ultrabithorax e Extradenticle in *Drosophila*.

Un semplicissimo meccanismo di regolazione che potrebbe modulare l'attività delle proteine con HD a livello di sequenze regolatrici è la presenza di competitori per il legame alla sequenza

bersaglio. La sequenza 5'-TAAT-3', motivo nucleotidico riconosciuto dagli HD, è un'ottima candidata per una interazione specifica anche con le proteine HMGI (vedi sezione 1.2.3). Pertanto, in una parte del lavoro riportato in questa tesi, si è voluto indagare se un'effettiva competizione potesse avvenire tra proteine contenenti un HD e membri della famiglia HMGI sia *in vitro* che *in vivo* a livello di un promotore naturale solitamente controllato dal fattore di trascrizione contenente un *omeodominio*.

I risultati hanno dimostrato che le proteine HMGI sono capaci di interagire con alcune sequenze di DNA bersagli dei fattori contenenti gli HD, e che il legame delle HMGI e degli HD all'elemento 5'-TAAT-3' è di mutua esclusione. In particolare, la presenza di proteine HMGI blocca l'interazione tra gli HD ed il DNA non solo attraverso una semplice competizione, ma apparentemente anche alterando la struttura del DNA a cui si legano. Si è scelto un modello dato dall'attivazione trascrizionale di HOXD9 su un elemento autoregolativo e si è dimostrato che HMGY è capace di competere con HOXD9 per il legame al promotore di HOXD9 *in vitro* e di ridurre l'attivazione trascrizionale specifica di HOXD9 in cellule trasfettate.

Nella conduzione della ricerca, come primo approccio, si è verificato, attraverso esperimenti di *ritardo di banda*, se le proteine HMGI-C ed HMGY fossero capaci di legare specificamente sequenze di riconoscimento specifiche degli HD. Per tale motivo, le sequenze codificanti per HMGI-C ed HMGY murine sono state amplificate mediante PCR (utilizzando oligonucleotidi specifici derivanti dalla sequenza N- e C-terminale rispettivamente) e sono state clonate "*in frame*" tra i due siti *NdeI* e *BamHI* del vettore di espressione batterico pAR3038 sotto il promotore del batteriofago T7. Le rispettive proteine ricombinanti sono state espresse in batteri, estratte con acido perclorico e purificate mediante HPLC. La bontà del prodotto è stata verificata mediante spettrometria di massa. (vedi sezioni 3.1.3, 3.1.9 e 3.3.1-3.3.5)

Gli oligonucleotidi usati negli esperimenti di *bandshift* (sezione 3.1.6) sono stati scelti per individuare una specificità di legame tipica delle HMGI essendo stato scelto come controllo positivo di legame specifico per le HMGI l'oligonucleotide PRDII derivato dal promotore umano dell'*INFβ* (Thanos 92):

PRDII: 5'-GGGAAATTCCTGGGAAATTCGAGCT-3' (Controllo positivo)

Oligonucleotidi derivati da sequenze specifiche riconosciute da fattori aventi un HD usati negli esperimenti di *bandshift* (sono sottolineati i nucleotidi centrali della sequenza di riconoscimento degli HD):

CAnt1: 5'-CACTGCCCCAGTTAAATTGTTCTTGA-3'

NP: 5'-TGATATTTAATTGATTTT-3'

CAnt4: 5'-CACTGCCCCAGTTAAGTGTTTCTTGA-3'

T109: 5'-GCACGCTAATTGAGACGC-3'

HCR11: 5'-GACACATTAATCTATAATCAAATAC-3'

Con tali esperimenti si è evidenziata una particolare specificità di HMGI-C ed HMGY nei confronti delle sequenze di riconoscimento degli HD. Innanzitutto entrambe le HMGI legano efficientemente CAnt1, che è un sito d'interazione dominante degli HD di Antennapedia ed Engrailed, e richiedono che l'elemento 5'-TAAT-3' sia integro, perchè CAnt4, dove il *core* di CAnt1 è stato mutato a 5'-TAAG-3', non viene più legato né da HMGI-C né da HMGY. Interagiscono inoltre efficacemente con NP che rappresenta anch'esso un sito di legame per molti tipi di HD. Invece nessuna interazione efficace si è potuta stabilire tra le proteine HMGI ed l'oligonucleotide T109, che invece è riconosciuto sia da Antennapedia che da Engrailed. Abbiamo concluso che è molto probabile che le proteine HMGI riconoscano uno specifico sottogruppo di sequenze legate dagli HD.

Data questa evidenza sperimentale è possibile che le proteine HMGI e quelle con un HD si influenzino a vicenda durante l'interazione con gli stessi siti sul DNA. Questa possibilità è valutata incubando assieme due rappresentanti di queste famiglie proteiche con la sonda marcata CAnt1. I risultati dimostrano chiaramente che in presenza sia di HMGY che di HMGI-C l'omeodominio di Engrailed (EnHD) non è più in grado di formare un complesso con CAnt1, mentre il complesso EnHD/T109 non è minimamente inficiato. Ciò significa che per bloccare l'interazione tra un HD con il suo sito di riconoscimento è necessario che le proteine HMGI si leghino al DNA e quindi molto probabilmente non entrano in gioco interazioni proteina-proteina, come invece è stato descritto per HMG1, che interagisce con proteine HOX (proteine umane con un HD) anche in assenza di DNA.

È interessante notare che l'effetto inibitore è osservato ad un rapporto molare di HMGI-C/EnHD pari a 1,8. Sapendo che le proteine HMGI sono presenti nella cellula in grandi quantità (141) si può concludere che nella cellula questo rapporto è sicuramente maggiore e che quindi l'interferenza osservata potrebbe avvenire anche *in vivo*.

Il meccanismo molecolare attraverso il quale le proteine HMGI inibiscono il legame dell'*omeodominio* al DNA potrebbe essere duplice: innanzitutto gli HD contattano il DNA sia a livello del solco maggiore, con l'elica principale di riconoscimento, sia nel solco minore con il cosiddetto "braccio" N-terminale, per cui potrebbe esserci una prima competizione con le proteine HMGI per il legame nel solco minore; il secondo effetto potrebbe essere dato dal fatto che l'AT-hook, una volta legatosi, ha la capacità di distorcere il DNA e quindi non permette più all'HD di legarsi nel solco opposto.

Esperimenti di cinetica di questa inibizione mostrano che la sequenzialità dell'aggiunta delle proteine nella miscela di incubazione con EnHD e l'oligonucleotide CAnt1 non ha importanza, infatti, mettendo HMGI-C messa prima o dopo l'EnHD inibisce comunque la formazione del complesso EnHD/CAnt1. Si è osservato che questa inibizione si completa in pochi secondi e che l'emivita del complesso EnHD/CAnt1 è molto ridotta in presenza di HMGI-C (in concentrazioni che permettono ancora la formazione del complesso stesso). Questi dati concordano con l'ipotesi che l'HMGI si lega al solco minore con maggiore efficienza spiazzando il braccio N-terminale di EnHD.

*Berenil* è un farmaco che lega il solco minore senza indurre cambiamenti della struttura del DNA ed altera l'interazione tra omeodomini ed il DNA solo modestamente se utilizzato a concentrazioni altissime (referenze 51 e 52 nell'Allegato 1). Inserendo anche questa sostanza negli esperimenti di *ritardo di banda* si è potuto ripristinare la formazione del complesso EnHD/CAnt1, inibito da HMGI-C. Infatti il berenile, legandosi nel solco minore compete con HMGI-C. Quindi la spiazza e permette all'EnHD di legarsi nuovamente a CAnt1, anche perchè il berenile per l'appunto non cambia la conformazione del DNA se viene usato alle stesse concentrazione come nei nostri esperimenti. Questa evidenza supporta il modello che cambiamenti strutturali del DNA mediano l'effetto inibitorio delle proteine HMGI sull'interazione tra HD e DNA.

Per verificare che questi cambiamenti strutturali avvengono sul DNA come conseguenza del legame con le proteine HMGI si sono effettuate delle analisi diicroismo circolare. Il confronto degli spettri CD dell'oligonucleotide CAnt1 da solo e con HMGI-C dimostrano che avvengono dei cambiamenti strutturali, molto simili a quelli osservati quando il DNA viene compattato nei nucleosomi. Quindi una molteplicità di esperimenti concordano nel asserire che una parte dell'interferenza di legame tra HMGI ed HD è dovuta a modificazioni strutturali del DNA.

Per rispondere alla domanda se l'effetto inibitorio delle HMGI sull'interazione HD-DNA possa essere rilevante anche *in vivo*, abbiamo allestito esperimenti di cotrasfezione in cellule NIH-3T3, per misurare l'effetto inibitorio di HMGY sull'attivazione trascrizionale dalla proteina HOXD9 a livello della regione autoregolativa, denominata HCRII, del proprio promotore. HCRII possiede più

motivi 5'-TAAT-3', di cui uno, se mutato, riduce di molto l'effetto trascrizionale di HOXD9. Controlli preliminari hanno dimostrato che HMGY è capace di legare HCRII *in vitro*, e che se si incuba assieme HMGY, HOXD9 ed HCRII si forma solo un complesso tra HMGY ed HCRII e non tra HOXD9 ed il DNA e nemmeno un complesso ternario HMGI/HCRII/HOXD9.

Per le trasfezioni è stato usato pTHCR, un plasmide *reporter* contenente il gene marcatore della luciferasi a valle dell'elemento HCRII fuso con il promotore della timidina chinasi del virus herpes simplex. Tale costrutto è stato co-trasfettato con quantità crescenti del plasmide di espressione per HOXD9 (pSGD9, recante il promotore dell'SV40) e si è potuto osservare un aumento dell'attività trascrizionale di base di HCRII di 4-5 volte. Lo stesso procedimento, con i dovuti controlli, in presenza di quantità fisse del plasmide pSVHMGY, che esprime HMGY sotto il controllo dello stesso promotore di SV40, porta ad un'inibizione dell'attivazione mediata da HOXD9 del 50%.

La scelta di questo modello non è casuale. Infatti, tranne che per pSVHMGY, si tratta di vettori già utilizzati per uno studio analogo con la proteina HMG 1 dove si era dimostrato che HMG 1 interagisce direttamente con HOXD9 rafforzandone sia l'attività di legame al DNA che il potere trascrizionale. (189) Questo effetto HMG 1 lo media solo attraverso interazione proteina-proteina, dato che non è in grado di legarsi ad HCRII, mentre con il nostro studio si dimostra che, usando gli stessi vettori e lo stesso sistema cellulare, HMGY agisce da repressore. L'attività repressiva delle HMGI(Y) nella trascrizione è già stata descritta per alcuni promotori come ad esempio il promotore dell'Ig  $\epsilon$  "germ line" o dell'IL-4 (vedi sezione HMGI e controllo trascrizionale). Nel caso dell'IL-4 la repressione che avviene in linfociti T può essere spiegata con il fatto che le HMGI(Y) sono espresse a buoni livelli e quindi potrebbero spiazzare i fattori NF-AT per un effetto "quantitativo".

L'ipotesi conclusiva di questo lavoro è molto affascinante e darebbe una spiegazione meccanicistica al fatto che le proteine aventi un HD richiedono degli eventi regolativi per funzionare correttamente. Ad esempio la stessa proteina Ultrabithorax espressa in contesti spazio-temporali diversi esplica differenti effetti negli embrioni di *Drosophila*. L'abilità delle proteine HMGI di impedire a fattori contenenti un HD di interagire con alcune loro sequenze bersaglio potrebbe essere un modo col quale uno stesso fattore può indurre effetti differenti a seconda della situazione contingente.

Inoltre, le HMGI potrebbero fornire la base degli "switch" regolatori tipici delle proteine con l'*omeodominio*. In particolare, attraverso cicli di fosforilazione e defosforilazione delle HMGI da parte della p34<sup>cdc2</sup>, che fanno allentare o rafforzare rispettivamente l'interazione col DNA, (vedi sezione Modificazioni post-traduzionali dell'HMGI) si potrebbe controllare l'accesso delle proteine con un HD alle loro sequenze bersaglio.

#### 4.1.2 Il fattore architetturale HMGI-C aumenta l'attivazione trascrizionale mediata da NF- $\kappa$ B (96 e Allegato n°2)

Nonostante l'importanza che riveste il fattore HMGI-C nello sviluppo embrionale e nella trasformazione neoplastica (vedi sezioni 1.2.6 e 1.2.7) non si sono studiate a fondo le sue proprietà biochimiche, come ad esempio la specificità di legame al DNA e il suo potere come modulatore della trascrizione. L'omologia con HMGI(Y) ovviamente suggerisce un comportamento analogo come fattore architetturale, ma nonostante aumenti il numero di elementi regolativi dove si dimostra un ruolo delle HMGI(Y) come modulatori della trascrizione attraverso cooperazione con altri fattori di trascrizione (sezione 1.2.5.2), non ci sono dati sufficienti che dimostrino che anche HMGI-C è capace di funzionare allo stesso modo.

In questo lavoro si è voluto dare una prima delucidazione sulle capacità di HMGI-C di modulare la trascrizione. Il modello scelto per questo studio è dato dall'elemento PRDII del promotore dell'*INF $\beta$*  (Thanos 1992) dove per la proteina HMGY è già stato dimostrato che interagisce con esso assieme all'eterodimero p50/p65 di NF- $\kappa$ B attivando la trascrizione. Siamo così riusciti a dimostrare che HMGI-C è in grado di legare PRDII con affinità simile ad HMGY e che accresce le proprietà attivanti del fattore di trascrizione NF- $\kappa$ B. Tale facilitazione dipende dal legame di HMGI-C con il DNA.

Gli oligonucleotidi utilizzati per analizzare e confrontare l'affinità di HMGI-C ed HMGY con diversi elementi di DNA sono i seguenti:

PRDII: 5'-GGGAAATTCCGTGGGAAATTCCGAGCT-3'

mPRDII: 5'-GGGAGATTCCGTGGGAGATTCCGAGCT-3'

(sono sottolineate le mutazioni inattivanti);

Ig $\kappa$ B: 5'-GGGACTTTCCGTGGGACTTTCCGAGCT-3'

(elemento derivato dall'enhancer della catena leggera delle Ig)

Sono state prodotte le proteine murine HMGI-C ed HMGY ricombinanti (come descritto nella sezione precedente e nei materiali e metodi sezioni 3.1.2 e 3.3.2-3.3.5) per effettuare esperimenti di *bandshift* preliminari, con i quali si è dimostrato che HMGI-C forma dei complessi con l'elemento PRDII *in vitro* con intensità analoga a quelli formati da HMGY. Inoltre, sempre in analogia ad HMGY, non si lega né all'elemento PRDII, mutato appositamente nel tratto ricco in AT, né ad



I $\kappa$ B, che costituisce sito di legame canonico per NF- $\kappa$ B ma non per HMGI(Y). Tutto ciò sta a dimostrare che vi è un'interazione sequenza specifica con PRDII anche da parte di HMGI-C. Sulla base di questi dati si sono allestiti esperimenti di facilitazione mediante *bandshift*, incubando quantità decrescenti di NF- $\kappa$ B con la sonda PRDII in assenza ed in presenza di HMGI-C, per vedere che effetto avesse HMGI-C sul legame di NF- $\kappa$ B all'elemento PRDII. Si è potuto constatare che a basse concentrazioni di NF- $\kappa$ B la proteina HMGI-C stimola la formazione del complesso in *bandshift*, similmente a quanto osservato per la proteina HMGY. (48)

Era d'interesse vedere, se anche *in vivo* potesse esserci cooperatività tra HMGI-C ed NF- $\kappa$ B mediante esperimenti di trasfezione. A tal proposito è stato costruito un vettore *reporter* recante un enhancer artificiale dato da due elementi PRDII contigui clonato a monte del promotore della timidina chinasi del virus herpes simplex che governa l'espressione del gene marcatore luciferasi di *Photinus pyralis*. (Materiali e metodi sezioni 3.1.1, 3.1.5, 3.4.2 e 3.4.3) Questo *reporter*, denominato PRDII TKluc, è stato innanzitutto saggiato in cellule PC Cl3 (sezione 3.4), prive di proteine HMGI (137), in modo tale da ottenere dati di attivazione trascrizionale di base dovuti solo ad NF- $\kappa$ B. La co-trasfezione di quantità fisse del *reporter* con quantità crescenti dei vettori di espressione di p50 e p65 di NF- $\kappa$ B dimostrano un aumento lineare fino a 14 volte. Si sono scelte a questo punto quelle quantità di p50 e p65 che dessero un'aumento dell'attivazione trascrizionale di 5 volte per effettuare trasfezioni in presenza di HMGI-C e confrontarle con gli stessi esperimenti senza la proteina. Una media di tre serie di esperimenti ha asserito che in presenza di HMGI-C o HMGY l'attivazione trascrizionale di NF- $\kappa$ B può incrementare fino ad 11-12 volte rispetto all'attività basale, confermando dati analoghi ottenuti da altri gruppi con diversi costrutti o modelli cellulari. Inoltre né HMGI-C né HMGY sono capaci da soli di transattivare il reporter, ed un reporter di controllo senza l'enhancer artificiale di PRDII non veniva in nessun modo inficiato dalla presenza di qualsiasi combinazione di vettori, dimostrando che tutti gli effetti osservati sono specifici per l'elemento PRDII.

In cellule NIH-3T3, nelle quali vi è espressione endogena di tutte e tre le proteine HMGI (Lanahan e nostre osservazioni), la cotrasfezione di PRDII TKluc con NF- $\kappa$ B e i plasmidi esprimenti HMGI-C ed Y mostra che l'iperespressione artificiosa delle due proteine è in grado di accrescere la trascrizione del reporter a livelli simili a quelli misurati in cellule PC Cl3.

Risulta essenziale la presenza del tratto AT intero dell'elemento PRDII nel costrutto PRDII TKluc, dato che analoghi costrutti reporter costruiti con mPRDII o con l'elemento I $\kappa$ B non mostrano nessun effetto accresciuto dell'attività trascrizionale di base ottenuto con NF- $\kappa$ B in cellule PC Cl3.

Gli stessi risultati sono stati ottenuti anche nella linea cellulare FRTL5 C12 (Sezione 3.4), anch'essa caratterizzata dall'assenza di fattori HMGI (137).

Il fatto che HMGI-C agisca come HMGY, almeno per quel che riguarda il sistema analizzato in questo lavoro, è coerente con l'omologia della struttura primaria che queste proteine presentano. Apre però una grande questione ancora irrisolta sull'utilità di avere almeno due fattori, HMGI-C ed HMGI(Y), che funzionano allo stesso modo.

Spesso si affronta tale questione nel momento in cui vengono utilizzate metodologie antisense per dimostrare l'essenzialità di uno di questi fattori. Thanos e Maniatis, ad esempio, dimostrarono che la proteina HMGI(Y) è essenziale per l'attività promotoriale di PRDII in seguito ad infezione virale e dipendente da NF- $\kappa$ B, attraverso l'utilizzo di un costrutto antisense di HMGI(Y) in esperimenti di trasfezione transiente. (48) La trasfezione di un plasmide antisense in cellule Jurkat o HeLa non solo inibisce l'attivazione del gene a partire da PRDII, dimostrando la necessità di HMGI(Y), ma potrebbe anche far notare che HMGI-C non può sostituirla e non tanto in cellule Jurkat, dove HMGI-C non è espressa (vedi Tabella V.1, sezione 1.), bensì nelle HeLa, in cui la proteina è presente (142). Inoltre in esperimenti con l'antisense della HMGI-C abbiamo osservato l'inibizione non solo di HMGI-C, ma anche della sitesi delle HMGI(Y). (146) L'alta omologia di sequenza anche a livello di RNA potrebbe rendere il costrutto antisense un arma a doppio taglio, inquanto non si ibridizza solo all'RNA nascente per il quale è stato progettato, ma anche per il trascritto di HMGI(Y), o nel primo caso anche per HMGI-C.

Resta comunque vero che non vi può essere ridondanza funzionale tra HMGI-C ed HMGI(Y), come dimostra il topo *pygmy*, nel quale l'*Hmgi-c* è deleto, ma vi è espressione normale dell'*Hmgi(y)*. (Sezione 1.2.6) Si pensa che la diversità di questi fattori risieda in interazioni con differenti sottogruppi di fattori trascrizionali e di bersagli genici, infatti la proteina HMGI-C alterata è coinvolta nella crescita anormale di cellule adipose che sviluppano lipomi (Sezione 1.2.8). Infatti i tratti di proteina che non presentano alta omologia tra HMGI-C ed HMGI(Y) sono oggetto di studio proprio per trovare queste diversità funzionali.

Anche la struttura del promotore bersaglio potrebbe essere importante nella selezione di HMGI(Y) o HMGI-C come fattore architetturale, in quanto possiedono regioni di lunghezza diversa tra i vari AT-hooks con i quali potrebbero contattare preferenzialmente sequenze con determinate spaziature dei tratti ricchi in AT.

## 4.2 Studio della regolazione dell'espressione dei fattori HMGI

### 4.2.1 Utilizzo delle proteine HMGI come marcatori nella diagnosi tumorale (153 e Allegato n°1)

Come anticipato nella sezione 1.2.7 alti livelli di espressione delle proteine HMGI sono una caratteristica propria di molti tessuti neoplastici e/o di cellule indifferenziate. Ad esempio, una loro aumentata espressione è associata all'insorgenza di un fenotipo altamente maligno in cellule differenziate di tiroide di ratto, dopo che erano state trasformate con un oncogene come *v-mos*, *v-ras-Ha* e *v-ras-Ki*. Le stesse, trasfettate con un costrutto antisense per HMGI-C, hanno perso la loro aggressività, dimostrando un ruolo causale delle proteine HMGI nella trasformazione neoplastica.

Avendo anche dimostrato un'associazione tra un'elevata espressione di queste proteine con la progressione tumorale in cellule mammarie del topo e con la capacità di cellule di un carcinoma prostatico di ratto di dare metastasi si è pensato di considerare queste proteine come delle costanti del fenotipo neoplastico e quindi come potenziali marcatori per la diagnosi ospedaliera.

Come estensione dello studio dell'espressione di proteine HMGI in tessuti tumorali e quindi anche della relativa regolazione abbiamo intrapreso un'analisi approfondita dell'espressione delle proteine HMGI(Y) in lesioni squamose intraepiteliali ed invasive della cervice uterina.

Utilizzando un anticorpo policlonale di coniglio contro il peptide SSSKSSQPLASKQ specifico dell'N-terminale di HMGI(Y) (Sez. 3.3.6 e 3.3.7) abbiamo analizzato mediante immunistochemical sia linee cellulari di carcinomi della cervice sia numerosi strisci, nonché campioni biotici inclusi in paraffina provenienti da tessuti sia normali che tumorali, in collaborazione con l'Istituto di Anatomia ed Istologia Patologica dell'Università di Trieste. In parallelo è stato usato sugli stessi campioni un marcatore diagnostico già in uso nella routine ospedaliera, noto come Ki-67, per poter confrontare i risultati ottenuti con l'anticorpo anti-HMGI(Y).

Gli studi di immunistochemical sono stati confrontati altresì con le diagnosi morfologiche dei tessuti, per trovare una correlazione tra il grado di alterazione displastica ed i livelli di espressione delle HMGI(Y), com'è stato fatto anche per tumori del seno, della tiroide e del colon.

La prima problematica affrontata in questo studio è stata la messa a punto dell'utilizzo dell'anticorpo anti-HMGI(Y) sia su cellule provenienti da colture preparate appositamente su vetrino mediante *cytospin*, sia poi su strisci e su tessuti inclusi in paraffina. A questo scopo sono state coltivate tre linee umane di carcinoma della cervice (C4-I, ME-180 e CaSki), scelte in base alle

caratteristiche simili a quelle trovate correntemente in casi clinici. Infatti le C4-I, pur essendo cellule tumorali, mantengono a lungo termine in coltura proprietà della differenziazione della cellula squamosa, mentre le CaSki sono derivate da un carcinoma epidermoide metastatico della cervice e le ME-180 da un carcinoma squamoso altamente invasivo. L'espressione delle HMGI(Y) è stata saggiata in queste cellule sia mediante *Northern blot* (sezione 3.2.3) che *Western blot* (sezioni 3.3.6 e 3.3.7), confrontandola con quella di opportuni controlli sia positivi (Carcinoma polmonare di Lewis, LLC) (140) che negativi (FRTL5-CI2) (137) e di quantità (Istone H1) nei Western blot. Tutte e tre le linee sono risultate positive per l'espressione sia dell'mRNA che della proteina. Quindi le stesse tre linee di carcinoma, i controlli già usati negli esperimenti appena descritti e strisci di cervice sane sono stati analizzati con tecniche di immunistoichimica con il medesimo anticorpo usato nel Western Blot, per saggiare la bontà dello stesso anche su tessuto. L'anticorpo riconosce specificamente ed intensamente l'antigene con localizzazione nucleare in tutte le cellule contenenti HMGI(Y), mentre gli strisci normali e le cellule di controllo negativo sono assolutamente prive di colorazione, permettendo quindi di definire questo anticorpo uno strumento molto adatto anche per studi di immunistoichimica.

Sono in seguito stati analizzati analogamente tessuti recanti diversi gradi di lesione provenienti da campioni inclusi in paraffina. Dopo deparaffinizzazione le sezioni sono state "immuno-colorate" con l'anticorpo anti-HMGI(Y) e controcolorate con ematossilina. I risultati ottenuti sono i seguenti: Nei 25 casi considerati come neoplasie intraepiteliali CIN 1, o di basso grado, solo pochi presentano cellule positive per la colorazione con l'anticorpo, mentre su 10 casi considerati come lesioni moderate o CIN 2, già si notano più cellule positive con una distribuzione preferenziale negli strati basali dell'epitelio. La stragrande maggioranza dei 25 casi diagnosticati per morfologia come lesioni di alto grado CIN 3 mostra una reattività cellulare intensissima all'anticorpo con una distribuzione in tutto l'epitelio. Foci sparsi di carcinoma invasivo presentano la totalità delle cellule immunoreattive. Anche alcune cellule nello stroma sottostante l'epitelio risultano positive. Questo potrebbe essere dovuto all'infiltrazione di lifociti che, pur essendo normali, esprimono le HMGI(Y). (142)

L'anticorpo anti-Ki-67 è usato frequentemente per valutare l'attività proliferativa di una cellula. Benché non si conoscano le funzioni della proteina nucleare da esso riconosciuta è comunque uno strumento utile per decisioni di importanza chirurgica. Il confronto dei risultati ottenuti dalle analisi delle stesse sezioni con l'anticorpo Ki-67 è confortante, in quanto il riconoscimento dei due anticorpi è molto simile e aumenta il numero di cellule positive con l'aumentare della gravità della lesione. Una minima differenza comunque c'è, in quanto più cellule nello strato basale risultano positive all'anticorpo anti-HMGI(Y) che non all'anticorpo anti-Ki-67, e questo forse si spiega col

fatto che i due anticorpi riconoscono antigeni che molto probabilmente rappresentano due stati cellulari diversi, ma strettamente connessi come la dedifferenziazione e la proliferazione, entrambe spesso contemporaneamente presenti in cellule tumorali.

In conclusione possiamo asserire che la definizione morfologica dei tre gradi di malignità CIN 1-3 nei carcinomi della cervice concorda con il tasso di positività all'anticorpo policlonale anti-HMGI(Y) e che vi è un riconoscimento simile a quello osservato con l'anticorpo contro Ki-67, suggerendo l'utilizzo combinato di due anticorpi nell'oncologia clinica di lesioni intraepiteliali ed invasive.

Vale la pena fare un piccolo inciso sulla presenza del genoma del papillomavirus umano HPV18 in queste lesioni. Infatti carcinomi genitali come quello della cervice hanno tipicamente un'eziologia virale, come gli epatomi che sono strettamente associati al virus dell'epatite B e/o C. Alcuni tessuti analizzati durante questo studio in effetti mostravano una coilocitosi superficiale, segno dell'avvenuta infezione da HPV. Non si possono escludere ipotesi che coinvolgono questo virus come evento scatenante la produzione di HMGI(Y) per diversi motivi. Uno è dato dalla proteina E6 dell'HPV 16 che sembra stimolare la produzione di HMGI(Y) almeno in cellule 10T1/2 E6 (136 e vedi Tabella VI) ed un altro motivo è il ritrovamento recente che la regione MAR nel locus 12q-13-15 è sito preferenziale di integrazione del genoma del virus HPV18 (190). Analisi molecolari del sito d'integrazione hanno dimostrato che in effetti si trova a 150 Kbase di distanza dal gene *HMGI-C* accanto ad un'isola CpG e che la trascrizione dei geni dell'HPV e di *HMGI-C* vanno in direzioni opposte.

L'essenzialità dell'infezione da HPV18 o 16 per ottenere un fenotipo maligno delle lesioni della cervice è ancora un fatto ricco di controversie. È vero che gran parte delle metastasi del carcinoma della cervice mostrano la cosiddetta *viral-gene persistence* come risultato dell'integrazione virale, ma ad esempio non tutti i campioni di carcinoma della cervice da noi analizzati possedevano il genoma virale integrato. È probabile che con l'integrazione virale sito-specifica si possano attivare meccanismi caratteristici che portano ad un specifico tumore, ma sembra piuttosto che tali integrazioni siano uno di tanti passaggi richiesti per generare il cancro.(190)

Nello studio di tre linee di carcinoma della cervice che presentavano integrazione del genoma di HPV18 a livello del locus 12q13-15 sono stati isolati diversi cloni cDNA di *HMGI-C*, ma in nessun caso vi erano mutazioni all'interno del gene.(190) Questo ci fa pensare che non vi sia collegamento diretto tra HPV ed espressione delle HMGI(Y), anche se l'integrazione di HPV18 nella regione cromosomica accanto ad *HMGI-C* potrebbe modularne l'espressione.

## 4.2.2 Caratterizzazione del promotore di *Hmgi-c*

La proteina HMGI-C presenta un'espressione molto più esclusiva rispetto ad HMGI(Y), in quanto non è mai stata rilevata in nessun tessuto o cellula normale adulti analizzati fin'ora, ma anche a livello di cellule tumorigeniche la sua presenza è associata a fenotipi particolari, spesso molto maligni e sdifferenziati. (128, 184, 137, 156) Inoltre abbiamo recentemente dimostrato che l'inibizione della sintesi di HMGI-C, mediante metodologie antisenso, inibisce la trasformazione in senso neoplastico di cellule epiteliali di tiroide di ratto trasformate con virus recanti gli oncogeni *v-mos* e *v-ras-Ki*, (146) palesando un ruolo causale di HMGI-C nella cancerogenesi.

Dal momento che è stato dimostrato che l'aumento della proteina HMGI-C in seguito a trasformazione è dovuto ad attivazione trascrizionale abbiamo iniziato a studiare la regolazione della sua espressione, isolando il gene corrispondente, (37) per poi definire la regione al 5' del gene coinvolta nella sua regolazione. La rivelazione dei meccanismi che regolano l'espressione di questo gene potrebbe essere di notevole interesse nella comprensione dell'insorgenza del fenotipo tumorale.

Lo studio della regolazione dell'espressione di *Hmgi-c* qui descritto è stato interamente eseguito nel Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole.

La prima parte della ricerca è consistita nell'isolare la regione minima in grado di attivare la trascrizione di *Hmgi-c* e di identificare i fattori agenti in *trans* responsabili di tale attività sia in cellule normali che nel contesto della trasformazione neoplastica. Parte di questo lavoro è stato pubblicato (191 vedi Allegato n°4).

### 4.2.2.1 I fattori Sp1 e CTF/NF-1 sono coinvolti nell'espressione basale del promotore prossimale di *Hmgi-c* (191 e Allegato n°4)

Il primo passo obbligato di questo studio è stato l'identificazione del sito d'inizio della trascrizione del gene *Hmgi-c*. Mediante esperimenti di Northern blot, (sezione 3.2) su RNA poly (A)+ estratto dal Carcinoma polmonare di Lewis con diverse sonde contigue, sono stati identificati 3 trascritti principali del gene, uno di 4.7, uno di 4.2 ed uno di 3.9 Kbasi, di cui il più corto è il trascritto

predominante. Con esperimenti di *Primer Elongation* e di *RACE* al 5' (sezione 3.2.5) il sito di inizio più abbondante dei trascritti di 3.9 Kbasi è stato mappato a livello di una Guanina a -201 pb dalla tripletta di inizio traduzione (vedi Figura 5).

Sono stati individuati altri cloni, derivanti dal *RACE*, meno abbondanti, che terminavano a livello di alcune pirimidine, posizionate a pochi nucleotidi di distanza dalla G -201 (vedi Figura 5, punti neri). La maggior parte di questi cloni recavano una extra-G (stabilito dal confronto con la sequenza genomica) all'estremità 5', derivante molto probabilmente da un'aggiunta della trascrittasi inversa eseguita al termine della polimerizzazione, una volta che lo stampo di RNA non c'è più, oppure sono stati descritti anche fenomeni di trascrizione inversa del cappuccio al 5' del messaggero, notoriamente una Guanina. (192) Comunque l'identificazione di queste extra-G dimostra chiaramente che tutti questi cloni sono copie complete dei trascritti di *Hmgi-c* e non derivanti da un distacco prematuro della trascrittasi inversa dallo stampo di RNA.

La sequenza genomica attorno al sito di inizio della trascrizione più abbondante è stata confrontata con due banche dati, denominate TFD e MatInspector, per cercare elementi regolatori putativi. La sequenza a monte della G +1 non possiede né un elemento consenso TATA né un elemento CCAAT, un aspetto tipico dei geni costitutivi, detti anche "house keeping". Questo dato concorda con il ritrovamento di siti multipli della trascrizione anziché uno solo come nella maggior parte dei geni che contengono una TATA box. Anche il gene che codifica per le altre due proteine HMGI non possiede tali elementi in nessuno dei suoi promotori (Friedman NAR). L'analisi dei profili CpG della regione 5' di *Hmgi-c* indica la presenza di una "CpG island" che si estende da -1300 a +500 (rispetto la G +1), ossia dalla regione 5' oltre il primo esone fino al primo introne (90 e sezione 1.2.5.1). Una CpG island di simile grandezza e posizione è presente anche nel gene *HMGI-C* umano. Molti potenziali siti di legame per fattori di trascrizione sono stati trovati e sono sottolineati nella Figura 5. Effettuando la stessa ricerca sul gene umano abbiamo inoltre identificato alcuni elementi che sono conservati tra uomo e topo e per questo sono stati considerati più verosimili. Tra questi vi sono: i fattori "zinc-finger" Sp1, MZF1 ed Ik-2, i fattori "ets-like" Elk-1 e c-Ets-1, ed i fattori Yi, Delta EF1, CTCF, AP2, AP4 e CTF/NF-1.

Di particolare importanza risulta essere un tratto polipirimidinico/polipurinico (ppyr/ppur) di 60 pb (da -83 a -25) nel topo e di 39 pb nell'uomo, costituito quasi interamente di una ripetizione del trinucleotide TCC. Simili sequenze con una composizione asimmetrica di filamento sono state trovate anche in altri geni come l'EGF-R, c-myc, I-R, c-Ki-ras, c-src, TGF  $\beta$ 3 e la catena A del PDGF, ed è stato dimostrato che esse sono essenziali per l'attività genica, che interagiscono con fattori nucleari capaci di regolare la trascrizione e di essere sensibili al trattamento con la nucleasi

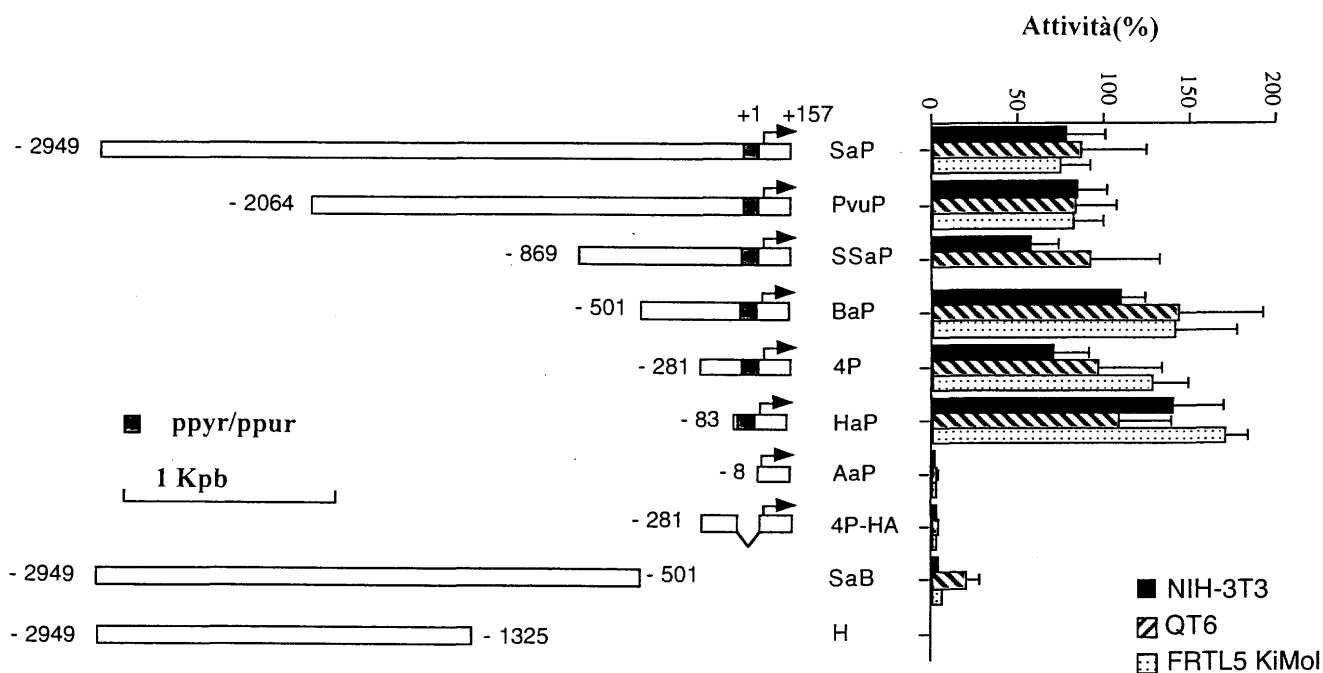




Mentre il gene umano *HMGI-C* (38) presenta molte omologie con quello murino qui studiato ed è molto conservato a livello di sequenza, quello di *HMGI(Y)* risulta avere un arrangiamento complesso di più promotori su esoni distinti, attraverso i quali, a seconda di quale promotore è attivato, si formeranno anche mRNA diversificati nella regione a monte della tripletta che codifica per l'amminoacido iniziale della proteina. (131) In particolare vi sono 4 siti distinti di inizio della trascrizione funzionanti, di cui ad esempio solo uno viene indotto dal TPA e infatti solo a livello del rispettivo promotore vi è un sito di riconoscimento per il fattore AP-1. (Figura 1 sezione 1.2, gli esoni I, II, II' e III e V possiedono i siti di inizio 1-4.) Il 5' di *Hmgi-c* non presenta *splicing* alternativo, come stabilito da esperimenti di RT-PCR (sezione 3.2) su tutta la lunghezza del 5' non tradotto (1351 pb) (37), ma i dati di Northern blot dimostrano la presenza di più trascritti, per cui non escludiamo anche l'esistenza di più promotori indipendenti.

Per caratterizzare l'attività trascrizionale del promotore *Hmgi-c* un frammento di 3,2 Kb della regione 5' è stata clonata nel vettore reporter (pGL2 Basic) a monte del gene della luciferasi di lucciola. Da questo costrutto sono stati ottenuti degli altri mediante delezioni progressive sia dal 5' sia al 3' sfruttando diversi siti di restrizione (vedi sezione 3.1.3 e Allegato n°4). Questi sono stati trasfettati in cellule NIH 3T3 (sezione 3.4.3), in modo da saggiare *in vivo* la loro attività trascrizionale e per definire la regione minima, al 5' del gene, capace di promuovere la trascrizione. La scelta della linea cellulare è motivata dal fatto che le cellule embrionali di fibroblasti di topo esprimono ad alti livelli la proteina HMGI-C, quindi sono dotate di un corredo completo di fattori nucleari endogeni in grado di stimolare la trascrizione anche del promotore *Hmgi-c* trasfettato.

Gli stessi costrutti sono stati saggiati in un linea cellulare di una specie diversa, cioè cellule QT6 (fibrosarcoma di quaglia), e FRTL5 KiMol (cellule epiteliali di tiroide di ratto trasformate con *v-Ki-ras*) sempre esprimenti l'*Hmgi-c*, ed essenzialmente è stato ottenuto lo stesso risultato, in quanto il massimo di attività trascrizionale, comparabile a quella di un promotore virale come quello dell'SV40, è contenuto nel frammento che va da -80 a + 157. E' interessante notare che le 80 basi al 5' della G+1 contengono l'elemento ppyr/ppur, facendo dedurre che anche nel gene *Hmgi-c*, come in altri geni correlati alla crescita, questa sequenza è importante ai fini dell'attività trascrizionale. Inoltre sono stati effettuati saggi di transfezione stabile con alcuni costrutti chiave in cellule NIH-3T3, per verificare che l'integrazione cromosomica non avesse qualche influenza sulla regolazione dell'espressione. I dati ottenuti in precedenza sono stati nuovamente confermati.



**Figura 6**

**Saggi di attività del promotore di *Hmgi-c* *in vivo*.**

Diversi costrutti di delezione del 5' di *Hmgi-c* sono stati saggiati per l'attività trascrizionale conferita ad un gene marcatore per la luciferasi. Le attività misurate sono state messe a confronto con quella di un promotore noto, dell'SV40, messo arbitrariamente pari al 100%. Le attività sono state normalizzate trasfettando uguali quantità del vettore SV- $\beta$ gal. (vedi sezione 3.4.3)

Questa sequenza, da noi definita come promotore minimo è stata estesamente studiata come segue.

Per identificare fattori coinvolti nella regolazione del gene *Hmgi-c*, quattro oligonucleotidi (A, B, C e D) (Fig. 5) a doppio filamento, che coprono tutta la regione del promotore minimo (-83 fino a +2) e che si sovrappongono parzialmente, sono stati utilizzati in esperimenti di ritardo di banda (EMSA) (sezione 3.3.6) con estratti proteici nucleari provenienti da diverse linee cellulari. Gli oligonucleotidi sono stati progettati di modo che A, B e C comprendano tutta la regione delle ripetizioni TCC, mentre oligo D copre la sequenza rimanente del promotore minimo fino al sito di inizio della trascrizione.

I dati ottenuti mediante bandshift usando estratto da diversi tipi cellulari (NIH-3T3 e FRTL5 KiMol) con le sonde A, B o C mostrano la formazione di tre complessi maggiori per tutte e tre le sonde con simili mobilità che vengono competuti dall'aggiunta di ognuno dei tre oligo freddi in eccesso, ma anche da un oligonucleotide con il consensus per il fattore Sp1, indicando che tali complessi contengono fattori Sp1-simili. Questo dato è in accordo con l'analisi computerizzata che ha rivelato un sito putativo di legame per Sp1 in posizione -42 a -36, e anche con altre documentazioni che riportano che simili lunghi tratti ppyr/ppur sono siti multipli di legame per il fattore di trascrizione Sp1. L'oligonucleotide D specifico di *Hmgi-c*, ma scorrelato dal punto di vista della sequenza, non compete per la formazione di nessun complesso.

Altre bande, di cui si cerca di identificare i componenti, sono formate con queste tre sonde. In particolare ve ne sono due, la cui intensità aumenta in seguito a competizione con l'oligo Sp1, suggerendo che sono composte da fattori nucleari che legano la stessa sequenza degli elementi Sp1-simili. Infatti vi sono sequenze consenso per tanti altri fattori *zinc-finger* oppure *Ets-like* in tutto il tratto ppyr/ppur.

L'identità dei fattori coinvolti nella formazione dei tre complessi maggiori è stata ulteriormente ricercata, usando degli anticorpi specifici in esperimenti di "super-shift" con la sonda C (sezione 3.X). Fin'ora sono stati identificati quattro membri appartenenti alla famiglia Sp1, denominati Sp1-Sp4; abbiamo centrato la nostra attenzione soltanto su Sp1 ed Sp3, perchè Sp2 ha una specificità di legame diversa per il DNA, mentre Sp4 è un fattore la cui espressione sembra essere ristretta solo al cervello. Incubando con anticorpi specifici per Sp1 si è ottenuto un "supershift" del complesso più alto, mentre con l'anticorpo anti-Sp3 si sono ulteriormente appesantiti i due complessi rimanenti. Un'anticorpo specifico contro il fattore Oct2 è stato usato come controllo negativo e infatti non intacca la formazione di nessun complesso. Gli stessi risultati sono stati ottenuti anche usando le sonde A e B. Questi dati rivelano che i due fattori di trascrizione, Sp1 ed Sp3, o proteine immunologicamente relate, sono capaci di legarsi probabilmente alcune volte al promotore minimo di *Hmgi-c* nel tratto ppyr/ppur.

La prova finale del fatto che il tratto ppyr/ppur è veramente sito multiplo per fattori della famiglia Sp1 è stata ottenuta mediante un esperimento di footprinting dove il tratto ppyr/ppur è totalmente coperto dalla proteina Sp1.

Fattori che possano legare a valle del tratto ppyr/ppur sono stati ricercati con l'uso dell'oligo D. Incubando con estratto nucleare proveniente sia da cellule NIH-3T3 che da FRTL5 KiMol si può osservare la formazione di una banda prominente, la quale viene efficacemente competuta dall'aggiunta dello stesso oligo D non marcato, ma non dall'oligo C e neppure da un oligonucleotide non relato, contenente il sito di legame per il fattore AP1, dimostrando la specificità

del complesso. E' stato quindi usato l'oligonucleotide NF1, che possiede un sito di legame ad alta affinità per il fattore di trascrizione CTF/NF-1 derivato dall'origine di replicazione di adenovirus, come competitore specifico, dato che attraverso l'analisi computerizzata è stato trovato metà sito palindromico per questo fattore nel promotore dell'*Hmgi-c* in posizione -28 fino -10. NF1 è risultato un competitore molto forte, in quanto già ad un eccesso di concentrazione molare di dieci volte compete con l'oligo D per la formazione del complesso in tutti gli estratti saggiati. L'oligo HPV, derivato dalla regione regolativa URR del papillomavirus umano 16, contenente metà sito di legame palindromico per CTF/NF-1, è capace anch'esso di competere con l'oligo D, anche se non così efficacemente come NF1. Al contrario il complesso non viene competuto dall'oligo mD, derivato dall'oligo D, nel quale era stata introdotta una singola mutazione capace di abolire il legame per CTF/NF-1. Inoltre è stata confermata l'inabilità di mD di formare il complesso usandolo come sonda.

Dato che anticorpi capaci di riconoscere CTF/NF-1 non erano disponibili abbiamo usato la proteina CTF/NF-1 purificata da cellule K-562 in esperimenti di *ritardo di banda*. La proteina CTF/NF-1 purificata ha prodotto un complesso ritardato assieme all'oligo D che assomiglia molto a quello ottenuto con l'intero estratto nucleare, che è stato efficientemente competuto dagli oligonucleotidi freddi D, NF1 ed HPV, ma non dall'oligo mD o AP1. La specificità di tale complesso è stata ulteriormente dimostrata usando l'oligo mD come sonda. Questo risultato dimostra quindi che oltre ad Sp1 ed Sp3 anche CTF/NF-1 è capace di legarsi al promotore di *Hmgi-c in vitro*.

Usando la proteina CTF/NF-1 purificata in esperimenti di footprinting sul tratto contenente il promotore minimo, si è dimostrato ulteriormente l'interazione tra questo fattore di trascrizione con il promotore di *Hmgi-c* a -28/-10 basi dal sito di inizio della trascrizione.

Per stabilire l'importanza del fattore CTF/NF-1 nell'attività del promotore abbiamo effettuato degli esperimenti di transfezione transiente in cellule NIH-3T3 con un costrutto derivato da HaP (che contiene il promotore minimo a monte della luciferasi), denominato NF1mut, nel quale era stato distrutto il sito di legame per CTF/NF-1, introducendo la stessa mutazione dell'oligo mD mediante mutagenesi sito-specifica. (vedi sezione 3.X) L'attività di NF1mut è stata quindi confrontata con quella di HaP, arbitrariamente posta come 100%. Nove esperimenti indipendenti, usando due cloni diversi di NF1mut, mostrano che CTF/NF-1 contribuisce all'attivazione trascrizionale, in quanto la mutazione, che ha mostrato di abolire il legame di CTF/NF-1 *in vitro*, è capace di ridurre l'attività del gene reporter del 35% *in vivo*.

Per analizzare l'effetto dei fattori Sp1 ed Sp3 sul promotore *Hmgi-c* sono stati allestiti esperimenti nel sistema cellulare delle cellule SL2 di *Drosophila*, notoriamente privi di questi fattori di trascrizione. Cotrasfezioni con il costrutto HaP, che possiede il promotore minimo di *Hmgi-c*, hanno

dimostrato un effetto attivatore da parte di Sp1 dose-dipendente, mentre Sp3 non attiva, ma al massimo coopera all'attività trascrizionale di Sp1.

Il gene *Hmgi-c* si colloca nella classe emergente di geni cosiddetti *TATA-less*; generalmente sono ricchi in GC, hanno multipli siti di inizio per la trascrizione ed usano sequenze iniziatore per indicare dove il complesso di inizio dovrà legarsi per dare il via alla trascrizione. L'esistenza di sequenze ppyr/ppur in posizione vicina a promotori, come abbiamo trovato nel gene *Hmgi-c*, è già stata riportata per tutta una serie di altri geni di mammifero, generalmente correlati con la crescita. Il ruolo preciso di elementi ppyr/ppur nella regolazione di questi promotori non è ancora molto chiara, ma dato che costituiscono siti multipli di legame per il fattore di trascrizione Sp1, il quale a sua volta sembra essere capace di reclutare e stabilizzare il complesso del fattore di trascrizione II D (TFIID), potrebbero avere un ruolo nell'iniziare la trascrizione. Abbiamo dimostrato, mediante esperimenti di competizione crociata e anticorpi specifici e con esperimenti di footprinting, che anche l'elemento ppyr/ppur del promotore *Hmgi-c* è sito multiplo di legame per Sp1 ed Sp3. Dato che la maggior funzione di Sp1 sembra essere l'attivazione della trascrizione, potrebbe quindi servire a reclutare i fattori basali della trascrizione e conferire così attività basale al promotore del gene *Hmgi-c* in analogia a quanto è stato riportato per altri geni correlati con la crescita come EGF-R, c-Ki-ras e c-myc. La funzionalità di questo elemento è inoltre confermata dai nostri esperimenti di trasfezione (vedi sezione 4.2.2.4) che mostrano come piccole delezioni successive del tratto ppyr/ppur riducono progressivamente l'attività del promotore, confermando quindi l'esistenza di tratti regolatori multipli.

Il sito putativo per CTF/NF-1 in *Hmgi-c* è conservato anche nel 5' del gene umano per HMGI-C. Si tratta di un fattore di trascrizione originariamente coinvolto nella replicazione del DNA dell'adenovirus in cellule HeLa, ma è in grado di regolare la trascrizione anche di molti altri geni sia virali che cellulari. Esso può attivare la trascrizione da solo, ma spesso coopera con altri fattori che si legano a sequenze vicine o compete con altri fattori per lo stesso sito. Dato che l'attivazione di *Hmgi-c* avviene in seguito a stimoli di diversa natura, sarà interessante indagare se ciò avviene tramite l'azione di CTF/NF-1, per il quale è stata dimostrata una stretta associazione con l'espressione adipocitaria e con la trasformazione mediante oncogeni nucleari.

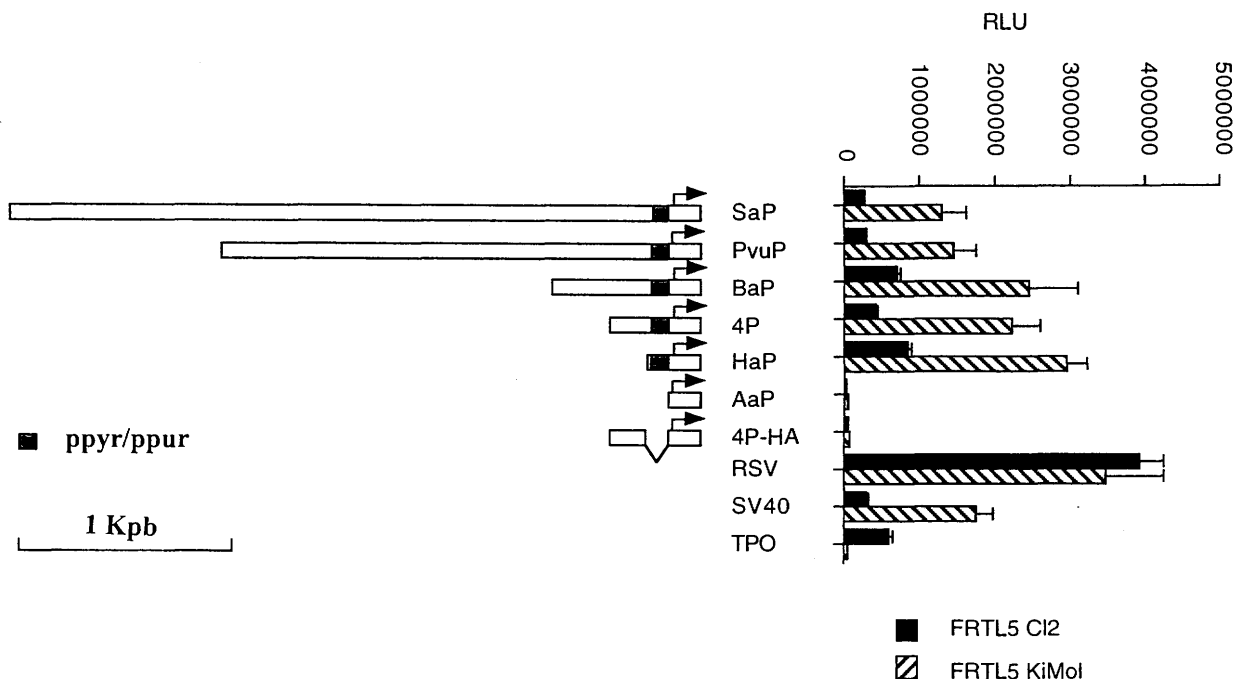
Altri siti putativi sono stati trovati nella sequenza del promotore minimo che potrebbero spiegare l'attivazione trascrizionale di *Hmgi-c* nella trasformazione neoplastica: Elk-1 sembra essere un buon candidato, dato che è un attivatore che agisce a valle di raf-1 che attiva la sintesi di HMGI-C, oppure PEA3, un elemento induttore comune di molti geni *TATA-less*, che svolge la sua funzione durante l'embriogenesi, angiogenesi e la metastatizzazione.

#### 4.2.2.2 Il promotore Hmgi-c è più attivo in cellule trasformate che nella loro controparte normale.

In contemporanea al lavoro descritto nella sezione 4.2.2.1 si è approfondito lo studio del promotore per identificare regioni coinvolte nell'attivazione della trascrizione del gene *Hmgi-c* nel contesto della trasformazione neoplastica. Per gli esperimenti di transfezione è stato utilizzato un modello cellulare di confronto, costituito da due linee ben caratterizzate (188, sezione 3.4):

- FRTL5 Cl2, linea differenziata di cellule epiteliali della tiroide di ratto ed
- FRTL5 KiMol, la sua controparte trasformata nella quale l'HMGI-C è abbondantemente espressa in seguito alla trasformazione indotta dall'oncogene *v-ras*-Ki.

L'attività più alta in cellule trasformate è nuovamente correlata con la presenza dell'elemento ppyr/ppur, in quanto costrutti contenenti questo elemento fanno aumentare l'attività del promotore *Hmgi-c* in esperimenti di trasfezione fino a 5 volte rispetto alle cellule normali. (Figura 7)



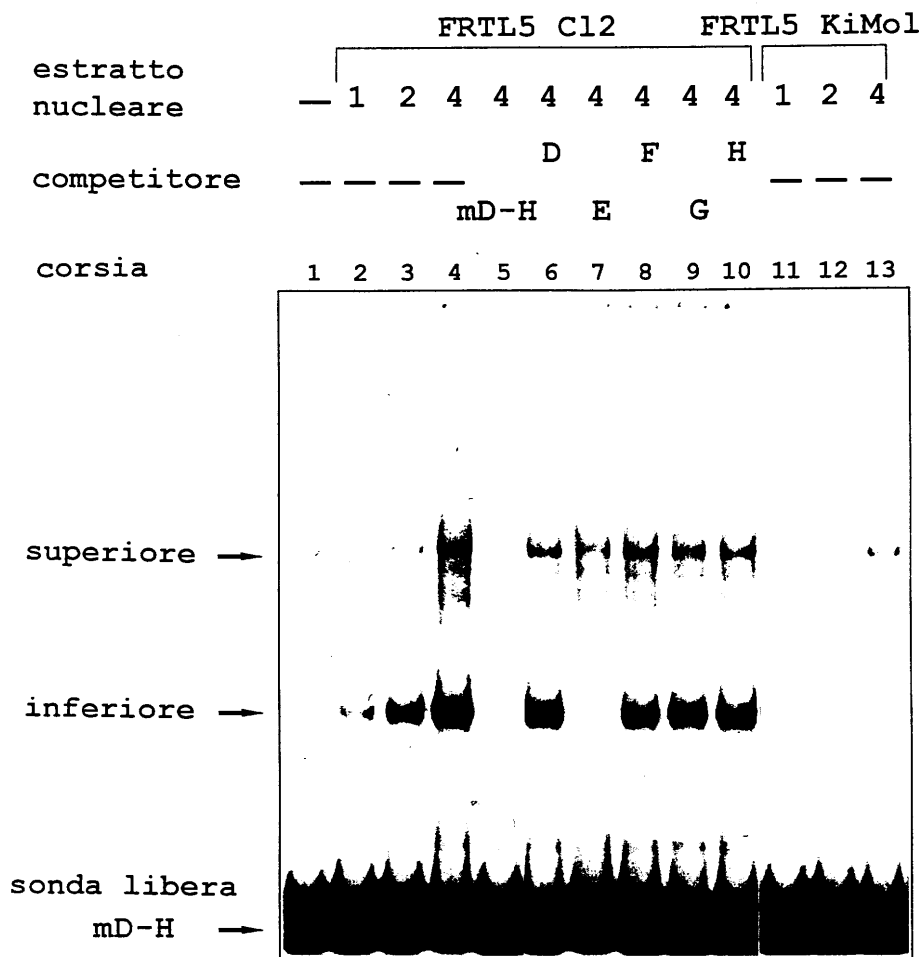
**Figura 7**

**Attività differenziale del promotore Hmgi-c in cellule trasformate e normali.**

L'attività trascrizionale di *Hmgi-c* è stata confrontata con quella di diversi promotori come controlli, come ad esempio l'RSV e quello della tiroperossidasi (TPO), la cui attività è stata ampiamente studiata in queste linee cellulari. (194) Come mostrato nella Figura 7, mentre il promotore RSV è attivo in entrambe le linee allo stesso livello, l'attività dei costrutti di *Hmgi-c* che contengono la ripetizione TCC, è più alta di quattro fino a sei volte nelle cellule trasformate che nelle normali. È interessante notare che il costrutto più corto, HaP, corrispondente al promotore minimo, mantiene attività differenziale. Un andamento analogo è dato dal promotore SV40, il quale contiene siti di legame per AP-1 che si sa essere più attivo nelle cellule trasformate da *v-Ki-ras*. (194) Il promotore TPO al contrario, è 16 volte più attivo nelle cellule FRTL5 Cl2 differenziate che nelle trasformate. (194)

La regione minima responsabile dell'attività differenziale di *Hmgi-c* è stata esaminata per la presenza differenziale di complessi nucleoproteici in *bandshift* (sezione 4.2.2.2), con estratti provenienti dalla linea FRTL5 Cl2 differenziata e la controparte trasformata FRTL5 KiMol. Non avendo trovato alcuna differenza tra le cellule KiMol e FRTL5 Cl2 per quel che riguarda i *bandshift* effettuati con i frammenti di DNA a monte del sito di inizio della trascrizione (oligonucleotidi A, B, C e D, Fig. 5) si è deciso di esaminare anche la parte attorno ad esso ed il DNA a valle, sempre mediante l'analisi preliminare del *bandshift* (sezione 3.3.6). Infatti, se si ipotizza un meccanismo di induzione della trascrizione di *Hmgi-c* in cellule trasformate, determinato da un aumento della quantità o attività o dalla comparsa di nuovi effettori, si dovrebbero individuare complessi con intensità diversa o complessi differenti in cellule trasformate rispetto a quelle normali. Non avendo riscontrato una situazione del genere, abbiamo rivolto la nostra attenzione ad altri meccanismi di controllo trascrizionale come ad esempio la repressione, che può coinvolgere segmenti di DNA a valle del sito di inizio della trascrizione lungo tutto il gene. È stata dunque prodotta mediante PCR (sezione 3.1.3) un frammento di DNA corrispondente alla regione di *Hmgi-c* da -32 a +181 (vedi Figura 4, oligonucleotidi D senso e H antisenso), marcato e in seguito sottoposto a *bandshift* incubandolo con estratti proteici nucleari provenienti da cellule FRTL5 Cl2 e FRTL5 KiMol.

Dato che il frammento iniziava con il nucleotide -32 è stata utilizzata una piccola strategia per far sì che non si leghi CTF/NF-1, perché il suo complesso abbondante avrebbe reso invisibili eventuali bande deboli provocate dalla presenza di fattori meno abbondanti: infatti per l'amplificazione è stato usato l'oligonucleotide mD (sezione 3.3.6) al 5' anziché l'oligonucleotide D. Il confronto degli esperimenti effettuati su cellule normali e cellule trasformate ci ha rivelato la presenza di un complesso proteico almeno quattro volte più intenso con estratti provenienti da cellule normali rispetto alle trasformate, legato alla sonda mD-H (Figura 8 complesso *basso*). Per restringere al minimo il sito di legame per questo fattore/complesso abbiamo effettuato un *bandshift* di competizione utilizzando come competitori gli oligonucleotidi D, E, F, G ed H che si sovrappongono tra loro e ricoprono tutto il frammento di PCR mD-H. (Figura 8)



**Figura 8**

**Attività di legame differenziale del promotore minimo *Hmgi-c***

Sono indicati i  $\mu\text{g}$  di estratto nucleare proteico in alto nella figura. I competitori sono stati usati in eccesso molare di 200 volte. Il frammento mD-H non marcato compete per tutti i complessi, indicando specificità di legame.

Abbiamo così potuto stabilire che l'oligonucleotide E è in grado di competere per la formazione di questo complesso ad un eccesso molare di 200 volte rispetto alla sonda marcata mD-H (Fig. 8 corsia 7), ma anche gli oligonucleotidi D ed F lo sono, anche se con meno efficacia (Fig. 8 corsie 6 e 8 rispettivamente), dimostrando che tale complesso risulta interagire col promotore di *Hmgi-c* nella regione tra -32 e +12. Un esperimento di *bandshift* effettuato utilizzando l'oligonucleotide E come sonda ha infine dimostrato che il sito di legame di questo ancora ignoto fattore/complesso è centrato proprio sulla sequenza dell'oligonucleotide E. Nello stesso tempo il fatto che, sia l'oligonucleotide D che F competino un minimo per la formazione di questo complesso, fa pensare che il contatto tra



il fattore/complesso ed il DNA avvenga su tutta la lunghezza dell'oligonucleotide E oppure necessiti anche di sequenze fiancheggianti.

Questo complesso risulta più intenso nei bandshift effettuati con estratti provenienti da altri modelli di confronto di cellule normali e delle stesse trasformate in diversi modi (Figura 9 e sezione 3.6). Tutti gli estratti proteici usati per confrontare cellule normali con la loro controparte trasformata, sono stati normalizzati per la presenza di attività di legame del fattore Sp1, che non risulta alterata dalla trasformazione.

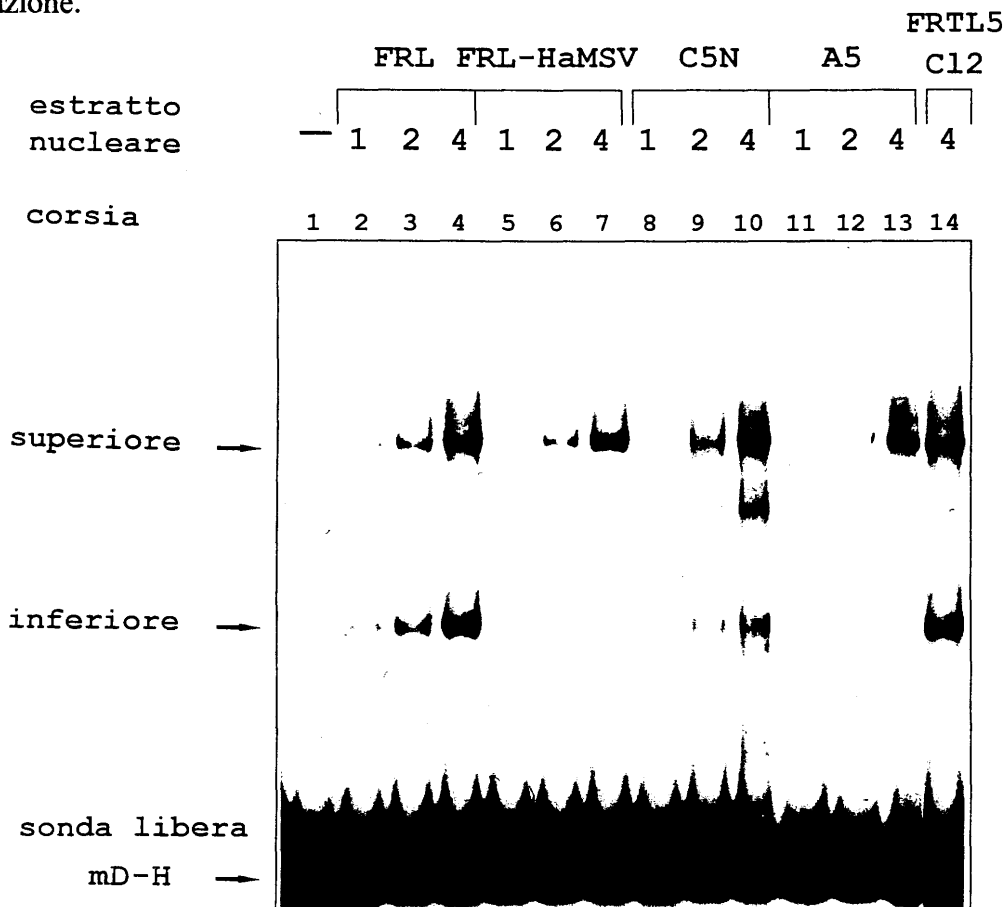


Figura 9

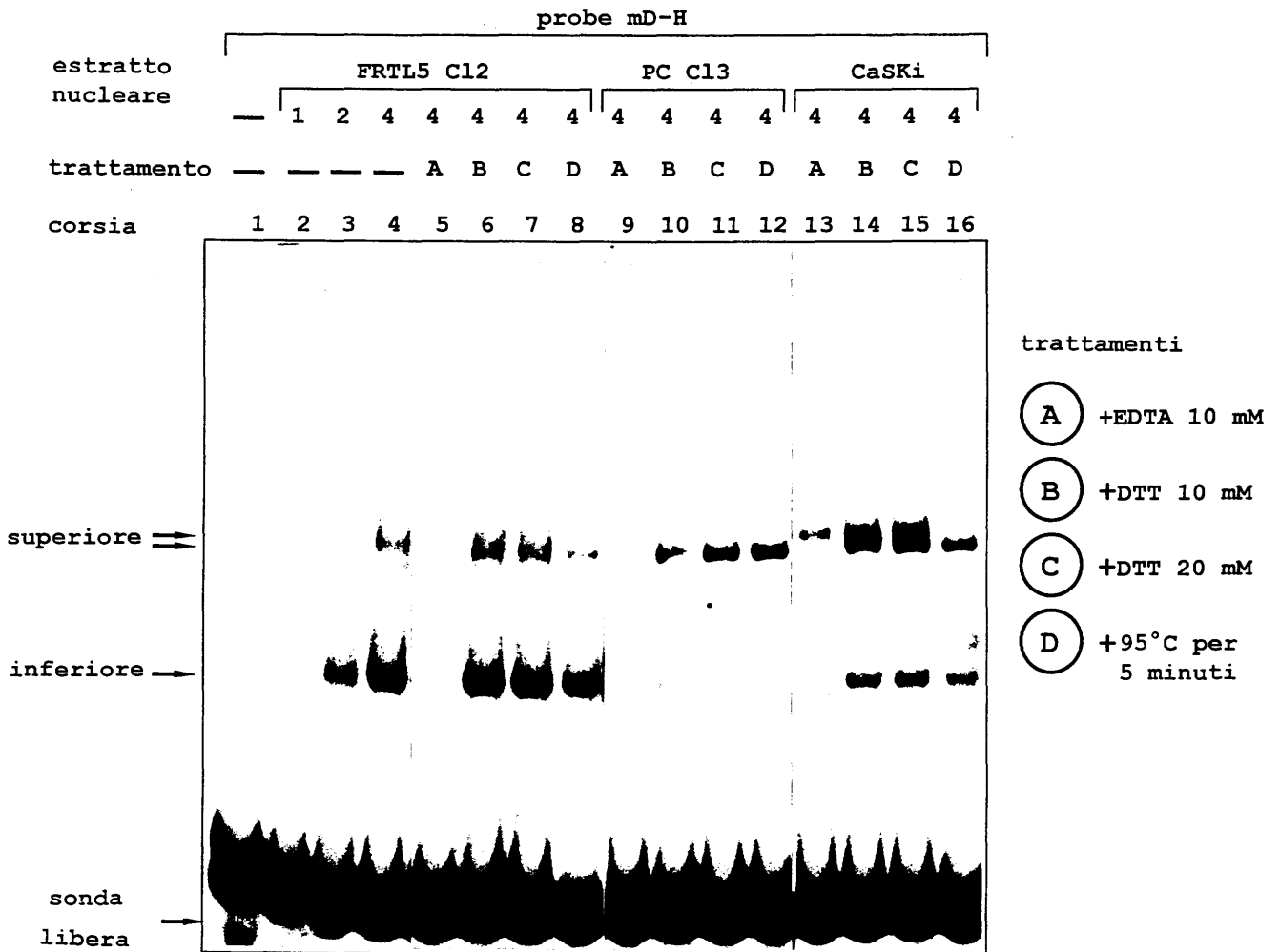
#### Attività differenziale del promotore *Hmgi-c* in diversi modelli sperimentali.

Le linee qui utilizzate (sezione 3.6) sono state analizzate per la presenza della proteina HMGI-C prima di effettuare gli esperimenti di *bandshift*. Infatti tutte e due le linee normali (FRL, C5N) non esprimono la proteina, mentre le trasformate sì (FRL-ras-V, A5). Come controllo positivo è stato caricato nello stesso gel anche l'esperimento effettuato con l'estratto delle FRTL5 KiMol.

Coerentemente all'ipotesi di un meccanismo di repressione in cellule normali, il complesso *basso* si forma a livello del sito di inizio della trascrizione, suggerendo quindi fortemente che l'attivazione di *Hmgi-c* in più sistemi cellulari derivi piuttosto da un meccanismo di derepressione che da attivazione trascrizionale.

Con alcuni estratti si formano anche un o due complessi più alti (*alto*) che sembrano in qualche modo associati alla presenza di quello *basso*. Non vediamo però alcuna regolarità nell'apparsa degli stessi, per cui saranno necessari altri esperimenti per chiarire la presenza di questi complessi.

Studi delle proprietà biochimiche dei fattori dei complessi *alto* e *basso* dimostrano che sono resistenti alla degradazione termica. Inoltre alte concentrazioni di ditiotreitolo non intaccano in nessun modo la formazione di entrambi i complessi, mentre l'aggiunta del chelante di ioni bivalenti EDTA impedisce la formazione di quest'ultimi. (Figura 10)



**Figura 10**  
Caratteristiche biochimiche dei complessi *basso* e *alto*.

Allo scopo di determinare l'identità del/dei complessi formati sono stati fatti esperimenti di competizione utilizzando oligonucleotidi contenenti il consensus per diversi fattori trascrizionali. L'analisi della sequenza di questo tratto di promotore aveva infatti indicato siti potenziali per AP4, Elk-1 e c-ets-1 ed E2F. Tutti questi fattori rivestono una qualche importanza per quel che riguarda

l'attività trascrizionale in sistemi cellulari trasformati, specialmente Elk-1, che sembra essere un effettore a valle della via di trasduzione del segnale attivata da Raf-1, il quale è stato dimostrato indurre la sintesi di HMGI-C (134). Anche alcuni membri della famiglia E2F presentano alcune caratteristiche identificate in questo studio, quali ad esempio un legame bipartito con le sequenze fiancheggianti l'oligonucleotide E e un'attività da repressore. Tutte queste osservazioni sono sotto attiva indagine.

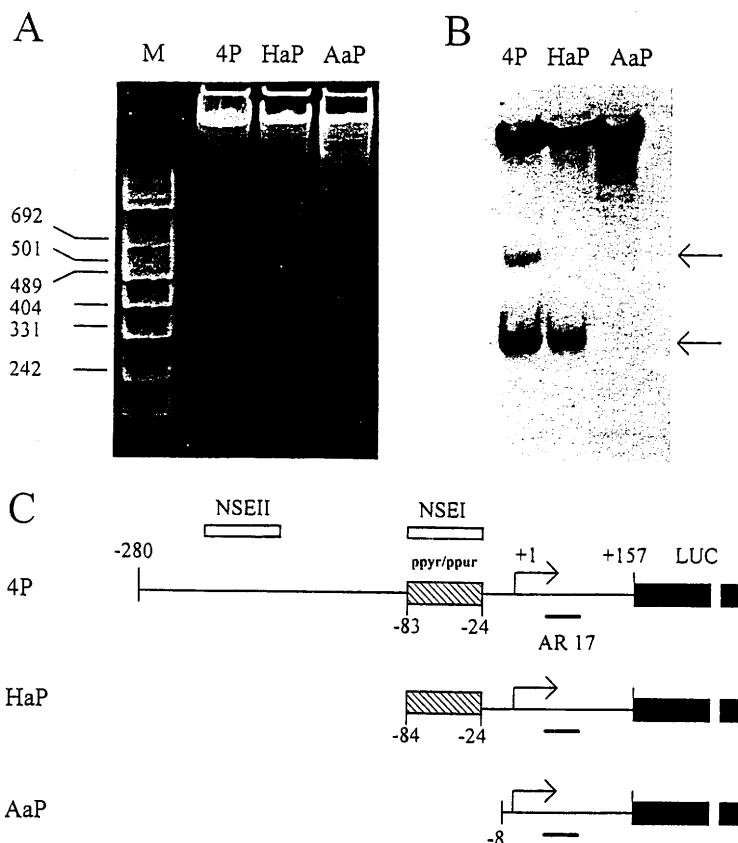
#### **4.2.2.3 Il tratto omopirimidinico/omopurinico del promotore minimo di *Hmgi-c* è una caratteristica comune di molti geni correlati con la crescita**

Il controllo dell'espressione genica è mediata in gran parte attraverso la regolazione dell'inizio della trascrizione. Mentre ormai è un dato di fatto che sia in procarioti che eucarioti le sequenze nucleotidiche del B-DNA costituiscono segnali di riconoscimento importanti per fattori di trascrizione e l'RNA polimerasi, si sono accumulate molte evidenze sul fatto che anche strutture non-B possono rappresentare foggie importanti di regioni di controllo trascrizionale. (192) Si è proposta l'esistenza di una varietà di diverse strutture di DNA non-B poste in regioni immediatamente a monte di sequenze codificanti. (193)

Dal momento che oltre alla forma B-DNA le sequenze ppyr/ppur hanno la potenzialità di adottare conformazioni poco usuali come cruciformi, triple eliche e anse, (194) si è iniziato uno studio molto dettagliato di questa regione così ricca in pirimidine, per evidenziare simili potenzialità. Spesso rappresentano infatti regioni funzionalmente importanti, com'è stato dimostrato per tutt'una serie di geni (EGF-R, catena A del PDGF, c-myc, A-R etc.), in quanto potrebbero fornire un ulteriore livello di controllo nella regolazione della loro espressione.

Il tratto ppyr/ppur di *Hmgi-c* ha una notevole omologia di sequenza con una regione presente nel promotore del gene che codifica per il recettore del fattore di crescita epidermico (EGF-R) (195) e degli androgeni (AR) (196) ma anche conservata in altri geni correlati alla crescita come il gene c-myc(197), il gene Ki-ras(198), il gene per il recettore dell'insulina (I-R)(199), il gene per la catena A del fattore di crescita derivato dalle piastrine (PDGF A chain) (200), il gene per il fattore di crescita trasformante beta (TGF  $\beta_3$ ) (201), e ancora altri.





**Figura 11**

**Analisi dei siti sensibili alla S1 nucleasi nel promotore *Hmgi-c*.**

Pannello (A) Gel d'agarosio dopo la migrazione dei plasmidi 4P, HaP ed AaP trattati con S1 nucleasi ed *Xba*I. Pannello (B) Southern blot del gel nel pannello A. I frammenti uscenti dalla digestione con S1 ed *Xba*I, riconosciuti dall'oligonucleotide AR17 come sonda, sono indicati a destra nel pannello B con 2 frecce. Pannello (C) Sono indicate schematicamente le regioni di *Hmgi-c* nei tre costrutti. Sono evidenziate con una barretta vuota i due elementi sensibili al trattamento con la S1 nucleasi, e con una barretta rigettata la regione ppyr/ppur, detta anche TCC.

Tre diversi costrutti plasmidici *Hmgi-c*/LUC nella forma superavvolta sono stati digeriti con la S1 nucleasi (sezione 3.1.3) e poi con l'enzima di restrizione *Xba*I che taglia una volta sola nel vettore. Si sono formati così alcuni frammenti (pannello A): Il costrutto 4P ha dato due frammenti di 280 e 520 pb rispettivamente, ed il costrutto HaP ne ha creato uno di 280 pb, mentre AaP è stato solo

linearizzato. La generazione della banda di 280 pb è coerente con il taglio della S1 nucleasi nella regione ppyr/ppur e indicata qui come regione NSE I. Una seconda regione sensibile è presente più a monte nel costrutto 4P, denominata NSE II. Non abbiamo focalizzato questo elemento, perchè mostra una sensibilità alla S1 nucleasi visibilmente più debole (Figura 11, pannello B, frammento alto nella corsia 4P). Nei pannelli B e C sono mostrate le ibridizzazioni di controllo del relativo Southern blot con due oligonucleotidi diversi AR4 ed AR17, schematizzati nel pannello D.

Per determinare in seguito i siti di taglio dell'elemento NSE I a livello nucleotidico abbiamo eseguito una cosiddetta mappatura ad alta risoluzione della regione ppyr/ppur. A questo scopo il plasmide superavvolto BaP (sezione 3.1.3), contenente 680 pb di 5' di *Hmgi-c* è stato parzialmente digerito con la S1 nucleasi per generare una popolazione di DNA tagliati solo su di un filamento ("nicked") ma non linearizzati. Poi la stessa preparazione è stata tagliata con *HindIII*, che taglia solo ad un'estremità del sito multiplo di clonaggio del vettore (pGL2 Basic-Promega). (sezione 3.1.3) I prodotti di digestione sono stati marcati o con la trascrittasi inversa sui filamenti non codificanti o con la PNK su quelli opposti e poi tagliati con *SpeI*, per ottenere frammenti di DNA *SpeI-HindIII* a doppio filamento ma con un *nick* in qualche posto (su uno o sull'altro frammento) tra le due estremità dello stesso. Sono poi stati risolti in un gel di poliacrilammide denaturante con gli opportuni controlli negativi e con sequenze Sanger come pesi molecolari. I risultati mostrano una collezione di tagli presenti su entrambi i filamenti, i quali, dopo un'attento calcolo dei loro pesi molecolari, sono stati schematizzati nella Figura 12.



Sono state infatti isolate proteine che legano il singolo filamento in modo specifico, stabilizzando la struttura non-B del DNA, a discapito del DNA a doppio filamento, prevenendo così il legame di fattori di trascrizione come Sp1 e interferendo in questo modo con l'inizio della trascrizione. (202, 203) Comunque il coinvolgimento di queste strutture nella regolazione dell'espressione genica rimane controversa, in quanto altri lavori dimostrano un effetto opposto, correlando la presenza di strutture non-B del DNA, evidenziata come sensibilità alla nucleasi S1, con attivazione trascrizionale, in quanto l'apertura del doppio filamento permette l'accesso facilitato all'RNA polimerasi II.

L'aspetto strutturale altamente dinamico del DNA assume sempre più importanza in questo gruppo di geni come nuovo modello della regolazione trascrizionale (194) in quanto contribuirebbe con locali decondensazioni della cromatina a permettere e facilitare l'entrata e l'assemblamento del complesso basale della trascrizione (204).

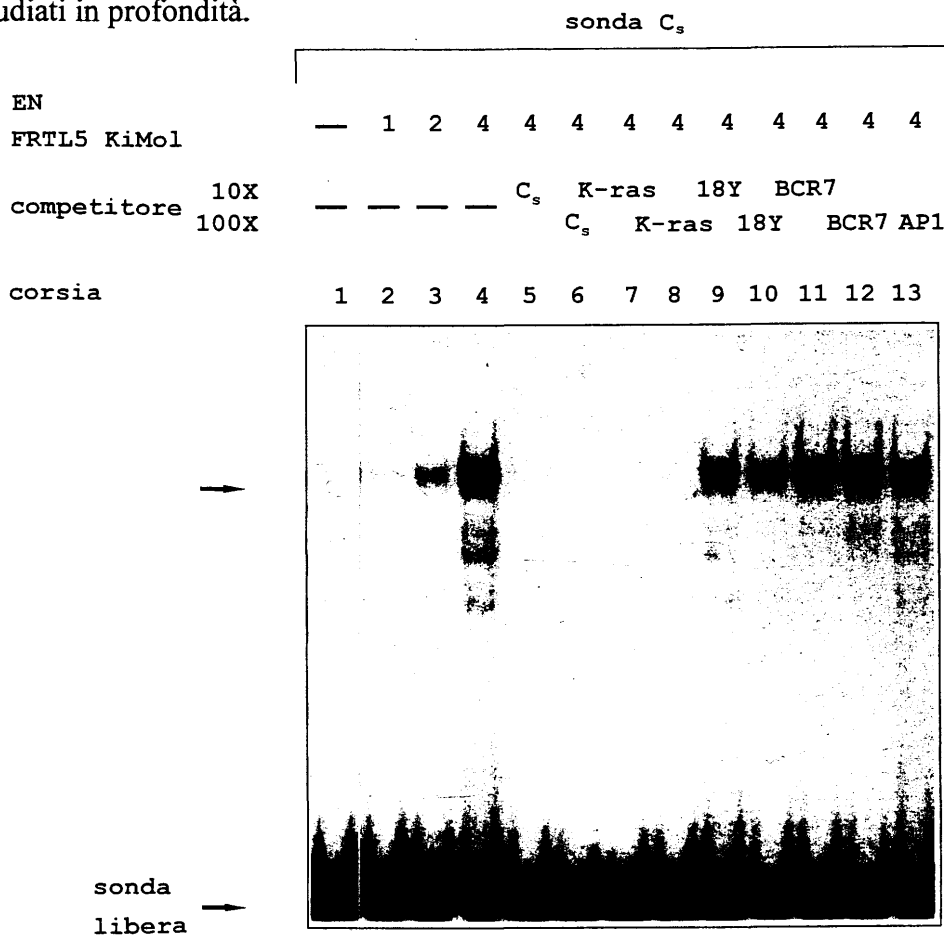
Come accennato precedentemente, sono stati individuati diversi fattori nucleari capaci di riconoscere e di legare questo tratto di sequenza. Alcuni interagiscono con la sequenza omopirimidinica-omopurinica strutturata a doppio filamento, ed è frequente il caso di Sp1, ma altri invece specifici al riconoscimento del DNA a singolo filamento.

Per esempio fattori nucleari come ERDBP-1 (205), NSEP-1 (206), THZif-1 (202) sono stati identificati in quanto interagiscono con il tratto omopirimidico-omopurino a singolo filamento. La funzione di questi fattori potrebbe essere coinvolta nella stabilizzazione della struttura atipica del DNA. Da un lato queste proteine potrebbero favorire l'evento trascrizionale in quanto stabilizzano le conformazioni "aperte" di tipo non-B facilitando l'entrata dell'RNA polimerasi II, dall'altro possono venir correlate ad una repressione trascrizionale. In tal caso l'effetto inibitorio potrebbe essere motivato dalla struttura atipica del DNA, la quale impedisce ai fattori trascrizionali di legarsi ai loro consensus, che normalmente riconoscono se presenti in una conformazione del DNA di tipo B (AR). Nel caso del recettore degli androgeni i siti per Sp1 presenti nel tratto con le ripetizioni TCC, potrebbero venir sottratti al riconoscimento, in quanto coinvolti nella struttura tridimensionale non B. Tuttavia la regolazione trascrizionale dipendente dalla dinamicità del DNA a livello di queste regioni, risulta tuttora poco chiara.

Sempre mediante bandshift con estratti nucleari derivanti da diversi tipi cellulari (NIH-3T3, FRTL5 Cl2, FRTL5 KiMol, PC Cl3, PC E1a/raf e QT6, vedi sezione 3.6) è stata saggiata la presenza di proteine nucleari capaci di interagire con l'oligonucleotide C senso o antisense marcato come sonda. Come mostra la Figura 13 si forma un complesso predominante specifico per l'oligonucleotide C senso, il quale non viene competuto da sequenze omopirimidiniche casuali (18Y), o sequenze omopirimidiniche icastonate in DNA a sequenza casuale (bcr-1) ma bensì soltanto da sequenze con



ripetizioni dirette dell'elemento TCCC come nell'oligonucleotide derivato dal promotore di c-Ki-ras, usato in un esperimento di *competizione* (Vedi sequenze oligo Sezione 3.1.3). (198) Stiamo tutt'ora ricercando fattori che, in comune con altri promotori qui elencati, possano avere un ruolo funzionale nella regolazione dell'espressione di *Hmgi-c* attraverso strutture a singolo filamento. Analogamente tre complessi specifici si formano con l'oligonucleotide antisenso, ma non sono ancora stati studiati in profondità.

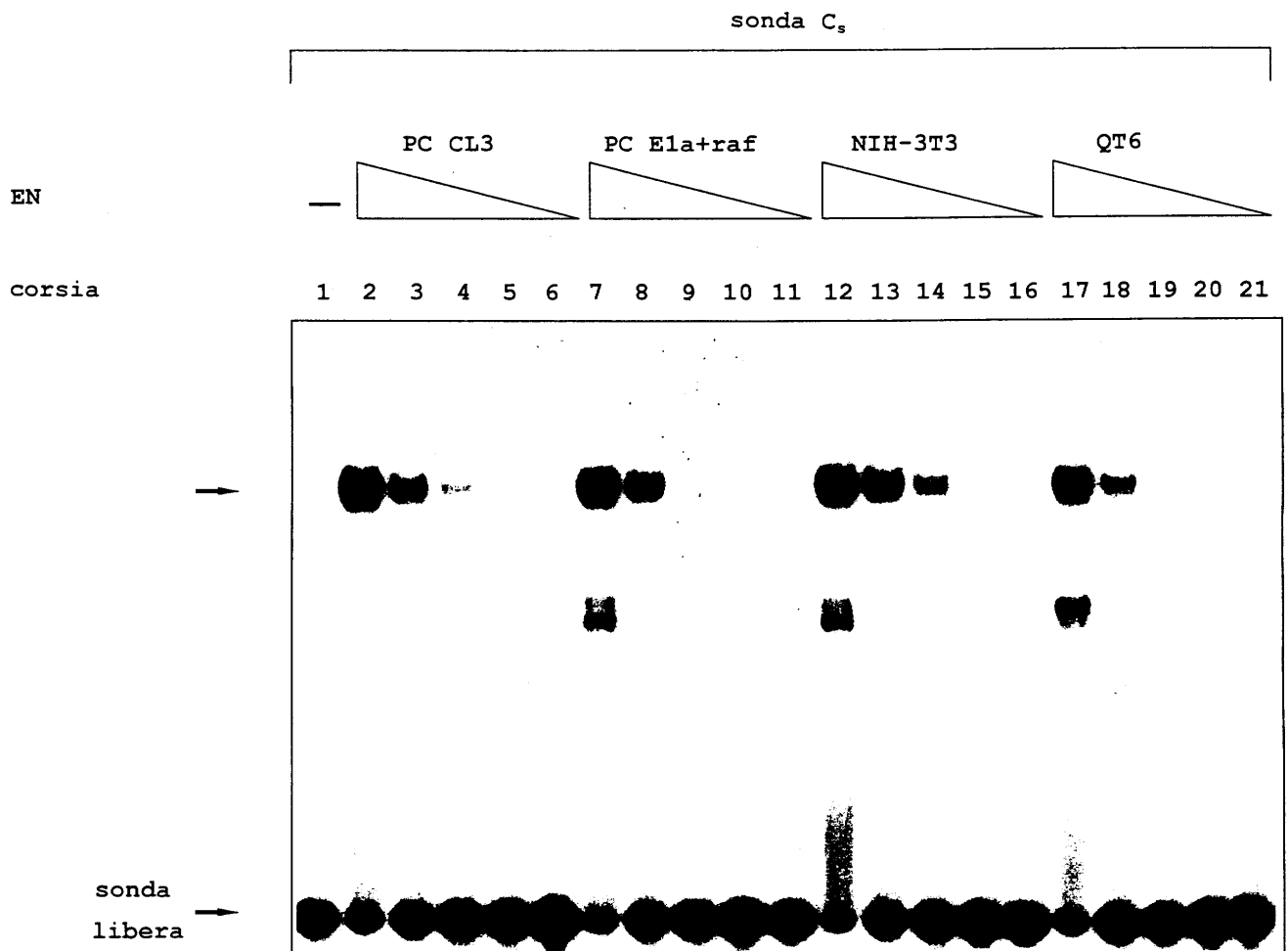


**Figura 13**

### Un fattore specifico lega il singolo filamento pirimidinico di *Hmgi-c*

Un complesso dominante specifico si forma tra fattori dell'estratto di cellule FRTL5 KiMol e l'oligonucleotide C senso. Sono mostrate diverse competizioni con un eccesso molare di 50 e 100 volte rispetto alla sonda dei singoli filamenti di altri consensus per fattori di trascrizione. Solo l'oligonucleotide corrispondente al promotore di c-Ki-ras murino e costituito anch'esso da una sequenza ppyr/ppur (con ripetizioni TCCC) è in grado di competere efficacemente (corsie K-ras), mentre non lo è AP1, completamente scorrelato e nemmeno BCR7, che però possiede alcune ripetizioni pirimidiniche ma in un contesto di purine. L'oligonucleotide 18Y è totalmente pirimidinico, ma non ha regolarità o ripetizioni, e compete solo inefficacemente anche ad un'eccesso molare di 100 volte.

Estratti provenienti da diverse linee cellulari usati in esperimenti di *bandshift* (Figura 14) mostrano la presenza dello stesso complesso maggiore con l'oligonucleotide C senso, indicando che questi fattori che legano il singolo filamento potrebbero essere ubiquitari. Inoltre non vi sono differenze apprezzabili con questo metodo sperimentale tra cellule che esprimono *Hmgi-c* (PC E1a+raf, NIH-3T3 e QT6) e cellule prive del fattore (PC Cl3). Ciò non implica per forza che non possano esistere meccanismi di regolazione trascrizionale diversi anche attraverso strutture a singolo filamento.

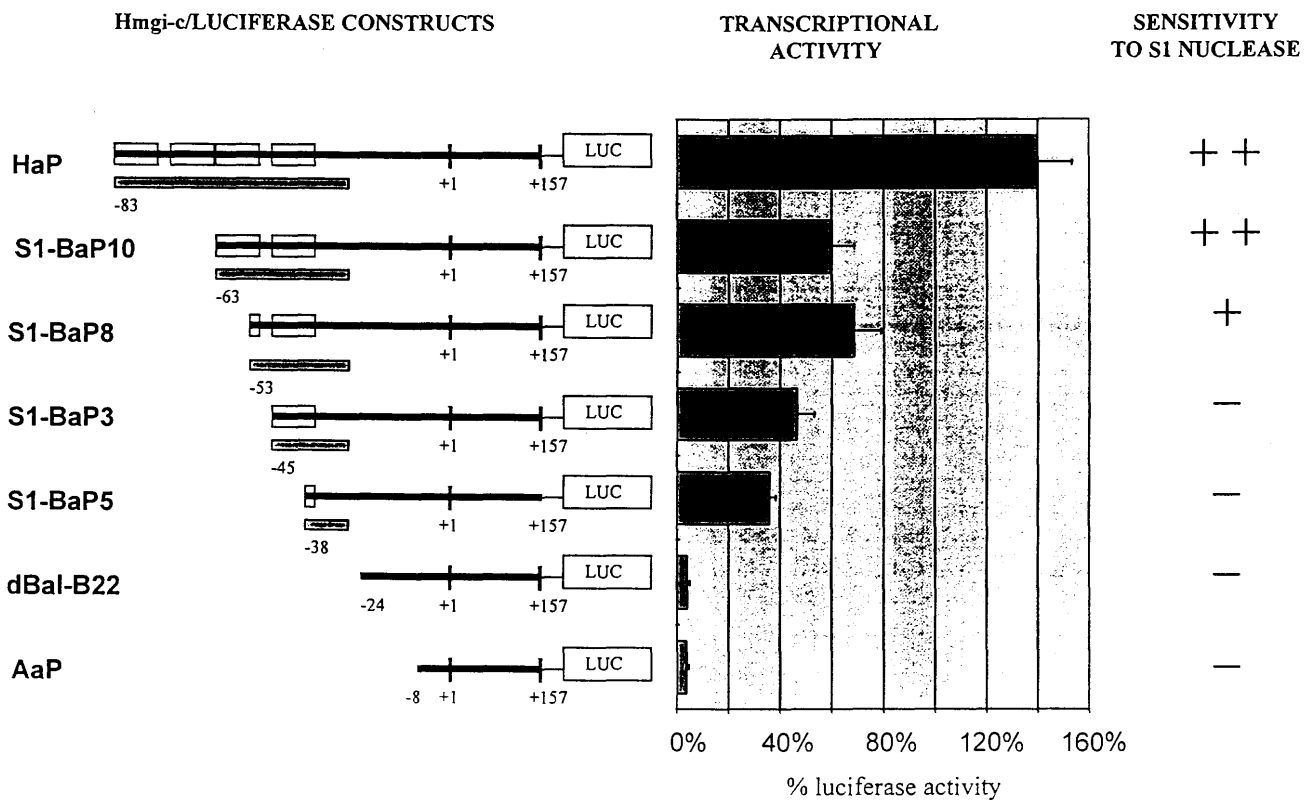


**Figura 14**

**Il fattore che lega il singolo filamento è presente in più linee cellulari, esprimenti e non esprimenti HMGI-C.**

Per rispondere alla domanda se la struttura non-B del tratto ppyr/ppur avesse un ruolo anche *in vivo* si sono prodotti dei costrutti recanti a monte del gene reporter della luciferasi delezioni progressive del promotore minimo e sono stati saggiati in esperimenti di trasfezione in cellule NIH-3T3 e successivamente ancora sono stati analizzati per la loro capacità di mantenimento della sensibilità verso la nucleasi S1. Nella Figura 15 sono indicati a sinistra i costrutti di delezione di HaP con schematizzati i quattro gruppi delle ripetizioni TCCTCCTCC dell'elemento ppyr/ppur che vanno via

via scomparendo con l'entità della delezione. In centro sono indicate le medie delle attività trascrizionali saggate in almeno tre esperimenti che dimostrano una caduta sequenziale dell'attività rispetto ad HaP, fino al costrutto dBal-B22 che non ha più né sequenza ppyr/ppur, né attività. Se si nota però l'andamento della sensibilità alla S1 nucleasi, allora non c'è correlazione tra la perdita della sensibilità alla S1 nucleasi e perdita di attività, in quanto il costrutto S1-BaP10, la cui attività si è abbassata di più della metà, mantiene ancora la completa sensibilità alla S1 nucleasi. Inoltre S1-Bap8 perde sensibilità ma non l'attività rispetto a S1-Bap10. Concludiamo dicendo che, almeno per quel che riguarda il sistema dei fibroblasti NIH-3T3, non sembra esserci una regolazione dell'espressione di *Hmgi-c* attraverso strutture non-B del promotore, ma piuttosto la perdita di sequenza concomitante ad una caduta di attività è più remissiva di una regolazione attraverso il doppio filamento e le sequenze di riconoscimento per Sp1, come schematizzato nella Figura 16.

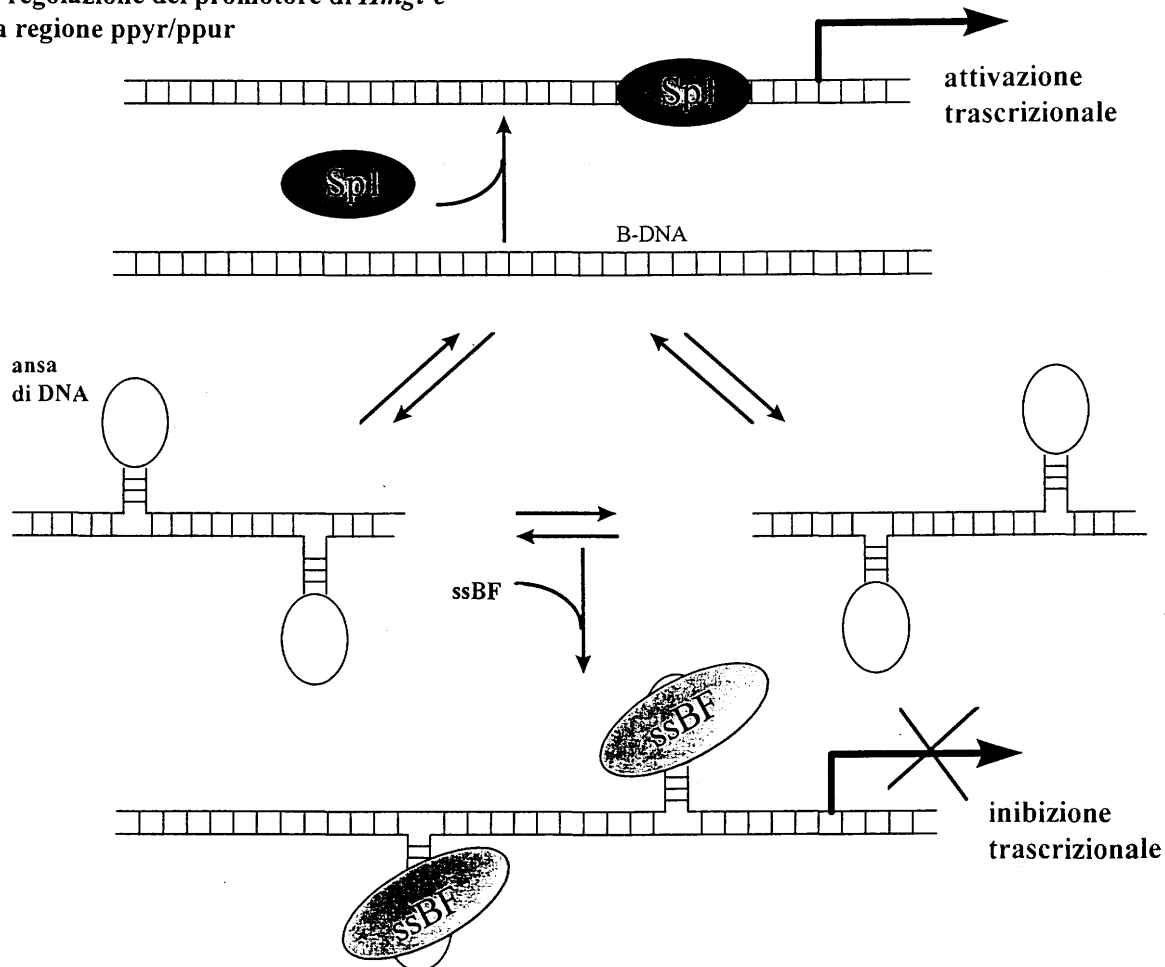


**Figura 15**

**Confronto tra attività trascrizionale e sensibilità alla S1 nucleasi di alcuni costrutti di delezione di HaP.**

Sono indicati a sinistra i costrutti utilizzati e al centro le loro attività trascrizionali saggate in esperimenti di trasfezione transiente in cellule NIH-3T3. A destra è indicata la rispettiva sensibilità alla S1 nucleasi.

Ipotesi sulla regolazione del promotore di *Hmgi-c*  
attraverso la regione ppyr/ppur



**Figura 16**

**Ipotesi di regolazione attraverso strutture a singolo filamento.**

Analisi preliminari di esperimenti *in vitro* con gli oligonucleotidi A, B e C sia senso che antisenso, indicano che con molta probabilità il promotore minimo di *Hmgi-c* è capace di formare triple eliche di vario tipo, ed inoltre essendoci molte ripetizioni all'interno di questa sequenza, in via teorica lo scivolamento di un filamento sull'altro è molto facile e potrebbe dare anse a singolo filamento, eventualmente stabilizzate da fattori nucleari specifici.

Dalle recenti scoperte sui riarrangiamenti genetici caratterizzanti alcuni fenotipi tumorali è stato rivelato un'interessante coinvolgimento delle regioni sensibili alla nucleasi S1. Infatti è stato riportato un caso di leucemia linfoblastica acuta a carico dei linfociti T che presenta una traslocazione t(10;14). Questa aberrazione cromosomica coinvolge una regione con frequenti punti di rottura, caratterizzata da un tratto oligopirimidinico-oligopurinic (207), facendo pensare ad un possibile coinvolgimento delle ripetizioni TCC (conservate anche nel promotore umano) del gene *HMGI-C*, nelle traslocazioni che interessano il 5' del gene, riscontrate in tre casi di amartoma polmonare condroide (PCH) (178).

#### 4.2.2.4 Analisi del promotore *in vivo*.

Allo scopo di analizzare la completezza e la funzionalità della sequenza promotore del gene *Hmgi-c* da noi identificate, sono stati creati dei topi transgenici in collaborazione con il gruppo della Prof.ssa F. Altruda del Dipartimento di Genetica, Biologia e Chimica Medica dell'Università di Torino. Questi topi portano un transgene costruito nel modo seguente: circa 4 Kbasi di 5' non tradotto del gene murino *Hmgi-c* sono state fuse al gene marcatore di *E. coli lacZ*, codificante per la  $\beta$ -galattosidasi di *Escherichia coli*, a valle. La sequenza del gene marcatore è stato opportunamente modificata, con l'aggiunta di una sequenza per la localizzazione nucleare, affinché la sua espressione non si confonda con quella endogena citoplasmatica. Con la generazione parentale di topi, sette linee in tutto, sono stati effettuati diversi incroci con altri topi normali per ottenere sia generazioni figlie che mantenessero nel tempo le diverse linee di topi transgenici, sia per poter effettuare la colorazione con il 5-Bromo-4-chloro-3-indolil  $\beta$ -D-galactopyranoside o X-GAL, substrato della  $\beta$ -galattosidasi codificata dal gene *lacZ*, su embrioni di topi di diversa età (10-16 gg). Quest'ultima dovrebbe dare una colorazione blu in ogni cellula dove il promotore del gene *Hmgi-c* è fisiologicamente attivo, ossia dove è stata descritta l'espressione del gene *Hmgi-c* endogeno (Hirning-FolzWilda M. *et al.* 97), posto che in quella sequenza di promotore usata per costruire il transgene sia contenuta effettivamente tutta l'informazione necessaria per la corretta espressione del gene *in vivo*.

Se questo non è il caso, significa che esistono sequenze regolatrici (*enhancers* o *silenziatori*) situate in posti diversi da quelle 4 Kbasi identificate da noi come sequenza promotore. L'analisi istologica delle sezioni degli embrioni, colorati nel nostro laboratorio, è stata affidata alla Dott.ssa P. Braghetta (Istituto di Istologia ed Embriologia dell'Università di Padova), la quale effettuerà anche l'inclusione in paraffina e/o la crioconservazione degli embrioni. Dati preliminari indicano che almeno due delle sette linee trasmettono il transgene alla progenie e lo fanno esprimere agli embrioni ottenuti dai loro incroci, e che quelle 4 kbasi di 5' non tradotto sono sufficienti per la corretta espressione del transgene in quei tessuti dove anche il gene *Hmgi-c* endogeno viene espresso.

Un secondo costrutto, denominato ApaP-*lacZ*, contenente solo 1.5 Kbasi di sequenza al 5' di *Hmgi-c* (corrispondente al costrutto ApaP per le trasfezioni), è stato usato per creare una seconda linea di topi transgenici, la cui analisi è tutt'ora in corso. Bisogna considerare comunque che più corta è la sequenza di *Hmgi-c* a monte del gene *lacZ*, più probabile è che il gene marcatore sia governato non dalla sequenza di *Hmgi-c*, bensì da sequenze che, a seconda del sito di integrazione del transgene, possono influenzare l'andamento dell'espressione di *lacZ* sia in senso positivo che negativo.

Su sette fondatori anche in questo caso vi sono stati almeno 3 che possiedono il gene silenziato, come stabilito dal confronto tra l'analisi di PCR sul DNA estratto dalle codine di topini e la colorazione degli embrioni delle generazioni figiali. (vedi sezione 3.5.1) Sulle altre linee, oltre alla colorazione degli embrioni, viene effettuata un'analisi mediante *Southern blot* del DNA estratto dal sacco vitellino (sezioni 3.5.2 e 3.5.3), che sarà ibridizzato con sonde specifiche del transgene, per stabilire, in confronto ad opportuni controlli quantitativi, quante copie del transgene si sono integrate nel genoma di questi topi in modo tale da correlare il numero di copie integrate con l'intensità di colorazione ottenuta nello stesso embrione. È possibile infatti che l'analisi dei topi transgenici solo con la PCR sia falsata dalla presenza di falsi positivi (allora il Southern genomico è il controllo discriminante) oppure dal fatto che la colorazione non viene rilevata a causa di un numero basso di copie integrate per cui potrebbero andare perse informazioni importantissime. Sono mostrati due esemplari, rappresentanti delle due linee di transgenici, nella figura 17.

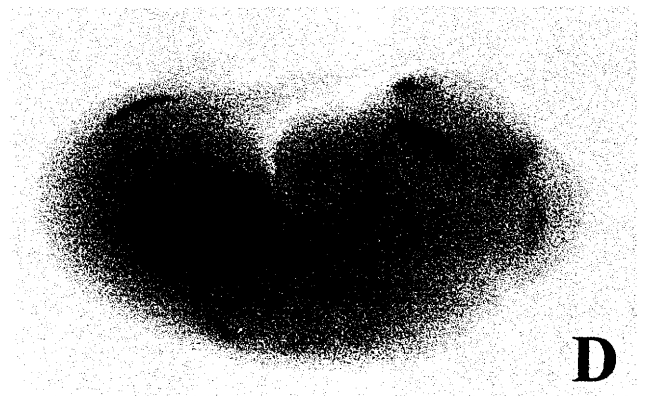
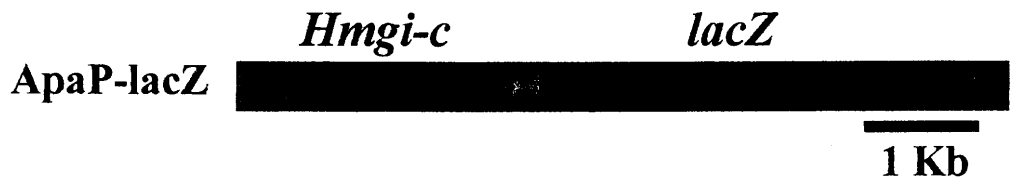


Fig. 17

A e B: Visione dorsale e laterale, rispettivamente, di un embrione a circa 12 giorni p.c. della prima serie di topi transgenici con il costrutto SalP-lacZ, dopo la colorazione con x-gal. Il costrutto del transgene è mostrato nello schema sotto le 2 immagini (pannello superiore).

C e D: Visione dorsale e laterale, rispettivamente, di un embrione a circa 12 giorni p.c. della seconda serie di topi transgenici con il costrutto ApaP-lacZ, dopo la colorazione con x-gal. Il costrutto del transgene è mostrato nello schema sopra le 2 immagini (pannello inferiore).

In verde le sequenze al 5' di *Hmgi-c*, in viola il segnale di localizzazione nucleare ed in blu la componente codificante per la  $\beta$ -galattosidasi.

## Conclusioni

Nei quattro anni di dottorato, il mio studio, ha dato un contributo alla conoscenza del funzionamento dei fattori HMGI in diversi processi biologici, come il controllo della trascrizione genica e l'espressione nella tumorigenesi.

In tal senso si è rivelata importantissima la produzione di un anticorpo anti-HMGI(Y) utilizzabile nella diagnosi ospedaliera, come dimostrano molte pubblicazioni in modelli tumorali di diversissima natura (carcinomi del colon, della tiroide e della cervice), come marcatore del fenotipo trasformato già a livello di lesioni moderate.

In special modo è stata estesa l'analisi del fattore HMGI-C, per poterlo definire fattore architetturale in analogia ad HMGI(Y), tenendo conto però delle sue diversità. Infatti almeno per quel che riguarda la regolazione della sua espressione, l'isolamento e la caratterizzazione del suo promotore nel topo e il ritrovamento che l'attività basale è dovuta a fattori della famiglia Sp1 e CTF/NF-1, hanno permesso di stabilire che ci sono differenze sostanziali già nella struttura al 5' dei geni *Hmgi-c* ed *HMGI(Y)*, che riflettono diversità di comportamento nell'attività genica degli stessi.

L'attenzione è comunque rivolta all'attività differenziale del promotore minimo di *Hmgi-c* in cellule normali e trasformate. I primi dati con una certa rilevanza e riproducibilità in più modelli cellulari indicherebbero che un meccanismo generale per questo fenomeno è dato dalla derepressione della trascrizione in cellule trasformate, e non da un'induzione della stessa.

L'analisi della sensibilità all'S1 nucleasi fornisce un ulteriore livello di conoscenza su meccanismi alternativi esistenti nella regolazione della trascrizione e potrebbe fornire un'ipotesi meccanicistica su molte traslocazioni che avvengono a livello del 5' del gene HMGI-C nel locus 12q14-15. Inoltre le strutture di DNA non convenzionali possono essere soggetto a terapie geniche mediante *TPO's* (triple forming oligonucleotides) che potrebbero bloccare il promotore impedendo l'inizio della trascrizione.

Lo studio del promotore *Hmgi-c* trova il suo apice nel topo transgenico, in quanto attraverso il confronto tra espressione fisiologica nel topo normale durante l'embriogenesi e la colorazione blu degli embrioni transgenici con il costrutto lungo e quello corto, si potranno identificare regioni essenziali nella regolazione dell'*Hmgi-c in vivo*.



Durante il periodo di svolgimento dell'attività di ricerca per il conseguimento del dottorato di ricerca in Biochimica ho lavorato presso il Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole dell'Università degli Studi di Trieste.

Ho anche avuto l'opportunità di visitare alcuni istituti universitari italiani, quali il Dipartimento di Genetica, Biologia e Chimica Medica dell'Università di Torino (Prof.ssa F. Altruda) e l'Istituto di Istologia ed Embriologia dell'Università di Padova (Dott.ssa P. Braghetta e Prof. G.M. Bressan) per il lavoro sui topi transgenici. Inoltre ho visitato il Dipartimento di Scienze Mediche dell'Università di Torino, sede di Novara, per l'acquisizione di tecniche e materiali riguardanti lo studio del fattore CTF/NF-1.

42. Zhou, X., Benson, K.F., Ashar, H.R., and Chada, K. (1995) *Nature* **376**, 771-774.
43. Ashar, H.R., Cerath, L., Przybysz, K.M., and Chada, K., (1996) *Genomics* **31**, 207-214.
44. Lund, T., Holtlund, J., Fredriksen, M., Laland, S.G. (1983) *FEBS Lett.* **152**, 163-167.
45. Holth, L.T., Thorlacius, A.E., and Reeves, R. (1997) *DNA Cell Biol.* **16**, 1299-1309.
46. Begum, N., Pas., J.M, and Bhorjee, J.S. (1990) *J. Biol. Chem.* **265**, 11936-11941.
47. Lanahan, A., Williams, J.B., Sanders, L.K., and Nathans, D. (1992) *Mol. Cell. Biol.* **12**, 3919-3929.
48. Thanos, D., and Maniatis, T. (1992) *Cell* **71**, 777-789.
49. Maher, J.F., and Nathans, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6716-6720.
50. Strauss, F., and Varshavsky, A. (1984) *Cell* **37**, 889-901.
51. Geierstanger, B.H., Bolkman, B.F., Kremer, W., and Wemmer, D.E. (1994) *Biochemistry* **33**, 5347-5355.
52. Kopka, M.L., Yoon, C., Goodseel, D., Pjura, P., and Dickerson, R. (1985) *J. Mol. Biol.* **183**, 553-563.
53. Frank, O. Schwanbeck, R, Wisniewsky, J.R. (1998) *J. Biol. Chem.* **273**, 20015-20020.
54. Huth, J.R., Beweley, C.A., Nissen, M., Evans, J.N.S., Reeves, R., Groneborn, A.M., and Clore, G.M. (1997) *Nature Struct. Biol.* **4**, 657-665.
55. Aravind, L. and Landsman, D. (1998) *Nucleic Acids Res.* **26**, 4413-4421.
56. Hill, D.A., and Reeves, R. (1997) *Nucleic Acids Res.* **25**, 3523-3531.
57. Yie, J., Merika, M., Munshi, N., Chen, G., and Thanos, D. (1997) *EMBO J.* **18**, 3074-3089.
58. Panagiotidis, C.A., and Silverstein, S. (1999) *J. Virol.* **256**, 64-74.
59. Falvo, V.J., Thanos, D., and Maniatis, T. (1995) *Cell* **83**, 1101-1111.
60. Strick, R., and Laemmli, U. (1995) *Cell* **83**, 1137-1148.
61. Elton, T.S., Nissen, M.S., and Reeves, R. (1986) *Anal. Biochem.* **157**, 53-62.
62. Giancotti, V., Bandiera, A., Sindici, C., Perissin, L., and Crane-Robinson, C. (1996) *Eur. J. Biochemistry* **317**, 865-870.
63. Hunter, T. (1995) *Cell* **80**, 225-236.
64. Linnala- Palvimo, J., and Kankkunen, A. (1989) *FEBS Lett.* **257**, 101-104.
65. Ferranti, P., Malorni, A., Marino, G., Pucci, P., Goodwin, G.H., Manfioletti, G., and Giancotti, V. (1992) *J. Biol. Chem.* **267**, 22486-22489.
66. Moreno, S., and Nurse, P. (1990) *Cell* **61**, 549-551.
67. Lund, T., Skålhegg, B., Holtlund, J., Blomhoff, H.K., and Laland, S.G. (1987) *Eur. J. Biol.* **166**, 21-26.
68. Lund, T., and Laland, S.G. (1990) *Biochem. Biophys. Res. Commun.* **171**, 342-347.
69. Meijer, L., Ostvold, A.C., Walaas, S.I., Lund, T., and Laland, S.G. (1991) *Eur. J. Biochem.* **196**, 557-567.
70. Reeves, R., Langan, T.A., and Nissen, M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1671-1675.
71. Ptashne, M. (1988) *Nature* **355**, 683-689.
72. Wang, D.Z., Ray, P., and Boothby, M. (1995) *J. Biol. Chem.* **270**, 22924-22932.
73. Wang, D., Zamorano, J., Keegan, A.D., and Boothby, M.A. (1997) *J. Biol. Chem.* **272**, 25083-25090.
74. Reeves, R., and Nissen, M. (1990) *J. Biol. Chem.* **265**, 8573-8582.
75. Nissen, M., Langan, T.A., and Reeves, R. (1991) *J. Biol. Chem.* **266**, 19945-19952.
76. Claus, P., Schulze, E., and Wiśniewsky (1994) *J. Biol. Chem.* **269**, 33042-33048.
77. Schwanbeck, R., and Wiśniewsky, J.R. (1997) *J. Biol. Chem.* **272**, 27476-27483.
78. Schwanbeck, R., Manfioletti, G., and Wiśniewsky, J.R (1999) *J. Biol. Chem.* in corso di stampa.
79. Munshi, N., Merika, M., Yie, J., Senger, K., Chen, G., and Thanos, D. (1998) *Mol. Cell* **2**, 457-467.
80. Giancotti, V., Bandiera, A., Sindici, C., Perissin, L., and Crane-Robinson C. (1996) *Biochem. J.* **317**, 865-870.
81. Reeves, R., and Nissen, M.S. (1993) *J. Biol. Chem.* **268**, 21137-21146.
82. Radic, M.Z., Saghbini, M., Elton, Y.S., Reeves, R., and Hamkalo, B.A. (1992) *Chromosoma* **101**, 602-608.
83. Disney, J.E., Johnson, K.R., Magnuson, N.S., Sylvester, S.R., and Reeves, R. (1989) *J. Cell. Biol.* **109**, 1975-1982.
84. Martelli, A.M., Riccio, M., Bareggi, R., Manfioletti, G., Tabellini, G., Baldini, G., Narducci, P., and V. Giancotti (1998) *J. Histochem. Citochem.* **46**, 863-864.
85. Verga-Weisz, P. Zlatanova, J., Leuba, S.H., Schroth, G.P., and van Holde K. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3525-3529.
86. Bianchi, M. (1988) *EMBO J.* **7**, 843-849.
87. Hill, D.A., Pedulla, M.L., and Reeves, R. (1999) *Nucleic Acids Res.* **27**, 2135-2144.
88. Farnet, C.M., and Bushman, F. (1997) *Cell* **88**, 483-492.
89. Zhao, K., Kaes, E., Gonzalez, E., and Laemmli, U. (1993) *EMBO J.* **12**, 3237-3247.

# Bibliografia

1. Goodrich, J.A., Cutler, G., and Tjian, R. (1996) *Cell*, **84**, 825-830.
2. Zawel, L., and Reinberg, D. (1995) *Annu. Rev. Biochem.* **64**, 533-561.
3. Weis, L., and Reinberg, D., (1992) *FASEB J.* **6**, 3300-3309.
4. Hernandez, N. (1993) *Genes Dev.* **7**, 1291-1308.
5. Dimitar, V. (1995) *Nature* **377**,
6. Smale, T., and Baltimore, D. (1989) *Cell* **57**, 103-113.
7. Goodrich, J.A., and Tjian, R. (1994) *Curr. Opin. Cell. Biol.* **6**, 403-409.
8. Treitzenberg, S. J. (1995) *Curr. Opin. Genet. Dev.* **5**, 190-196.
9. McKnight, S., and Yamamoto, Y. (1992) eds. Transcriptional Regulation, *Cold Spring Harbor Laboratory Press*
10. Klein, C., and Struhl, K. (1994) *Science* **266**, 280-282.
11. Liebermann, P. M., and Berk, A. J. 1994 *Genes. Dev.* **8**, 995-1006
12. Tjian, R., and Maniatis, T. (1994) *Cell* **77**, 5-8.
13. Serizawa, H., Makela, T.P., Conaway, R. C., Weinberg, R. A., and Young, R.A. (1995) *Nature* **374**, 280-282.
14. Liao, S.M., Zhang, J., Jeffery, D.A., Koleske, A.J., Thompson, C.M., Chao, D.M., Viljoen, M., van Vuuren, H.J., and Young, R.A. (1995) *Nature* **374**, 193-196.
15. Dikstein, R., Ruppert, S., and Tjian, R. (1996) *Cell* **84**, 781-790.
16. Ptashne, M. (1992) A Genetic Switch Second Edition, Cambridge, Massachusetts: *Blackwell Scientific Publications*
17. Wolffe, A.P. (1994) *Cell* **77**, 13-16.
18. Felsenfeld, G. (1992) *Nature* **355**, 219-224.
19. Kadonaga, J.T. (1998) *Cell* **92**, 307-313.
20. Ito, T., Tyler, J.K., and Kadonaga, J.T. (1997a) *Genes Cells* **2**, 593-600.
21. Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R., and Kadonaga, J.T. (1997b) *Cell* **90**, 145-155.
22. Bustin, M., and Reeves, R., (1996) In: W.E. Cohn and K. Moldave (eds.) *Progr. Nucleic Acid Res.* **54**, 35-100.
23. Bustin, M., Lehn, D.A., and Landsman, D. (1990) *Biochim. Biophys. Acta* **1049**, 231-243.
24. Bianchi, M.E., and Lilley, D.M.J. (1995) *Nature* **375**, 532-536.
25. Varga-Weisz, P. van Holde, K.E., and Zlatanova, J. (1994) *Biochem. Biophys. Res. Commun.* **203**, 1904-1911.
26. Ner, S. (1992) *Curr. Biol.* **2**, 208-210.
27. Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.S., Foster, J.W., Frischauf, A.M., Lowell-Badge, R, and Goodfellow, P.N. (1990) *Nature* **346**, 240-244.
28. Ferrari, S., Harley, V.R., Pontiggia, A., Goodfellow, P.N., Lovell-Badge, R., and Bianchi, M.E. (1992) *EMBO J.* **11**, 4497-4504.
29. Sandeen, G., Wood, W.J., and Felsenfeld, G. (1988) *Nucleic Acids Res.* **8**, 3757-3778.
30. Postnikov, Y.V., Trieschmann, L., Rickers, A., and Bustin, M. (1995) *J. Biol. Chem.* **252**, 423-430.
31. Alfonso, P.J., Crippa, M.P., Hayes, J.J., and Bustin, M. (1994) *J. Biol. Chem.* **236**, 189-195.
32. Ding, H.F., Rimsky, S., Batson, S.C., Bustin, M., and Hansen, U. (1994) *Science* **265**, 796-799.
33. Johnson, K.R., Lehn, D.A., Elton, T.S., Barr, P.J., and Reeves, R. (1988) *J. Biol. Chem.* **263**, 18338-18342.
34. Johnson, K.R., Cook, S.A., and Davidson, M.T. (1991) *Genomics* **12**, 503-509.
35. Cmarik, J.L., Li, Y., Ogram, S.A., Min, H., Reeves, R., and Colburn, N.H. (1998) *Oncogene* **16**, 3387-3396.
36. Manfioletti, G., Giacotti, V., Bandiera, A., Buratti, E., Sautière P., Cary, P., Crane-Robinson, C., Coles, B., and Goodwin, G.H. (1991) *Nucleic Acids Res.* **19**, 6793-6797.
37. Manfioletti, G., Rustighi, A., Mantovani, F., Goodwin, G.H. and Giacotti, V., (1995) *Gene* **167**, 249-253.
38. Patel, U., Bandiera, A., Manfioletti, G., Giacotti, V., Chau, K.Y., and Crane-Robinson, C. (1994) *Biochem Biophys. Res. Commun.* **201**, 63-70.
39. Friedman, M., Holth, L., Zoghbi, H.Y., and Reeves, R., (1998) *Nucleic Acids Res.* **21**, 4259-4267.
40. Ashar, H.R., Schoenberg Fejzo, M., Tkachenko, A., Zhou, X., Fletcher, J.A., Weremowicz, S., Morton, C.C., and Chada, K. (1995) *Cell* **82**, 57-65.
41. Geurts, J.M.W., Schoenmakers, E.F.P.M., Van de Ven, W.J.M., (1997) *Cancer Genet. Cytogenet.* **95**, 198-205.

90. Bird, A.P. (1987) *Trends Genet.* **3**, 342-347.
91. Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., and Bird, A. (1992) *Cell* **69**, 905-914.
92. Bonnefoy, E., Bandu, M.T., and Doly, J. (1999) *Mol. Cell. Biol.* **19**, 2803-2816.
93. Weintraub, H. (1985) *Cell*, **42**, 705-711.
94. Thanos, D., and Maniatis, T. (1995) *Cell* **83**, 1091-1100.
95. Goodbourn, S.E.Y., and Maniatis, T. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1447-1451.
96. Mantovani, F., Covaceuszach, S., Rustighi, A., Sgarra, R., Heath, C., Goodwin, G.H., and Manfioletti, G. (1998) *Nucleic Acids Res.* **26**, 1433-1439.
97. Du, W., Thanos, D., and Maniatis, T. (1993) *Cell* **74**, 887-898.
98. Merika, M., Williams, A.J., Chen, G., Collins, Y., and Thanos, D. (1998) *Mol. Cell* **1**, 277-287.
99. Kim, T.K., and Maniatis, T. (1997) *Mol. Cell* **1**, 119-129.
100. Yuan, W., Condorelli, G., Caruso, M., Felsani, A., and Giordano, A. (1996) *J. Biol. Chem.* **271**, 9009-9013.
101. Kwok R.P.S., Lundbland, J.R., Chrivia, J.C., Richards, J.P., Bachinger, H.P., Brennan, R.G., Roberts, S.G.E., Green, M.R., and Goodman, R.H. (1994) *Nature* **370**, 223-226.
102. Bannister, A., and Kouzarides, T. (1996) *Nature* **384**, 641-643.
103. Halle, J.P., and Meisterernst, M. (1996) *Trends Genet.* **12**, 161-163.
104. Thanos, D., and Du, W., (1993) In: *Cold Spring Harb. Symp. Quant. Biol.* **58**, 73-81.
105. Zhang, X.M., and Verdine, G.L. (1999) *J. Biol. Chem.* **274**, 20235-20243.
106. Du, W., and Maniatis, T. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11318-11322.
107. Whitley, M., Thanos, D., Read, M.A., Maniatis, T. and Collins, T. (1994) *Mol. Cell. Biol.* **14**, 6464-6475.
108. Wood, L.D., Farmer, A.A., and Richmond, A. (1995) *Nucleic Acids Res.* **23**, 4210-4219.
109. Himes, S.R., Coles, L.S., Reeves, R., and Shannon, M.F. (1996) *Immunity* **5**, 479-489.
110. John, S., Reeves, R., Lin, J.X., Child, R., Leiden, J.m., Thompson, C.B., and Leonard, W.J. (1995) *Mol. Cell. Biol.* **15**, 1786-1796.
111. Reeves, R., Elton, T.S., Nissen, M., Lehn, D., and Johnson, K. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6531-6535.
112. Khadadoust, M.M., Khan, K.D., Park, E., and Bothwell, A.L.M. (1998) *Blood* **92**, 2399-2409.
113. Abulkadir, S.A., Krishna, S., Thanos, D., Maniatis, T., Strominger, J., and Ono, S.J. (1995) *J. Exp. Med.* **182**, 487-500.
114. Fashena, S.J., Reeves, R., Ruddle, N.H. (1992) *Mol. Cell. Biol.* **12**, 894-903.
115. Leger, H., Sock, E., Renner, K., Grummt, F., and Wegner, M. (1995) *Mol. Cell. Biol.* **15**, 3738-3747.
116. French, S., Schmidt, M.C., and Glorioso, J. (1996) *Mol. Cell. Biol.* **16**, 5393-5399.
117. Perrella, M.A., Pellacani, A., Wiesel, P., Chin, M.T., Foster, L.C., Ibanez, M., Hsieh, C.M., Reeves, R., Yet, S.F., and Lee, M.E. (1999) *J. Biol. Chem.* **274**, 9045-9052.
118. Pellacani, A., Chin, M.T., Wiesel, P., Ibanez, M., Patel, A., Yet, S.F., Hsieh, C.M., Paulaskis, J.D., Reeves, R., Lee, M.E., and Perrella, M. (1999) *J. Biol. Chem.* **274**, 1525-1532.
119. Chuvpilo, S., Schomberg, C., Gerwig, R., Heinfling, A., Reeves, R., Grummt, F., and Serfling, E. (1993) *Nucleic Acids Res.* **21**, 5694-5704.
120. Skalnik, D.G., and Neufeld, E.J. (1992) *Biochem. Biophys. Res. Commun.* **187**, 563-569.
121. Kim, J., Reeves, R., Rothman, P., and Boothby, M. (1995) *Eur. J. Immunol.* **25**, 798-808.
122. Chin, M.T., Pellacani, A., Wang, H., Lin, S.S.J., Jain, M.K., Perrella, M.A., and Lee, M.E. 1998 *J. Biol. Chem.* **273**, 9755-9760
123. Currie, R.A. (1997) *J. Biol. Chem.* **272**, 30880-30888.
124. Arlotta, P., Rustighi, A., Mantovani, F., Manfioletti, G., Giancotti, V., Tell, G., and Damante G., (1997) *J. Biol. Chem.* **272**, 29904-29910.
125. Chiappetta, G., Avantiaggiato, V., Visconti, R., Fedele, M., Battista, S., Trapasso, F., Merciai, B.M., Fidanza, V., Giancotti, V., Santoro, M., Simeone, A., and Fusco, A. (1996) *Oncogene* **13**, 2439-2446.
126. Rogalla, P., Drechsler, K., Frey, G., Hennig, Y., Helmke, B., Bonk, U., and Bullerdiek, J. (1996) *Am. J. Pathol.* **149**, 775-779.
127. Hirning-Folz, U., Wilda, M., Rippe, V., Bullerdiek, J., and Hameister, H. (1998) *Genes Chromosomes and Cancer* **23**, 350-357.
128. Benson, K. And Chada, K.K. (1994) *Genet. Res. Camb.* **64**, 27-33.
129. Ruyter-Spyra, C.P., de Groof, A.J.C., van der Poel, J.J., Herbergs, J., Masabanda, J., Fries, R., and Groenen, M.A.M. (1998) *J. Hered.* **89**, 295-300.
130. Ogram, S.A., and Reeves, R. (1995) *J. Biol. Chem.* **270**, 14235-14242.

131. Ram, T.G., Reeves, R., and Hosick, H.L. (1993) *Cancer Res.* **53**, 2655-2660.
132. Ji, Y.S., Xu, Q., Schmedtje Jr, J.F. (1998) *Circ. Res.* **83**, 295-304.
133. Li, D., Lin, H., McMahon, M., Ma, H., and Ann, D.K. (1997) *J. Biol. Chem.* **272**, 25062-25070.
134. Lokuta, M.A., Maher, J., Noe, K.H., Pitha, P.M., Shin, M.L., and Shin, H.S. (1996) *J. Biol. Chem.* **271**, 13731-13738.
135. Kinoshita, T., Shirasawa, H., Shino, Y., Shimizu, K., Morija, H., and Simizu, B. (1996) *Virus Res.* **42**, 119-125.
136. Giaccotti, V., Pani, B., D'Andrea, P., Berlingieri, M.T., Di Fiore, P.P., Fusco, A., Vecchio, G., Philip, R., Crane-Robinson, C., Nicolas, R.H. Wright, C.A., and Goodwin, G.H. (1987) *EMBO J.* **6**, 1981-1987.
137. Johnson, K.R., Lehn, D.A., and Reeves, R. (1989) *Mol. Cell. Biol.* **9**, 2114-2123.
138. Giaccotti, V., Berlingieri, M.T., Di Fiore, P.P., Fusco, A., Vecchio, G., and Crane-Robinson, C. (1985) *Cancer Res.* **45**, 6051-6057.
139. Giaccotti, V., Buratti, E., Perissin, L., Zorzet, S., Balmain, A., Portella, G., Fusco, A., and Goodwin, G.H. (1989) *Exp. Cell Res.* **184**, 538-545.
140. Giaccotti, V., Bandiera, A., Ciani, L., Santoro, D., Crane-Robinson, C., Goodwin, G.H., Boiocchi, M., Dolcetti, R., Casetta, B. (1993) *Eur. J. Biochem.* **213**, 825-832.
141. Giaccotti, V., Bandiera, A., Buratti, E., Fusco, A., Marzari, R., Coles, B., and Goodwin, G.H. (1991) *Eur. J. Biochem.* **198**, 216-211.
142. Goodman, S.D., Nicholson, S.C., and Nash, H.A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **119**, 11910-11914.
143. Bussemakers, M. J.G., van de Ven, W.J.M., Debruyne, F.M.J., and Schalken, J.A. (1991) *Cancer Res.* **51**, 606-612.
144. Tamimi, Y., van der Poel, H.G., Denyn, M., Umbas, R., Karthaus, H.F.M., Debruyne, F.M.J., and Schalken, J.A. (1993) *Cancer Res.* **53**, 5512-5516.
145. Vertiainen, E., Palvimo, J., Mahonen, A., Linnala-Kankkunen, A., and Maenpaa, P. (1988) *FEBS Lett.* **228**, 45-49.
146. Johnson, K.R., Disney, J.E., Wyatt, C., and Reeves, R. (1990) *Exp. Cell Res.* **187**, 69-76.
147. Chiappetta, G., Bandiera, A., Berlingieri, M.T., Visconti, R., Manfioletti, G., Battista, S., Martinez-Tello, F.J., Santoro, M., Giaccotti, V., and Fusco, A. (1995) *Oncogene* **10**, 1307-1313.
148. Chiappetta, G., Tallini, G., De Biasio, M.C., Manfioletti, G., Martinez-Tello, F.J., Pentimalli, F., de Nigris, F., Mastro, A., Botti, G., Fedele, M., Berger, N., Santoro, M., Giaccotti, G., and Fusco, A. (1998) *Cancer Res.* **58**, 4193-4198.
149. Abe N., Watanabe, T., Sugiyama, M., Uchimura, H., Chiappetta, G., Fusco, A., and Atomi, Y. (1999) *Cancer Res.* **59**, 1169-1174.
150. Rogalla, P., Drechsler, K., Kazmierczak, B., Rippe, V., Bonk, U., and Bullerdiek, J. (1997) *Mol. Carcinog.* **19**, 153-156.
151. Bandiera, A., Bonifacio, D., Manfioletti, G., Mantovani, F., Rustighi, A., Zanconati, F., Fusco, A., Di Bonito, L. and Giaccotti, V. (1998) *Cancer Res.* **58**, 426-431.
152. Kim, D.H., Park, Y.S., Park, C.J., Son, K.C., Nam, E.s., Shin, H.S., Ryu, J.W., Kim, D.S., Park, C.K., and Park, Y.E. (1999) *Int. J. Cancer* **84**, 376-380.
153. Fedele, M., Bandiera, A., Chiappetta, G., Battista, S., Viglietto, G., Manfioletti, G., Casamassimi, A., Santoro, M., Giaccotti, V., and Fusco, A. (1996) *Cancer Res.* **56**, 1896-1901.
154. Giannini, G., Di Marcotullio, L., Ristori, E., Zani, M., Crescenzi, M., Scarpa, S., Piaggio, G., Vacca, A., Peverali, F.A., Diana, F., Screpanti, I., Frati, L., and Gulino, A. (1999) *Cancer Res.* **59**, 2484-2492.
155. Vallone, D., Battista, S., Pierantoni, G.M., Fedele, M., Casalino, L., Santoro, M., Viglietto, G., Fusco, A., and Verde, P. (1997) *EMBO J.* **16**, 5310-5321.
156. Berlingieri, M.T., Manfioletti, G., Santoro, M., Bandiera, A., Visconti, R., Giaccotti, V., and Fusco, A., (1992) *Mol. Cell. Biol.* **15**, 1545-1553.
157. Fedele, M., Berlingieri, M.T., Scala, S., Chiarotti, L., Viglietto, G., Rippe, V., Bullerdiek, J., Santoro, M., and Fusco, A. (1998) *Oncogene* **17**, 413-418.
158. Rabbits, T.H. 1994 (*Nature*) **372**, 143-149.
159. Schoenmakers, E.F.P.M., Wanschura, S., Mols, R., Bullerdiek, B., Van den Berghe, H., and Van de Ven, W.J.M. (1995) *Nature Genetics* **10**, 436-444.
160. Ashar, H.R. Schoenberg-Fejzo, M., Tkachenko, A., Zhou, X., Fletcher, J.A., Weremowicz, S., Morton, C.C., and Chada, K.K. (1995) *Cell* **82**, 57-65.
161. Petit, M.M.R., Mols, R., Schoenmakers, E.F.P.M., Mandahl, N., and Van de Ven, W.J.M. (1996) *Genomics* **36**, 118-129.
162. Rogalla, P., Kazmierczak, B., Meyer-Bolte, K., Tran, K.H., and Bullerdiek, J. (1998) *Genes, Chromosomes & Cancer* **22**, 100-104.

163. Geurts, J.M.W., Schoenmakers, E.F.P.M., Roeijer, E., Stenman, G., and Van de Ven, W.J.M. (1997) *Cancer Res.* **57**, 13-17.
164. Petit, M.M.R., Schoenmakers, E.F.P.M., Huysmans, C., Geurts, J.M.W., Mandahl, N., and Van de Ven, W.J.M. (1999) *Genomics* **57**, 438-441.
165. Kazmierczak, B., Rosigkeit, J., Wanschura, S., Meyer-Bolte, K., Van de Ven, W.J.M., Kayser, K., Kriehoff, B., Kastendiek, H., Bartnitzke, S., and Bullerdiek, B. (1996) *Oncogene* **12**, 515-521.
166. Schoenberg Fejzo, M., Yoon, S.J., Montgomery, K.T., Rein, M.S., Weremovicz, S., Krauter, K.S., Dorman, T.E., Fletcher, J.A., Mao, J., Moir, D.T., Kucherlapati, R.S., and Morton, C.C. (1994) *Genomics* **26**, 265-271.
167. Williams, A.J., Lee Powell, W., Collins, T., and Morton, C.C. (1997) *Am. J. Pathol.* **150**, 911-918.
168. Geurts, J.M.W., Schoenmakers, E.F.P.M., Roeijer, E., Astroem, A.K., Stenman, G., and Van de Ven, W.J.M. (1998) *Oncogene* **16**, 865-872.
169. Kazmierczak, B., Hennig, Y., Wanschura, S., Rogalla, P., Bartnitzke, S., Van de Ven, W.J.M., and Bullerdiek, J. (1995) *Cancer Res.* **55**, 6038-6039.
170. Tkachenko, A., Ashar, H.R., Meloni, A.M., Sandberg, A.A., and Chada, K.K. (1997) *Cancer Res.* **57**, 2276-2280.
171. Rabbits, T.H., and Boehm, T. (1991) *Adv. Immunol.* **50**, 119-146.
172. Tkachuk, D.C., Kohler, S., Cleary, M.L. (1992) *Cell* **71**, 691-700
173. Tallini, G., Dal Cin, P., Rhoden, K., Chiappetta, G., Manfioletti, G., Giancotti, V., Fusco, A., Van de Berge, H., and Sciot, R. (1997) *Am. J. Pathol.* **151**, 37-43.
174. Staats, B., Bonk, U., Wanschura, S., Hanisch, P., Schoenmakers, E.F.P.M., Van de Ven, W.J.M., Bartnitzke, S., and Bullerdiek, J. (1996) *Breast Cancer Res. Treat.* **38**, 299-303.
175. Bol, S., Wanschura, S., Thode, B., Deichert, U., Van de Ven, W.J.M., Bartnitzke, S., and Bullerdiek, J. (1996) *Cancer Genet Cytogenet.* **90**, 88-90.
176. Berner, J.M., Meza-Zepeda, L.A., Kools, P.F.J., Forus, A., Schoenmakers, E.F.P.M., Van de Ven, W.J.M., Fodstad, Ø., and Myklebost, O. (1997) *Oncogene* **14**, 2935-2941.
177. Kottickal, L.V., Sarada, B., Ashar, H., Chada, K., and Nagarajan, L. (1998) *Biochem. Biophys. Res. Commun.* **242**, 452-456.
178. Wanschura, S., Kazmierczak, B., Pohnke, Y., Meyer-Bolte, K., Bartnitzke, S., Van de Ven, W.J.M., and Bullerdiek, J. (1996) *Cancer Letters* **102**, 17-21.
179. Melillo, R.M. manoscritto inviato per la pubblicazione.
180. Battista, S., Fidanza, V., Fedele, M., Klein-Szanto, A.J.P., Outwater, E., Brunner, H., Santoro, M., Croce, C.M., and Fusco, A. (1999) *Cancer Res.* **59**, 4793-4797.
181. Arlotta, P., Tai, A.K.F., Manfioletti, G., Clifford, C., Jay, G., and Ono S.J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* in corso di stampa.
182. Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) *Molecular Cloning 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press.*
183. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156-159.
184. Del Sal, G., Ruaro, M.E., Philipson, L., and Schneider, C. (1992) *Cell* **70**, 595-607.
185. Ambesi-Impiomato, F.S., Parks, L.A.M., and Coon, H.G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3455-3459.
186. Zappavigna, V., Falciola, L., Catterich, M.H., Mavilio, F., and Bianchi, M.E. (1996) *EMBO J.* **15**, 4981-4991.
187. Gallego, M.I., Schoenmakers, E.F.P.M., Van de Ven, W.J.M., and Lazo, P.A. 1997 *Molecular Carcinogenesis* **19**, 114-121
188. Rustighi, A., Mantovani, F., Fusco, A., Giancotti, V., and Manfioletti, G. (1999) *Biophys. Biochem. Res. Commun.* **19**, 265, 439-447.
189. Kulpa, D., Topping, R., and Telesnitsky, A. (1997) *EMBO J.* **16**, 856-865.
190. Tripathi, J., and Brahmachari, S.K. (1991) *J. of Biomolecular Structure & Dynamics* **9**, 387-397.
191. Francis-Lang, H., Zannini, M., De Felice, M., Berlingieri, M.T., Fusco, A. and Di Luro, R. (1992) *Mol. Cell. Biol.* **12**, 5793-5800.
192. Crick, F.H.C. (1971) *Nature* **234**, 25-27.
193. Larsen, A., and Weintraub, H. (1982) *Cell*, **29**, 609-622.
194. Wells, R.D., Collier, D.A., Hanvey, J.C., Shimizu, M., and Wohlrab, F. (1988) *FASEB J.* **2**, 2939-2949.
195. Johnson, A., Jinno, Y., and Merlino, G.T. (1988) *Mol. Cell. Biol.* **8**, 4174-4184.
196. Chen, S., Supakar, P.C., Vellanoweth, R.L. Song, C.S., Chatterjee, B., and Roy, A.K. (1997) *Mol. Endocrinol.* **11**, 3-15.
197. Kinniburgh, A.J., Firulli, A.B., and Kolluri, R. (1994) *Gene* **149**, 93-100.

198. Pestov, D.G., Dayn, A., Siyanova, E.Y., George, D.L., and Mirkin, S.M. (1991) *Nucleic Acids Res.* **19**, 6527-6532.
199. Seino, S, Seino, M., Nishi, S., and Bell, G.I. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 114-118.
200. Wang, Z.Y., Qiu, Q.Q., Enger, K.T., and Deuel, T. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8896-8900.
201. Lafyatis, R., Denhez, F., Williams, T., Sporn, M., and Roberts, A. (1991) *Nucleic Acids Res.* **19** 64419-6425.
202. Sakatsume, O., Tsutsui, H., Wand, Y., Gao, H., Tang, X., Yamauchi, T., Murata, T., Keiishi, I., Yokoyama, K.K. (1996) *J. Biol. Chem.* **271**, 31322-31333.
203. Grossmann, M.E., and Tindall, D.J. (1995) *J. Biol. Chem.* **270**, 10968-10975.
204. Paranjape, A. (1993) *Ann. Rev. Biochem.* **63**, 265-297.

## Appendice

La maggior parte del lavoro descritto in questa tesi è stata pubblicata su riviste internazionali di biochimica e di biologia molecolare ed è stato presentato a Congressi sia nazionali che internazionali. Di seguito sono elencate le pubblicazioni più significative che sono state allegate a questa tesi:

### Lista delle pubblicazioni allegate:

- n° 1 Arlotta, P., Rustighi, A., Mantovani, F., Manfioletti, G., Giancotti, V., Tell, G., and Damante G. (1997) "High Mobility Group I Proteins Interfere with the Homeodomains Binding to DNA." *J. Biol. Chem.* 272, 29904-29910.
- n°2 Mantovani, F., Covaceuszach, S., Rustighi, A., Sgarra, R., Heath, C., Goodwin, G.H., and Manfioletti, G. (1998) "NF- $\kappa$ B mediated transcriptional activation is enhanced by the architectural factor HMGI-C." *Nucleic Acids Res.* 26, 1433-1439.
- n°3 Bandiera, A., Bonifacio, D., Manfioletti, G., Mantovani, F., Rustighi, A., Zanconati, F., Fusco, A., Di Bonito, L., and Giancotti V. (1998) "Expression of HMGI(Y) Proteins in Squamous Intraepithelial and Invasive Lesions of the Uterine Cervix." *Cancer Res.* 58, 426-431.
- n°4 Rustighi, A., Mantovani, F., Fusco, A., Giancotti, V., and Manfioletti, G. (1999) "Sp1 and CTF/NF-1 Transcription Factors Are Involved in the Basal Expression of the Hmgi-c Proximal Promoter." *Biochem. Biophys Res. Commun.* 265, 439-447.èè



## High Mobility Group I Proteins Interfere with the Homeodomains Binding to DNA\*

(Received for publication, April 25, 1997, and in revised form, August 1, 1997)

Paola Arlotta, Alessandra Rustighi, Fiamma Mantovani‡, Guidalberto Manfioletti, Vincenzo Giancotti, Gianluca Tell§, and Giuseppe Damante§¶

From the Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Università di Trieste, 34100 Trieste, Italy and the §Dipartimento di Scienze e Tecnologie Biomediche, Università di Udine, 33100 Udine, Italy

Homeodomains (HDs) constitute the DNA binding domain of several transcription factors that control cell differentiation and development in a wide variety of organisms. Most HDs recognize sequences that contain a 5'-TAAT-3' core motif. However, the DNA binding specificity of HD-containing proteins does not solely determine their biological effects, and other molecular mechanisms should be responsible for their ultimate functional activity. Interference by other factors in the HD/DNA interaction could be one of the processes by which HD-containing proteins achieve the functional complexity required for their effects on the expression of target genes.

Using gel-retardation assay, we demonstrate that two members of the high mobility group I (HMGI) family of nuclear proteins (HMGI-C and HMGY) can bind to a subset of HD target sequences and inhibit HDs from binding to the same sequences. The inhibition of the HD/DNA interaction occurs while incubating HMGI-C with DNA either before or after the addition of the HD.

The reduced half-life of the HD-DNA complex in the presence of HMGI-C, and the shift observed in the CD spectra recorded upon HMGI-C binding to DNA, strongly suggest that structural modifications of the DNA are responsible for the inhibition of the HD-DNA complex formation. Moreover, by co-transfection experiments we provide evidence that this inhibition can occur also *in vivo*.

The data reported here would suggest that HMGI proteins may be potential regulators of the function of HD-containing proteins and that they are able to interfere with the access of the HD to their target genes.

Homeodomains (HDs)<sup>1</sup> are 61-amino acid-long structures that are able to interact with DNA in a sequence-specific manner (1). They represent the DNA binding domain of a large number of transcription factors that control cell fate decisions in a wide range of organisms, including yeast, insects, and

vertebrates (2, 3). The structure of HDs and their mode of DNA interaction are conserved (1, 4, 5). With a few exceptions, sequences recognized by HDs possess a 5'-TAAT-3' core motif (6–8). Though recognizing similar DNA sequences, when expressed in the same temporal and spatial context, distinct HD-containing proteins may show different biological activities (9, 10). These findings indicate that the DNA binding specificity of these proteins cannot be the only molecular event that determines their biological effect. In fact, some other molecular mechanisms have been identified, in addition to the DNA binding specificity, that allow the selection of different target genes by distinct HD-containing proteins (11, 12). One of these mechanisms is based on the differential interaction of HD-containing proteins with accessory factors that mediate the contacts with the basal transcriptional machinery. For example, the Pou-domain-containing proteins Oct-1 and Oct-2 bind to the same DNA sequence, yet only the former is able to interact with the acidic transcriptional activator VP16 (11). Another mechanism is based on specific protein-protein interactions with other factors that bind to DNA tracts located near the sequences recognized by an HD-containing protein. A clear example of this phenomenon is the interaction between Ultrabithorax and Extradenticle proteins (12).

A simple molecular mechanism to control the binding of HD-containing transcription factors to a DNA regulatory element could be the presence of competitors for the interacting DNA sequence. In this paper we show that the high mobility group I proteins (HMGI, HMGY, and HMGI-C) may have this function. The HMGI proteins are a family of low molecular mass non-histone nuclear proteins rich in both basic and acidic residues, which comprise an important component of the active chromatin structure (reviewed in Ref. 13) and that have been shown to bind in the minor groove of AT-rich sequences of DNA through a DNA binding motif called AT-hook (14–17). This family includes three polypeptides: HMGI and HMGY derived from alternative splicing of the same gene and differing from each other by 11 amino acids (18) and a third protein called HMGI-C that is the product of a separate gene (19, 20).

HMG(Y) have been demonstrated to contribute to the transcriptional regulation of several promoters by interacting with different transcription factors (21–27). The human interferon  $\beta$  gene provides a particularly well characterized example of how HMGI(Y) could interact with specific DNA sequences on the promoter and cooperate with other proteins such as ATF-2, NF- $\kappa$ B, and c-Jun in activating this gene by stimulating the binding of these transcription factors to their target sites (28, 29).

An elevated expression of all these proteins has been correlated with malignant phenotypes in epithelial and fibroblastic rat thyroid cells, and in many experimental carcinomas (30–32). In addition, at least one of them, HMGI-C, has been found

\* This work was supported by grants from the Associazione Italiana per la Ricerca sul Cancro (AIRC), Progetto Finalizzato ACRO of the Consiglio Nazionale delle Ricerche (CNR), Università degli Studi di Udine, Università degli Studi di Trieste, and Ministero della Ricerca Scientifica e Tecnologica. 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.

‡ Recipient of a fellowship from the Fondazione Italiana per la Ricerca sul Cancro.

¶ To whom correspondence should be addressed: Dipartimento di Scienze e Tecnologie Biomediche Università di Udine, Via Gervasutta 48, 33100 Udine, Italy. Tel.: 39-432-547301; Fax: 39-432-547210; E-mail: giuseppe.damante@dstb.uniud.it.

<sup>1</sup> The abbreviation used are: HD(s), homeodomain(s); HMG, high mobility group; En, Engrailed; GST, glutathione S-transferase.

to be important for the retrovirally induced transformation of rat thyroid cells (33) and many groups have recently reported a correlation between rearrangements of the HMGI-C gene and some benign human mesenchymal tumors (34, 35). Furthermore, the inactivation of the *Hmgi-c* gene has been demonstrated to cause the pygmy phenotype in the mouse (36), postulating an important role of this protein in cell proliferation and differentiation.

The DNA binding specificity of HMGI proteins (preference for AT-rich sequences) would suggest that these molecules may interact with the 5'-TAAT-3'-containing sequences, specific HDs target sequences. In this way, HMGI proteins might interfere with the DNA binding function of HD-containing transcription factors and therefore influence the biological effect of these proteins.

In this study, we demonstrate that HMGI proteins are indeed able to interact with some HD binding sequences. Specifically, we show that the binding of HDs and HMGI proteins to 5'-TAAT-3' motif-containing sequences is mutually exclusive. In particular, the presence of HMGI proteins blocks HDs from binding to 5'-TAAT-3' sequences. Apparently this effect is mediated by structural modifications of the DNA induced by HMGI proteins and not only by a simple competition for the same target sequence. As a model to study the functional role of these interactions, we utilize the transcriptional activation of HOXD9 on an autoregulatory element (37). We show that HMGI is able to compete for the binding of HOXD9 to the DNA *in vitro* and to reduce the target-specific transcriptional activation of HOXD9 in transfected cells.

#### EXPERIMENTAL PROCEDURES

**Oligonucleotides**—The following oligonucleotides were used in gel-retardation assay experiments (only upper strand sequence shown): C<sub>Ant1</sub>, 5'-CACTGCCAGTTAATTGTTCTTGA-3'; PRDII, 5'-GGGAAATTCCGTGGGAAATCCGAGCT-3'; NP, 5'-TGATATTTAATTGATTTT-3'; C<sub>Ant4</sub>, 5'-CACTGCCAGTTAAGTGTCTTGA-3'; T109, 5'-GCACGCTAATTGAGACGC-3'; HCRII, 5'-GACACATTAATCTATAATCAAATAC-3'.

**Protein Expression and Purification**—The plasmid coding for Engrailed (En) has been already described (38). In this plasmid the transcription of the HD-coding sequence is driven by the T7 RNA polymerase promoter. The murine HMGI-C and HMGY coding sequences were amplified by polymerase chain reaction using the appropriate primers deduced from the amino and carboxyl termini, and cloned between the *NdeI* and *BamHI* sites of the bacterial expression vector pAR3038 under the bacteriophage T7 promoter (39). The resulting clones were verified by sequencing. All proteins were expressed using the BL21 (DE3) *Escherichia coli* strain, which contains the T7 RNA polymerase under lacUV5 promoter control (39). Cultures were grown to A<sub>600</sub> = 0.6, induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside, grown overnight, and harvested.

EnHD used for DNA-binding assays was partially purified by using Econo Pac S cartridges (Bio-Rad). The concentration of the active protein was measured by oligonucleotide saturation assay. A gel-retardation assay (see below) was performed with increasing amounts of oligonucleotides (from 0.3 to 50 nM), then the protein-bound and free oligonucleotide concentration values were subjected to a Scatchard plot analysis. The value of the active protein percentage over the total amount was usually 40–80%.

HMGI proteins were selectively extracted from bacterial cells with 5% (v/v) perchloric acid and precipitated with acetone-HCl as described previously (30). The proteins were purified by reverse-phase high performance liquid chromatography on a Bio-Rad RP304 column using a Waters apparatus as elsewhere described (32). The consistency between purified recombinant HMGI proteins and calculated molecular mass from sequence was checked by mass spectrometry (Perkin-Elmer API 1 spectrometer). The construct encoding GST-HOXD9 was described previously (40), and the GST fusion protein was expressed and purified according to standard procedures (41).

**Gel-retardation Assay and Quantitation of the Binding Activity**—Double-stranded oligonucleotides were labeled at the 5'-end with <sup>32</sup>P and used as probes in the gel-retardation assays. Gel-retardation assay was performed incubating protein and DNA in a buffer containing 20

mM Tris-HCl, pH 7.6, 75 mM KCl, 55 μg/ml bovine serum albumin, 3 mM dithiothreitol, 13% glycerol for 30 min at room temperature. Protein-bound DNA and free DNA were separated on native 7.5% polyacrylamide gel run in 0.5 × TBE (Tris borate-EDTA), 300 V at 4 °C. Gels were dried, exposed to x-ray films, and bands were quantitated by densitometric scanning of the autoradiogram using a LKB laser densitometer. The monoclonal antibody against GST protein was purchased from Santa Cruz Biotechnology and used at a final concentration of 10 μg/ml.

**Circular Dichroism and Structure Analysis**—Jacketed cells of 2–5 mm were used, and typically 10 spectra were accumulated, averaged, and base line-corrected with a spectrum of buffer alone on a Jasco J-600 spectropolarimeter interfaced to an Olidata personal computer. Calibration of the instrument was performed with D(+)-10-camphorsulfonic acid at 290 nm. All spectra were collected at 10 ± 0.1 °C. Thermostability was controlled by a Haake F3 water bath. Spectra were recorded in 10 mM sodium phosphate buffer, pH 7.5, KCl 50 mM.

**Cell Culture and Transfection**—The expression construct producing HOXD9, as well as pT81luc and pTHCR reporter constructs, were described elsewhere (40). The pSVHMGY expression construct was obtained subcloning a 1.7-kilobase pair *EcoRI* fragment containing the full-length murine HMGI cDNA (18) into the polylinker of pGDSV7 expression vector (42). NIH-3T3 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 100 units/ml penicillin, and 100 μg/ml streptomycin, at 37 °C in a humidified 5% CO<sub>2</sub> incubator. Cells were transfected by the calcium phosphate co-precipitation procedure (43). Briefly, cells were plated at a density of 0.5 × 10<sup>6</sup>/60-mm diameter culture dish. Precipitates, containing 1 μg of reporter plasmid, 7 μg of expression constructs, and 0.5 μg of pTK β-galactosidase expression vector as internal control, were applied to subconfluent cells for 16 h before changing medium. Total amounts of each plasmid were kept constant in all experiments, by adding the corresponding empty vectors.

**β-Galactosidase and Luciferase Assays**—48 h after transfection cells were harvested and lysed with Promega's lysis buffer. Luciferase enzyme assays were performed using the luciferase assay system (Promega) in a luminometer (Lumat LB 9501, EG & G Berthold) according to manufacturer's instructions; fluorimetric β-galactosidase assays were performed using 4-methylumbelliferyl-β-D-galactoside (MUG-Sigma) as substrate (44). Luciferase activity of the reporter vectors was then normalized for transfection efficiency using β-galactosidase activity.

#### RESULTS

**HMGI Proteins Recognize HD Binding Sequences**—To test whether HMGI proteins are able to interact with sequences recognized by HDs, the binding activity of recombinant HMGI-C and HMGY to several oligonucleotides containing the 5'-TAAT-3' motif was evaluated. Results are shown in Fig. 1. The oligonucleotide PRDII contains two copies of the human interferon β promoter HMGI(Y) binding sequence (28) and therefore provided a positive control. Both HMGI-C and HMGY efficiently recognize the oligonucleotide C<sub>Ant1</sub>, which is a dominant interaction site for both Antennapedia and EnHDs (45). The binding activity of both proteins requires the integrity of the 5'-TAAT-3' motif. In fact, the oligonucleotide C<sub>Ant4</sub>, in which the 5'-TAAT-3' motif has been changed to 5'-TAAG-3', is no longer bound by either HMGI-C or HMGY. Both proteins are also able to interact with the NP oligonucleotide, which is a binding site for several HD-containing proteins (Ref. 46 and references therein). HMGI proteins may recognize only a subset of HDs binding sequences. In fact, neither HMGI-C nor HMGY are able to establish an efficient interaction with the oligonucleotide T109, which is recognized by both Antennapedia and EnHDs (data not shown).

**Interference between HDs and HMGI Proteins in DNA Binding**—The evidence that HMGI proteins interact with some HD binding sequences would predict that this class of proteins and HDs may influence each other during the interaction to common sites. We have evaluated this possibility incubating together HDs and HMGI proteins with C<sub>Ant1</sub> oligonucleotide. Results are shown in Fig. 2. The presence of either HMGI-C or HMGY completely blocks the binding of EnHD.

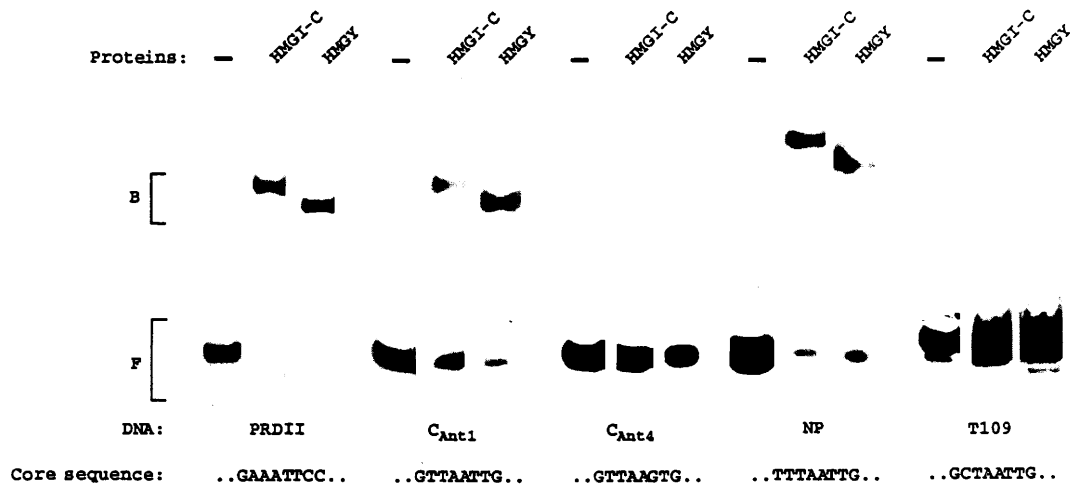


FIG. 1. HMGI proteins recognize HD binding sites. The full oligonucleotide sequences are shown under "Experimental Procedures." Here only the putative HMGI/HD binding sequence is shown. Oligonucleotides were labeled at the 5'-end with the same specific activity and used at a final concentration of  $0.05 \mu\text{M}$ . Both HMGY and HMGI-C proteins were used at the final concentration of  $0.6 \mu\text{M}$ . *F*, free DNA; *B*, protein bound to DNA.

It has been demonstrated recently that HMGI1, another member of the HMG family, is able to interact with HOX proteins in absence of DNA binding (40). Therefore, we have tested whether the effect of HMGI proteins on HD binding requires HMGI/DNA interaction. As shown in Fig. 3, HMGI-C is able to prevent the EnHD/Cant1, but not the EnHD/T109 interaction. These data, along with our experimental observation that HMGI-C is not able to efficiently recognize the T109 oligonucleotide (see Fig. 1), indicate that the HMGI/DNA interaction is necessary to block the EnHD binding to DNA. These results indicate that the effect of HMGI proteins on HD/DNA binding is not due to direct protein-protein interactions. It is important to note that the inhibitory effect is observed at a HMGI-C protein/EnHD molar ratio of 1.8. Since HMGI proteins are present in the cell in large amounts (47), the HMGI/EnHD molar ratio used should be lower than in the nucleus. This estimation suggests that the interference that we observe may occur *in vivo*.

**Mechanism by Which HMGI Proteins Inhibit HD Binding to DNA**—HDs contact DNA both in the major groove (by the recognition helix) and in the minor groove (by the  $\text{NH}_2$ -terminal arm). The inhibitory effect of HMGI proteins, which interact with DNA only by minor groove contacts, could occur through a complex mechanism comprised of two molecular events. First, HMGI proteins gain access to DNA competing with the HDs  $\text{NH}_2$ -terminal arm for the establishment of contacts in minor groove. Subsequently, the interaction of HMGI proteins modifies the DNA structure so that the HD/major groove interaction is less efficient, leading to the release of the whole HD from the DNA. The structural modifications of DNA which occur upon interaction with HMGI proteins could support the existence of the second mechanism (48–50). To test this hypothesis several experiments were performed. The gel, shown in A of Fig. 4, demonstrates that the inhibitory effect of HMGI-C on the EnHD binding to CAnt1 occurs with the same efficiency when HMGI-C is added before or after the formation of the EnHD-DNA complex. Kinetic experiments revealed that HMGI-C inhibits EnHD binding in less than 10 s (data not shown). Panel B of Fig. 4 shows that when HMGI-C protein is used at a concentration in which the EnHD-CAnt1 oligonucleotide complex is still detectable, the half-life of the EnHD-CAnt1 complex is greatly reduced relative to that observed in the absence of HMGI-C protein. These data are compatible with a mechanism in which HMGI-C binds to the minor groove more efficiently than the  $\text{NH}_2$ -terminal arm of EnHD, which is re-

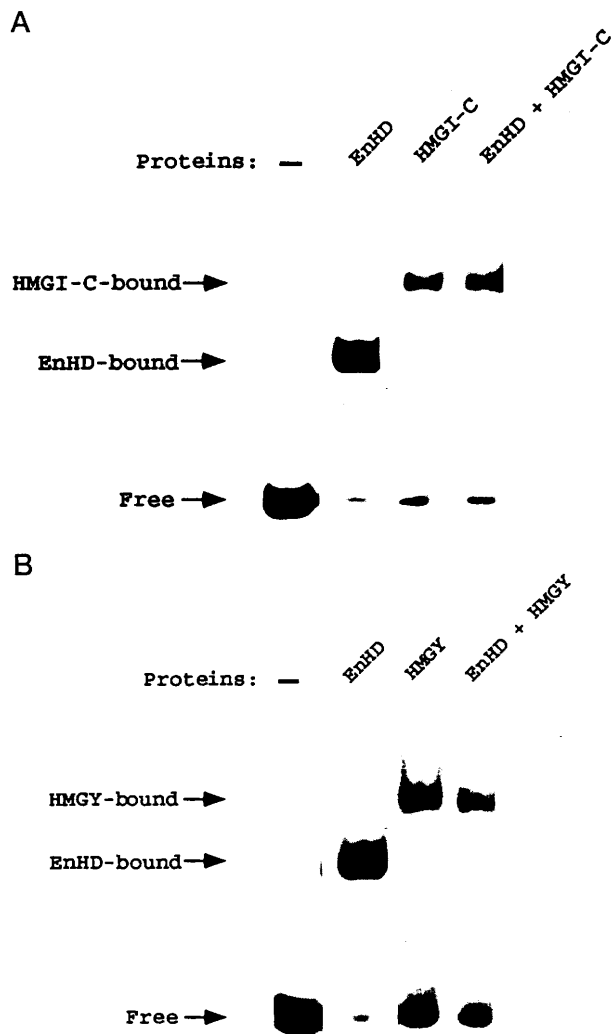


FIG. 2. The presence of HMGI proteins inhibits the binding of EnHD. The gel-retardation assay was performed as described under "Experimental Procedures." Proteins were used at the following final concentrations: HMGI-C,  $1.1 \mu\text{M}$ ; HMGY,  $1.1 \mu\text{M}$ ; EnHD,  $0.6 \mu\text{M}$ . *A*, effect of HMGI-C; *B*, effect of HMGY.

elled by the presence of HMGI-C. Consequently, we hypothesize that the HMGI-C binding to the minor groove influences the structure of the major groove in a way that EnHD is no

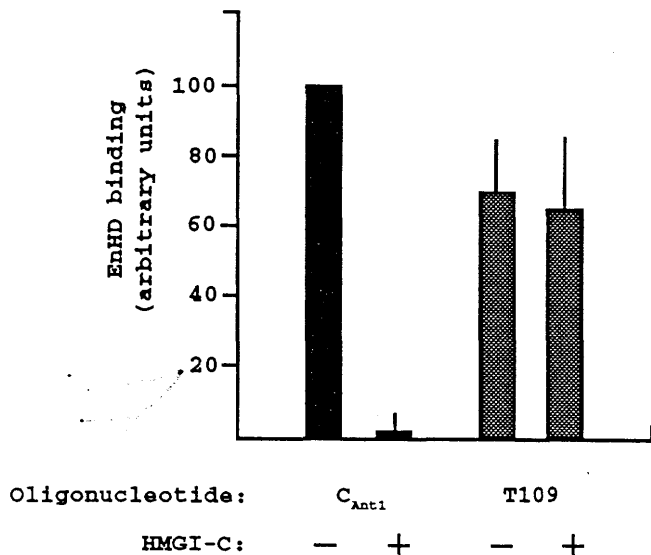


FIG. 3. The HMGI-C-DNA interaction is required for the inhibitory effect on HD-binding. Experimental conditions were those used in the experiments shown in Fig. 2. The values of EnHD binding to DNA have been obtained by densitometric scanning of gel-retardation assay autoradiograms. Columns and bars represent the mean values and the standard deviations, respectively, of three independent experiments.

longer able to bind to the DNA.

The drug berenil (Sigma) is able to interact with DNA into the minor groove but is not able to induce structural modifications on the bound DNA (51). Moreover, it has been demonstrated that, even at very high concentrations, this drug is not able to affect HDs/DNA interactions (52). Therefore, berenil should compete with HMGI-C for the minor groove interaction and reduce the inhibitory effect of this protein on HD binding to DNA. The effect of berenil on the inhibition of EnHD binding to DNA by HMGI-C protein is shown in Fig. 5. When used at a concentration of 800  $\mu$ M, berenil determines only a modest reduction of EnHD binding to the C<sub>Ant1</sub> sequence (compare lanes 2 and 4). However, the abolition of EnHD-binding by HMGI-C protein is significantly reduced by the presence of berenil (compare lanes 3 and 5). These results indicate that the establishment of a minor groove interaction without modification of the DNA structure (such as what occurs with berenil treatment) is not able to prevent HDs from binding to DNA, supporting the model that structural changes of DNA mediate the inhibitory effect of HMGI proteins on HD-DNA interaction.

To verify the structural modifications of DNA that could occur as a result of interaction with HMGI proteins, CD spectra of the HMGI-C-C<sub>Ant1</sub> complex were recorded and compared with those of the C<sub>Ant1</sub> oligonucleotide alone. Results are shown in Fig. 6. Following the interaction with HMGI-C the long-wavelength positive CD band of DNA decreases with a shift in the 258 nm crossover to longer wavelength. These modifications are likely due to significant structural alterations of the C<sub>Ant1</sub> oligonucleotide, since the CD spectrum of the protein shows no bands of absorption in the region above 240 nm (data not shown). Similar CD changes have been observed in compaction of the DNA into particles *in vitro*, into phage heads or into nucleosomes (53–55). While we have demonstrated that structural modifications of the C<sub>Ant1</sub> sequence occur upon HMGI-C binding, a more complete understanding of the reaction mechanism involved will necessitate higher resolution structural studies. Nevertheless, our CD data are compatible with the binding data shown in Figs. 4 and 5, suggesting that part of the HMGI-C inhibition of the EnHD-C<sub>Ant1</sub> interaction is due to a structural modification of the DNA.

*HMGI Is Able to Inhibit in Vivo HOXD9-mediated Tran-*

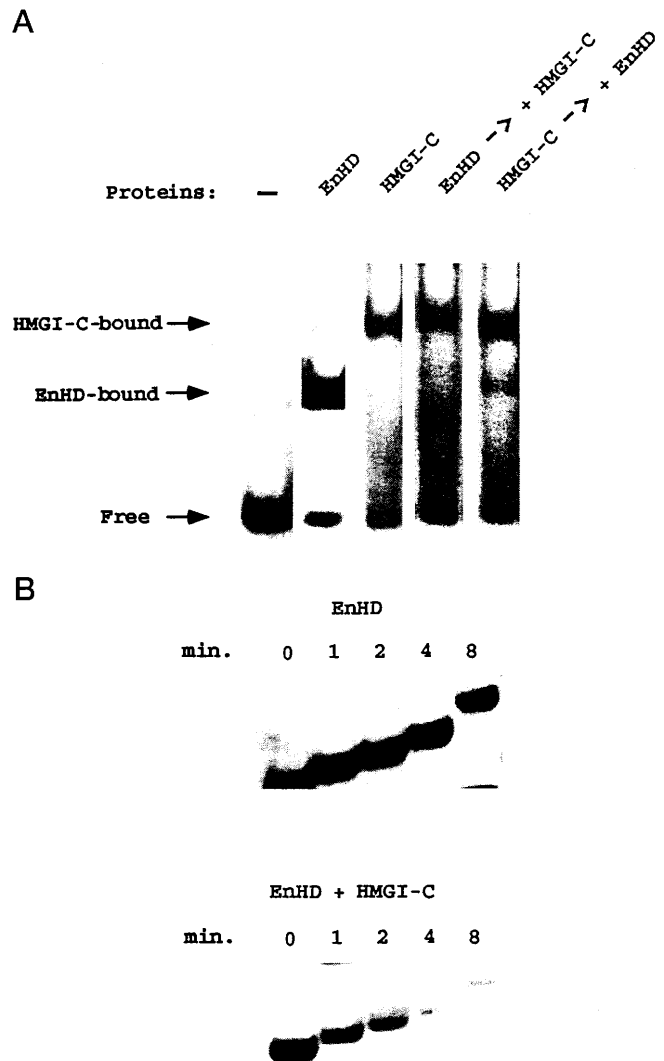


FIG. 4. Kinetics of HMGI-C inhibitory effect on HD-C<sub>Ant1</sub> interaction. **A**, gel-retardation assay demonstrating that the effect of HMGI-C on HD-DNA interaction occurs adding HMGI-C to DNA either before or after the HD-DNA complex formation. *EnHD*  $\rightarrow$  + *HMGI-C* indicates that the addition of HMGI-C occurs after formation of EnHD-DNA complex, while *HMGI-C*  $\rightarrow$  + *EnHD* indicates the reverse sequence of addition. **B**, half-life of the EnHD-C<sub>Ant1</sub> complex in absence or in presence of HMGI-C protein. EnHD and EnHD plus HMGI-C proteins were incubated for 30 min with the <sup>32</sup>P-labeled C<sub>Ant1</sub> oligonucleotide. Subsequently, a 100-fold excess of cold C<sub>Ant1</sub> oligonucleotide was added to the incubation mixture and then aliquots were loaded in a running native polyacrylamide gel at the indicated times.

*scriptional Activation* – To test whether the inhibitory effect of HMGI proteins on the HD-DNA interaction could have an *in vivo* relevance, we measured the effect of expression of HMGY on transcriptional activation induced by a HOX protein by co-transfection assay. As reported by Zappavigna *et al.* (37), the *HOXD9* gene product activates transcription binding to an autoregulatory element (HCR) deriving from its own promoter. The ~100-base pair HCR element contains several 5'-TAAT-3' motifs, which are binding sites for HOXD9 protein, possibly representing a target for a positive autoregulatory expression of HOX gene products. The oligonucleotide HCRII contains a 5'-TAAT-3' motif whose mutation reduces the transcriptional effect of HOXD9 (40). **A** of Fig. 7 demonstrates that HMGY protein is able to interact with HCRII (lane 4). The retarded band shows a mobility higher than that due to the interaction of the GST-HOXD9 fusion protein with the HCRII oligonucleotide (lane 2). When both HMGY and GST-HOXD9 proteins are incubated with the oligonucleotide HCRII, only the high mo-

bility band due to HMGY protein is detected (*lane 5*). To exclude that the retarded band observed in the presence of both HMGY and GST-HOXD9 is due to a trimeric complex containing both proteins, a co-incubation with a monoclonal antibody raised against the GST protein was performed. This antibody, which is able to induce a supershift when only the GST-HOXD9 protein is present (*lane 3*), is not able to produce any super-shifted band in the presence of both GST-HOXD9 and HMGY proteins (*lane 6*). This result demonstrates that the retarded band, observed in the presence of both HMGY and GST-HOXD9 proteins, is due only to the former. To test if there could be a functional effect of HMGY binding to the 5'-TAAT-3' motifs toward HOXD9-mediated activation of HCR, we then used an HCR-containing luciferase reporter: pTHCR (40), where the HCR element is fused to the herpes simplex virus thymidine kinase promoter. When pTHCR was co-transfected with different amounts of a construct expressing HOXD9 under the control of SV40 promoter, pSGD9, transcriptional activity was enhanced 4–5-fold over the basal level. Co-transfection of increasing amounts of pSGD9 in the presence of a fixed amount of pSVHMGY expressor construct, which produces HMGY protein under control of SV40 promoter, led to an inhibition, up to 50%, of HOXD9-mediated activation of HCR (Fig. 7B), while transfection of pSVHMGY alone did not significantly reduce

the basal activity of pTHCR. The control reporter pT81luc, lacking HCR sequence, was not affected by transfection of HOXD9 and HMGY constructs, alone or in combination.

## DISCUSSION

HMGI proteins do not require a precise sequence to efficiently interact with DNA, but they do require the presence of an AT-rich stretch (14–17). This functional flexibility in terms of DNA binding could be explained by the evidence that HMGI proteins contact DNA in the minor groove, where the steric/electrostatic differences between base pairs are much less evident than those existing in the major groove (56). Due to this DNA binding flexibility, it is not surprising that HMGI proteins are able to interact with DNA sequences that are binding sites for other DNA-binding proteins. Therefore, some of the effects of HMGI proteins on gene expression may be due to binding interference with other DNA-binding proteins, through a competition for common sites.

Based on our results, the inhibitory effect of HMGI proteins on HD-DNA interaction could be described by a model in which three essential events occur: (i) HMGI proteins are able to interact with a DNA sequence already occupied by a HD-containing protein, displacing the NH<sub>2</sub>-terminal arm of the HD from the minor groove; (ii) during the interaction, HMGI proteins induce conformational changes to the DNA; (iii) these conformational changes are not compatible with an efficient HD/DNA interaction in the major groove, therefore the whole HD-containing protein is released from the DNA. A similar model has been suggested to explain the effect of distamycin A on the Antennapedia HD-DNA complex (52). Interestingly, it has already been proposed that distamycin A and HMGI DNA binding domains have a similar planar crescent-shaped structure and can compete for the binding in the minor groove of AT-rich sequences (16). The observation that berenil, although interacting in minor groove, is not able to induce structural modifications to DNA and is not efficient in inhibiting the HD-DNA interaction, further supports our model. Crystallographic or NMR studies on HMGI proteins-DNA complexes are required to definitively prove the existence of HMGI protein-induced DNA modifications and to further test the validity of the model.

Although several experimental results indicate HMGI proteins as factors able to enhance the DNA binding of other DNA-interacting proteins (28–29), some data have been published showing an inhibitory effect of HMGI proteins on DNA binding of other factors. A clear example of this phenomenon

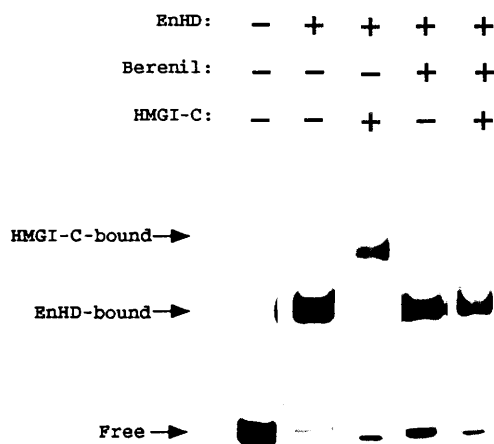


FIG. 5. Berenil blocks the inhibitory effect of HMGI-C on EnHD-DNA interaction. Gel-retardation assay was performed as described under "Experimental Procedures." Berenil (Sigma), HMGI-C, and EnHD were used at the concentration of 800, 0.48, and 0.25  $\mu\text{M}$ , respectively.

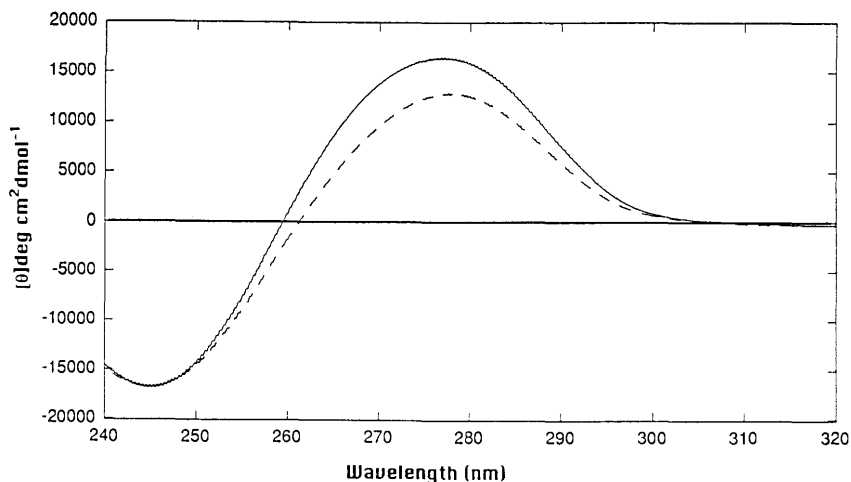
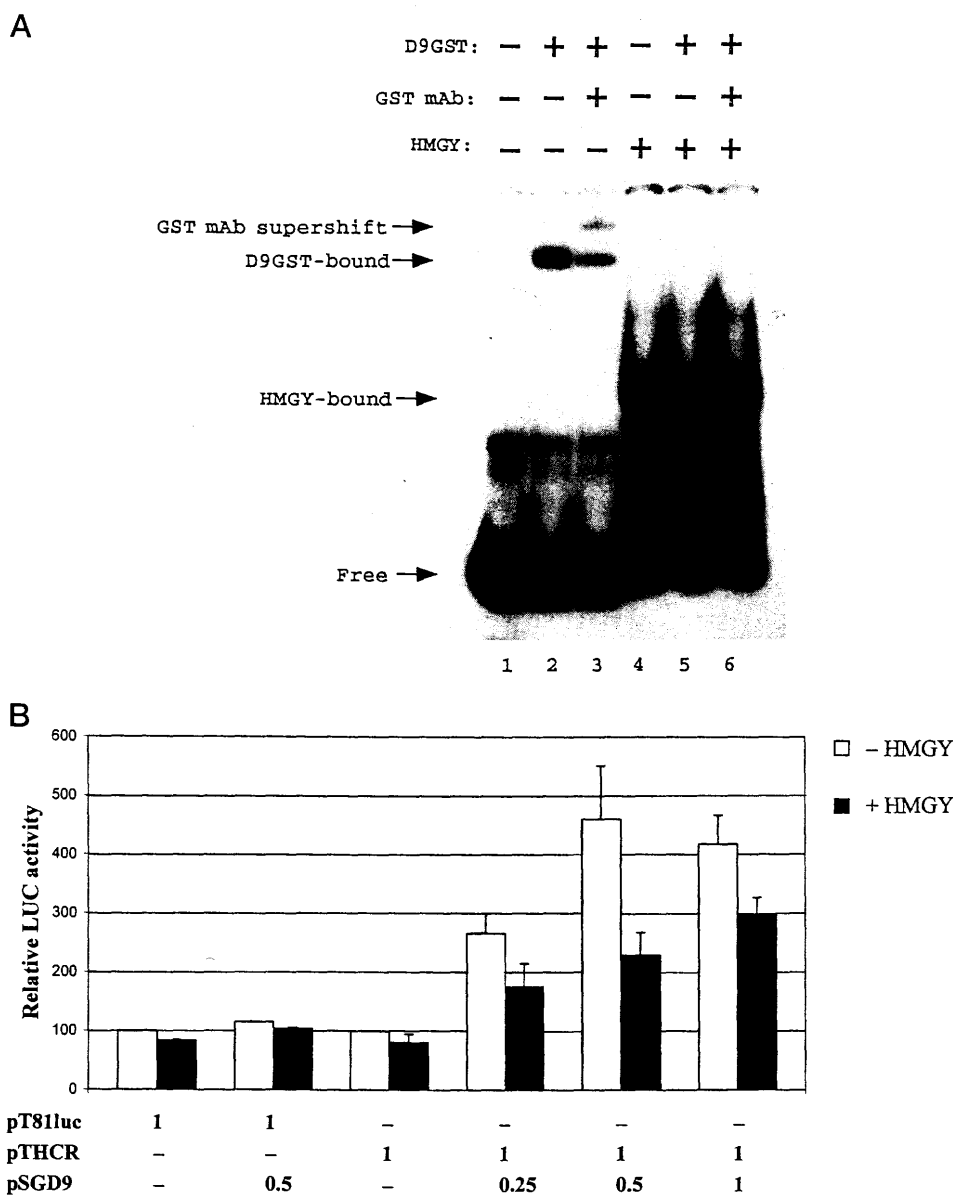


FIG. 6. Circular dichroism spectroscopy of HMGI-C binding to DNA. Near-UV CD spectrum of 2  $\mu\text{M}$  double-stranded CAnt1 oligonucleotide  $\pm$  2  $\mu\text{M}$  HMGI-C (1:1 molar ratio). The ellipticity values are given in terms of the molar concentration of oligonucleotides. The solid line represents the oligonucleotide alone, while the dashed line represents the oligonucleotide plus protein, after correction for the contribution of the protein.



**FIG. 7. HMGI protein interferes with HOXD9 binding to DNA and inhibits its transcriptional activity.** *A*, HMGI protein binds to the oligonucleotide HCRII and abolishes the binding of GST-HOXD9 fusion protein. Gel-retardation assay was performed as described under "Experimental Procedures" using double-stranded HCRII oligonucleotide as DNA probe. HCRII was used at a final concentration of  $0.02 \mu\text{M}$ , the final concentrations of HMGI and GST-HOXD9 proteins were  $0.5$  and  $0.1 \mu\text{M}$ , respectively. *B*, NIH-3T3 cells were transiently transfected with either pT81luc or with pTHCR along with pSGD9 expression vector. The amounts indicated are expressed in micrograms. The same experiment was done in the absence (*open columns*) and in the presence (*filled columns*) of a fixed amount ( $5 \mu\text{g}$ ) of pSVHMGI expression construct. In all experiments, total amounts of pSV and pSG vectors were kept constant by using pGDSV7 and pSG5 empty vectors, and  $0.5 \mu\text{g}$  of pTK  $\beta$ -galactosidase plasmid were co-transfected as internal standard. Transcriptional activity is indicated as percentage of that measured for pTHCR reporter alone. *Columns* and *bars* represent the means and the S.D. values, respectively, of at least four independent experiments.

occurs in T cells at the level of the interleukin-4 promoter (57). In fact, a high HMGI(Y) concentration, typical of peripheral blood T lymphocytes, is able to displace NF-AT factors from interleukin-4 promoter, reducing both its constitutive and inducible transcriptional activity. A negative effect of HMGI(Y) on the Ig heavy chain  $\epsilon$  germ line RNA promoter has been also reported (27). In this case, the inhibitory effect induced by HMGI(Y) appears to be mediated by the interaction of these proteins with a binding sequence for a STAT-like factor. A direct inhibition of Oct-1 protein DNA binding by HMGI(Y) has been also described (22). Interestingly, HMGI(Y) appears to exert distinct effects on the interaction of Oct-1 and Oct-2 proteins with the octamer sequence. In fact, while Oct-1 binding is inhibited, the presence of HMGI(Y) enhances the Oct-2/octamer sequence interaction. Such a difference may suggest that the final outcome of HMGI proteins on the transcriptional

activity of factors interacting with the same or nearby sequences is due to a combination of protein/DNA and protein/protein interactions. With regard to HD-containing proteins, Zappavigna *et al.* (40) have recently reported that HMG1 is able to interact with HOX proteins, enhancing their DNA binding and transcriptional activity. This effect is mediated only via protein-protein interaction, since HMG1 alone is not able to bind to the HOX target sequence. We instead, using the same cellular system and the same vectors, show that HMGI is able to compete with HOXD9, hence mediating repression. Therefore, although opposite effects of HMG proteins on the DNA binding activity of transcriptional regulators have been reported, taken together these data point out the relevance of these architectural factors in the regulation of gene transcription.

Several studies indicate that HD-containing proteins require

regulatory events for a proper action. For instance, Castelli-Gair and Akam (58) have demonstrated that the Ultrabithorax protein elicits distinct effects in different spatial-temporal contexts of *Drosophila* embryos. The ability of HMGI proteins to negatively affect the HD binding to some DNA target sequences could be one of the mechanisms by which the same HD-containing proteins are able to induce different effects depending on the context. One intriguing possibility is that HMGI proteins provide HD-containing proteins with a sort of regulatory switch. In fact, it has been demonstrated that the binding of HMGI proteins to DNA is regulated by phosphorylation (59, 60). In particular, phosphorylation of HMGI proteins induced by p34<sup>cdc2</sup> kinase greatly reduces their binding to DNA. Based on our experimental results, we hypothesize that the phosphorylation/dephosphorylation of HMGI proteins controls the access of HD-containing proteins to their DNA target sequences.

**Acknowledgments**—We are grateful to Vincenzo Zappavigna (DIBIT, Milano, Italy) for kindly providing the HOXD9 expression vectors as well as the reporter constructs. We thank Luca Cattarossi for assistance with computer work and Wayne Newhauser and Andrea Keane-Mayers for their editorial comments.

## REFERENCES

- Gehring, W. J., Affolter, M., and Burglin, T. (1994) *Annu. Rev. Biochem.* **63**, 487–526
- Gehring, W. J. (1987) *Science* **236**, 1245–1252
- Kaufman, T. C., Seeger, M. A., and Olsen, G. (1990) *Adv. Genet.* **27**, 309–362
- Pabo, C. O., and Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61**, 1053–1095
- Hirsch, J. A., and Aggarwal, A. K. (1995) *EMBO J.* **14**, 6280–6291
- Laughon, A. (1991) *Biochemistry* **30**, 11357–11367
- Ekker, S. C., Jackson, D. G., Von Kessler, D. P., Sun, B. I., Young, K. E., and Beachy, P. A. (1994) *EMBO J.* **13**, 3551–3560
- Damante, G., Fabbro, D., Pellizzari, L., Civitareale, D., Guazzi, S., Polycarpou-Schwartz, M., Cauci, S., Quadrifoglio, F., Formisano, S., and Di Lauro, R. (1994) *Nucleic Acids Res.* **22**, 3075–3083
- Chan, S. K., and Mann, R. S. (1993) *Genes Dev.* **7**, 796–811
- Sreenath, T., Pollock, R. A., and Bieberich, C. J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 9636–9640
- Stern, S., Tanaka, M., and Herr, W. (1989) *Nature* **341**, 624–630
- Chan, S. K., Jaffe, L., Capovilla, M., Botas, J., and Mann, R. S. (1994) *Cell* **78**, 603–615
- Bustin, M., and Reeves, R. (1996) *Prog. Nucleic Acid Res. Mol. Biol.* **54**, 35–100
- Solomon, M., Strauss, F., and Varshavsky, A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 1276–1280
- Maher, J. F., and Nathans, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6716–6720
- Reeves, R., and Nissen, M. S. (1990) *J. Biol. Chem.* **265**, 8573–8582
- Geierstanger, B. H., Volkman, B. F., Kremer, W., and Wemmer, D. E. (1994) *Biochemistry* **33**, 5347–5355
- Johnson, K. R., Lehn, D. A., Elton, T. S., Barr, P. J., and Reeves, R. (1988) *J. Biol. Chem.* **263**, 18338–18342
- Manfioletti, G., Giaccotti, V., Bandiera, A., Buratti, E., Sautiere, P., Cary, P., Crane-Robinson, C., Coles, B., and Goodwin, G. H. (1991) *Nucleic Acids Res.* **19**, 6793–6797
- Manfioletti, G., Rustighi, A., Mantovani, F., Goodwin, G. H., and Giaccotti, V. (1995) *Gene (Amst.)* **167**, 249–253
- Leger, H., Sock, E., Renner, K., Grummt, F., and Wegner, M. (1995) *Mol. Cell. Biol.* **15**, 3738–3747
- Abdulkadir, S. A., Krishna, S., Thanos, D., Maniatis, T., Strominger, J. L., and Ono, S. J. (1995) *J. Exp. Med.* **182**, 487–500
- Whitley, M. Z., Thanos, D., Read, M. A., Maniatis, T., and Collins, T. (1994) *Mol. Cell. Biol.* **14**, 6464–6475
- Lewis, H., Kaszubska, W., DeLamar, J. F., and Whelan, J. (1994) *Mol. Cell. Biol.* **14**, 5701–5709
- John, S., Reeves, R. B., Lin, J. X., Child, R., Leiden, J. M., Thompson, C. B., and Leonard, W. J. (1995) *Mol. Cell. Biol.* **15**, 1786–1796
- Chuvpilo, S., Schomberg, C., Gerwig, R., Heinfli, A., Reeves, R., Grummt, F., and Serfling, E. (1993) *Nucleic Acids Res.* **21**, 5694–5704
- Kim, J., Reeves, R., Rothman, P., and Boothby, M. (1995) *Eur. J. Immunol.* **25**, 798–808
- Thanos, D., and Maniatis, T. (1992) *Cell* **71**, 777–789
- Du, W., Thanos, D., and Maniatis, T. (1993) *Cell* **74**, 887–898
- Giaccotti, V., Berlingieri, M. T., DiFiore, P. P., Fusco, A., Vecchio, G., and Crane-Robinson, C. (1985) *Cancer Res.* **45**, 6051–6057
- Giaccotti, V., Pani, B., D'Andrea, P., Berlingieri, M. T., Di Fiore, P. P., Fusco, A., Vecchio, G., Philp, R., Crane-Robinson, C., Nicolas, R. H., and Goodwin, G. H. (1987) *EMBO J.* **6**, 1981–1987
- Giaccotti, V., Buratti, E., Perissin, L., Zorzet, S., Balmain, A., Portella, G., Fusco, A., and Goodwin, G. H. (1989) *Exp. Cell Res.* **184**, 538–545
- Berlingieri, M. T., Manfioletti, G., Santoro, M., Bandiera, A., Visconti, R., Giaccotti, V., and Fusco, A. (1995) *Mol. Cell. Biol.* **15**, 1545–1553
- Ashar, H. R., Fejzo, M. S., Tkachenko, A., Zhou, X., Fletcher, J. A., Weremowicz, S., Morton, C. C., and Chada, K. (1995) *Cell* **82**, 57–65
- Schoenmakers, E. F., Wanschura, S., Mols, R., Bullerdiek, J., Van den Berghe, H., and Van de Ven, W. J. (1995) *Nat. Genet.* **10**, 436–444
- Zhou, X., Benson, K. F., Ashar, H. R., and Chada, K. (1995) *Nature* **376**, 771–774
- Zappavigna, V., Renucci, A., Izpisua-Belmonte, J. C., Urier, G., Peschle, C., and Duboule, D. (1991) *EMBO J.* **10**, 4177–4187
- Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., and Desplan, C. (1993) *Genes Dev.* **7**, 2120–2134
- Studier, F. W., and Moffatt, B. A. (1986) *J. Mol. Biol.* **189**, 113–130
- Zappavigna, V., Falciola, L., Catterich, M. H., Mavilio, F., and Bianchi, M. E. (1996) *EMBO J.* **15**, 4981–4991
- Smith, D. B., and Johnson, K. S. (1988) *Gene (Amst.)* **67**, 31–40
- Del Sal, G., Manfioletti, G., Gustincich, S., Ruaro, E., and Schneider, C. (1994) *BioTechniques* **16**, 134–138
- Graham, F. L., and Van der Eb, A. J. (1973) *Virology* **52**, 456–467
- Jain, V. K., and Magrath, I. T. (1991) *Anal. Biochem.* **199**, 119–124
- Damante, G., Pellizzari, L., Esposito, G., Fogolari, F., Viglino, P., Fabbro, D., Tell, G., Formisano, S., and Di Lauro, R. (1996) *EMBO J.* **15**, 4992–5000
- Hayashi, S., and Scott, M. P. (1990) *Cell* **63**, 883–894
- Giaccotti, V., Bandiera, A., Ciani, L., Santoro, D., Crane-Robinson, C., Goodwin, G. H., Boiocchi, M., Dolcetti, R., and Casetta, B. (1993) *Eur. J. Biochem.* **213**, 825–832
- Nissen, M. S., and Reeves, R. (1995) *J. Biol. Chem.* **270**, 4355–4360
- Falvo, J. V., Thanos, D., and Maniatis, T. (1995) *Cell* **83**, 1101–1111
- Brown, D. G., Sanderson, M. R., Skelly, J. V., Jenkins, T. C., Brown, T., Garman, E., Stuart, D. I., and Neidle, S. (1990) *EMBO J.* **9**, 1329–1334
- Gray, D. M., Taylor, T. N., and Lang, D. (1978) *Biopolymers* **17**, 145–156
- Dorn, A., Affolter, M., Muller, M., Gehring, W. J., and Leupin, W. (1992) *EMBO J.* **11**, 279–286
- Dorman, B. P., and Maestre, M. F. (1973) *Proc. Natl. Acad. Sci. U. S. A.* **70**, 255–266
- Ivanov, V. I., Minchenkova, L. E., Minyat, E. E., Frank-Kamenetskii, M. D., and Schyolkina, A. K. (1974) *J. Mol. Biol.* **87**, 817–833
- Mandel, R., and Fasman, G. D. (1976) *Nucleic Acids Res.* **3**, 1839–1848
- Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73**, 804–808
- Klein-Henling, S., Schneider, C., Heinfli, A., Chuvpilo, S., and Serfling, E. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15311–15316
- Castelli-Gair, J., and Akam, M. (1995) *Development (Camb.)* **121**, 2973–2982
- Nissen, M. S., Langan, T. A., and Reeves, R. (1991) *J. Biol. Chem.* **266**, 19945–19952
- Reeves, R., Langan, T. A., and Nissen, M. S. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 1671–1675

# NF- $\kappa$ B mediated transcriptional activation is enhanced by the architectural factor HMGI-C

Fiamma Mantovani, Sonia Covaceuszach, Alessandra Rustighi, Riccardo Sgarra, Carol Heath<sup>1</sup>, Graham H. Goodwin<sup>1</sup> and Guidalberto Manfioletti\*

Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Università di Trieste, via Giorgieri, 1 34127 Trieste, Italy and <sup>1</sup>The Institute of Cancer Research, Haddow Laboratories, Sutton, UK

Received November 28, 1997; Revised and Accepted February 2, 1998

## ABSTRACT

High mobility group 1 proteins (HMGI, HMGY and HMGI-C) are a family of low molecular mass non-histone nuclear proteins which constitute an important component of the active chromatin structure. Two members of this family, HMGI and HMGY, have been demonstrated to contribute to the transcriptional regulation of several promoters by interacting with the DNA and with different transcription factors. On the contrary, very little is known about the third member, HMGI-C, which plays an important role during embryonic growth and in the process of cell transformation, its gene being rearranged in a large number of mesenchymal tumors. In this paper we show for the first time that HMGI-C is also able to function as architectural factor, enhancing the activity of a transcription factor, NF- $\kappa$ B, through the PRDII element of the  $\beta$ -interferon enhancer. Moreover we show that this enhancement is absolutely dependent on the binding of HMGI-C to its target sequence. The demonstration that HMGI-C is able to modulate transcription is thus an important initial step in the identification of genes regulated by this factor.

## INTRODUCTION

The high mobility group proteins HMGI-C, HMGI and HMGY belong to a family of non-histone nuclear proteins collectively termed as HMGI proteins. These architectural factors are able to modulate DNA conformation and are involved in fundamental processes such as transcription and recombination (1). Although encoded by a different gene (2), HMGI-C shares some structural homologies with HMGI and HMGY, which are produced by differential splicing from a single gene, and therefore also referred as HMGI(Y) (3). All three proteins in fact have three short basic domains termed AT-hooks which recognise AT-rich DNA sequences via the minor groove, and highly acidic C-terminal tails (1).

HMGI expression levels are very low or undetectable in adult tissues, but after oncogenic transformation their expression is up-regulated reaching levels comparable to those observed during early embryo development (4,5). The existence of a correlation between the expression levels of HMGI proteins and the degree of malignant transformation has suggested their use as tools for tumor diagnosis (6,7).

Direct evidence that HMGI-C participates in the oncogenic process was first provided when the expression of antisense HMGI-C RNA was shown to prevent retrovirally induced neoplastic transformation of rat thyroid cells (8). Moreover, the human HMGI-C gene, mapped to a 'multiple aberration region' (MAR) on chromosome 12, has been reported to be a target for translocations in eight different benign solid tumor types (9-11) and to be amplified and rearranged in a number of malignant sarcomas (12). Because of the extremely high incidence of these tumors, this rearrangement can be considered one of the most frequent aberrations in human tumors (13).

It has been shown that disruption of the *Hmgi-c* gene in mice results in growth retardation and *pygmy* phenotype characterized by a marked inhibition of adipose tissue development (4). Since HMGI(Y) expression remains unaltered in these mice, this implies that HMGI(Y) cannot compensate for the lack of HMGI-C, suggesting that their functions are not redundant and therefore implying a specific role for each of the HMGI proteins. It could therefore be possible that HMGI-C regulates a different set of genes than HMGI(Y).

Although increasing evidence is accumulating on the biological relevance of the HMGI-C protein in proliferation and development, very little is known about its biochemical properties, such as its DNA-binding specificity and potentiality to modulate transcription. Its homology to HMGI(Y) suggests that it could have a similar function as architectural transcription factor, but despite the increasing number of gene regulatory elements where HMGI(Y) have been shown to cooperate with other proteins to modulate transcription (positively or negatively) (14-20), there are still no data showing that HMGI-C is able to act in a similar way.

In this paper we start to address this point by asking whether HMGI-C is able to modulate transcription. To this aim we have used a simple system consisting of the well characterised PRDII element of the  $\beta$ -interferon promoter (14,21,22). Previous work by Thanos and co-workers has shown that the HMGY protein binds to this element together with the p50/p65 NF- $\kappa$ B heterodimer to activate transcription from the promoter. In this paper we demonstrate that HMGI-C is able to bind the PRDII element with similar affinity as HMGY and that it can enhance the activating properties of the transcription factor NF- $\kappa$ B. Moreover, we demonstrate that this enhancement is dependent on the binding of the protein to its target sequence.

\*To whom correspondence should be addressed. Tel: +39 40 6763675; Fax: +39 40 6763694; Email: manfiole@univ.trieste.it



## MATERIALS AND METHODS

### Oligonucleotides

The following oligonucleotides were used in gel-retardation assays and for reporter constructs (only the upper strand sequence is shown): PRDII, 5'-GGGAAATTCCGTGGGAAATTCCGAGCT-3'; mPRDII, 5'-GGGAGATTCCGTGGGAGATTCCGAGCT-3'; I $\kappa$ B, 5'-GGGACTTTCGTGGGACTTTCGAGCT-3'. For the band-shift analysis of the effect of HMGI-C on the binding of the NF- $\kappa$ B to the PRDII element, a shorter oligonucleotide (16mer), containing only one PRDII element, was used.

### Protein expression and purification

Recombinant HMGI proteins were expressed as previously described (20) using the murine HMGI-C and HMGY cDNAs cloned in the vector pAR3038, expression being under control of the bacteriophage T7 promoter. Briefly, proteins were expressed using the BL21 (DE3) *Escherichia coli* strain which contains the T7 RNA polymerase under lacUV5 promoter control. Cultures were grown to OD<sub>600</sub> = 0.6, induced with 0.5 mM isopropyl  $\beta$ -D-thiogalactopyranoside, grown overnight and harvested. HMGI proteins were selectively extracted from bacterial cells with 5% (v/v) perchloric acid and precipitated with acetone-HCl (23). The proteins were purified by reverse-phase HPLC on a BioRad RP304 column using a Waters apparatus as described elsewhere (24). The purified recombinant HMGI proteins were analyzed by mass spectrometry (Perkin-Elmer API 1 spectrometer) to confirm that they had the correct molecular masses. The NF- $\kappa$ B subunits, p50 and p65, were provided by Dr Ron Hay (University of St Andrews). These subunits were expressed as GST fusion proteins. After purification the proteins were cleaved with thrombin and further purified on a DNA-Sephrose column. The p50/p65 heterodimer was formed by mixing the subunits in an equimolar ratio and incubating for 30 min at 37°C, followed by the addition of 20 mM DTT and incubation on ice for 15 min. Protein concentrations were determined by the Bradford assay using a known protein mixture as a standard.

### Band-shift analyses

Double-stranded oligonucleotides were labeled at the 5'-end with [ $\gamma$ -<sup>32</sup>P]ATP and 10 U of T4 polynucleotide kinase (Epicentre Technologies) or by filling in the 5' overhang with Klenow enzyme and [ $\alpha$ -<sup>32</sup>P]dCTP. Unincorporated nucleotides were removed by chromatography through a Sephadex G-50 (Pharmacia) spun-column equilibrated in TEN buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 100 mM NaCl). One hundred fmol of each probe (30 000 c.p.m.) were then separately incubated with increasing concentrations of HMGY or HMGI-C protein (5–200 ng, as indicated) in 20  $\mu$ l reactions containing 20 mM Tris-HCl, pH 7.5, 75 mM KCl, 5  $\mu$ g/ $\mu$ l BSA, 1 mM DTT, 13% glycerol for 20 min at room temperature. After incubation, protein-bound DNA and free DNA were separated on native 7% polyacrylamide gel run in 0.5 $\times$  TBE, at 15 V/cm at 4°C. Gels were then dried and exposed to X-ray films (Hyperfilm MP, Amersham). For band-shifts of NF- $\kappa$ B and HMGI-C, 12.5 ng of HMGI-C and a range from 50 to 0.5 ng of NF- $\kappa$ B were used.

### Plasmid construction

Basic TKluc reporter vector was constructed as follows. A *Bam*HI-*Xba*I fragment, containing the Herpes Simplex Virus (HSV) Thymidine Kinase (TK) promoter, was obtained from pT109luc plasmid (25) and inserted into pGL2basic luciferase reporter vector (Promega), cleaved with *Bgl*II-*Xba*I. PRDII TKluc, mPRDII TKluc and I $\kappa$ B TKluc reporter plasmids were constructed by inserting the corresponding double-stranded (ds) oligonucleotides listed above into the *Sma*I-*Sac*I sites of BasicTKluc, upstream of the TK promoter.

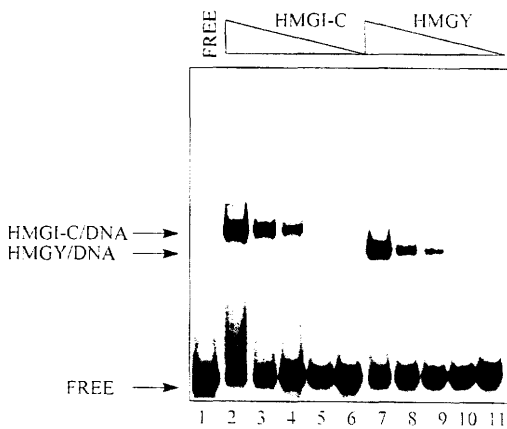
The pSVHMGY expression construct containing the full-length murine HMGY cDNA under the control of SV40 promoter has been described previously (20). pSVHMGI-C was constructed by inserting into the pGDSV7 expression vector (26) a 1.8 kb *Eco*RI-*Hind*III fragment containing the full-length murine HMGI-C cDNA (2). cDNAs for p50 and p65 NF- $\kappa$ B components, kindly provided by Dr Neil Perkins (Dundee University), were under control of the RSV promoter (27).

### Cell culture and transfections

NIH-3T3 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum at 37°C in a humidified 5% CO<sub>2</sub> incubator. PC Cl3 and FRTL-5 Cl2 cells were maintained in Coon's modified Ham's F12 medium supplemented with 5% fetal calf serum and six growth factors: the thyrotropic hormone (TSH), insulin, transferrin, somatostatin, hydrocortisone and the tripeptide glycyl-histidyl-lysine as described (28). Cells were plated at a density of 0.5  $\times$  10<sup>6</sup> cells per 60 mm diameter culture dish and transfected by the calcium phosphate co-precipitation procedure. For NIH-3T3 transfection, precipitates containing 2  $\mu$ g of reporter plasmid, 5.2  $\mu$ g of expression constructs and 0.4  $\mu$ g of pSV  $\beta$ -galactosidase expression vector (Promega) as internal control, were applied to subconfluent cells 16 h before changing medium. Total amounts of each expression plasmid were kept constant in all the experiments by adding the corresponding empty vectors. For PC Cl3 and FRTL-5 Cl2 transfection, precipitates were left on cells for 1 h, followed by glycerol shock treatment and medium replacement. Due to the lower transfection efficiency of thyroid cells, internal control was provided by adding 0.25  $\mu$ g of pRL-TK *Renilla* luciferase expression vector (Promega), which allows a more sensitive detection of enzyme activity.

### $\beta$ -Galactosidase and luciferase assays

After transfection (48 h), cells were harvested and lysed with lysis buffer (Promega). For NIH-3T3 cells, firefly luciferase enzyme assays were performed using the Luciferase Assay System (Promega) in a luminometer (Lumat LB 9501, EG&G Berthold) according to the manufacturer's instructions. Colorimetric  $\beta$ -galactosidase assays were performed using *o*-nitrophenol- $\beta$ -D-galactoside (ONPG-Sigma) as substrate. Luciferase activity of the reporter vector was then normalized for transfection efficiency using  $\beta$ -galactosidase activity. For PC Cl3 and FRTL-5 Cl2 cells, reporter (firefly luciferase) and control (*Renilla* luciferase) enzyme assays were performed using the Dual Luciferase Reporter Assay System (Promega).



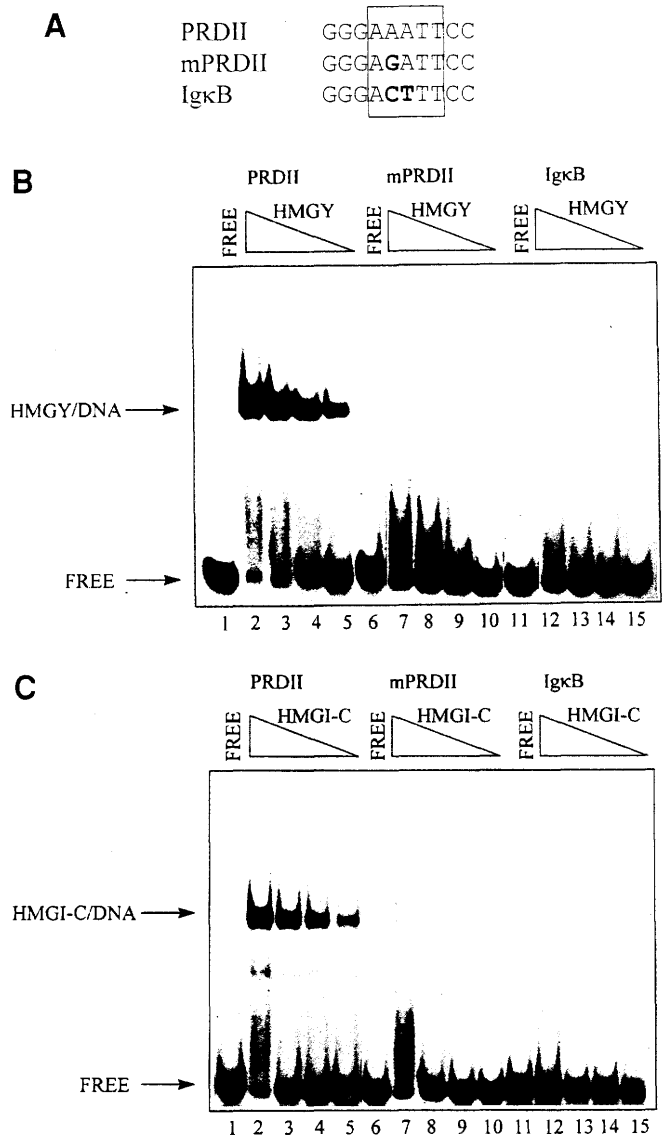
**Figure 1.** HMGI-C and HMGY bind to PRDII with the same affinity. The autoradiogram of an EMSA analysis is shown, in which binding affinity of HMGI-C and HMGY to the PRDII oligonucleotide was compared. <sup>32</sup>P-labeled ds-PRDII oligonucleotide was incubated in separate reactions with corresponding amounts of the two purified recombinant proteins. Lane 1, free probe; lanes 2–6, decreasing amounts (100, 50, 25, 10 and 5 ng, respectively) of HMGI-C; lanes 7–11, decreasing amounts (100, 50, 25, 10 and 5 ng, respectively) of HMGY.

**RESULTS**

To investigate if HMGI-C could play a similar role as HMGY in the assembly of a complex able to activate transcription from a combinatorial promoter, the well characterized model of the human β-interferon enhancer was chosen. Previous work has shown that HMGY plays a structural role by facilitating the binding of different transcription factors to their consensus sequences and this contributes to the viral inducibility of the entire enhancer (14). Part of the enhancer consists of the PRDII element, which binds NF-κB in the major groove and HMGY in the minor groove of the central AT-rich region (14). Binding of the HMGY protein has been shown to potentiate the transcriptional activation by the NF-κB protein and we have sought to investigate whether the HMGI-C protein will function in a similar manner.

**HMGI-C binds to the PRDII element *in vitro* with the same affinity as HMGY**

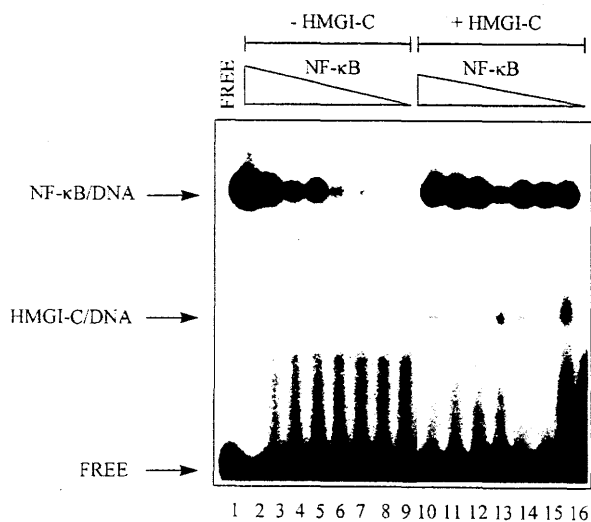
As a first step, band-shift analyses were performed to test the ability of HMGI-C to interact with the PRDII element. A ds-PRDII oligonucleotide, which contains two copies of the PRDII element, was incubated with different concentrations of HMGY and HMGI-C recombinant proteins and subjected to EMSA. As shown in Figure 1, incubation with either HMGY or HMGI-C produces a protein–DNA complex, whose mobility is higher for HMGY than for HMGI-C (in accordance with the different molecular weights of the proteins), and whose intensity increases with protein concentration. Interestingly, no significant differences of intensity can be noted when comparing identical concentrations of HMGY and HMGI-C, suggesting a very similar binding affinity of the two proteins for the PRDII sequence. It should be noted that dimer band-shifts are not observed; this could be because binding of one HMGI molecule to one AT element precludes the binding of a second molecule or, more likely, because of the stronger binding of one HMGI molecule utilising two AT-hooks to occupy both AT elements simultaneously, as suggested for DNA sequences with multiple AT tracts (22,29).



**Figure 2.** HMGY and HMGI-C bind to the wild-type PRDII site but not to the mutated mPRDII or IgκB sites. (A) Comparison of PRDII, mPRDII and IgκB sequences. The bases which are different in mPRDII and IgκB with respect to PRDII are indicated in bold character, the central region recognized by HMGI proteins is boxed. (B) Binding of HMGY to PRDII, mPRDII and IgκB oligonucleotides. EMSA analyses were performed incubating decreasing amounts of recombinant HMGY protein with either PRDII, mPRDII or IgκB labelled oligonucleotide probes. Lanes 1, 6 and 11, free PRDII, mPRDII and IgκB probes; lanes 2–5, 7–10 and 12–15, PRDII, mPRDII and IgκB incubated with 200, 100, 50 and 25 ng of HMGY, respectively. (C) Binding of HMGI-C to PRDII, mPRDII and IgκB oligonucleotides. EMSAs were performed incubating decreasing amounts of recombinant HMGI-C protein with the same probes as in (B). Lanes 1, 6 and 11, free PRDII, mPRDII and IgκB probes; lanes 2–5, 7–10 and 12–15, PRDII, mPRDII and IgκB incubated with 200, 100, 50 and 25 ng of HMGI-C, respectively.

**Disruption of the AT-rich core of PRDII abolishes the binding of both HMGY and HMGI-C**

In order to test the sequence-specificity of HMGI-C binding to the PRDII element, two different oligonucleotides, mPRDII and IgκB, were synthesized where the AT-rich central core of PRDII is disrupted. As shown in Figure 2A, mPRDII differs from PRDII



**Figure 3.** HMGI-C enhances the binding of NF- $\kappa$ B to the PRDII element. The autoradiogram of an EMSA analysis is shown where a range of concentrations of NF- $\kappa$ B were incubated with the PRDII oligonucleotide in the absence or in the presence of 12.5 ng of HMGI-C. Lane 1, free probe; lanes 2–9, decreasing amounts (50, 1.6, 1.3, 1.0, 0.8, 0.7, 0.6 and 0.5 ng, respectively) of NF- $\kappa$ B without HMGI-C. Lanes 10–16, decreasing amounts of NF- $\kappa$ B, as in lanes 3–9, with 12.5 ng of HMGI-C.

by a single base pair substitution (A–G), while Ig $\kappa$ B oligonucleotide bears the NF- $\kappa$ B consensus found in the immunoglobulin light chain gene enhancer. Both sequences have been shown to have lower HMGY binding affinity when compared to PRDII, while the binding of NF- $\kappa$ B is not significantly affected (14). As shown in Figure 2B, band-shift analyses in which corresponding amounts of HMGY protein were incubated with either PRDII, mPRDII or Ig $\kappa$ B probes confirmed this finding, since only the wild-type PRDII oligonucleotide gives a specific band-shift, while neither mPRDII nor Ig $\kappa$ B show detectable retarded bands, although at high protein concentration, some weak and unstable complexes are generated in the case of mPRDII, indicating that a labile interaction can occur under these conditions.

When the same experiment was performed with HMGI-C (Fig. 2C), a similar behaviour was observed, since this protein also fails to bind both mPRDII and Ig $\kappa$ B probes at the concentrations which produce a strong shift on the wild-type sequence. Again, a weak binding to the mPRDII sequence can be observed only at high protein concentrations.

#### HMGI-C enhances the ability of NF- $\kappa$ B to bind to the PRDII element *in vitro*

To investigate the possibility that HMGI-C, like HMG(Y), is able to enhance the binding of NF- $\kappa$ B to the PRDII element, band-shift experiments were carried out using similar conditions to those described by Thanos and Maniatis (14). The PRDII oligonucleotide was incubated with decreasing amounts of recombinant NF- $\kappa$ B heterodimer with and without a constant amount of recombinant HMGI-C protein. The results in Figure 3 show that at low concentrations of NF- $\kappa$ B, the HMGI-C protein strongly stimulates the formation of the band-shift complex (compare lanes 6–9 and 13–16). The effect seen is similar to that described by Thanos and Maniatis (14) for HMG(Y).

#### Both HMGY and HMGI-C are able to enhance NF- $\kappa$ B mediated transcriptional activation through the PRDII element

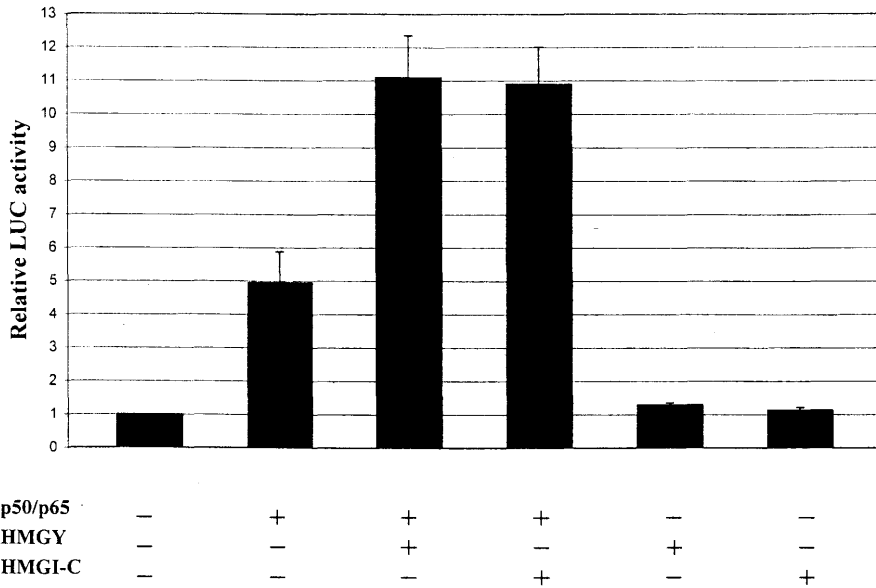
Having demonstrated that HMGI-C, like HMGY, is capable of sequence-specific binding to the IFN- $\beta$  PRDII element and enhances the binding of NF- $\kappa$ B *in vitro*, it was of interest to verify if it could also have a similar function in cooperating *in vivo* with NF- $\kappa$ B in transcriptional activation through this enhancer element. A reporter construct, bearing an artificial enhancer containing two copies of PRDII, was then generated by subcloning the PRDII oligonucleotide, previously employed in band-shift assays, upstream of the HSV TK promoter driving the expression of the firefly luciferase gene. This reporter, named PRDII TKluc, was first assayed for transactivation by NF- $\kappa$ B in the differentiated rat epithelial cell line PC Cl3. This cell line was chosen because of its total lack of endogenous HMGI expression (30); it was thus possible to establish the transactivation levels achieved solely by NF- $\kappa$ B. Cotransfection of a fixed amount of reporter vector with increasing amounts of an equimolar mixture of p50 and p65 expression vectors led to a linear increase (up to 14-fold) of luciferase activity (data not shown). The PRDII TKluc reporter was then cotransfected with a fixed amount of p50/p65 in order to achieve a transactivation level of ~5-fold, and the experiment was performed in the presence of either the HMGY or HMGI-C expression vectors, or the empty SV40 expression vector. Cotransfection of HMGY resulted in a further increase of luciferase activity, up to 11–12-fold over the basal level (Fig. 4), confirming the results obtained by other groups, with different constructs and cellular systems (22,31). Interestingly, very similar levels of enhancement were observed also for the HMGI-C protein (Fig. 4). These results demonstrate that, at least in this model system, HMGI-C is able to act as a factor enhancing the activity of a transcription factor, NF- $\kappa$ B.

Neither of the two HMGI proteins is capable of direct transactivation, since transfection of either HMGY or HMGI-C with PRDII TKluc in the absence of p50/p65 expressors did not produce any detectable effect on luciferase activity (Fig. 4). The control reporter vector Basic TKluc, lacking PRDII element, was not affected by cotransfection with the expression vectors in all combinations, demonstrating that the effects are specific for the PRDII element (not shown).

NIH-3T3 cells, which express endogenous HMGI proteins (32, and our observations), were also used in the assay, cotransfecting PRDII TKluc reporter with NF- $\kappa$ B and HMGY or HMGI-C plasmids. Figure 5 shows that in this different system overexpression of these two proteins is able to enhance transcription to similar levels.

#### Mutations that decrease the binding affinity of HMGY and HMGI-C to the PRDII sequence *in vitro* abolish HMGY/HMGI-C-dependent transcription enhancement *in vivo*

To determine whether binding to the AT-rich core was required for HMGI proteins to facilitate transactivation of the PRDII by NF- $\kappa$ B, two reporter constructs were generated in which the oligonucleotides mPRDII and Ig $\kappa$ B, which are not bound by HMGI proteins as demonstrated above, were inserted upstream of the TK promoter. The constructs mPRDII TKluc and Ig $\kappa$ B TKluc, as well as the wild-type reporter PRDII TK luc, were then cotransfected with p50/p65 expression vectors alone or in combination with either HMGY or HMGI-C. Figure 6 shows the

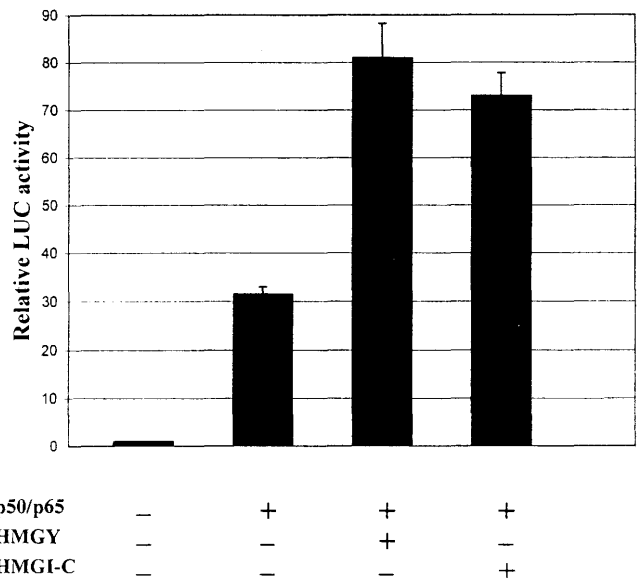


**Figure 4.** HMGY and HMGI-C show similar abilities in enhancing NF- $\kappa$ B mediated activation of PRDII Tkluc. PC Cl3 cells were transiently transfected with 2  $\mu$ g of PRDII TKluc reporter vector alone or in various combinations with 200 ng of an equimolar mixture of p50/p65 expression vectors and 5  $\mu$ g of either pSVHMGY or pSVHMGI-C expression plasmids, as indicated in the lower part of the figure. Total amounts of each expression plasmid were kept constant in all the experiments by adding the corresponding empty vectors. pRL-TK *Renilla* luciferase expression vector (0.25  $\mu$ g) was included as internal standard. Transcriptional activation is indicated as percentage of the luciferase activity measured for reporter alone; columns and bars represent the mean and standard deviation, respectively, of at least five independent experiments.

results of parallel experiments in PC Cl3 cells. Expression of HMGY and HMGI-C was not able to cause a transcriptional enhancement above the levels achieved by NF- $\kappa$ B when the mPRDII TKluc and the Ig $\kappa$ B TKluc reporters were used, in contrast to the enhancement observed with the PRDII TKluc reporter. These results were confirmed using another rat thyroid cell line, FRTL-5 Cl2, that also lacks HMGI proteins (30), and again HMGI-dependent enhancement was seen only with reporters that had the AT-core sequence (not shown).

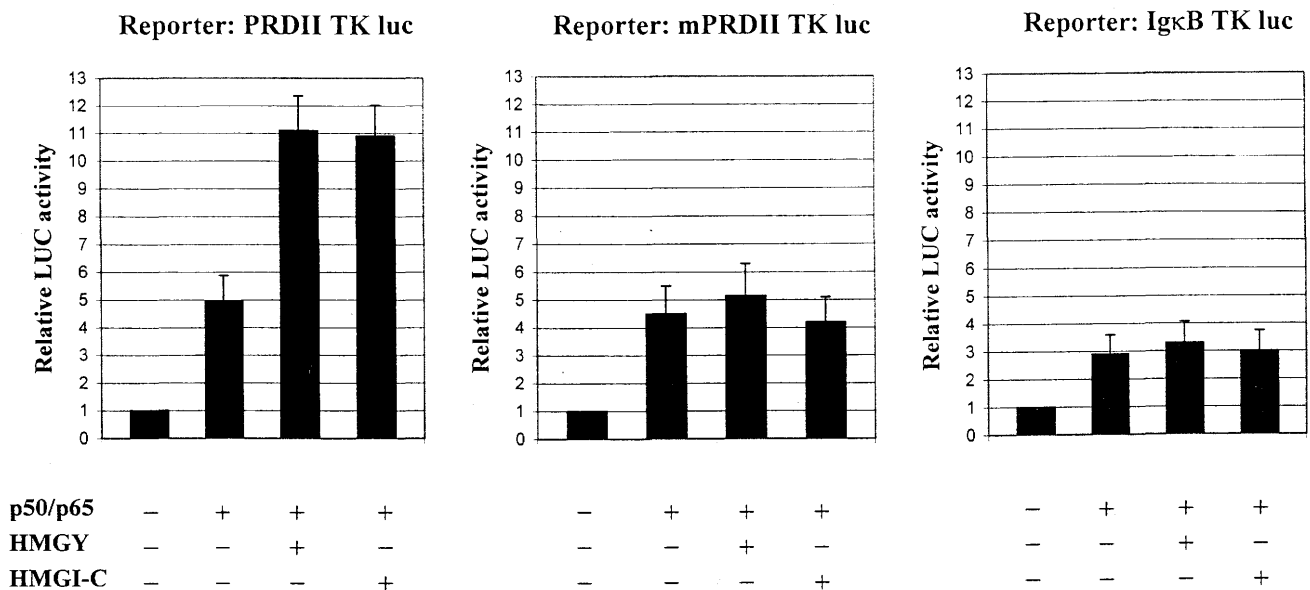
**DISCUSSION**

Eukaryotic cells are able to respond to a large number of different extracellular signals and environmental stresses, leading to the activation of specific sets of genes. The understanding of this process is complicated by the fact that many transcription factors like homeodomain-containing proteins and members of the Rel family, once activated can bind specifically to regulatory sequences upstream of a large number of different genes (33), but only a subset of such genes are activated in response to a given signal. One mechanism through which this specificity is achieved is the interaction with accessory factors that can bind nearby, altering the DNA conformation and consequently facilitating or inhibiting the interaction of a transcription factor with DNA. HMGI proteins are a class of non-histone chromosomal proteins that have been shown to fulfill this role, acting as ancillary or architectural factors. By doing so they can increase the affinity (14-18) or can compete for the binding (19,20) of the transcription factor to the DNA. Probably the best characterized example is the virus-inducible enhancer of the human interferon- $\beta$  (IFN- $\beta$ ) gene (14,21,22) where four positive regulatory domains (PRDI, PRDII, PRDIII and PRDIV) are present. HMGI(Y) are required for the transcriptional activities of both NF- $\kappa$ B and ATF-2/c-Jun,



**Figure 5.** Enhancement of transcriptional activity by HMGY and HMGI-C occurs also in NIH-3T3 cells. NIH-3T3 cells were transiently transfected with 2  $\mu$ g of PRDII TKluc reporter vector alone or in combination with 200 ng of p50/p65 mixture and 5  $\mu$ g of either pSVHMGY or pSVHMGI-C expression vectors, as indicated in the scheme. Total amounts of each expression plasmid were kept constant in all the experiments by adding the corresponding empty vectors. Samples (400 ng) of pSV  $\beta$ -galactosidase expression vector were included as internal standard. Three separate experiments were performed and means and standard deviations calculated. Transcriptional activity is indicated as percentage of that obtained for reporter alone.

binding respectively to the PRDII and PRDIV elements of the promoter. For this reason we chose the PRDII element to



**Figure 6.** HMGY and HMGI-C enhance NF- $\kappa$ B activation of PRDII, but not of mPRDII or I $\kappa$ B elements. Samples (2  $\mu$ g) of each reporter plasmid: PRDII TKluc, mPRDII TKluc and I $\kappa$ B TKluc were cotransfected in separate experiments with 200 ng of p50/p65 equimolar mixture of expression vectors, in the presence of 5  $\mu$ g of either pSVHMGY or pSVHMGI-C, as shown in the scheme. Total amounts of each expression plasmid were kept constant in all the experiments by adding the corresponding empty vectors. pRL-TK *Renilla* luciferase expression vector (0.25  $\mu$ g) was included as an internal standard. Transcriptional activity is indicated as percentage of that measured for each reporter alone, columns and bars representing means and standard deviations, respectively, of five independent experiments.

investigate whether another member of the HMGI family, HMGI-C, is able to function as architectural factor facilitating the transactivating ability of NF- $\kappa$ B.

HMGI(Y) proteins contain a DNA-binding motif, the AT-hook, which is repeated three times in the protein sequence. It has been shown that this motif is able to specifically bind to the minor groove of AT-rich DNA sequences (1). HMGI-C also contains three AT-hooks suggesting that this protein binds to similar sequences to those bound by HMGI(Y), but since the proteins differ in amino acid sequence outside the AT-hooks and the acidic C-terminals, it is possible that HMGI-C functions differently to HMGI(Y) in the regulation of transcription. There are in fact many examples of nuclear factors sharing the same type of DNA binding domain which, however, display quite different transacting capabilities.

In this report we show that HMGI-C is able to bind to the PRDII element with the same affinity as HMGY. The integrity of the AT-rich core present in the PRDII element is necessary for the binding of both HMGY and HMGI-C since a single base pair substitution is sufficient to drastically decrease their affinity for the DNA. We also demonstrate that HMGI-C is able to enhance the binding of NF- $\kappa$ B to this element. Furthermore like HMGY, HMGI-C enhances the NF- $\kappa$ B mediated transcriptional activation of the PRDII element in different cell lines. Since it was not able to affect the basal activity of the PRDII TKluc reporter in the absence of NF- $\kappa$ B, we can conclude that all the HMGI proteins do not possess transactivation capabilities on their own, but rather act by influencing the activities of classical transcription factors. It is apparent that the HMGI proteins need to bind to DNA in order to exert their effect since HMGY and HMGI-C are not able to enhance NF- $\kappa$ B mediated transactivation when transfected with the reporter vectors mPRDII TKluc and I $\kappa$ B TKluc which,

although able to bind NF- $\kappa$ B, do not bind HMGY and HMGI-C strongly. It has been shown that HMGI(Y) are able to interact with NF- $\kappa$ B in the absence of DNA and our data would suggest that such protein-protein interactions might stabilize the ternary complex. This mechanism is thus different to that described for other nuclear proteins like the viral tax protein that can facilitate the binding of transcription factors without binding to the DNA (34).

Thanos and Maniatis originally demonstrated that the HMGI(Y) protein was required for the virally-induced and NF- $\kappa$ B-dependent promoter activity of the PRDII element, by using an antisense HMGI(Y) construct in transient transfection experiments (14). Transfection of the antisense plasmid into HeLa and Jurkat T cells inhibited the PRDII element, demonstrating the requirement for HMGI(Y), but furthermore may suggest that HMGI-C cannot substitute for HMGI(Y). However T cells do not express HMGI-C (35) and HeLa express very low levels, as judged by northern analysis of HeLa cell RNA (unpublished data). Also, in experiments with antisense HMGI-C we found that the expression of HMGI(Y) as well as HMGI-C is inhibited (8) and so it is possible that an HMGI(Y) antisense would inhibit the expression of HMGI-C as well. Thus, the antisense experiments of Thanos and Maniatis do not exclude the possibility that HMGI-C can bind to and activate the PRDII element in the similar manner as HMGI(Y).

HMGI(Y) act as accessory factors on a number of other inducible promoter elements, interacting with the transcription factors ATF-2 (36), IRF-1 (21), E1f-1 (16) and Oct-2A (17), but it is not known whether HMGI-C interacts with any of these proteins. Whilst it is possible that the three HMGI proteins interact with some common partners, such as NF- $\kappa$ B, it is likely that the functional specificity of these proteins resides in their interactions with different subsets of transcription factors. Thus,

since an altered HMGI-C protein is involved in the abnormal growth and development of fat cells resulting in lipomas (9–11), and that HMGI-C knock-out is responsible in the *pygmy* mouse for a disproportionate reduction in body fat (14) which is not compensated by HMGI(Y), it is feasible that HMGI-C might specifically interact with transcription factors involved in adipocyte growth and development.

Promoter structure may also play a role in determining whether HMGI(Y) or HMGI-C is recruited. In fact, regulatory elements which contain multiple AT sites interact with HMGI proteins with higher affinity, this cooperative binding is possible since two or three AT-hooks of the same molecule are simultaneously employed in the interaction with DNA (22,29). In this context the different lengths of the spacer regions between the AT-hooks of the HMGI proteins might be relevant (2); AT elements that are differently spaced within promoters might preferentially recruit only one of the HMGI. Thus, an important line of future research will be the identification of target genes involved in cell growth and development that are activated by HMGI-C and the characterisation of the promoters and the transcription factors that it interacts with.

## ACKNOWLEDGEMENTS

This study was supported by Associazione Italiana per la Ricerca sul Cancro (AIRC), Milano, Italy, Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Roma, Italy and Università degli Studi di Trieste, Italy. F.M. was a recipient of a fellowship from Fondazione Italiana per la Ricerca sul Cancro (FIRC). G.H.G. is supported by the Cancer Research Campaign. We thank V.Giancotti for continuous encouragement and advice during this work.

## REFERENCES

- Bustin,M. and Reeves,R. (1996) *Progr. Nucleic Acid Res. Mol. Biol.*, **54**, 35–100.
- Manfioletti,G., Giancotti,V., Bandiera,A., Buratti,E., Sautiere,P., Cary,P., Crane-Robinson,C., Coles,B. and Goodwin,G.H. (1991) *Nucleic Acids Res.*, **19**, 6793–6797.
- Johnson,K.R., Lehn,U.A., Elton,I.S., Barr,P.J. and Reeves,R. (1988) *J. Biol. Chem.*, **263**, 18338–18342.
- Zhou,X., Benson,K.F., Ashar,H.R. and Chada,K. (1995) *Nature*, **376**, 771–774.
- Chiappetta,G., Avantiaggiato,V., Visconti R., Fedele,M. Battista,S., Trapasso,F., Merciai,B.M., Fidanza,V., Giancotti,V., Santoro,M., *et al.*, (1996) *Oncogene*, **13**, 2439–2446.
- Chiappetta,G., Bandiera,A., Berlingieri,M.T., Visconti.R., Manfioletti,G., Battista,S., Francisco,J., Martinez-Tello,F.J., Santoro,M., Giancotti,V. and Fusco,A. (1995) *Oncogene*, **10**, 1307–1314.
- Fedele,M., Battista,S., Manfioletti,G., Chiappetta,G., Viglietto,G., Casamassimi,A., Bandiera,A., Santoro,M., Giancotti,V. and Fusco,A. (1996) *Cancer Res.*, **56**, 1896–1901.
- Berlingieri,M.T., Manfioletti,G., Santoro,M., Bandiera,A., Visconti,R., Giancotti,V. and Fusco,A. (1995) *Mol. Cell. Biol.*, **15**, 1545–1553.
- Fejzo,M.S., Yoon,S., Montgomery,K.T., Rein,M.S., Weremowicz,S., Krauter,K.S., Dorman,T.E., Fletcher,J.A., Mao,J., Moir,D.T. *et al.*, (1995) *Genomics*, **26**, 265–271.
- Ashar,H.R., Fejzo,M.S., Tkachenko,A., Zhou,X., Fletcher,J.A., Weremowicz,S., Morton,C.C. and Chada,K. (1995) *Cell*, **82**, 57–65.
- Schoenmakers,E.F., Wanschura,S., Mols,R., Bullerdiek,J., Van den Berghe,H. and Van de Ven,W.J. (1995) *Nature Genet.*, **10**, 436–444.
- Berner,J., Meza-Zepeda,L.A., Kools,P., Forus,A., Schoenmakers,E., Van de Ven,W., Fodstad,O. and Myklebost,O. (1997) *Oncogene*, **14**, 2935–2941.
- Kazmierczak,B., Pohnke,Y. and Bullerdiek,J. (1996) *Genomics*, **38**, 223–226.
- Thanos,D. and Maniatis,T. (1992) *Cell*, **71**, 777–789.
- Whitley,M.Z., Thanos,D., Read,M.A., Maniatis,T. and Collins,T. (1994) *Mol. Cell. Biol.*, **14**, 6464–6475.
- John,S., Reeves,R.B., Lin,J.X., Child,R., Leiden,J.M., Thompson,C.B. and Leonard,W.J. (1995) *Mol. Cell. Biol.*, **15**, 1786–1796.
- Abdulkadir,S.A., Krishna,S., Thanos,D., Maniatis,T., Strominger,J.L. and Ono,S.J. (1995) *J. Exp. Med.*, **182**, 487–500.
- Himes,S.R., Coles,L.S., Reeves,R. and Shannon,M.F. (1996) *Immunity*, **5**, 479–489.
- Klein-Hessling,S., Schneider,G., Heinfling,A., Chuvpilo,S. and Serfling,E. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 15311–15316.
- Arlotta,P., Rustighi,A., Mantovani,F., Manfioletti,G., Giancotti,V., Tell,G. and Damante,G. (1997) *J. Biol. Chem.*, **272**, 29904–29910.
- Thanos,D. and Maniatis,T. (1995) *Cell*, **80**, 529–532.
- Yie,J., Liang,S., Merika,M. and Thanos,D. (1997) *Mol. Cell. Biol.*, **17**, 3649–3662.
- Giancotti,V., Berlingieri,M.T., DiFiore,P.P., Fusco,A., Vecchio,G. and Crane-Robinson,C. (1985) *Cancer Res.*, **45**, 6051–6057.
- Giancotti,V., Buratti,E., Perissin,L., Zorzet,S., Balmain,A., Portella,G., Fusco,A. and Goodwin,G.H. (1989) *Exp. Cell. Res.*, **184**, 538–545.
- Nordeen,S.K. (1988) *BioTechniques*, **6**, 454–457.
- Del Sal,G., Manfioletti,G., Gustincich,S., Ruaro,E. and Schneider,C. (1994) *BioTechniques*, **16**, 134–136.
- Gorman,C, Padmanabhan,R. and Howard,B.H. (1983) *Science*, **221**, 551–553.
- Ambesi-Impombato,F.S., Parks,L.A.M. and Coon,H.G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3455–3459.
- Maher,J.F. and Nathans,D. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 6716–6720.
- Giancotti,V., Pani,B., D'Andrea,P., Berlingieri,M.T., Di Fiore,P.P., Fusco,A., Vecchio,G., Philp,R., Crane-Robinson,C., Nicolas,R.H. and Goodwin,G.H. (1987) *EMBO J.*, **6**, 1981–1987.
- Lehming,N., Thanos,D., Brickman,J.M., Ma,J., Maniatis,T. and Ptashne,M. (1994) *Nature*, **371**, 175–179.
- Lanahan,A., Williams,J.B., Sanders,L.K. and Nathans,D. (1992) *Mol. Cell. Biol.*, **12**, 3919–3929.
- Mc Knight,S.L. and Yamamoto,K.R. (1992) *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Wagner,S. and Green,M.R. (1993) *Science*, **262**, 395–399.
- Giancotti,V., Bandiera,A., Buratti,E., Fusco,A., Marzari,R., Coles,B. and Goodwin,G.H. (1991) *Eur. J. Biochem.*, **198**, 211–216.
- Du,W., Thanos,D. and Maniatis,T. (1993) *Cell*, **74**, 887–898.

## Expression of HMGI(Y) Proteins in Squamous Intraepithelial and Invasive Lesions of the Uterine Cervix<sup>1</sup>

Antonella Bandiera, Daniela Bonifacio, Guidalberto Manfioletti, Fiamma Mantovani, Alessandra Rustighi, Fabrizio Zanconati, Alfredo Fusco, Luigi Di Bonito, and Vincenzo Giancotti<sup>2</sup>

Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Università degli Studi di Trieste, 34127 Trieste, Italy [A. B., G. M., F. M., A. R., V. G.]; Istituto di Anatomia ed Istologia Patologica, Università degli Studi di Trieste, Ospedale Maggiore, 34125 Trieste, Italy [D. B., F. Z., L. D. B.]; and Dipartimento di Medicina Sperimentale e Clinica, Facoltà di Medicina e Chirurgia di Catanzaro, Università degli Studi di Reggio Calabria, 88100 Catanzaro, Italy [A. F.]

### ABSTRACT

The expression of nuclear proteins high mobility group (HMG) I and HMGY was investigated in intraepithelial and invasive lesions of the uterine cervix. Human carcinoma cell lines C-4I, ME-180, and CaSki were used for testing protein expression in neoplastic cells from the cervix. Morphological grading of the dysplasias (CIN I, CIN 2, and CIN 3) and invasive carcinomas from formalin-fixed paraffin-embedded samples parallels the degree of nuclear immunostaining obtained using a polyclonal antibody raised against the amino-terminal region of HMGI(Y) proteins. The immunostaining obtained with HMGI(Y) antibody was compared with that observed using the antibody Ki-67, and the results were similar. We suggest the use of HMGI(Y) antibody in clinical oncology as a useful marker of intraepithelial lesions and invasive carcinomas.

### INTRODUCTION

Since the discovery of HMG<sup>3</sup> I proteins as components of the chromatin of mammalian cells (1-3), the attention of researchers has been attracted by the fact that high levels of expression of these proteins are observable in neoplastic transformed cells by comparison with their normal counterparts, in which HMGI proteins are absent or detectable at very low levels (3, 4). The group is formed by three polypeptides named HMGI (which gives the name to the entire group), HMGY, and HMGI-C (1-5), all constituted by about 100 amino acid residues and showing a preferential binding to A-T-rich DNA sequences (6-11), mainly due to a repeating stretch of positive amino acid residues contained three times and highly conserved in the sequence of the three proteins (12, 13). Proteins HMGI and HMGY derive from the same gene by alternative splicing (12) and are currently indicated as HMGI(Y) proteins. However, HMGI-C is the product of a different gene that expresses a protein having about 50% sequence homology with HMGI(Y) (13, 14).

It has been demonstrated that HMGI(Y) proteins participate as structural factors in protein complexes that stimulate the transcription of genes such as *IFN-β*, *tumor necrosis factor β*, and *interleukin 2 receptor α chain* (15-19). On the other hand, it has also been reported that chromosomal aberrations in the region 12q14-15 involve *HMGI-C* gene rearrangements that correlate to the onset of benign solid tumors such as lipomas and pulmonary chondroid hamartomas

(20-22), whereas constitutional knockout of the *Hmgi-c* mouse gene induces the *pygmy* phenotype (23).

Finding new markers that could aid in the early detection of cancer, in the evaluation of the metastatic potentiality of cancerous cells, and in therapy monitoring is a general aim of oncological research. To this end, tumor markers such as carcinoembryonic antigen, proliferating cell nuclear antigen, nuclear proliferative antigen Ki-67, and AgNOR proteins are widely used (24, 25). Each of the suggested markers refers to an antigen involved in a different aspect of the cellular activity; therefore, a higher reliability of the assessment of cell proliferation and/or neoplastic transformation could be achieved by the use of more than one marker. Due to the high level of expression of HMGI proteins found in both transformed cell lines and tumors having different origins (4, 5, 26, 27), we and others have suggested HMGI(Y) proteins as new markers of the transformed cellular phenotype and used them as a progression marker for prostate cancer metastasis (28, 29) and malignant phenotype assessment of both human thyroid and colorectal neoplasias (30, 31).

As a further extension of the study of the expression of HMGI proteins in tumor tissues, here we report the use of an anti-HMGI(Y) antibody as a diagnostic aid in the grading of intraepithelial and invasive lesions of the uterine cervix. Immunohistochemical results were compared on one hand with the morphological diagnoses and, on the other hand, with parallel immunohistochemical data obtained using the Ki-67 proliferating marker. Our results allow us to suggest anti-HMGI(Y) antibody as a helpful aid in clinical oncology.

### MATERIALS AND METHODS

**Cell Cultures and Tissues.** The three human cervix carcinoma cell lines (C-4I, ME-180, and CaSki) used in this study for testing HMGI(Y) protein expression were purchased from American Type Culture Collection (Rockville, MD) and grown in DMEM containing 10% FCS. Epithelial rat thyroid FRTL-5 cells were cultured as reported previously (3). Tumors from Lewis lung carcinoma cells were induced in mice as described previously (26). Normal and pathological specimens from human cervix tissue for immunohistochemical analyses were obtained from formalin-fixed paraffin-embedded samples of the tissue file of the Istituto di Anatomia ed Istologia Patologica of the University of Trieste.

**Antibodies.** A polyclonal antibody was raised in rabbits using the synthetic peptide SSSKSSQPLASKQ specific for the HMGI(Y) proteins and affinity purified. The antibody was tested for activity by Western blot analysis using high-performance liquid chromatography-purified samples of recombinant HMGI(Y) proteins. Ki-67 monoclonal antibody was purchased from DAKO.

**Western and Northern Blot Analyses.** Total HMG proteins were selectively extracted by 5% (v/v) perchloric acid from cultured cells and tumors and recovered by precipitation with acetone as reported previously (3, 26). Resolution of the two electrophoretic bands due to proteins HMGI and HMGY was achieved by polyacrylamide (15%) SDS/Tris/Tricine gel electrophoresis (5) run for 17 h at 200 V. For Western analysis, gels were soaked in 750 mM Tris base and 0.1% SDS (pH 11) and transferred for 30 min at 300 mA to polyvinylidene difluoride membrane (Immobilon P) using a trans-blot SD semidry transfer cell (Bio-Rad) in 75 mM Tris base and 0.01% SDS. Membranes were then treated as reported previously (30), and proteins were detected by chemiluminescence using enhanced chemiluminescence Western

Received 6/19/97; accepted 11/21/97.

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.

<sup>1</sup>Supported by grants from Associazione Italiana per la Ricerca sul Cancro (Milan, Italy), Consiglio Nazionale delle Ricerche (Rome, Italy), Ministero della Ricerca Scientifica e Tecnologica (Rome, Italy), Università degli Studi di Trieste, and Progetto Finalizzato Applicazioni Cliniche della Ricerca Oncologica. F. M. was a recipient of a fellowship awarded by the Fondazione Italiana per la Ricerca sul Cancro.

<sup>2</sup>To whom requests for reprints should be addressed, at Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Università degli Studi di Trieste, via L. Giorgieri 1, 34127 Trieste, Italy. Phone: 39-40-676-3676; Fax: 39-49-676-3694; E-mail: GIANCOT@univ.trieste.it.

<sup>3</sup>The abbreviations used are: HMG, high mobility group; CIN, cervical intraepithelial neoplasia; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; DAB, 3,3'-diaminobenzidine.

blotting detection reagent (Amersham). Equal amounts of total protein from various sources loaded on the gels were assured by checking that each sample contained the same amount of histone H1 that can be assumed to be constantly expressed in the different cell lines (27). Total RNA extraction and Northern analyses were carried out as described previously (13). The probes used were pHMGI(Y) (31) and GAPDH (32).

**Immunohistochemistry.** Cultured cell suspension (200  $\mu$ l) was centrifuged (Cytospin) for 5 min at 500 rpm, obtaining a 5-mm diameter area on a slide. Cells were then fixed with acetone/methanol (1:1, v/v) for 10 min at  $-20^{\circ}\text{C}$  and soaked in PBS.

The following smears and formalin-fixed paraffin-embedded cases of normal and pathological tissues were selected: (a) 3 smears from normal epithelia; (b) 28 normal cervixes; (c) 25 low-grade lesions (CIN 1); (d) 10 moderate-grade lesions (CIN 2); (e) 25 high-grade lesions (CIN 3); and (f) 27 invasive squamous carcinomas. About 3- $\mu$ m sections were cut with a microtome and dried overnight at  $37^{\circ}\text{C}$  on a slide for immunohistochemical analysis. Samples were deparaffinized by xylene at room temperature for 30 min and washed with graded ethanol/water mixtures (100, 90, 70, and 50% ethanol) before washing with tap water and then with distilled water. Samples were then put in the microwave for 30 min and subsequently left at room temperature for 30 min. After washing with tap water and distilled water, samples were soaked in PBS.

Both cellular and tissue specimens were treated with 3% hydrogen peroxide for 5 min and washed as described above before immunoperoxidase staining. Samples were incubated in normal horse serum for 10 min and then treated for 60 min at room temperature with the antibody diluted 1:100 or 1:200 in PBS. The slides were washed in PBS three times for 5 min and then incubated with biotinylated universal secondary horse antibody (Vector Laboratories). After washing, slides were incubated for 10 min in streptavidin/peroxidase complex. Immunostaining was performed by incubation of the slides in DAB (Abbott Laboratories) solution or aminoethylcarbazole (Biogenex). Slides were then washed with water and counterstained with hematoxylin for 30 s. Finally, slides from tissues were washed, dehydrated, and mounted with permanent mounting medium (Entellam; MERCK); cell culture slides were directly mounted in aqueous mounting medium (Glycergel; DAKO). Micrographs were taken on Kodak Ektachrome film with a photomicroscope (Leitz).

## RESULTS

**HMGI Gene Expression in Cell Lines Obtained from Squamous Carcinomas of the Uterine Cervix.** The expression of the *HMGI(Y)* gene in carcinomas from the uterine cervix was tested by both Northern and Western blot analyses on three cell lines (C4-I, CaSki, and ME-180). These lines were chosen because they resemble the most common features currently found in clinical cases. In fact, C-4I cells are cervical carcinoma cells that retain properties of squamous cell differentiation in long-term culture, CaSki cells initiated from metastatic epidermoid carcinoma of the cervix, and ME-180 cells are epidermoid cervical cells established from a highly invasive squamous carcinoma. In Fig. 1, the negative control was a sample derived from the FRTL-5 cell line that does not express HMGI(Y) proteins (3, 4), whereas the positive control was Lewis lung carcinoma (26). Data reported in Fig. 1 show that all three analyzed cervix carcinoma cell lines express HMGI(Y) proteins; therefore, we decided to verify the usefulness of the antibody raised against HMGI(Y) proteins in a histopathological study of squamous intraepithelial lesions of the uterine cervix.

**Immunohistochemical Analysis of Normal and Neoplastic Cells of the Uterine Cervix.** Before analyzing protein expression in cervix tissues carrying various dysplastic alterations, we tested the used immunohistochemical method on cell lines. Negative and positive controls of antigen expression were constituted by the rat thyroid FRTL-5 normal cell line and by the same cell line infected with Kirsten murine sarcoma virus. The FRTL-5 cell line does not express HMGI(Y) proteins, whereas the virus-infected counterpart is neoplastically transformed and expresses high levels of antigen (3, 4). Con-

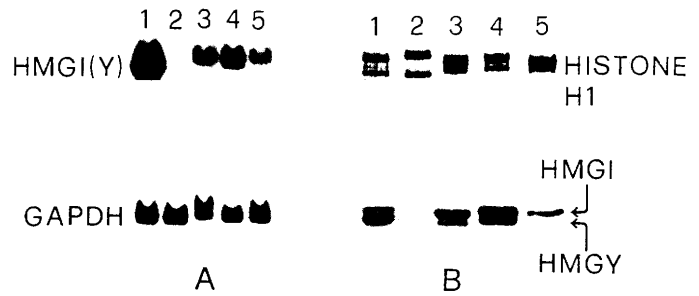


Fig. 1. Analysis of *HMGI(Y)* gene expression. A. Northern blot analysis of the RNA from three human cell lines derived from cervix carcinomas. About 20  $\mu$ g of total RNA were loaded in each slot; GAPDH was used as an internal control for uniform RNA loading. B. Western blot analysis of HMGI(Y) proteins extracted from the same cell lines. About 10  $\mu$ g of total extracted protein were loaded in each slot, and equal loading was assured by checking the amount of histone H1 coextracted with HMGI(Y) proteins and stained with Coomassie Blue. Lanes 1, Lewis lung carcinoma cells, positive control; Lanes 2, FRTL-5 cell line, negative control; Lanes 3, C-4I cervix carcinoma cell line; Lanes 4, CaSki cervix carcinoma cell line; Lanes 5, ME-180 cervix carcinoma cell line. Results show that HMGI(Y) proteins are expressed in all three cervix carcinoma cell lines.

sequently, immunostaining with anti-HMGI(Y) antibody was negative for normal cells and highly positive for the transformed cell line (data not shown).

Using the same antibody, we subsequently carried out immunohistochemical analyses on both normal human cells from cervical smears and on the above-reported cervix carcinoma cell lines, assuming that normal cells from smears and malignant cells from cervix carcinoma should represent the extremes in the neoplastic transformation process. Fig. 2, A1 and A2, shows superficial cells of normal squamous epithelium from cervical smears characterized by a large cytoplasm and a small nucleus; no antigen recognition was observed in these cells that, at the same time, are also poorly stained by hematoxylin. On the other hand, the cell lines derived from cervix carcinomas (Fig. 2, B-D) show strong nuclear positivity consistent with the high level of gene expression detected by both Northern and Western blot analyses (Fig. 1).

All examined tissues carrying various degrees of alteration were derived from formalin-fixed paraffin-embedded specimens. After the deparaffinizing procedure, slides were immunostained using the anti-HMGI(Y) antibody and counterstained with hematoxylin. Results are summarized in Table 1, and representative cases are shown in Figs. 3 and 4. As shown in Fig. 3A, most of the cells from normal stratified epithelia do not stain or stain faintly. In CIN 1 cervical intraepithelial neoplasias (that very frequently showed superficial koilocytosis), there was a marked immunostaining increase in the lower third of the epithelium, as in Fig. 3B, but strongly stained positive cells were also found in the remaining part of the epithelium. CIN 3 cervical intraepithelial neoplasias (Fig. 3C) are characterized by a lack of cell polarity, cellular atypia and mitotic figures can be found up to the superficial layer; most of the cells are strongly immunostained. Finally (Fig. 3D), nests of malignant cells from invasive carcinomas infiltrate the connective tissue and are strongly immunostained by the anti-HMGI(Y) antibody. In many analyses, such as that of Fig. 3C, we noted positively immunostained cells in the stroma underlying the epithelium. This fact could be due to an infiltration of lymphocytes that are recognized by the antibody anti-HMGI(Y) in accordance with the previously reported finding that these cells, even if normal, express HMGI(Y) proteins (5).

**Comparison between Immunostaining of Nuclear Antigens Using HMGI(Y) Antibody and Ki-67 Antibody.** The antibody Ki-67 is frequently used in the assessment of proliferation activity. Although it recognizes a nuclear antigen whose function is still unknown (33), positive staining with this antibody is considered of significant im-



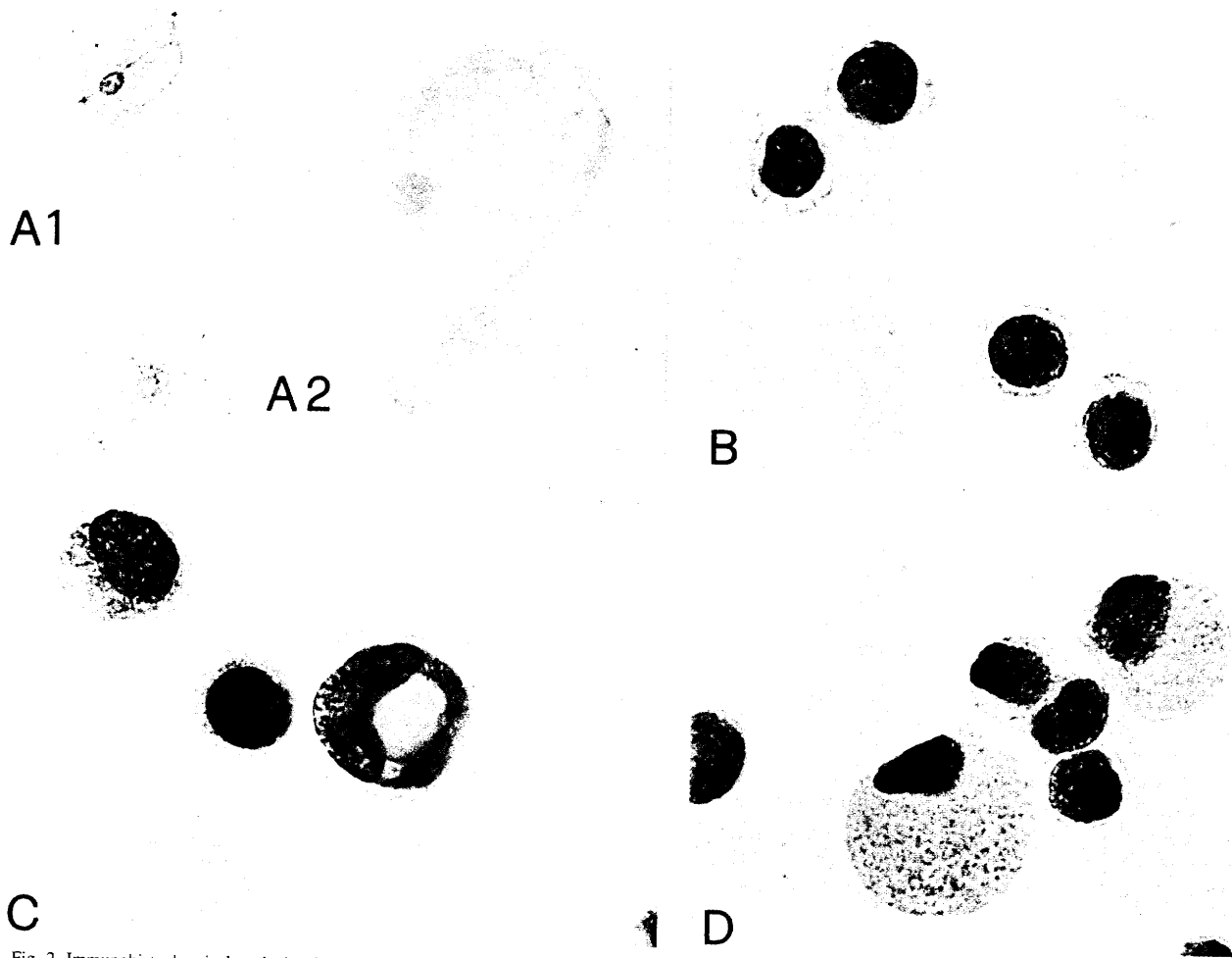


Fig. 2. Immunohistochemical analysis of HMGI(Y) protein expression in normal squamous superficial epithelial cells from cervical smears (A1 and A2) and in cervix carcinoma cell lines C-41 (B), ME-180 (C), and CaSki (D). Objective magnification:  $\times 100$ , A2 and B-D.;  $\times 40$ , A1. Cells were immunostained with streptavidin/peroxidase/aminocarbazole and counterstained with hematoxylin. No immunostaining was observed in normal cells (A1 and A2), whereas strong nuclear positivity was detected in carcinoma cells (B-D).

portance in surgical pathology, because antigen recognition correlates with the proliferation degree of the cells (24, 33-37). Therefore, we thought that it could be of some interest to compare the immunostaining of the same tissues with both Ki-67 and HMGI(Y) antibodies. Fig. 4, A and B, shows one of the cases of CIN2 moderate dysplasias that was immunostained in distinct experiments with the two antibodies. The lower two-thirds of the epithelium, if considered as a whole,

show strong positivity in both experiments. However, a more careful inspection reveals that basal and parabasal layers are more stained by the anti-HMGI(Y) antibody in comparison with Ki-67 immunostaining. This could be due to the fact that the two antibodies recognize antigens that should refer to different biological cell states (see "Discussion"). Analyses similar to those of Fig. 4, A and B, were also carried out on samples from invasive carcinomas; the results are shown in Fig. 4, C and D, and it is possible to note that immunostaining of cervix carcinomas with HMGI(Y) antibody is very similar to that obtained using the Ki-67 antibody.

Table 1 Expression of HMGI(Y) proteins in squamous intraepithelial and invasive lesions of the uterine cervix

Type of tissue (no. of cases)	Positivity (no. of cases)		
	- <sup>a</sup>	+ <sup>b</sup>	++ <sup>c</sup>
Normal stratified epithelium (28)			
Basal and parabasal layers	24/28	4/28	0/28
Remaining part of epithelium	27/28	1/28	0/28
CIN 1 (25)			
Lower third of epithelium	0/25	20/25	5/25
Upper 2/3 of epithelium	14/25	8/25	3/25
CIN 2 (10)			
Lower 2/3 of epithelium	0/10	1/10	9/10
Upper third of epithelium	0/10	4/10	6/10
CIN 3 (25)			
All epithelium	0/25	5/25	20/25
Invasive carcinoma (27)			
Nests	0/27	0/27	27/27

<sup>a</sup> Seventy to ninety % of cells are not immunostained, the remaining cells show faint immunostaining.

<sup>b</sup> About 50% of cells show no immunostaining or faint immunostaining, whereas the remaining 50% show intermediate or strong immunostaining.

<sup>c</sup> At least 80% of cells show intermediate or strong immunostaining.

## DISCUSSION

HMGI is a family of three mammalian nuclear proteins (HMGI, HMGY, and HMGI-C) involved *in vivo* in the positive or negative regulation of gene transcription (38), and a close correlation between the increase of their expression and neoplastic transformation has been reported (3-5, 26-31, 39, 40). A more direct relationship between HMGI expression and malignant phenotype has been demonstrated by an antisense methodology: blocking the expression of HMGI-C protein prevents the transformation of rat thyroid cell lines by viral oncogenes such as *v-mos* and *v-ras*-Ki (41). The high constitutive levels of HMGI proteins of neoplastic cells prompted us and other laboratories to suggest the use of HMGI expression as a marker of cellular transformation (28-31). In particular, we demonstrated that high concentrations of HMGI(Y) proteins could be used as markers of the malignant phenotype of human thyroid neoplasias as well as

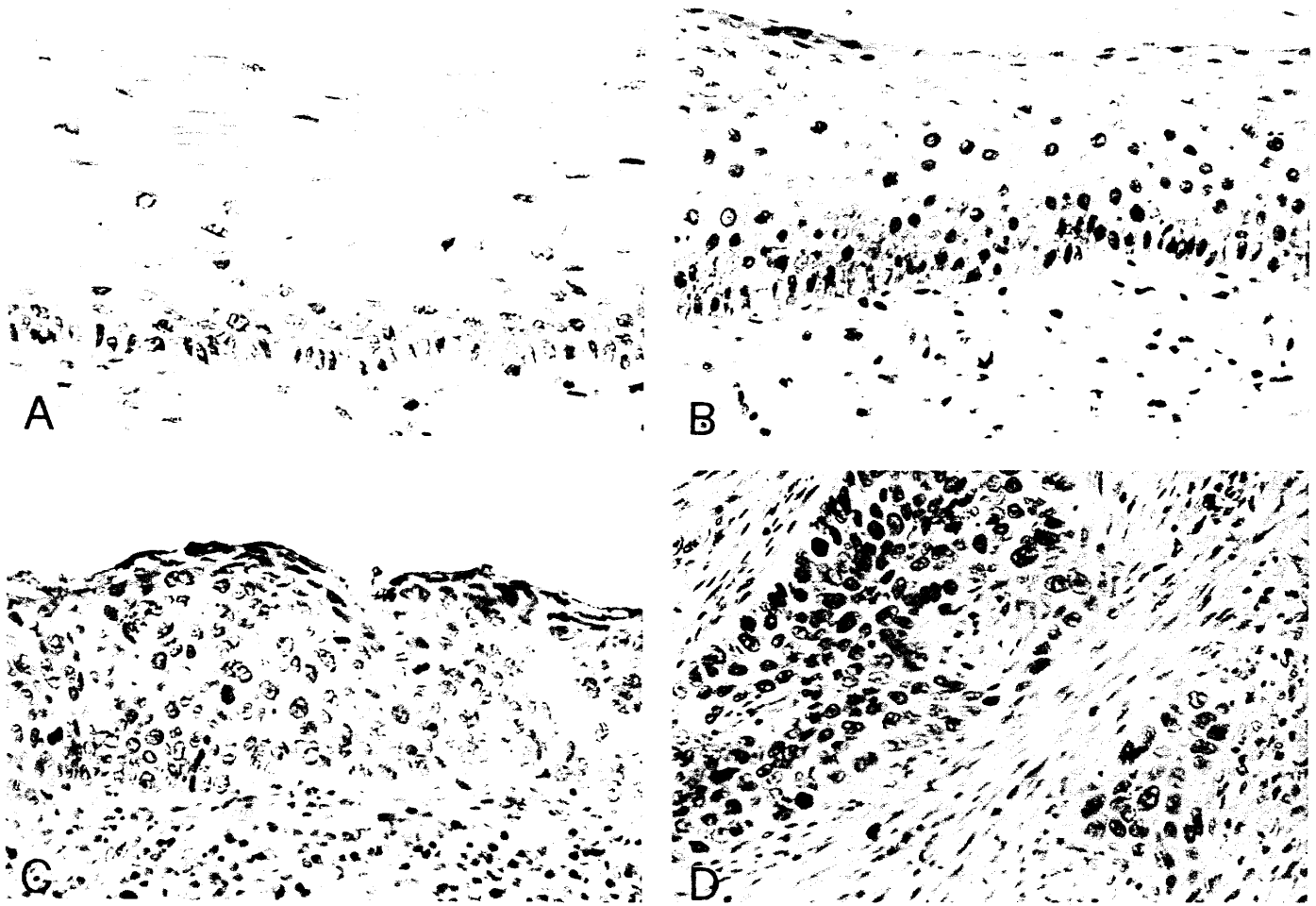


Fig. 3. Immunohistochemical analysis of HMGI(Y) protein expression in squamous intraepithelial and invasive lesions of the uterine cervix. A, normal stratified epithelium; B, CIN 1 lesion; C, CIN 3 lesion; D, invasive carcinoma. Objective magnification,  $\times 40$ . Tissues were immunostained with streptavidin/peroxidase/DAB and counterstained with hematoxylin. No immunostaining or faint immunostaining was observed in cells of normal epithelium (A). Cells of the lower third of the CIN 1 lesion (B) show intermediate or strong positivity. Most of the nuclei of the cells of the CIN 3 lesion (C) are immunostained, and strong positivity is observable in the nuclei of the invasive carcinoma (D).

colorectal tumorigenesis (30, 31). More recently, high levels of HMGI-C expression have been detected in atypical lipomatous tumors (42).

Further extension of the study of the expression of HMGI proteins in tumor tissues is reported in this paper dealing with squamous intraepithelial and invasive lesions of the uterine cervix. This research moves toward a wider use of the HMGI(Y) antibodies as a diagnostic tool, because the aim is to correlate the level of antigen recognition with the degree of tissue alteration. With this view, we took into consideration the ectocervix of adult women, which is formed by a stratified squamous epithelium composed of basal, parabasal, intermediate, and superficial cell layers. Maturation of the epithelium takes place from the basal layer toward the superficial layer, with a progressive modification of the cellular morphology. Dysplastic processes introduce more or less severe alterations in both the epithelium and cellular morphology. Grading of dysplastic cervical epithelia is based thus far on the proportion of the epithelial thickness formed by dedifferentiated basal layer-like cells or single-cell criteria such as nuclear atypia. In our opinion, the recognition of nuclear antigens such as HMGI proteins, whose expression is related to neoplastic transformation, could supply molecular information to be considered in clinical diagnoses together with the morphological observation. To this aim, we selected an antibody able to recognize the two nuclear HMGI(Y) proteins on the basis of preliminary analyses carried out on cultured carcinoma cells of the uterine cervix that, according to both Northern and Western blot experiments, express the

HMGI(Y) gene. The results reported here show that in the analyzed samples, the level of antigen recognition parallels the CIN grade. In fact, no positivity has been detected in normal cells from smears, whereas only low levels of positivity have been found in samples from formalin-fixed paraffin-embedded normal epithelia. A progressive increase of positivity matches the progression of the epithelium lesion up to the immunohistochemical detection of all cells of the invasive carcinoma.

In this study, we also report that the antibody anti-HMGI(Y) gives immunohistochemical data similar to those obtained using the antibody Ki-67. Rationalizing this result is not straightforward, because the two antibodies recognize different antigens, even if both are nuclear. Ki-67 is a nuclear antigen formed by two large polypeptides (345 and 395 kDa) that are expressed in proliferating cells but absent in resting cells (33). The function of these polypeptides, both of which are recognized by Ki-67 antibody, is still unknown, but the relationship between their expression and cell proliferation is well established (24, 33–36, 43–47). Antigen recognition by Ki-67 of tissues from the uterine cervix increases with increasing lesion grade (34, 35); our data of Table 1 and Fig. 3 show similar behavior for HMGI(Y) antibody. HMGI(Y) proteins are low molecular weight (about 11 kDa) nuclear proteins that are involved in both gene expression regulation and cancer (38). High levels of HMGI(Y) proteins are constitutive of tumor cells and independent of cellular growth rates (3, 4, 26). Moreover, high levels of HMGI(Y) proteins have been detected in both rat embryos and thyroid fibroblasts (2, 4), and more recently,

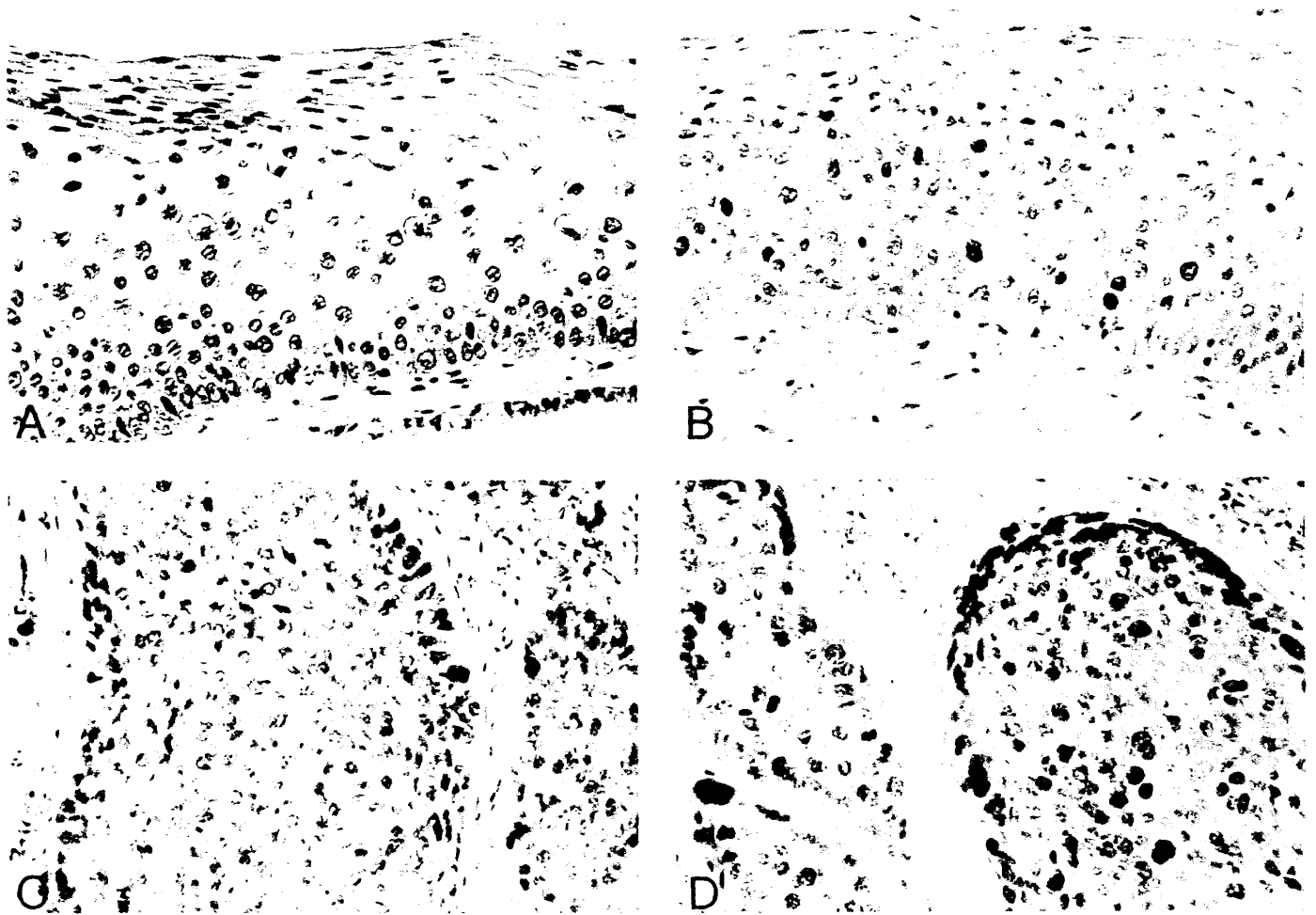


Fig. 4. Comparison between the HMGI(Y) and Ki-67 antibodies in the immunostaining of a CIN 2 intraepithelial lesion and of an invasive carcinoma of the uterine cervix. *A* and *B*, two samples derived from the same CIN 2 tissue immunostained in *A* with the HMGI(Y) antibody and in *B* with the Ki-67 antibody. *C* and *D*, two samples derived from the same invasive carcinoma immunostained in *C* with the HMGI(Y) antibody and in *D* with the Ki-67 antibody. Objective magnification,  $\times 40$ . Tissues were immunostained with streptavidin/peroxidase/DAB and counterstained with hematoxylin. The two antibodies gave similar results; strong nuclear positivity interesting the lower two-thirds of the epithelium is observable in both *A* and *B*, whereas most of the malignant invasive cells are recognized in both *C* and *D*.

the same results have been reported in a detailed study on the embryonic development of the mouse (48). From overall consideration of the above data, we infer that the expression of the two classes of proteins [Ki-67 polypeptides and HMGI(Y) proteins] should refer to different cellular events that, however, could be connected. Overlapping of the expression of both antigens could be present in tumor-proliferating cells or undifferentiated/dedifferentiated proliferating cells. On the other hand, it is possible that phenotype-altered cells recognized at high positivity by the anti-HMGI(Y) antibody are detected at a low level by the Ki-67 antibody because they are not actively proliferating. This could be the case of the basal layers shown in Fig. 4, *A* and *B*. Having this scheme in mind, the cells of the cervix epithelium that recognition by Ki-67 indicates as proliferating should be formed by an altered phenotype if they are also recognized by the anti-HMGI(Y) antibody. In conclusion, we think that immunohistochemical analyses of tissues carried out using both the HMGI(Y) and Ki-67 antibodies can increase our understanding of the biological alterations that are present in dysplastic epithelia.

## REFERENCES

- Lund, T., Holtlund, J., Fredriksen, M., and Laland, S. G. On the presence of two new high mobility group-like proteins in HeLa S3 cells. *FEBS Lett.*, *152*: 163-167, 1983.
- Goodwin, G. H., Cockerill, P. N., Kellam, S., and Wright, C. A. Fractionation by high-performance liquid chromatography of the low-molecular-mass high-mobility-group (HMG) chromosomal proteins present in proliferating rat cells and an investigation of the HMG proteins present in virus-transformed cells. *Eur. J. Biochem.*, *149*: 47-51, 1985.
- Giancotti, V., Berlingieri, M. T., Di Fiore, P. P., Fusco, A., Vecchio, G., and Crane-Robinson, C. Changes in nuclear proteins on transformation of rat epithelial thyroid cells by a murine sarcoma retrovirus. *Cancer Res.*, *45*: 6051-6057, 1985.
- Giancotti, V., Pani, B., D'Andrea, P., Berlingieri, M. T., Di Fiore, P. P., Fusco, A., Vecchio, G., Philp, R., Crane-Robinson, C., Nicolas, R. H., Wright, C. A., and Goodwin, G. H. Elevated levels of a specific class of nuclear phosphoproteins in cells transformed with *v-ras* and *v-mos* oncogenes and by co-transfection with *c-myc* and polyoma middle T genes. *EMBO J.*, *6*: 1981-1987, 1987.
- Giancotti, V., Bandiera, A., Buratti, E., Fusco, A., Marzari, R., Coles, B., and Goodwin, G. H. Comparison of multiple forms of the high mobility group I proteins in rodent and human cells. Identification of the human high mobility group I-C protein. *Eur. J. Biochem.*, *198*: 211-216, 1991.
- Strauss, F., and Varshavsky, A. A protein binds to a satellite DNA repeat at three sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell*, *37*: 889-901, 1984.
- Solomon, M. J., Strauss, F., and Varshavsky, A. A mammalian high mobility group protein recognizes any stretch of six A-T base pairs in duplex DNA. *Proc. Natl. Acad. Sci. USA*, *83*: 1276-1280, 1986.
- Elton, T. S., Nissen, M. S., and Reeves, R. Specific A-T DNA sequence binding of RP-HPLC purified HMG-I. *Biochem. Biophys. Res. Commun.*, *143*: 260-265, 1987.
- Reeves, R., Elton, T. S., Nissen, M. S., Lehn, D., and Johnson, K. R. Post-translational gene regulation and specific binding of the nonhistone protein HMG-I by 3' untranslated region of bovine interleukin 2 cDNA. *Proc. Natl. Acad. Sci. USA*, *84*: 6531-6535, 1987.
- Disney, J., Johnson, K. R., Magnuson, N. S., Sylvester, S. R., and Reeves, R. High-mobility group protein HMG-I localizes to G/Q- and C-bands of human and mouse chromosomes. *J. Cell Biol.*, *109*: 1975-1982, 1989.
- Radic, M. Z., Saghbini, M., Elton, T. S., Reeves, R., and Hamkalo, B. Hoechst 33258, distamycin A, and high mobility group I (HMG-I) compete for the binding to mouse satellite DNA. *Chromosoma (Berl.)*, *101*: 602-608, 1992.

12. Johnson. K. R., Lehn. D. A., Elton, T. S., Barr, P. J., and Reeves, R. Complete murine cDNA sequence, genomic structure, and tissue expression of the high mobility group protein HMGI(Y). *J. Biol. Chem.*, *263*: 18338-18342, 1988.
13. Manfioletti, G., Giacotti, V., Bandiera, A., Buratti, E., Sautière, P., Cary, P., Crane-Robinson, C., and Goodwin, G. H. cDNA cloning of the HMGI-C phosphoprotein, a nuclear protein associated with neoplastic and undifferentiated phenotypes. *Nucleic Acids Res.*, *19*: 6793-6797, 1991.
14. Patel, U. A., Bandiera, A., Manfioletti, G., Giacotti, V., Chau, K. Y., and Crane-Robinson, C. Expression and cDNA cloning of the human high-mobility-group protein HMGI-C. *Biochem. Biophys. Res. Commun.*, *201*: 63-70, 1994.
15. Thanos, D., and Maniatis, T. NF- $\kappa$ B: a lesson in family values. *Cell*, *80*: 529-532, 1995.
16. Thanos, D., and Maniatis, T. The high mobility group protein HMGI(Y) is required for NF- $\kappa$ B-dependent virus induction of the human IFN- $\beta$  gene. *Cell*, *71*: 777-789, 1992.
17. Du, W., Thanos, D., and Maniatis, T. Mechanisms of transcriptional synergism between distinct virus-inducible enhancer elements. *Cell*, *74*: 887-898, 1993.
18. Fashena, S. J., Reeves, R., and Ruddle, N. H. A poly(dA-dT) upstream activating sequence binds high-mobility-group I protein and contributes to lymphotoxin (tumor necrosis factor- $\beta$ ) gene regulation. *Mol. Cell. Biol.*, *12*: 894-903, 1992.
19. John, S., Reeves, R., Lin, J. X., Child, R., Leiden, J. N., Thompson, C. B., and Leonard, W. J. Regulation of cell-type-specific interleukin-2 receptor  $\alpha$ -chain gene expression: potential role of physical interactions between E1f-1, HMGI(Y), and NF- $\kappa$ B family proteins. *Mol. Cell. Biol.*, *15*: 1786-1796, 1995.
20. Schoenmakers, E. F. P. M., Wanshura, S., Mols, R., Bullerdiek, J., Van den Berghe, H., and Van de Ven, W. J. M. Recurrent rearrangements in the high mobility group gene, HMGI-C, in benign mesenchymal tumours. *Nat. Genet.*, *10*: 436-444, 1995.
21. Wanshura, S., Belge, G., Stenman, G., Kools, P., Dal Cin, P., Schoenmakers, E. F. P. M., Huysmans, C., Bartnitzke, S., Van de Ven, W. J. M., and Bullerdiek, J. Mapping of the translocation breakpoints of primary pleomorphic adenomas and lipomas within a common region of chromosoma 12. *Cancer Genet. Cytogenet.*, *86*: 39-45, 1996.
22. Kazmierczak, B., Risgkeit, J., Wanshura, S., Meyer-Bolte, K., Van de Ven, W. J. M., Kayser, K., Krieghoff, B., Kastendiek, H., Bartnitzke, S., and Bullerdiek, J. HMGI-C rearrangements as the molecular basis for the majority of pulmonary chondroid hamartomas: a survey of 30 tumors. *Oncogene*, *12*: 515-521, 1996.
23. Zhou, X., Benson, K. F., Ashar, H. R., and Chada, K. Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature (Lond.)*, *376*: 771-774, 1995.
24. Hofstädter, F., Knüchel, R., and Rüschoff, J. Cell proliferation assessment in oncology. *Virchows Arch.*, *427*: 323-341, 1995.
25. Palmieri, R. J., and Crawford, D. R. Tumor markers: an update. *Med. Clin. North Am.*, *80*: 185-199, 1996.
26. Giacotti, V., Buratti, E., Perissin, L., Zorzet, S., Balmain, A., Portella, G., Fusco, A., and Goodwin, G. H. Analysis of the HMGI nuclear proteins in mouse neoplastic cells induced by different procedures. *Exp. Cell Res.*, *184*: 538-545, 1989.
27. Giacotti, V., Bandiera, A., Ciani, L., Santoro, D., Crane-Robinson, C., Goodwin, G. H., Boiocchi, M., Dolcetti, R., and Casetta, B. High-mobility-group (HMG) proteins and histone H1 subtypes expression in normal and tumor tissues of mouse. *Eur. J. Biochem.*, *213*: 825-832, 1993.
28. Bussemakers, M. J. G., Van de Ven, W. J. M., Debruyne, F. M. J., and Schalken, J. A. Identification of high mobility group protein I(Y) as potential progression marker for prostate cancer by differential hybridization analysis. *Cancer Res.*, *51*: 606-611, 1991.
29. Tamimi, Y., Van der Poel, H. G., Karthaus, H. F. M., Debruyne, F. M. J., and Schalken, J. A. A retrospective study of high mobility group protein I(Y), a progression marker for prostate cancer determined by *in situ* hybridization. *Br. J. Cancer*, *74*: 573-578, 1996.
30. Chiappetta, G., Bandiera, A., Berlingieri, M. T., Visconti, R., Manfioletti, G., Battista, S., Martinez-Tello, F. J., Santoro, M., Giacotti, V., and Fusco, A. The expression of the high mobility group HMGI(Y) proteins correlates with the malignant phenotype of human thyroid neoplasias. *Oncogene*, *10*: 1307-1314, 1995.
31. Fedele, M., Bandiera, A., Chiappetta, G., Battista, S., Viglietto, G., Manfioletti, G., Casamassimi, A., Santoro, M., Giacotti, V., and Fusco, A. Human colorectal carcinomas expresses high levels of high mobility group HMGI(Y) proteins. *Cancer Res.*, *56*: 1896-1901, 1996.
32. Manfioletti, G., Gattei, V., Buratti, E., Rustighi, A., De Iulii, A., Aldinucci, D., Goodwin, G. H., and Pinto, A. Differential expression of a novel proline-rich homeobox gene (Prh) in human hematolymphopoietic cells. *Blood*, *85*: 1237-1245, 1995.
33. Schlüter, C., Duchrow, M., Wohlenberg, C., Becker, M. H. G., Key, G., Flad, H.-D., and Gerdes, J. The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins. *J. Cell Biol.*, *123*: 513-522, 1993.
34. Al-Saleh, W., Delvenne, P., Greimers, R., Fridman, V., Doyen, J., and Boniver, J. Assessment of Ki-67 antigen immunostaining in squamous intraepithelial lesions of the uterine cervix. *Am. J. Clin. Pathol.*, *104*: 154-160, 1995.
35. Isacson, C., Kessis, T. D., Hedrick, L., and Cho, K. R. Both cell proliferation and apoptosis increase with lesion grade in cervical neoplasia but do not correlate with human papillomavirus type. *Cancer Res.*, *56*: 669-674, 1996.
36. Kube, D., Janda, J., Johannsmeyer, K. D., and Bergamann, S. Characterization of human papillomavirus type 16 activity in separate biopsies from a carcinoma of the cervix uteri. *Virchows Arch.*, *425*: 473-480, 1994.
37. Gerdes, J., Li, L., Schlüter, C., Duchrow, M., Wohlenberg, C., Gerlach, C., Stahmer, I., Kloth, S., Brandt, E., and Flad, H. D. Immunobiochemical and molecular biologic characterization of the cell proliferation-associated nuclear antigen that is defined by monoclonal antibody Ki-67. *Am. J. Pathol.*, *138*: 867-873, 1991.
38. Bustin, M., and Reeves, R. High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *In*: W. E. Cohn and K. Moldave (eds.), *Progress in Nucleic Acid Research and Molecular Biology*, Vol. 54, pp. 35-100. San Diego, CA: Academic Press, 1996.
39. Ram, T. G., Reeves, R., and Hosick, H. L. Elevated high mobility group-I(Y) gene expression is associated with progressive transformation of mouse mammary epithelial cells. *Cancer Res.*, *53*: 2655-2660, 1993.
40. Tamimi, Y., Van der Poel, H. G., Denyn, M.-M., Umbas, R., Karthaus, H. F. M., Debruyne, F. M. J., and Schalken, J. A. Increased expression of high mobility group protein I(Y) in high-grade prostatic cancer determined by *in situ* Hybridization. *Cancer Res.*, *53*: 5512-5516, 1993.
41. Berlingieri, M. T., Manfioletti, G., Santoro, M., Bandiera, A., Visconti, R., Giacotti, V., and Fusco, A. Inhibition of HMGI-C protein synthesis suppresses retrovirally induced neoplastic transformation of rat thyroid cells. *Mol. Cell. Biol.*, *15*: 1545-1553, 1995.
42. Tallini, G., Dal Cin, P., Rhoden, K. J., Chiappetta, G., Manfioletti, G., Giacotti, V., Fusco, A., Van den Berghe, H., and Sciot, R. Expression of HMGI-C and HMGI(Y) in ordinary lipoma and atypical lipomatous tumors: immunohistochemical reactivity correlates with karyotypic alterations. *Am. J. Pathol.*, *151*: 37-43, 1997.
43. Choong, P. F. M., Åkerman, M., Willén, H., Anderson, C., Gustafson, P., Baldetorp, B., Fernö, M., Alvegård, T., and Rydholm, A. Prognostic value of Ki-67 expression in 182 soft tissue sarcomas. Proliferation-a marker of metastasis? *APMIS*, *102*: 915-924, 1994.
44. Takano, Y., Saegusa, M., Ikenaga, M., Mitomi, H., and Okayasu, I. Apoptosis of colon cancer: comparison with Ki-67 proliferative activity and expression of p53. *J. Cancer Res. Clin. Oncol.*, *122*: 166-170, 1996.
45. Okayasu, I., Saegusa, M., Fujiwara, M., Hara, Y., and Rose, N. R. Enhanced cellular proliferative activity and cell death in chronic thyroiditis and thyroid papillary carcinoma. *J. Cancer Res. Clin. Oncol.*, *121*: 746-752, 1995.
46. Katoh, R., Bray, C. E., Suzuki, K., Komiyama, A., Hemmi, A., Kawaoi, A., Oyama, T., Sugai, T., and Sasou, S. Growth activity in hyperplastic and neoplastic human thyroid determined by an immunohistochemical staining procedure using monoclonal antibody MIB-1. *Hum. Pathol.*, *26*: 139-146, 1995.
47. Healy, E., Angus, B., Lawrence, C. M., and Rees, J. L. Prognostic value of Ki67 antigen expression in basal carcinomas. *Br. J. Dermatol.*, *133*: 737-741, 1995.
48. Chiappetta, G., Avantaggiato, V., Visconti, R., Fedele, M., Battista, B., Trapasso, F., Merciai, B. M., Fidanza, V., Giacotti, V., Santoro, M., Simeone, A., and Fusco, A. High level expression of the HMGI(Y) gene during embryonic development. *Oncogene*, *13*: 2439-2446, 1996.

**Sp 1 and CTF/NF-1 transcription factors are involved in the basal expression of the *Hmgi-c* proximal promoter.**

Alessandra Rustighi<sup>1</sup>, Fiamma Mantovani<sup>1,2</sup>, Alfredo Fusco<sup>3</sup>, Vincenzo Giaccotti<sup>1</sup> and Guidalberto Manfioletti<sup>1,4</sup>

<sup>1</sup>Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole. Università di Trieste, Trieste, Italy. <sup>3</sup>Dipartimento di Medicina Sperimentale e Clinica. Università di Reggio Calabria, Catanzaro, Italy.

<sup>2</sup> Present address: International Centre for Genetic Engineering and Biotechnology. AREA Science Park, Padriciano 99, 34012 Trieste, Italy.

<sup>4</sup>To whom correspondence should be addressed: Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole. via Giorgieri, 1-34127 Trieste, Italy. Tel.: 39-040-6763675; Fax: 39-040-6763694; e-mail: manfiolo@univ.trieste.it

Abbreviations used: HMGI, High Mobility Group I; ppyr/ppur, polypyrimidine/polypurine; LLC, Lewis lung carcinoma; ORF, open reading frame; bp, base pair; RACE, rapid amplification of cDNA ends; Kbp, kilobase pair; EMSA, electrophoretic mobility shift assay, UTR, untranslated region.

<sup>4</sup>This information should appear on the title page of the printed article.

## ABSTRACT

HMGI-C is a nuclear architectural factor which is expressed during embryogenesis but not in adult tissues while it becomes re-expressed following neoplastic transformation. In this paper we identify the promoter region of the mouse *Hmgi-c* gene and by stepwise deletion of the 5' sequences we map the promoter activity of the most abundant transcript to a very short fragment containing a long polypyrimidine/polypurine (ppyr/ppur) tract. We demonstrate that this tract is a multiple binding site for the transcription factors Sp1 and Sp3 and that in *Drosophila* SL2 cells Sp1 activates the *Hmgi-c* promoter. In addition another transcription factor, CTF/NF-1, binds the proximal promoter immediately downstream of this region and its mutation decreases transcription in NIH-3T3 cells. This study identifies factors responsible for the basal activity of *Hmgi-c* gene and provides a foundation for further analysis of the mechanism of its regulation.

## INTRODUCTION

Transcriptional regulation involves the assembly of multiprotein complexes at enhancers and promoters. In the assembly of these higher-order nucleoprotein complexes several DNA-binding proteins act as architectural elements; among these there is the high mobility group (HMG) I family of nuclear factors (1). This group is composed by three members: HMGI and HMGY, together called HMGI(Y), which result from alternative splicing of the same gene (2, 3) and HMGI-C that, instead, is encoded by a different gene (4, 5). These factors have no transcriptional activity *per se* but they can orchestrate the assembly of nucleoprotein structures through a complex network of protein-DNA and protein-protein interactions resulting in transcription potency. The enhanceosome of the interferon  $\beta$  promoter is probably the best studied example (6), but a large number of genes such as cytokines, cell adhesion molecules, growth factors and transcription factors have been demonstrated to be regulated, either positively or negatively, by HMGI proteins, therefore demonstrating an important role for these cofactors in chromatin architecture and in regulation of gene expression (1, 7, 8).

A strong association between the transformed phenotype and the high expression of HMGI proteins is well established (9, 10). In fact they have been proposed as useful markers in the evaluation of the neoplastic phenotype of different tumors including thyroid, colorectal and cervical carcinoma (11, 12, 13). For HMGI-C in particular, a causal role in cell transformation has been demonstrated since the inhibition of its protein synthesis prevents thyroid cell transformation by different oncogenes such as *v-Ki-ras* and *v-mos* (14). In addition the dysregulation of the *HMGI-C* gene structure is correlated to the onset of several human benign tumors of mesenchymal origin such as leiomyomas and lipomas (15, 16).

The physiological expression of *Hmgi-c* is restricted to early embryogenesis and it decreases with the beginning of organogenesis in the mouse embryo. A role in development is in fact

underscored by the finding that the inactivation of the *Hmgi-c* gene in mouse is responsible for the *pygmy* phenotype, which shows a growth retardation and a particular reduction of body fat (17).

No HMGI-C has been detected in normal adult cells, but the protein becomes re-expressed in transformed cells and in experimental tumors (4, 9, 10). Run-on experiments with rat thyroid cells transformed with different oncogenes have provided evidence that the regulation of the expression of HMGI proteins occurs at the transcriptional level (14), but the molecular mechanisms that are responsible for this re-expression are not known. To elucidate these mechanisms we have previously cloned the mouse *Hmgi-c* gene (5) and in this study we report the characterization of *Hmgi-c* gene promoter sequences identifying nuclear factors responsible for its basal activity.



## MATERIALS AND METHODS

### RNA extraction and Northern blot analysis.

Total cellular RNA was isolated from Lewis lung carcinoma (LLC) and from mouse liver by the guanidine isothiocyanate method (18). For poly(A)<sup>+</sup> preparation, total RNA was selected by affinity chromatography using PolyA Tract mRNA Isolation system (Promega). For Northern blot analysis equal amounts of poly(A)<sup>+</sup> RNA were loaded on a 1% agarose/formaldehyde gel, electrophoresed and transferred to GeneScreen membrane (Du Pont). Northern blots were probed with labelled fragments derived from mouse *Hmgi-c* cDNA sequences (5). Probe ORF is a 437bp *PstI-PstI* fragment (from +157 to +593, numbers refer to the major transcription initiation site determined in this paper) containing the ORF, probe A is a 164 bp *AatII-PstI* fragment (from -8 to +156), probe B is a 261 bp *EcoRI-AatII* fragment (from -270 to -9).

### Determination of the transcription initiation site.

RACE experiments were performed essentially as described (19). 1 µg of poly(A)<sup>+</sup> RNA from LLC was reverse-transcribed for 1h at 42° C using 200 U of Moloney murine leukaemia virus (M-MLV) reverse transcriptase (Gibco BRL) and the *Hmgi-c* specific antisense primer AR3 (5'-CCGCTACCACTGCCTCTA-3', from +187 to +170). Amplification proceeded using a nested primer, AR15 (5'-CTCGGCCTCCGCCGCCCGAAT-3', from +148 to +127) and the 15CEco primer (5'-GATGGAATTCCCCCCCCCCCCCCC-3') complementary to the polyG tail that was added at the 5' end of the cDNA. Conditions for PCR were 1 min at 94° C, 1 min at 55° C and 1 min 30 sec at 72° C for 40 cycles. The PCR products were analyzed on a 1.5 % agarose gel and by Southern blotting using another internal primer AR16 (5'-TTCGGGAGACGGGATGTA-3', from +113 to +96) as a probe. One AR16 positive band of approximately 160 bp was identified and subcloned into pBluescript KS+ cloning vector (Stratagene). The resulting clones were sequenced.

For primer extension analysis, 1 ng of end-labelled AR16 ( $2.5 \times 10^5$  cpm) was hybridized for 1h at 50° C with 10 µg of total RNA extracted either from LLC or from mouse liver. Primer extension was carried out with 200 U of M-MLV reverse transcriptase (Gibco BRL) for 90 min at 42° C. The synthesized product was electrophoresed on a 7 % denaturing polyacrilamide gel. The gel was dried under vacuum and exposed to x-ray film (Kodak XAR-5).

### **Construction of *Hmgi-c*/Luciferase Reporter Vectors.**

*Hmgi-c*/Luciferase reporter plasmids were constructed starting from the previously isolated genomic clone  $\lambda$  I-C.5 (5). The 0.65 Kbp *Bam*HI-*Pst*I restriction fragment, containing the transcription start, was cloned into pBluescript KS+. The same fragment was cut out with *Sac*I and *Hind*III and cloned into the polylinker of the luciferase reporter plasmid pGL2-Basic (Promega). This construct was called BaP. The longest construct of 3.2 Kbp (SaP) was obtained by subcloning the *Sac*I-*Ssp*I genomic fragment into the BaP construct using *Sac*I of the polylinker and *Ssp*I internal to the *Bam*HI-*Pst*I genomic fragment. Sequential 5' deletion constructs were prepared starting from SaP and using restriction sites *Pvu*II, *Sac*II, *Hae*II and *Aat*II, internal to the *Hmgi-c* DNA obtaining the constructs PvuP, SSaP, HaP and AaP respectively. The 4P construct was produced by ligating the amplification product of the oligonucleotides AR4 and AR3 (two 18-mers complementary to nucleotides -282 to -265 and +187 to +170 respectively) cut with *Pst*I at the 3' extremity into a vector obtained digesting BaP with *Sma*I and with *Pst*I. 4P-HA is a 75 bp internal deletion mutant of 4P, made by digesting 4P with the restriction enzymes *Hae*II and *Aat*II, cutting internally in the *Hmgi-c* DNA. NF1mut was obtained from HaP by PCR amplification using oligo A sense and mD antisense (see next chapter), which carries the CTF/NF-1 specific mutation. Construct H was obtained digesting SaP with *Hind*III, which cuts at -1325 in the promoter and in the polylinker of pGL2-Basic, and rejoining. SaB was originated by subcloning the 2.4 Kbp genomic fragment *Sac*I-*Bam*HI in the polylinker of the pGL2 Basic vector.

## Electrophoretic Mobility Shift Assay (EMSA), Supershift and DNaseI-footprinting.

Nuclear extracts from confluent cell cultures were prepared essentially as described (20). Protein concentration was measured using the Bradford Microassay procedure (Biorad). Oligonucleotides were radiolabeled at the 5' ends using [ $\gamma$ - $^{32}$ P]ATP (3000Ci/mmol, Du Pont) and 10 U of T4 polynucleotide kinase (Epicentre Technologies). 50 fmoles of oligonucleotide (30.000 cpm) were incubated with 1-4  $\mu$ g of protein nuclear extract or 5, 10 or 30 picomoles of CTF/NF-1 protein, purified from K562 cells (kind gift of C. Santoro), in 20  $\mu$ l reactions containing 20mM Tris/HCl pH 7.5, 75 mM KCl, 100 mM NaCl, 0.2  $\mu$ M ZnCl<sub>2</sub>, 3.5 mM DTT, 1  $\mu$ g/ $\mu$ l BSA, 6% glycerol, 0.5  $\mu$ g of poly(dI-dC), 0.3mM MgCl<sub>2</sub>, 0.175 mM EDTA, 0.02 mM EGTA. EMSAs for CTF/NF-1 were done in the same conditions, but without ZnCl<sub>2</sub> and with 250 mM NaCl. For competition binding reactions the unlabeled competitor was included in the reaction in 100-fold molar excess of the labeled probe. After incubation for 20 min. at RT, samples were loaded onto a native 5% polyacrilamide gel in 0.5 X TBE buffer, electrophoresed at 4° C, at 15 V/cm, dried and exposed to x-ray film (Hyperfilm MP, Amersham). For supershift experiments 0.1  $\mu$ g of  $\alpha$ -Sp3,  $\alpha$ -Sp1 or  $\alpha$ -Oct-2 antibody (Santa Cruz Biotechnology, Inc.) were incubated for 20 min. at RT with the nuclear extracts prior to addition of the labeled probe. The following oligonucleotides (only the upper strand is shown) were used in EMSA experiments: A (5'-CAGCGCCTCCTCCTTTCCCTCCTCCTC-3'; from - 89 to - 64); B (5'-TCCTCCTCTCCTCCTCCTCTCT-3'; from - 68 to - 47); C (5'-TCTCTTCCTCCTCCCCTCTCTTTTT-3'; from - 51 to - 25); D (5'-TCTTTTTGGCAGCCGCTGGACGTCCGGTGTTGA; from -31 to + 2); mD (mutation is indicated in bold print) (5'-TCTTTTT**T**GCAGCCGCTGGACGTCCGGTGTTGA-3'); Sp1 (5'-ATTCGATCGGGGCGGGGCGAGC-3'); AP1 (5'-CGCTTGATGAGTCAGCCGGAA-3'); NF1 (5'-CTAGCTATTTTGGATTGAAGCCAATATG-3'). For DNaseI-footprinting the *Hmgi-c* specific

*HaeII-HindIII* fragment obtained from BaP, was dephosphorilated, labelled by T4 polynucleotide kinase and digested with *PstI*. The PAGE-purified labelled *HaeII-PstI* DNA fragment (20.000 c.p.m.) was assayed using recombinant Sp1 (Promega) as suggested by the manufacturer or purified CTF/NF-1. DNaseI-footprinting competition experiments were carried out in the presence of a 100-fold molar excess of cold oligonucleotide.

### **Cell cultures and transfections.**

NIH-3T3 and QT6 (quail fibrosarcoma) (21) were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum at 37°C in a humidified 5% CO<sub>2</sub> incubator. FRTL5 KiMol cells were maintained in Coon's modified Ham's F12 medium supplemented with 5% fetal calf serum (14). For transfections 5 µg of reporter construct and, as internal control, 0.4 µg of pSV β-galactosidase expression vector, for NIH-3T3 and QT6, or 0.25 µg of pRL-TK *Renilla* luciferase expression vector (Promega) for FRTL5 KiMol were used. Cells were transfected by the calcium phosphate coprecipitation procedure and assayed as already described (7).

*Drosophila* Schneider SL2 cells were grown at 25° C in Schneider's medium (Life Technologies Inc.) supplemented with 10% heat inactivated fetal calf serum and transfected by the calcium phosphate method. The pPac empty vector was included in the transfections to compensate for different amounts of the pPacSp1 and pPacSp3 used (22, 23). The β-galactosidase expression vector pPacβ-gal was included to normalize for transfection efficiency. Luciferase and colorimetric β-galactosidase assays were performed as for NIH-3T3 cells as already described (7).

## RESULTS

### Identification of the transcription start sites.

Three *Hmgi-c* transcripts of 4.7, 4.2 and 3.9 Kb are revealed by Northern blot analysis, using poly (A)+ RNA from Lewis lung carcinoma (LLC), a spontaneously derived mouse tumor that we previously used for *Hmgi-c* cDNA cloning (4), with a probe derived from the open reading frame (ORF) of the *Hmgi-c* cDNA (Fig.1). The two longer transcripts have been shown to contain longer 5'UTR (5) and probably derive from upstream promoters. We focused our attention on the lower transcript because it clearly represents almost the entire amount of the *Hmgi-c* specific mRNA. In order to map the transcription initiation site of the most abundant transcript, we have performed additional Northern blot analysis using two contiguous short probes, A and B, located upstream of probe ORF. Mouse liver poly (A)+ RNA (L) was used as negative control. The pattern observed using probe A was the same of that obtained with probe ORF; on the contrary the signal corresponding to the 3.9 Kb transcript almost disappears using probe B (Fig. 1). These results indicate that the major transcription start should be positioned between probe A and B, around the recognition site of the restriction enzyme *AatII* used for the generation of the two probes.

The transcription start site was precisely determined by 5' RACE experiments. Primers complementary to sequences of the 5'UTR were selected for cDNA synthesis (AR3) and for PCR amplification (AR15), the last one being about 100 bp downstream from the *AatII* restriction site. One prominent band, corresponding to the expected length was obtained and the DNA was subcloned. Twelve independent clones were sequenced, six of these started at the same G located 201 bp upstream from the translation initiation codon and the remaining clones originated at different positions only at a few nucleotides distance (Fig. 2A). All RACE products reside in the vicinity of the *AatII* restriction site and are thus in full agreement with the Northern blot data.

To further validate the results obtained by RACE, a primer elongation was carried out on total RNA from LLC using a different oligonucleotide, AR16. One elongation product was obtained, corresponding to the nucleotide +1 already identified by the RACE approach (Fig.2B).

In conclusion, all these results show that the major *Hmgi-c* transcript arises from a narrow window, located about 200 bp upstream of the starting methionine and that there is one nucleotide which is preferentially used, that we identify as +1.

### **Analysis of the promoter sequence.**

The genomic sequence surrounding the major transcription start site was analyzed for putative regulatory elements using TFD and MatInspector data banks. The sequence upstream of the major initiation transcription site lacks TATAA and CCAAT box consensus elements. This is similar to the gene coding for the other two HMGI proteins (HMGI and HMGY) (3) and can explain why the transcription start site is heterogeneous. The analysis of the CpG profile of the mouse *Hmgi-c* 5' region indicates the presence of a CpG island that extends from the 5' flanking region through the first exon till the first intron (-1300 to +500). Several potential transcription factor binding sites were found, but only those which we found to be conserved also in the human *HMGI-C* gene (24) have been considered potentially important and shown in Fig.2A. Of particular interest is a ppyr/ppur tract, 60 bp long (from -84 to -25), consisting almost completely of a repetition of the trinucleotide TCC (TCC repeat). Similar sequences with asymmetrical strand composition have been found in the promoters of other genes such as EGF-R, c-Ki-*ras* and PDGF A-chain and have been shown to be essential for gene activity (25-27).

### **Identification of *Hmgi-c* minimal promoter.**

To examine the role of isolated putative promoter sequences, a fragment of 3.2 Kbp of the *Hmgi-c* 5' region was cloned upstream of a luciferase reporter gene, in the pGL2-Basic vector. This construct (SaP) was transiently transfected in NIH-3T3 mouse fibroblasts, where the pattern of expression of *Hmgi-c* is the same as that observed in LLC (28 and our observations). This 3.2 Kbp region is able to drive transcription with an efficiency comparable to the SV40 early promoter (pGL2-Promoter), which was arbitrarily set to 100% (Fig. 3). In order to define the minimal promoter, several 5' deletion constructs have been assayed. All constructs containing the first 83 bp of the 5' flanking sequence result in high levels of transcription. Promoter activity was almost abolished with a further deletion to nucleotide -8, which eliminates completely the ppyr/ppur tract.

To confirm these results, an internal deletion of 75 bp, which deletes the TCC repeat, was introduced into construct 4P. In the resulting construct, 4P-HA, the activity dropped again to very low levels. Therefore, an element between -83 to -3, containing the TCC repeat, confers the activity to the promoter of *Hmgi-c*.

The same constructs were also transfected in other cell lines such as FRTL5 KiMol and QT6, both of which express HMGI-C (10 and data not shown). The first are rat thyroid epithelial cells transformed with v-Ki-*ras* while the others were derived from a quail fibrosarcoma. The results obtained in all three cell types are very similar indicating that the regulation of this promoter seems to be conserved in different species and in different cell types (Fig. 3).

Deletions from the 3' end were also tested in the same cell lines in transient transfection assays (Fig. 3). Removal of 658 bp from the 3' end, construct SaB, reduces promoter activity almost completely, while a further deletion, construct H, totally abolishes the luciferase activity even though 1.5 Kbp of 5' flanking region are still present. The residual activity of construct SaB suggests that the 5' region present in this construct could be responsible for the generation of the two longer transcripts detected in Northern blot. These transfection data therefore strengthen the evidence that, in addition to the proximal promoter that we have identified, the mouse *Hmgi-c* gene harbours two other distal promoters, although very weak in all cell lines we analyzed.

### **Sp1 and Sp3 transcription factors bind to the ppyr/ppur region in the *Hmgi-c* proximal promoter.**

Several reports indicate that long ppyr/ppur stretches containing TCC repeats, as that found in the *Hmgi-c* minimal promoter, are multiple weak binding sites for the transcription factor Sp1 (25, 26, 29) and moreover, computer analysis showed a consensus binding site for Sp1 at position -42 to -36. Therefore, to test the possibility that Sp1 could bind in the minimal promoter, three oligos A, B and C, encompassing the TCC repeat region from -89 to -25, were employed in band-shift experiments. A typical experiment using probe C and nuclear extracts from FRTL5 KiMol cells is shown in Fig. 4A, but similar results were obtained with probe A and B and also using NIH-3T3 nuclear extract (not shown). Three major specific complexes (a, b and c) are formed which are

competed using oligos A and B as well as by the self-competing cold probe C (Fig. 4A, lanes 5-7). The three complexes are also competed by an oligo with the consensus for Sp1 (Fig. 4A, lane 9), therefore indicating that they contain Sp1-like factors, while oligo D, derived from the region -31 to +2 downstream the TCC repeats, is unable to compete for the formation of any of the complexes (Fig. 4A, lane 8).

The identity of the factors involved in the formation of complexes a, b and c was further analyzed using specific antibodies in supershift experiments (Fig. 4B). Incubation with Sp1 specific antibodies resulted in supershift of complex a (Fig. 4B, lane 3) while Sp3 specific antibodies supershifted complexes b and c (Fig. 4B, lane 4). The supershifts are specific since both antibodies were not able to produce any supershift when tested in the absence of nuclear extract (Fig.4B, lanes 6, 7) while an unrelated antibody against the transcription factor Oct2 was not able to affect the formation of any complex (Fig.4B, lane 5). Same results were obtained using probe A and B (data not shown).

Specific interaction of Sp1 to the *Hmgi-c* proximal promoter was confirmed by DNaseI-footprinting analysis using purified Sp1 protein (Fig.4C): a clear protection over the whole ppyr/ppur tract was evidenced.

### **Sp1 activates the *Hmgi-c* proximal promoter via the TCC repeat.**

To further investigate the ability of Sp1 and Sp3 transcription factors to bind and regulate the *Hmgi-c* promoter, we carried out co-transfection experiments using the *Drosophila melanogaster* Schneider SL2 cells, which are devoid of endogenous Sp1-like activity and therefore suitable to test Sp1-mediated activation *in vivo* (22). SL2 cells were transfected with HaP and AaP reporters together with expression vectors for Sp1 (pPacSp1) and Sp3 (pPacSp3) that were shown to express Sp1 and Sp3 at about the same level (22, 23). Activity of the HaP construct increased in a dose dependent manner when co-transfected with increasing amounts of pPacSp1. In contrast the AaP construct, where the ppyr/ppur tract is deleted, is not activated at all (Fig.5). Co-transfection with Sp3 expression vector alone was not able to induce any activation.

Since it was observed that Sp3 can either repress or enhance Sp1-mediated transcriptional activation, co-transfections with a variable amount of pPacSp3 and a fixed amount of pPacSp1 were



carried out. Experiments reported in Fig. 5 show that Sp3 does not repress, but rather slightly enhances Sp1-mediated transcriptional activation.

**CTF/NF-1 transcription factor binds from -28 to -10 downstream the ppyr/ppur tract.**

Factors binding downstream the TCC repeat region were also investigated using oligo D that covers the region from -31 to +2. Using nuclear extract from NIH-3T3 cells (Fig. 6A, lanes 2-4) (or from FRTL5 KiMol cells, data not shown) one prominent band is observed in band-shift experiments. The formation of this complex is competed by an excess of cold probe D (Fig. 6A, lane 5) or by the oligonucleotide NF1, containing a high affinity binding site for the CTF/NF-1 transcription factor (30) derived from the adenovirus origin of replication (Fig. 6A, lane 7). This competitor was used since a half-palindromic binding site for CTF/NF-1 transcription factor was identified by computer analysis from position -28 to -10 in the *Hmgi-c* promoter. On the contrary, the complex is not competed by a mutated form of oligo D (mD) (Fig. 6A, lane 6) where a single mutation, known to drastically reduce the binding for CTF/NF-1, was introduced (31), by an unrelated oligo with the consensus of the AP-1 binding site (Fig. 6A, lane 8) or by probe C (Fig. 6A, lane 9).

Since antibodies capable of supershifting CTF/NF-1 were not available we used purified CTF/NF-1 in a series of band-shift experiments. The purified CTF/NF-1 protein produces a retarded complex with oligo D (Fig. 6A, lanes 11-13) very similar to that obtained with the entire nuclear extracts and moreover it behaves exactly in the same way when assayed with the same competitors (Fig. 6A, lanes 14-18). Moreover, footprinting analysis using the purified protein shows the protection of the region from -28 to -12 which encompasses the TGGC motif that represents the target half-site for CTF/NF-1 (data not shown).

Finally, in order to assess the importance of CTF/NF-1 in *Hmgi-c* promoter activity, we performed transient transfection experiments in NIH-3T3 cells with a construct derived from HaP, called NF1mut, in which the CTF/NF-1 binding site has been destroyed, introducing the same

mutation as in oligo mD by site-specific mutagenesis. Transfection experiments show that CTF/NF-1 contributes *in vivo* to transcriptional activation since the activity of NF1mut construct is about 60% of the wild type construct HaP (Fig. 6B).

## DISCUSSION

The aim of this study was to identify cis-elements and transacting factors important for transcriptional regulation of the mouse *Hmgi-c* gene. We have therefore mapped the initiation of transcription for the most abundant transcript and, using transient transfection assays we provide evidence that the minimal promoter region is contained in a very short fragment of 83 bp upstream the major transcription start site. This small region is as effective as the 3 Kbp sequence (SaP construct) in conferring transcriptional activity to the reporter gene in a number of different cell lines derived from different species. A striking feature of this region is the presence of a very long ppyr/ppur tract. The existence of ppyr/ppur stretches in promoter-proximal locations has been reported for a number of growth related genes (25-27). These elements most likely have a role in transcription initiation, since they are multiple binding sites for the transcription factor Sp1, which has been shown to recruit and stabilize the TFIID complex (32). By band-shift, competition, supershift and footprinting experiments we have demonstrated that the ppyr/ppur tract in the *Hmgi-c* promoter is a multiple binding site for Sp1 and Sp3 and that, at least in the *Drosophila* SL2 cells, Sp1 activates the *Hmgi-c* proximal promoter through the ppyr/ppur tract. Therefore we suggest that, in analogy to other TATA-less genes, the binding of several Sp1 molecules could have a role in transcription initiation from the *Hmgi-c* proximal promoter.

Sp3, which can act both as a repressor or an activator, either competing (23) or synergizing with Sp1 (33) is able to bind the same sequence. Our results with SL2 cells clearly demonstrate that Sp3 by itself has no activity on the promoter and it can only moderately enhance the activity of Sp1 in co-transfection experiments.

The minimal promoter region contains a putative binding site, conserved also in the human gene, for the trans-acting regulatory protein CTF/NF-1. This transcription factor was originally described to be involved in the replication of adenovirus DNA in HeLa cells (34), but was later

found to regulate the transcription of several viral and cellular genes. CTF/NF-1 can stimulate transcription by itself, but binding sites for this factor are frequently found in or near other nuclear factor binding sites and both functional co-operation and competition have been reported (35, 36). CTF/NF-1 binds DNA through a palindrome containing two TGGC sequences on opposite strands, but several reports indicate that it is also able to bind half-palindromes in many eukaryotic promoters. Half CTF/NF-1 consensus sequence is present in the *Hmgi-c* promoter very close to the Sp1 binding sites; by means of different competition experiments and the use of purified protein, we show that it is able to bind the CTF/NF-1 factor *in vitro* and, moreover, that this binding site is functional *in vivo* since its mutation reduces the activity of the reporter construct in NIH-3T3 cells. This result thus excludes competition between CTF/NF-1 and Sp1 but further experiments will be required to assess if there is any co-operation between these factors in the regulation of *Hmgi-c* promoter.

HMGI-C is not expressed in normal cells of adult animals, but only in embryos and transformed cells (4, 17). It is thus of interest to understand the mechanism(s) through which this differential expression is achieved. The minimal promoter identified in this study is responsible for basal activity but, experiments in our laboratory indicate that it is also much more active in transformed cells than in their normal counterpart (unpublished data), and work is in progress to identify cis-elements and trans-acting factors responsible for this differential activity. Although both Sp1 and CTF/NF-1 are considered constitutive transcription factors it is becoming increasingly clear that their activity can be modulated by a variety of stimuli, therefore we do not exclude that they could be involved in the differential expression of *Hmgi-c* gene or its induction by different growth factors (28). For example, a role in directing tissue-specific, developmental and hormonal regulated gene expression has been demonstrated for Sp1 (37, 38) while CTF/NF-1 has been involved in adipocyte specific gene expression (39), in transformation by nuclear oncogenes (40) and in gene regulation by insulin (41). Moreover, other putative binding sites have been found in the

*Hmgi-c* minimal promoter, and some of them deserve particular attention. Particularly interesting are Elk-1, which is a transcriptional activator acting downstream the activated pathway of Raf-1 that has been shown to induce the *Hmgi-c* gene (42) and PEA3, which is a common enhancer element found near the transcriptional start sites of many TATA-less promoters that play important roles in embryogenesis, angiogenesis and tumor metastasis (43).

In conclusion, we believe that the definition of the promoter region able to confer basal activity to the *Hmgi-c* gene and the identification of functional binding sites for Sp1, Sp3 and CTF/NF-1, reported in this paper, represent an important step in understanding the molecular events that regulate the expression of the *Hmgi-c* gene during the processes of cell transformation and development.

## ACKNOWLEDGEMENTS

We are particularly indebted to C. Santoro (University of Turin) for kindly providing the purified CTF/NF-1 and for helpful suggestions, to B. Maiello (IIGB, Naples) for SL2 cells, pPacSp1 and pPacSp3 vectors, and to R. Tjian (MCB, Berkeley) for pPac $\beta$ -gal vector. Acknowledgements are also due to Fulvia Vascotto and Sonia Covaceuszach for technical assistance. This study was supported by Associazione Italiana per la Ricerca sul Cancro (AIRC), Milano, Italy; Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Roma, Italy (grant n° 9706274625) and Università degli Studi di Trieste, Italy. F.M. was a recipient of a fellowship from Fondazione Italiana per la Ricerca sul Cancro (FIRC).

## REFERENCES

1. Bustin, M., and Reeves, R. (1996) In: W.E. Cohn and K. Moldave (eds.) *Progr. Nucleic Acid Res.* **54**, 35-100.
2. Johnson, K.R., Lehn, D.A., and Reeves, R. (1989) *Mol. Cell. Biol.* **9**, 2114-2123.
3. Friedmann, M., Holth, L.T., Zoghbi, H.Y., and Reeves, R. (1993) *Nucleic Acids Res.* **21**, 4259-4267.
4. Manfioletti, G., Giancotti, V., Bandiera, A., Buratti, E., Sautière, P., Cary, P., Crane-Robinson, C., Coles, B., and Goodwin, G.H. (1991) *Nucleic Acids Res.* **19**, 6793-6797.
5. Manfioletti, G., Rustighi, A., Mantovani, F., Goodwin, G.H., and Giancotti, V. (1995) *Gene* **167**, 249-253.
6. Thanos, D., and Maniatis, T. (1992) *Cell* **71**, 777-789.
7. Mantovani, F., Covaceuszach, S., Rustighi, A., Sgarra, R., Heath, C., Goodwin, G.H., and Manfioletti, G. (1998) *Nucleic Acids Res.* **26**, 1433-1439.
8. Arlotta, P., Rustighi, A., Mantovani, F., Manfioletti, G., Giancotti, V., Tell, G., and Damante, G. (1997) *J. Biol. Chem.* **272**, 29904-29910.
9. Giancotti, V., Buratti, E., Perissin, L., Zorzet, S., Balmain, A., Portella, G., Fusco, A., and Goodwin, G.H. (1989) *Exp. Cell Res.* **184**, 538-545.
10. Giancotti, V., Pani, B., D'Andrea, P., Berlingieri, M.T., Di Fiore, P.P., Fusco, A., Vecchio, G., Philip, R., Crane-Robinson, C., Nicolas, R.H., Wright, C.A., and Goodwin, G.H. (1987) *EMBO J.* **6**, 1981-1987.
11. Chiappetta G., Bandiera, A., Berlingieri, M.T., Visconti, R., Manfioletti, G., Battista, S., Martinez-Tello, F.J., Santoro, M., Giancotti, V., and Fusco, A. (1995) *Oncogene* **10**, 1307-1314.

12. Fedele, M., Battista, S., Manfioletti, G., Chiappetta, G., Viglietto, G., Casamassimi, A., Bandiera, A., Santoro, M., Giancotti, V. and Fusco A. (1996) *Cancer Res.* **56**, 1896-1901.
13. Bandiera, A., Bonifacio, D., Manfioletti, G., Mantovani, F., Rustighi, A., Zanconati, F., Fusco, A., Di Bonito, L., and Giancotti, V. (1998) *Cancer Res.* **58**, 426-431.
14. Berlingieri, M.T., Manfioletti, G., Santoro, M., Bandiera, A., Visconti, R., Giancotti, V., and Fusco, A. (1995) *Mol. Cell. Biol.* **15**, 1545-1553.
15. Ashar, H.R., Fejzo, M.S., Tkachenko, A., Zhou, X., Fletcher, J.A., Weremowicz, S., Morton, C.C., and Chada, K. (1995) *Cell* **14**, 57-65.
16. Schoenmakers, E.F., Wanschura, S., Mols, R., Bullerdiek, J., Van den Berghe, H., and Van de Ven, W.J.N. (1995) *Nature Genetics* **10**, 436-444.
17. Zhou, X., Benson, K.F., Ashar, H.R., Chada, K. (1995) *Nature* **376**, 771-774.
18. Chomczynsky, P., and Sacchi, N.. (1987) *Anal. Biochem.* **162**,156-159.
19. Del Sal, G., Ruaro, M.E., Philipson, L., and Schneider, C. (1992) *Cell* **70**, 595-607.
20. Tell, G., Scaloni, A., Pellizzari, L., Formisano, S., Pucillo, C., and Damante, G. (1998) *J. Biol. Chem.* **273**, 25062-25072.
21. Moscovici, C., Moscovici, M.G., Jimenez, H., Lai, M.M., Hayman, M.J., and Vogt, P.K. (1977) *Cell* **11**, 95-103.
22. Majello, B., De Luca, P., Suske, G., and Lania, L. (1995) *Oncogene* **10**, 1841-1848.
23. Hagen, G., Mueller, S., Beato, M., and Suske, G. (1994) *EMBO J.* **13**, 3843-3851.
24. Chau, K.Y., Patel, U.A., Lee, K.L.D., Lam, H.Y.P., and Crane-Robinson, C. (1995) *Nucleic Acids Res.* **23**, 4262-4266.
25. Johnson, A., Jinno, Y., and Merlino, G.T. (1988) *Mol. Cell. Biol.* **8**, 4174-4184.
26. Raghu. G., Tevosian, S., Anant, S., Subramanian, K.N., George, D.L., and Mirkin, S.M. (1994) *Nucleic Acids Res.* **22**, 3271-3279.

27. Wang, Z., and Deuel T.F. (1996) In: W.E. Cohn and K. Moldave (eds.) *Prog. Nucleic Acid Res. Mol. Biol.* **55**, 227-244.
28. Lanahan, A., Williams, J.B., Sanders, L.K., and Nathans, D. (1992) *Mol. Cell. Biol.* **12**, 3919-3929.
29. Chen, S., Supakar, P.C., Vellanoweth, R.L., Song, C.S., Chatterjee, B., and Roy, A.K. (1997) *Mol. Endocrinol.* **11**, 3-15.
30. Santoro, C., Mermod, N., Andrews, P.C., and Tjian, R. (1988) *Nature* **334**, 218-224.
31. Chen, A., Beno, D.W., and Davis, B.H. (1996) *J. Biol. Chem.* **271**, 25994-25998.
32. Pugh, B.F. and Tjian, R. (1991) *Genes Dev.* **5**, 1935-1945.
33. Jiang, J.G., Chen, Q., Bell, A., and Zarnegar, R.. (1997) *Oncogene* **14**, 3039-3049.
34. Nagata, K., Guggenheimer, R.A., Emoto, T., Lichy, J.H., and Hurwitz, J. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6438-6442.
35. Martinez, E., Dusserre, Y., Wahli, W., and Mermod, N. (1991) *Mol. Cell. Biol.* **11**, 2937-2945.
36. Fisher, K.D., Haese, A., and Nowock, J. (1993) *J. Biol. Chem.* **268**, 23915-23923.
37. Marin, M., Karis, A., Visser, P., Grosveld, F., and Philipsen, S. (1997) *Cell* **89**, 619-628.
38. Ahlgren, R., Suske, G., Waterman, M.L., and Lund, J. (1999) *J. Biol. Chem.* **274**, 19422-19428.
39. Greaves, R.A., Tontonoz, P., Ross, S.R., and Spiegelman, B.M. (1991) *Genes Dev.* **5**, 428-437.
40. Schuur, E.R., Kruse, U., Iacovoni, J.S., and Vogt, P.T. (1995) *Cell Growth Differ.* **6**, 219-227.
41. Cooke, D., and Lane, M.D. (1999) *J. Biol. Chem.* **274**, 12917-12924.
42. Li, D., Lin, H., McMahon, M., Ma, H., and Ann, D.K. (1997) *J. Biol. Chem.* **272**, 25062-25070.



43. Yu, M., Yang, X.Y., Schmidt, T., Chinenov, Y., Wang, R., and Martin, M.E. (1997) *J. Biol. Chem.* **272**, 29060-29067.

## FIGURE LEGENDS

**FIG. 1.** Northern blot analysis of the 5' end of murine *Hmgi-c*. Five micrograms of poly(A)<sup>+</sup> RNA from Lewis lung carcinoma (LLC) and mouse liver (L), as negative control, were used for Northern blot analysis. Probes B (from -270 to -9, numbers refer to the initiation site determined in this paper), A (from -8 to +156), and ORF (from +157 to +593) were derived from the *Hmgi-c* cDNA. The arrows on the left indicate the three transcripts of the murine *Hmgi-c* gene. 28S and 18S ribosomal RNAs are indicated on the right. The same blots were also probed with a GAPDH cDNA fragment.

**FIG. 2.** Mapping of the transcription initiation site. (A) 340 bp of the 5' flanking region of the murine *Hmgi-c* gene are shown. The major transcription initiation site is indicated with a black arrow and numbering refers to it. The previously detected murine *Hmgi-c* cDNA 11a1 (4) is given by a bent arrow. The 5'RACE products are marked with an arrowhead. Putative binding sites for transcription factors are underlined and the corresponding name is indicated below the sequence. The ppyr/ppur region is indicated in bold and italicized. Restriction sites relevant for the generation of the *Hmgi-c* promoter containing constructs are shown above the sequence. (B) Primer extension analysis of the 5' end of the *Hmgi-c* mRNA is shown. cDNA synthesis was carried out by priming 10 µg of total RNA from Lewis lung carcinoma (LLC) and mouse liver (L) with labeled oligonucleotide AR16. The size of the primer extension product was determined by comparing it with the sequence of the genomic *Hmgi-c* primed with the same oligonucleotide. The arrow indicates the major transcription initiation site.

**FIG. 3.** Functional analysis of the *Hmgi-c* promoter in NIH-3T3, QT6 and FRTL5 KiMol cells. NIH-3T3, QT6 and FRTL5 KiMol cells were transiently transfected with the promoter construct

SaP and a series of its deletion mutants, as represented schematically on the left part of the figure. The construct SSaP was assayed only in NIH-3T3 and QT6 cells. For each construct, both the 5' and the 3' end of the promoter fragments, relative to the major transcription start site (bent arrow), are indicated. The transcription efficiency is expressed as percentage of that measured in cells transfected with the SV40 early promoter (pGL2-Promoter) and set arbitrarily equal to 100%. Columns and bars represent the mean and standard deviation, respectively, of at least five independent experiments. A vector expressing  $\beta$ -galactosidase from an SV40 promoter was included as internal control.

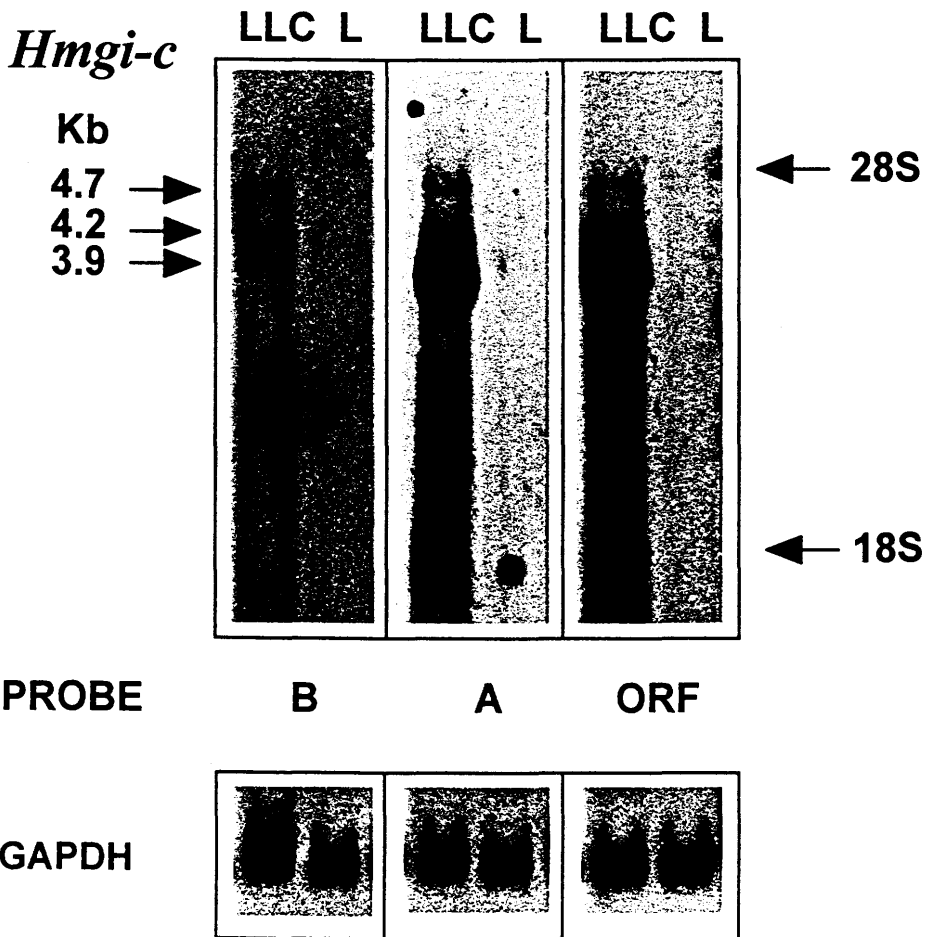
**FIG. 4.** Sp1 and Sp3 transcription factors bind to the ppyr/ppur region of the *Hmgi-c* gene promoter. (A) Probe C was incubated with increasing amounts of nuclear extracts (N.E.) (lanes 2-9), indicated in  $\mu$ g, from FRTL5 KiMol cells in EMSA experiments. The shifted complexes, discussed in the text, are shown with an arrow. Different competitors (lanes 5-9) at a 100-fold molar excess concentration were included to establish the specificity and identity of the shifted complexes. A, B, C, D oligos derive from the sequence of the *Hmgi-c* promoter while oligo Sp1 (Promega) contains the canonical Sp1 binding site. Lane 1 is free probe. (B) A supershift analysis on probe C is shown. Free probe is in lane 1, 6 and 7 while 4  $\mu$ g of FRTL5 KiMol nuclear extract were incubated with probe C in lanes 2-5. Specific  $\alpha$ -Sp1 and  $\alpha$ -Sp3 antibodies were added in lanes 3 and 4 respectively. Complex a is supershifted by anti-Sp1 specific antibody, while complexes b and c by the  $\alpha$ -Sp3 antibody.  $\alpha$ -Oct2 specific antibody was added in lane 5 as a negative control. In lanes 6 and 7  $\alpha$ -Sp1 and  $\alpha$ -Sp3 were incubated in the absence of nuclear extract. (C) A DNaseI-footprinting analysis of Sp1 binding to the *Hmgi-c* promoter is shown. The DNA fragment was incubated with 2 footprinting forming units of purified recombinant Sp1 (Promega) in the absence (lane 4) or in the presence (lane 5) of a 100 fold specific competitor oligo Sp1. Lane 3 is the control with BSA. The

sense strand is shown in lanes 2-5 while the antisense is in lane 1. Lanes 1 and 2 represent G+A and T+C Maxam-Gilbert sequencing ladders. The protected ppyr/ppur region is shown on the right.

**FIG. 5.** Sp1 activates *Hmgi-c* promoter in *Drosophila* Schneider cells. The reporter plasmids HaP and AaP (4 $\mu$ g) were transfected into *Drosophila* SL2 cells along with different amounts ( $\mu$ g) of either pPacSp1 and/or pPacSp3 expression vectors as indicated. The empty vector pPac was added to compensate for amount differences. Values represent the average of three independent transfections, standard deviations are shown by vertical bars. The vector pPac $\beta$ -gal was introduced to correct for transfection efficiency.

**FIG. 6.** CTF/NF-1 transcription factor binds to the *Hmgi-c* minimal promoter downstream the TCC repeat. (A) A band-shift experiment, using nuclear extract from NIH-3T3 cells and purified CTF/NF-1 is shown. Oligo D corresponding to the sequence downstream the TCC repeat was used as a probe. It was incubated with increasing amounts of NIH-3T3 nuclear extract, expressed in  $\mu$ g, as indicated (lanes 2-9) or with increasing amounts, expressed in picomoles, of purified CTF/NF-1 (lanes 11-18). The resulting complex is indicated with an arrow. Competition experiments (lanes 5-9 and 14-18) were performed using different oligonucleotides at 100-fold molar excess in order to assess the specificity of the retarded complex. mD oligo is the mutated oligo D where the consensus sequence TGGC for CTF/NF-1 has been altered in TTGC (31), NF1 oligo contains a binding site for CTF/NF-1 (30), AP1 (Promega) contains a binding site for AP1 and oligo C is derived from the *Hmgi-c* minimal promoter. Lanes 1 and 10 are free probe. (B) The constructs HaP and NF1mut, containing a mutation able to abolish the binding of CTF/NF-1 *in vitro*, were transiently transfected in NIH-3T3 cells. The luciferase activity of NF1mut is expressed as percentage of that obtained for HaP, that was set arbitrarily equal to 100%. A vector expressing  $\beta$ -galactosidase from an SV40

promoter was included as internal control to normalize for transfection efficiency. Values represent the average of four independent transfections, standard deviation is shown by vertical bar.



**A**

*Haell*

-118 CCACCTCTACCTCCGCCACCCACTGCCCGCAGCGCCTCCTCCTTTCTCC  
 DeltaEF1 AP2 Sp1 AP4 PEA3  
 PU.1

-68 TCCTCCTCCTCCTCCTCTCTTCCTCCTCCCCTCTCTCTTTTGGCAGCCG  
 PU.1 Sp1 Yi CTF/NF-1  
 MZF1 IK2 CTCF AP4

+1  
 AaII

-16 CTGGACGTCCGGTGTGATGGTGGCAGCGGGCGGCAGCCTAAGCAGCAGC  
 AP4 Elk-1 AP4 AP4 AP4  
 c-ets-1/c-ets-1

+34 AGTAGTCCCCGGCGCACTCGCCAGCTCGCCTGCTCTCGCCGCTTTGCC  
 IK2 MZF1 AP2 YY1 AP4

+84 TCTCCAGCTCTCTACATCCCGTCTCCCGAAAGGTGCTGGGCGGATTCGGG  
 PstI 11a1

+134 GCGGCGGAGGCCGAGCGGCTGCAGCGGCGGTACCGGTAGAGGCAGTGG

Met Ser Ala Arg Gly Glu Gly

+182 TAGCGGCGGCGGGAGGCAGG ATG AGC GCA CGC GGT GAG GGC

**B**

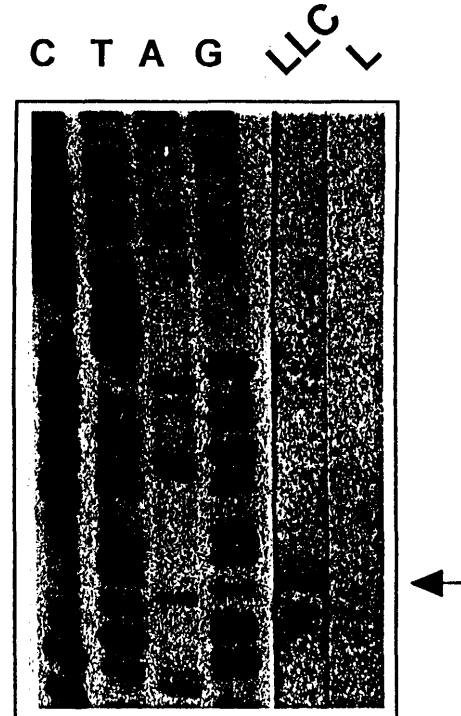
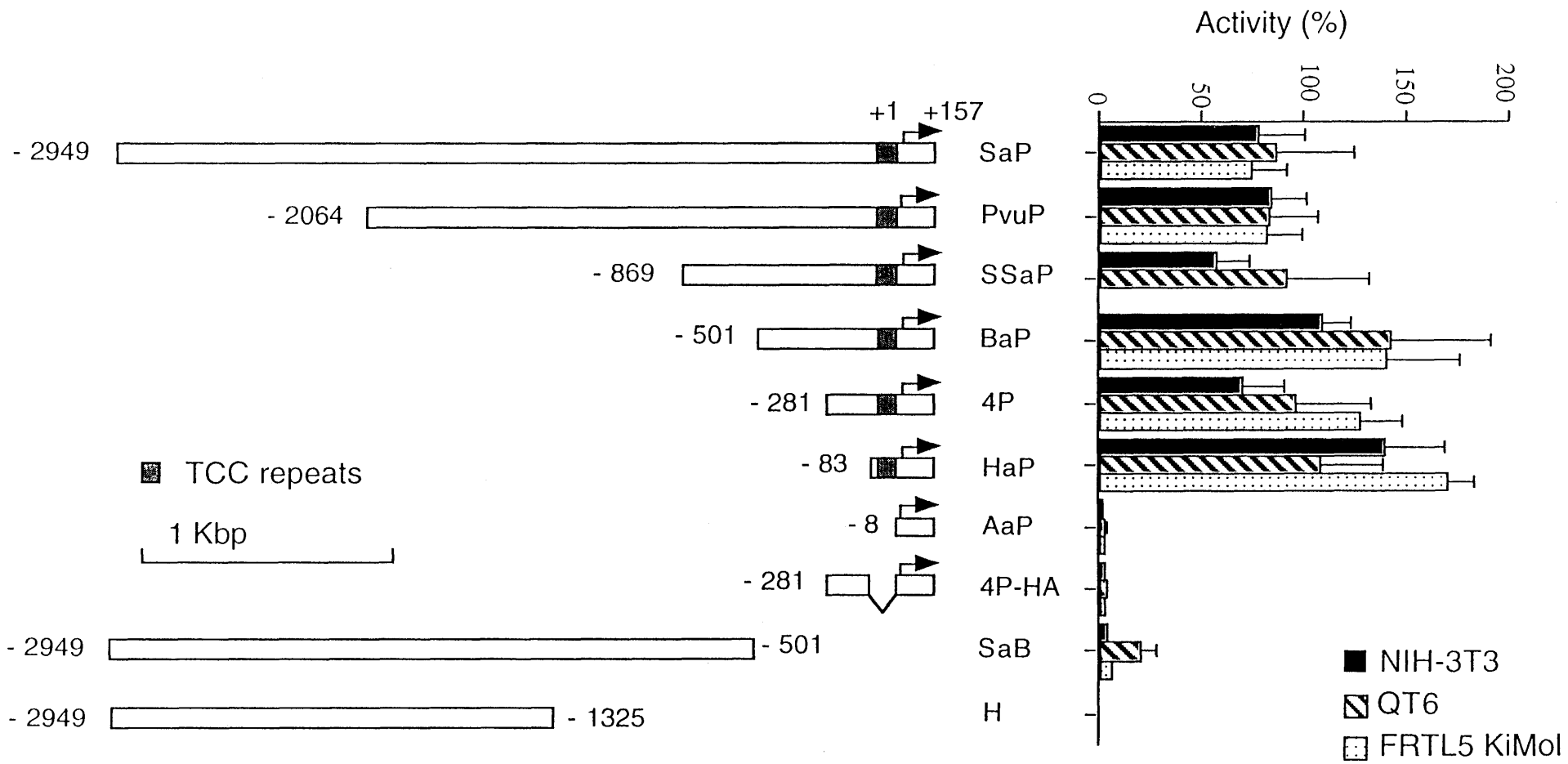
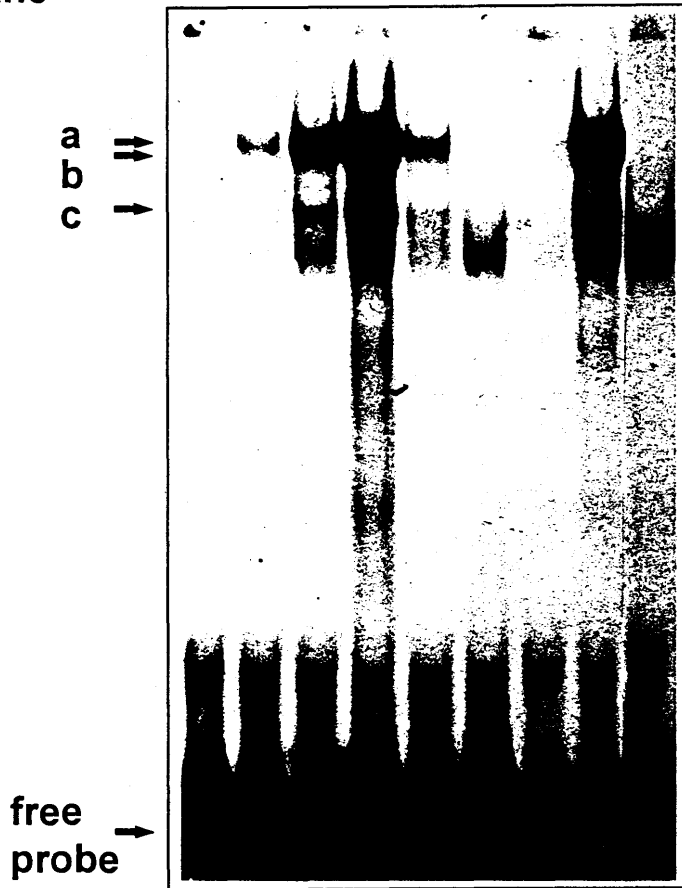


Fig 2

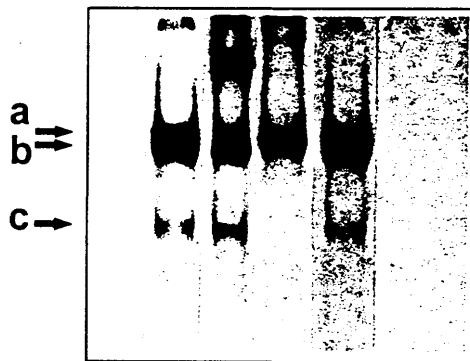




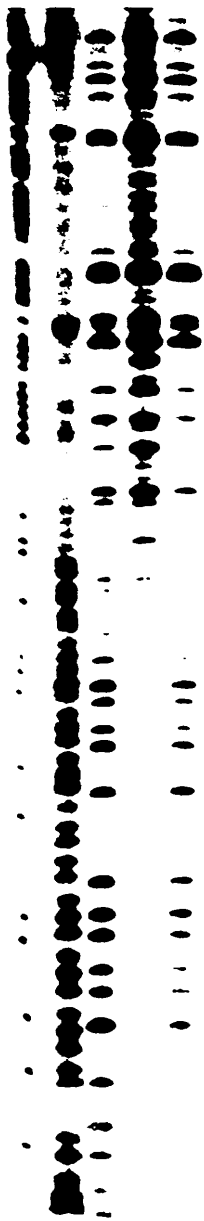
N.E.	-	1	2	4	4	4	4	4	4
Competitor	-	-	-	-	A	B	C	D	Sp1
lane	1	2	3	4	5	6	7	8	9



Antibody	-	-	$\alpha$ -Sp1	$\alpha$ -Sp3	$\alpha$ -Oct2	$\alpha$ -Sp1	$\alpha$ -Sp3
N.E.	-	4	4	4	4	-	-
lane	1	2	3	4	5	6	7



G+A  
T+C  
+BSA  
+Sp1  
+Sp1  
+comp.



-30

ppyr/ppur

1

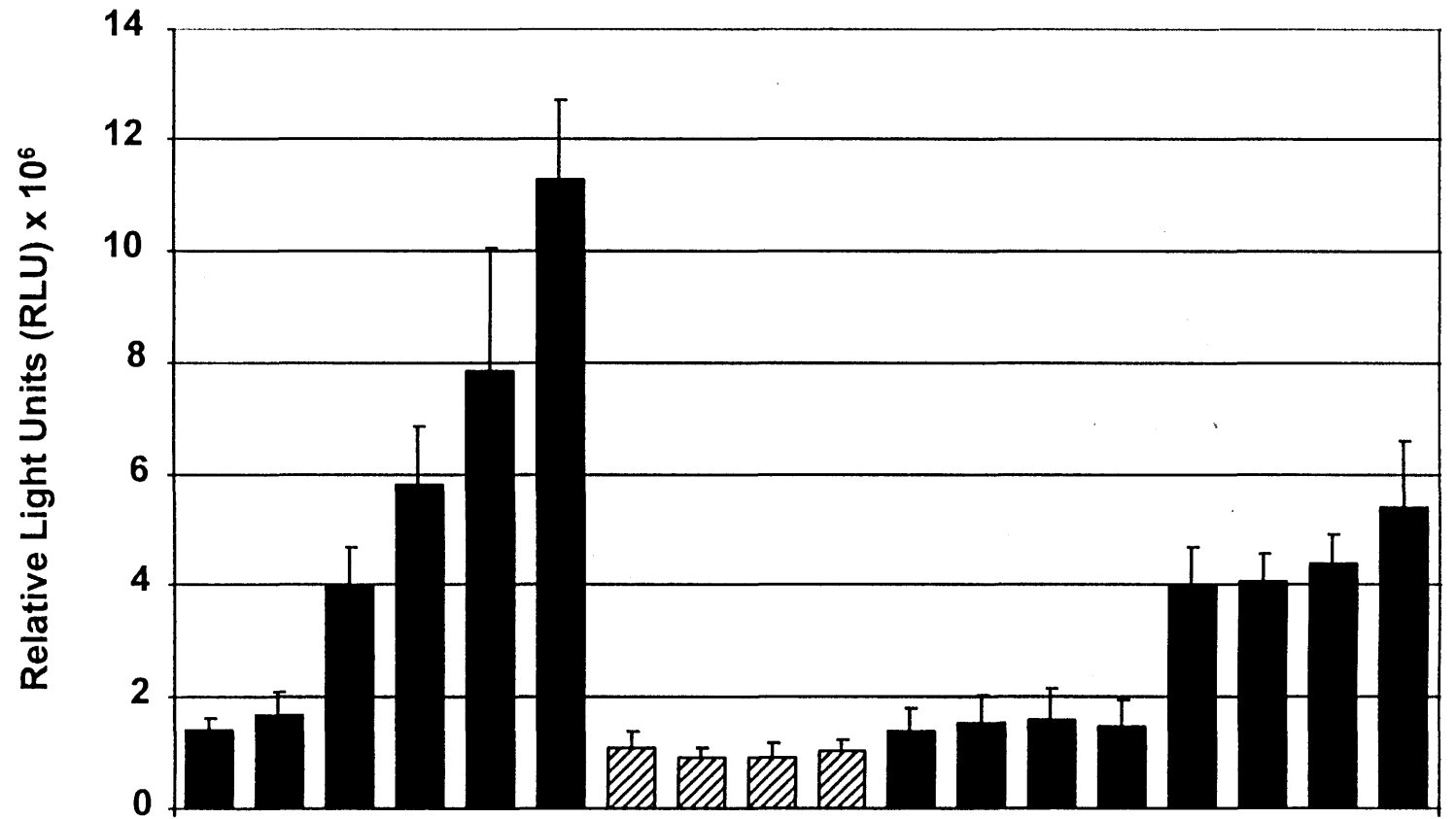
2

3

4

5

-68



HaP	4	4	4	4	4	4	-	-	-	-	4	4	4	4	4	4	4	4
AaP	-	-	-	-	-	-	4	4	4	4	-	-	-	-	-	-	-	-
pPacSp1	-	0.25	0.5	1	1.5	3	-	0.5	1	1.5	-	-	-	-	0.5	0.5	0.5	0.5
pPacSp3	-	-	-	-	-	-	-	-	-	-	0.5	1.5	2.5	3	-	0.1	0.5	2.5

