



Universitat Autònoma de Barcelona

**Facultad de Biociencias.**  
**Departamento de Bioquímica y Biología Molecular.**

**Estudios estructurales y caracterización de la unión  
al DNA del dominio C-terminal de la histona H1.  
Efecto de la fosforilación.**

Memoria presentada por Alicia Roque Córdova para obtener el grado de  
Doctor en Bioquímica y Biología Molecular.

Trabajo realizado en la Unidad de Biociencias del Departamento de  
Bioquímica y Biología Molecular de la Universidad Autònoma de Barcelona  
bajo la dirección de la Dra. Inma Ponte Marull y el Dr. Pedro Suau León.

---

Alicia Roque Córdova

---

Dra. Inma Ponte Marull

---

Dr. Pedro Suau León

Bellaterra, noviembre de 2007.



<b><u>ABREVIATURAS.....</u></b>	<b><u>1</u></b>
<b><u>INTRODUCCIÓN.....</u></b>	<b><u>3</u></b>
<b>LA CROMATINA.....</b>	<b>3</b>
<b>HISTONAS INTERNAS.....</b>	<b>5</b>
MODIFICACIONES POSTRADUCCIONALES DE LAS HISTONAS INTERNAS.....	7
Acetilación.....	9
Fosforilación.....	11
Metilación.....	12
Deiminación.....	13
ADP-ribosilación.....	14
Ubiquitinación.....	14
Sumoilación.....	15
<b>HISTONAS H1.....</b>	<b>15</b>
CLASES DE H1.....	17
DIVERSIDAD Y EXPRESIÓN.....	18
H1 somáticas.....	19
H1°.....	19
H1t.....	21
INTERACCIÓN DE LA HISTONA H1 CON LA CROMATINA Y EL DNA.....	23
Preferencia por las regiones SAR.....	25
MOTIVO SPKK.....	28
<b>DOMINIOS DE LA HISTONA H1.....</b>	<b>30</b>
DOMINIO GLOBULAR.....	30
DOMINIO N-TERMINAL.....	32
DOMINIO C-TERMINAL.....	34
<b>MODIFICACIONES POSTRADUCCIONALES DE LA HISTONA H1.....</b>	<b>38</b>
ACETILACIÓN.....	38
METILACIÓN.....	40
ADP-RIBOSILACIÓN.....	41
FORMILACIÓN.....	41
DEIMIDACIÓN.....	41
UBIQUITINACIÓN.....	42
FOSFORILACIÓN.....	42
<b>LA HISTONA H1 EN EL NÚCLEO CELULAR.....</b>	<b>49</b>
<b>OTROS EFECTOS DE LA FOSFORILACIÓN DE LA HISTONA H1.....</b>	<b>52</b>
<b>LA HISTONA H1 Y EL SISTEMA INMUNE.....</b>	<b>54</b>
<b><u>OBJETIVOS.....</u></b>	<b><u>59</u></b>

**CAPÍTULO 1.....61**

**INTRODUCTION .....63**  
**MATERIALS AND METHODS.....65**  
PEPTIDES AND PROTAMINE .....65  
SEPARATION OF HISTONE H1 SUBTYPES.....65  
CLONING, EXPRESSION AND PURIFICATION OF THE GLOBULAR AND C-TERMINAL DOMAINS OF  
H1 HISTONE SUBTYPES .....65  
PREPARATION OF DNA FRAGMENTS.....67  
BINDING ASSAYS .....67  
**RESULTS .....68**  
ALL HISTONE SUBTYPES ARE SAR BINDING PROTEINS.....68  
THE PREFERENTIAL BINDING OF HISTONE H1 IS DETERMINED BY ITS C-TERMINAL DOMAIN ....69  
PROTAMINES ALSO BIND PREFERENTIALLY TO SARs .....73  
DISTAMYCIN ABOLISHES THE SAR BINDING CHARACTER OF THE C-TERMINAL DOMAIN OF H1  
AND PROTAMINE .....74  
**DISCUSSION .....75**

**CAPÍTULO 2.....79**

**INTRODUCTION .....81**  
**MATERIALS AND METHODS.....82**  
CLONING, EXPRESSION AND PURIFICATION OF THE CARBOXY-TERMINAL DOMAINS OF HISTONE  
H1 SUBTYPES.....82  
CIRCULAR DICHROISM SPECTROSCOPY .....83  
INFRARED SPECTROSCOPY MEASUREMENTS .....83  
**RESULTS .....84**  
CD AND IR SPECTROSCOPY ANALYSIS OF THE CARBOXY-TERMINAL DOMAIN OF HISTONE H1 IN  
AQUEOUS SOLUTION AND IN TFE.....84  
IR SPECTROSCOPY OF THE CARBOXY-TERMINAL DOMAIN OF HISTONE H1 BOUND TO DNA ....86  
MELTING EXPERIMENTS.....90  
**DISCUSSION .....91**

**CAPÍTULO 3.....95**

**INTRODUCTION .....97**  
**MATERIALS AND METHODS.....98**  
CLONING, EXPRESSION AND PURIFICATION OF THE C-TERMINAL DOMAINS OF HISTONE H1  
SUBTYPES .....98  
CIRCULAR DICHROISM SPECTROSCOPY .....98  
INFRARED SPECTROSCOPY MEASUREMENTS .....99  
SMALL-ANGLE X-RAY SCATTERING.....99  
THERMAL MELTING .....100  
ANS FLUORESCENCE ASSAYS .....100  
DIFFERENTIAL SCANNING CALORIMETRY .....100

<b>RESULTS .....</b>	<b>101</b>
CD AND IR SPECTROSCOPY ANALYSIS OF THE C-TERMINAL DOMAIN OF HISTONE H1 IN THE PRESENCE OF CROWDING AGENTS .....	101
SMALL-ANGLE X-RAY SCATTERING.....	104
ANS BINDING.....	106
THERMAL MELTING AND DSC .....	106
<b>DISCUSSION .....</b>	<b>108</b>

**CAPÍTULO 4.....111**

<b>INTRODUCTION .....</b>	<b>113</b>
<b>MATERIALS AND METHODS.....</b>	<b>114</b>
CLONING, EXPRESSION AND PURIFICATION OF THE C-TERMINAL DOMAINS. ....	114
<i>IN VITRO</i> PHOSPHORYLATION ASSAY. ....	114
CONSTRUCTION OF MUTANTS IN CDK2 SITES.....	115
MALDI-TOF MASS SPECTROMETRY. ....	115
IR SPECTROSCOPY. ....	115
<b>RESULTS .....</b>	<b>116</b>
<i>IN VITRO</i> PHOSPHORYLATION OF CDK2 SITES. ....	116
EFFECTS OF HYPERPHOSPHORYLATION ON THE STRUCTURE OF THE FREE C-TERMINAL DOMAIN OF HISTONE H1. ....	117
EFFECTS OF HYPERPHOSPHORYLATION ON THE SECONDARY STRUCTURE OF THE C-TERMINAL DOMAIN OF HISTONE H1 BOUND TO DNA.....	118
EFFECTS OF PARTIAL PHOSPHORYLATION ON THE SECONDARY STRUCTURE OF THE C-TERMINAL DOMAIN OF HISTONE H1° BOUND TO DNA.....	119
EFFECTS OF CHARGE NEUTRALIZATION ON THE FOLDING OF THE C-TERMINUS OF H1°. ....	121
<b>DISCUSSION .....</b>	<b>123</b>
<b>SUPPLEMENTARY MATERIAL. ....</b>	<b>126</b>

**DISCUSIÓN GENERAL. ....127**

<b>EL DOMINIO C-TERMINAL DETERMINA LA PREFERENCIA DE LA HISTONA H1 POR LAS REGIONES SAR.....</b>	<b>127</b>
<b>ESTRUCTURA DEL DOMINIO C-TERMINAL DE LA HISTONA H1.....</b>	<b>129</b>

**CONCLUSIONES. ....137**

**BIBLIOGRAFÍA. ....141**



*A mis padres, hermana y abuelas.  
A Noel.*





## *Agradecimientos:*

*Ya sé que ser muy explícita no es mi estilo, pero lo haré lo mejor que pueda.*

*Quisiera agradecer a tanta gente, que espero no se me quede ninguna:*

*A mis directores de tesis Pere Suau e Inma Ponte por haber confiado en mí sin conocerme y darme una oportunidad. Por apoyarme y ayudarme durante todos estos años.*

*A mis compañeros de laboratorio, Mary, Marta, Xavi y Núria, de Tercer Ciclo y en general a todos los miembros de la Unidad, por hacerme sentir a gusto.*

*A Helena porque siempre estás dispuesta a escuchar, por todos los favores impagables y por hacer que siempre quiera venir el martes por la mañana.*

*Al Salva y el Jordi, porque siempre he podido contar con vosotros.*

*Al Dr. Josep Cladera, Marta Cortijo y Núria Benseny de la Unidad de Biofísica por su ayuda con el DSC y el IR.*

*Al Dr. José Luis Arrondo, Ibon Iloro e Igor, de la Universidad del País Vasco, Bilbao, por su ayuda con el IR.*

*Al Dr. Jordi Esquena y Jaime Caellas del CSIC, Barcelona, por su ayuda con el SAXS.*

*A mi familia en Barcelona: Miguel, Roberto, Joedmi, Ana, Bubby, Rodolfo, Emilio, Yudelsis, Yanet, Jose, Omarilys e Ismael. Por haber llenado un vacío muy importante en mi vida y estar siempre dispuestos a compartir buenos y malos momentos.*

*A mi familia, por su amor incondicional y por estar siempre a mi lado a pesar de las distancias.*

*A Noel, por compartir conmigo todos estos años y animarme a seguir adelante, pero sobretodo por hacerme feliz.*

*Gracias a todos, por hacer esta tesis posible.*



## Abreviaturas.

ADP: Difosfato de adenosina.  
ANS: 1-anilino-naphthaleno-8-sulfonato.  
ATP: Trifosfato de adenosina.  
cAMP: Monofosfato de adenosina cíclico.  
CAP: Proteína activadora por catabolito.  
CD: Dicroísmo circular.  
ChIP on CHIP: Inmunoprecipitación de cromatina más hibridación en chip de DNA.  
C-terminal: Carboxilo terminal.  
DNA: Ácido desoxirribonucleico.  
DSC: Calorimetría diferencial de barrido.  
FRAP: Recuperación de la fluorescencia tras el fotoblanqueo.  
FTIR: Espectroscopía de infrarrojo de transformada de Fourier.  
GAPDH: Gliceraldehído 3-fosfato deshidrogenasa.  
HIV: Virus de inmunodeficiencia humana.  
HMG: Proteínas del grupo de alta movilidad.  
HP1: Proteína de heterocromatina.  
HPLC: Cromatografía líquida de alta resolución.  
IDP: Proteínas intrínsecamente desordenadas.  
IPTG: Isopropil-β-D-tiogalactósido.  
MEL: Células de leucemia eritrocitaria de ratón.  
MMTV: Virus de tumor mamario de ratón.  
mRNA: Ácido ribonucleico mensajero.  
NK: Células citotóxicas naturales (Natural killer).  
N-terminal: Amino terminal  
OD: Densidad óptica.  
PCR: Reacción en cadena de la polimerasa.  
PEG: Polietilenglicol.  
PKA: Proteína quinasa A.  
RB: Proteína de retinoblastoma.  
rDNA: DNA ribosomal.  
RMN: Resonancia magnética nuclear  
SAR: Regiones asociadas a la matriz nuclear o "scaffold".  
SAXS: Dispersión de rayos X a bajo ángulo.  
SLE: Lupus eritematoso.  
TFE: 2,2,2-trifluoroetanol.  
TR: Hormona tiroidea.  
UBF: Factor de unión corriente arriba.  
UCE: Elemento de control corriente arriba.



## Introducción.

### **La cromatina.**

En los organismos eucariotas el DNA se encuentra en el núcleo de la célula en forma de cromatina. La cromatina es una fibra formada por complejos de DNA y proteínas básicas denominadas histonas y otras proteínas no histonas. La unidad básica de la fibra de cromatina es el nucleosoma<sup>1</sup>.

El conjunto formado por las histonas internas, el DNA enrollado sobre ellas y la H1 se conoce como nucleosoma (Fig. 1). Dos copias de cada una de las histonas internas H2A, H2B, H3 y H4 forman el núcleo proteico que junto con 146 pb de DNA, se conoce como partícula núcleo. La histona H1 se une a la partícula núcleo y al DNA de unión o "linker", protegiendo 20 pb adicionales de la digestión con nucleasa micrococcal, formando el cromatosoma. El DNA linker más el cromatosoma forman el nucleosoma<sup>1</sup>.

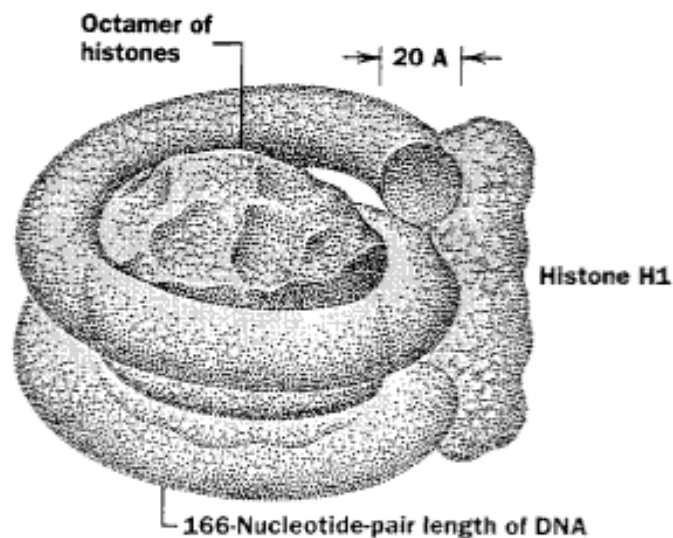


Figura 1. El nucleosoma.

Los nucleosomas unidos por el DNA internucleosomal forman la fibra de cromatina de 10 nm. La histona H1, que se une a la partícula núcleo y al DNA internucleosomal, permite la compactación de la fibra de 10 nm en una fibra de 30 nm, que constituye el estado en que se encuentra la mayoría de la cromatina durante la interfase (Fig. 2). Existen varios modelos de la estructura de esta fibra, pero éste es un tema sobre el que existen

discordancias importantes. Entre los modelos de la cromatina podemos citar el modelo de Thoma y col.<sup>2</sup>, conocido como modelo del solenoide, el modelo helicoidal de los nucleosomas en zig-zag propuesto por Woodcock y col.<sup>3</sup>, sustentado por estudios realizados por Richmond y col.<sup>4</sup> y también el modelo del solenoide interdigitado postulado por Dabán y col.<sup>5</sup>.

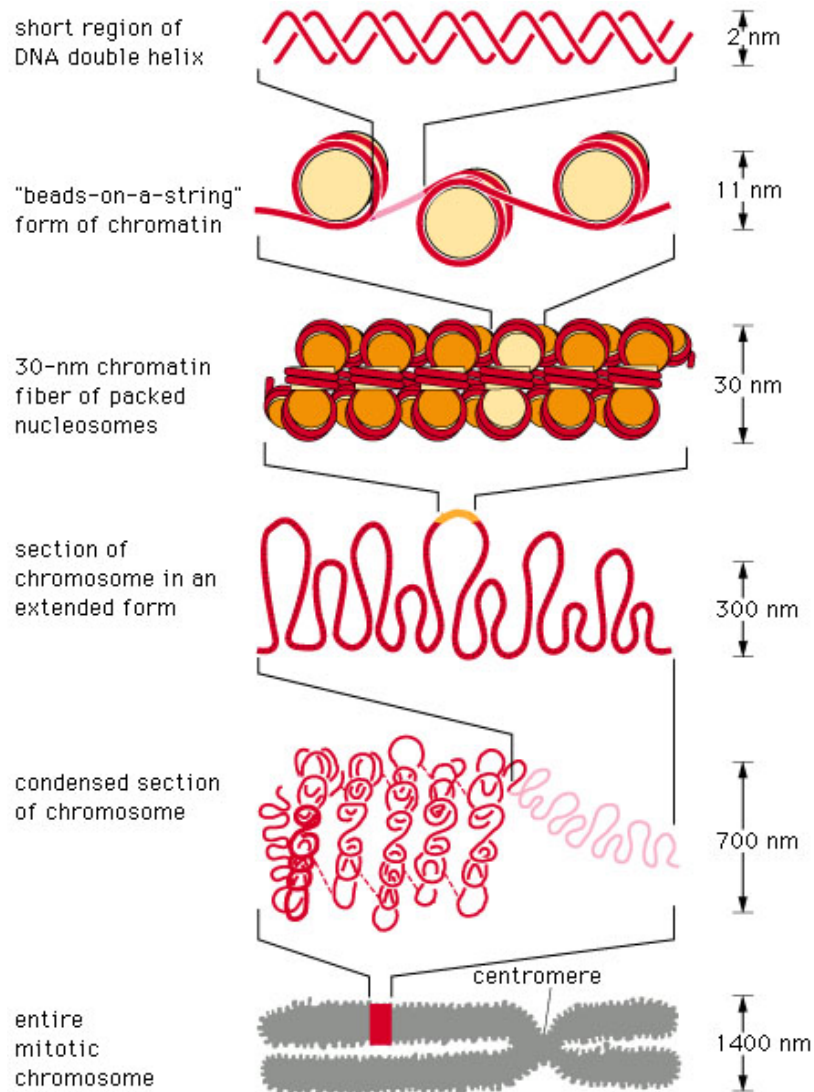


Figura 2. Empaquetamiento de la cromatina<sup>6</sup>.

La fibra de 30 nm se organiza en grandes dominios llamados lazos que están anclados a un esqueleto proteico o scaffold (Fig. 2). Su longitud varía entre 30 y 100 kb y se piensa que

podrían constituir grandes dominios transcripcionales dinámicos regulados topológicamente<sup>7</sup>. El modo y el grado de empaquetamiento de los lazos puede cambiar durante el ciclo celular y en relación con la expresión génica.

La estructura de la cromatina aporta la condensación y organización necesarias para que el empaquetamiento del DNA en el núcleo sea compatible con la realización de procesos como la replicación, la transcripción y la regulación de la expresión génica<sup>8</sup>.

En el núcleo eucariota encontramos dos tipos fundamentales de cromatina: eucromatina y heterocromatina. La eucromatina está compuesta por regiones más relajadas donde se encuentran genes activamente transcritos. Por el contrario, la heterocromatina son regiones más compactas que se encuentran generalmente cercanas al centrómero y los telómeros. El mantenimiento de las regiones heterocromáticas es necesario para la estabilidad genómica y el control de la expresión génica<sup>9</sup>.

### ***Histonas internas.***

El núcleo central proteico de los nucleosomas, alrededor del cual se enrolla el DNA contiene dos moléculas de cada una de las histonas internas o del core, H2A, H2B, H3 y H4. Las histonas internas son proteínas de bajo peso molecular (11-16 kDa) con más de un 20% de aminoácidos básicos. Las histonas H2A y H2B contienen más lisina mientras que las histonas H3 y H4 contienen más arginina<sup>10</sup>.

Todas las histonas internas están sumamente conservadas en longitud y en secuencia aminoacídica a lo largo de la evolución. Las histonas H3 y H4 son las más altamente conservadas<sup>11</sup> y tienen un papel central en el nucleosoma y en muchos procesos cromosómicos; estas restricciones funcionales y estructurales explican la conservación de su secuencia. Las histonas H2A y H2B están algo menos conservadas.

Las cuatro histonas presentan un dominio globular del tipo denominado precisamente “plegamiento de histona” (*histone-fold*) en el extremo carboxi-terminal de la proteína, a través del cual se producen las principales interacciones histona-histona y histona-DNA. Además, presentan una cola amino (N-) terminal básica que contiene la mayoría de los residuos básicos<sup>12</sup> (Fig. 3).

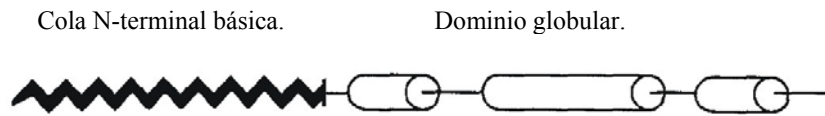


Figura 3. Estructura de las histonas internas<sup>13</sup>.

El dominio tipo “plegamiento de histona” C-terminal contiene tres hélices  $\alpha$ ; la central es la de mayor tamaño y está flanqueada por dos lazos que la conectan con las dos hélices más pequeñas en los extremos del dominio. Cada uno de los lazos contiene cierta estructuración en cadena  $\beta$ . La larga hélice central actúa como una zona de dimerización. Este dominio está conservado evolutivamente, debido a las restricciones que le confieren su importancia estructural en el nucleosoma; así mismo, la secuencia de la cola N-terminal cargada también se encuentra conservada. Estas colas N-terminales sufren varias modificaciones post-traduccionales en aminoácidos específicos, a menudo por parte de proteínas reguladoras de la transcripción<sup>14</sup>. Esta es una de las explicaciones de la conservación evolutiva de la secuencia de las colas, junto con su importancia en la interacción con otros componentes estructurales de la cromatina<sup>15, 16</sup>. Las colas N-terminales constituyen la diana para los diferentes enzimas que modifican la estructura de la cromatina. Se ha observado que estos dominios terminales son capaces de estructurarse ligeramente en hélice en presencia de inductores de estructura secundaria como el 2,2,2-trifluoroetanol (TFE). Parece, además, que la acetilación de estos dominios incrementa el porcentaje de hélice que pueden adquirir<sup>17, 18</sup>. La disposición de las histonas internas en el octámero consiste en un tetrámero (H3, H4)<sub>2</sub> central, flanqueado por dos dímeros (H2A, H2B). Esta estructura presenta una serie de surcos en su superficie que crean una rampa helicoidal levógira sobre la que se enrolla el DNA. Dentro de esta rampa hay expuestos 16 lazos de las histonas internas, que quedan formando 8 puentes  $\beta$  ( $\beta$ -bridges) paralelos en el momento de la dimerización. Cada uno de estos puentes  $\beta$  presenta al menos dos residuos cargados positivamente que son capaces de interactuar con los fosfatos del DNA en la superficie del octámero de histonas<sup>19</sup>. También el extremo N-terminal de la primera hélice de las histonas queda apareado en la dimerización y parece que se une al DNA (Fig. 4).



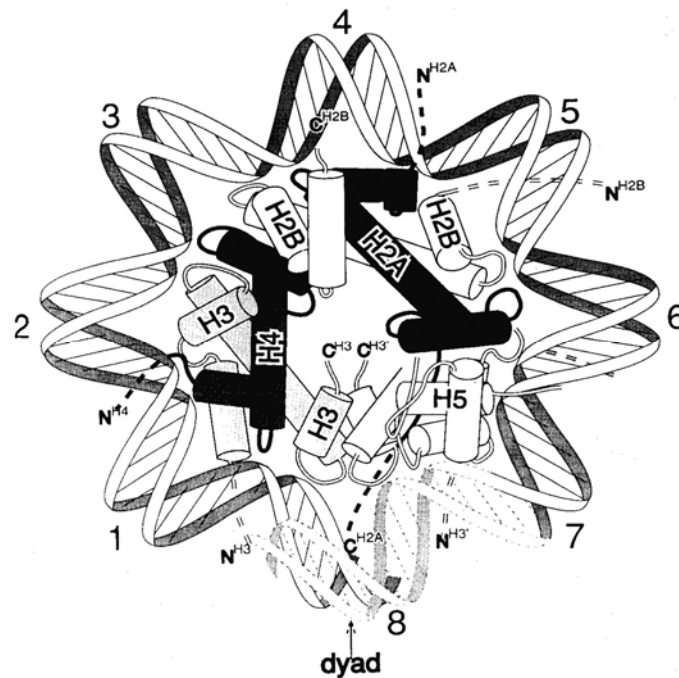


Figura 4. Dibujo del nucleosoma en el que se aprecia la disposición de las histonas internas, del DNA y un posible lugar de unión de la histona H5<sup>13</sup>.

Varios estudios indican que las interacciones electrostáticas de los residuos de arginina de las histonas internas con la cadena fosfodiéster del DNA son las más importantes en la organización del DNA en el nucleosoma. Efectivamente, sólo 14 residuos concretos de arginina son suficientes para mantener el DNA enrollado en el nucleosoma. El papel fundamental de las argininas podría venir dado por su capacidad de formar puentes de hidrógeno, además de interacciones electrostáticas con los fosfatos del DNA<sup>20</sup>. Varios estudios indican que las colas N-terminales altamente cargadas de las histonas internas no contribuyen de forma significativa al enrollamiento inicial del DNA en el nucleosoma<sup>21</sup> y que tampoco tienen un papel esencial en mantener la integridad del nucleosoma<sup>22</sup>. Sin embargo las modificaciones post-traduccionales de estas colas pueden alterar aspectos más sutiles de la conformación del nucleosoma y de las interacciones histona-DNA.

### Modificaciones postraduccionales de las histonas internas.

En el nucleosoma existen al menos 14 puntos de contacto entre el DNA y las histonas<sup>23</sup>, lo que confiere una gran estabilidad al complejo en condiciones fisiológicas<sup>24</sup>. La superficie de los nucleosomas en general se encuentra salpicada por múltiples modificaciones

postraduccionales. Las modificaciones pueden actuar delimitando los dominios de la cromatina: eucromatina y heterocromatina. Su presencia en las histonas determina la superestructura de la cromatina y puede orquestar el reclutamiento ordenado de complejos enzimáticos efectores de los procesos metabólicos relacionados con el DNA, incluyendo la transcripción, la replicación, la reparación y la condensación de los cromosomas. El término “código de las histonas” describe el papel de las modificaciones en estos procesos<sup>25</sup>.

Las modificaciones postraduccionales están bastante estudiadas en las histonas del core y ocurren fundamentalmente en las colas N-terminales que sobresalen del nucleosoma<sup>26, 27</sup> (Fig. 5).

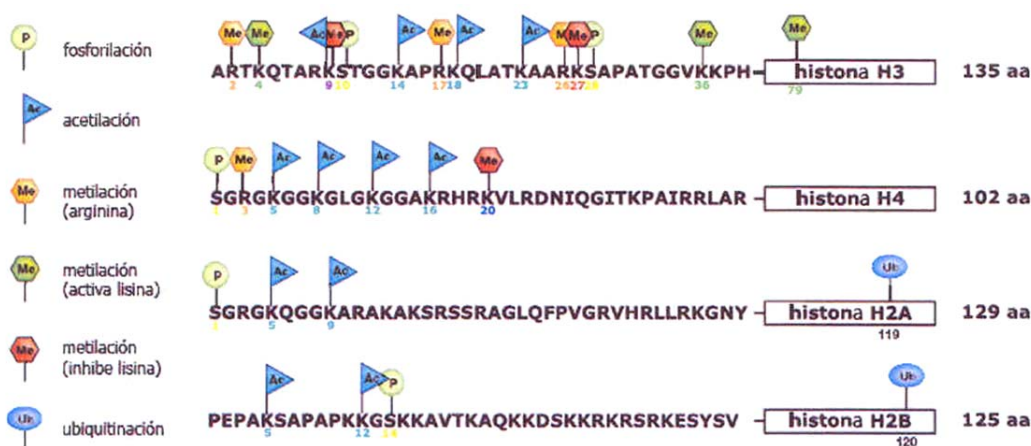


Figura 5. Principales modificaciones postraduccionales de las histonas internas en mamíferos<sup>28</sup>.

Estudios mediante espectrometría de masas y de “ChIP on CHIP” (Chromatin Immunoprecipitation on DNA chip) han permitido la identificación de acetilación, metilación, fosforilación, ubiquitinación, sumoilación, ADP-ribosilación y deiminación<sup>25, 29, 30, 31</sup>. Algunas de estas modificaciones son específicas para el tipo de histona, así como de las etapas del ciclo celular (Fig. 6).

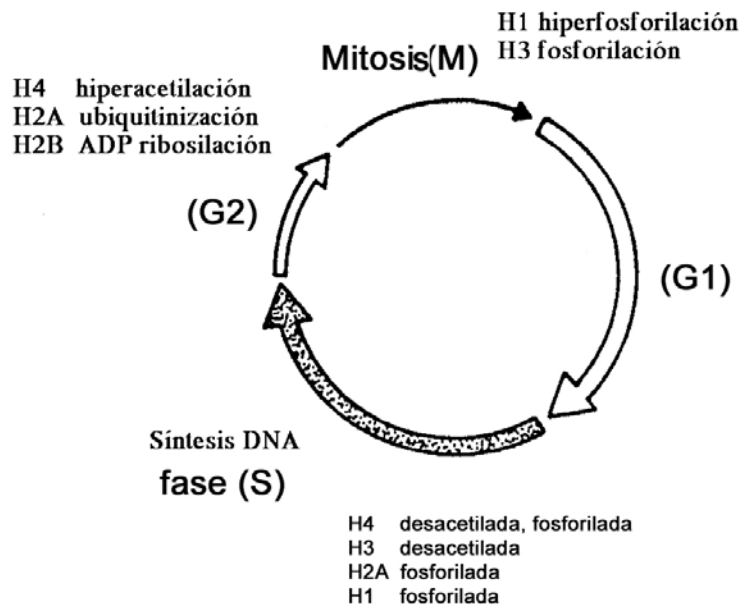


Figura 6. Modificaciones postraduccionales de las histonas durante el ciclo celular<sup>13</sup>.

### Acetilación.

La acetilación se produce en los residuos de lisina de las colas amino-terminales cargadas, aunque recientemente también se ha encontrado acetilada la posición K56 en el dominio globular de la H3<sup>25</sup>. Cada grupo acetato incorporado a la histona reduce su carga neta positiva en 1. El número de residuos lisina acetilados en una molécula de histona depende del equilibrio de las actividades acetilasa y desacetilasa. En el núcleo celular hay dos poblaciones de histonas acetiladas: histonas hiperacetiladas e hipoacetiladas. El patrón de residuos de lisina específicos que se acetila en las colas N-terminales varía entre especies y apunta a la existencia de especificidad de secuencia en las acetilasas y desacetilasas<sup>32</sup>.

La hiperacetilación de las colas de las histonas internas da lugar a pequeños cambios en la conformación del nucleosoma<sup>33</sup>, pero al parecer la consecuencia más importante de esta modificación es la afectación de las interacciones proteína-proteína, entre nucleosomas, bien con la histona H1 o con proteínas no histonas. La acetilación, así como la eliminación de las colas N-terminales de las histonas, facilitan el acceso de los factores de transcripción a sus dianas<sup>34, 35</sup>.

Observaciones *in vitro* sugieren que la acetilación de las histonas internas no influye directamente en la asociación de la histona H1 con los nucleosomas<sup>36</sup>, aunque al parecer las

regiones acetiladas son muy pobres en H1 y la acetilación inhibe la completa condensación de la cromatina en la interfase<sup>37</sup>.

Existe una posible relación entre la acetilación de la H3 y la histona H1. En experimentos con nucleosomas reconstituidos se observó que la presencia de la histona H1 promueve la interacción de la H3 con otra fibra de cromatina a partir de 1.3 mM de MgCl<sub>2</sub>. Este efecto se pierde si la cola N-terminal de la H3 está acetilada<sup>38</sup>.

Existe una correlación entre la acetilación de las histonas y la actividad transcripcional de la cromatina<sup>26, 39</sup>. Los dominios de la cromatina que se transcriben activamente, o que son potencialmente activos, se encuentran enriquecidos en histonas hiperacetiladas, mientras que la cromatina transcripcionalmente inactiva contiene histonas hipoacetiladas. También existen zonas específicas de mayor o menor acetilación en los genes activamente transcritos<sup>25</sup>.

La acetilación parece aumentar el contenido de hélice  $\alpha$  que los dominios N-terminales de las histonas internas son capaces de adquirir en presencia de TFE<sup>17, 40</sup>. En este sentido, se ha postulado que el espacio entre las lisinas acetilables de las colas de las histonas H3 y H4 marca el paso de rosca de una hélice  $\alpha$ <sup>41</sup>.

La acetilación también tiene un papel en la reparación del DNA. En levaduras la acetilación en K56 de la histona H3 ocurre en las histonas recién sintetizadas durante la fase S y desaparece en G2. En presencia de daño en el DNA esta modificación persiste<sup>42</sup>. En los sitios de reparación se encuentra la acetilasa Hat1 que acetila la H4 en la lisina 12<sup>43</sup>.

En la replicación se ha encontrado que la acetilación de la H4 por el enzima HBO1 es necesaria para la iniciación de la fase S y la señalización de los orígenes de replicación<sup>44</sup>.

La acetilación de la H4 en la lisina 16 podría estar relacionada con la descondensación de la cromatina<sup>45</sup>. La deacetilación de esta posición por un enzima específico, SirT2, que aparece en la cromatina en la transición G2/M, podría inducir la condensación de la cromatina *in vivo*<sup>46</sup>.

En resumen, el papel de la acetilación de las histonas tanto en la transcripción como en otros procesos, indica una importancia biológica crucial para esta modificación post-traducciona, comparable a la fosforilación en las cadenas de transducción de señales<sup>25</sup>.

## **Fosforilación.**

La fosforilación afecta a todas las clases de histonas del core. La histona H3 se fosforila rápidamente en residuos serina de su dominio básico amino-terminal, en respuesta a señales extracelulares como factores de crecimiento o ésteres de forbol que estimulan la proliferación celular<sup>47</sup>. La fosforilación de la H3 es máxima en mitosis, momento en el cual se fosforila la serina 10 del N-terminal que por ello se ha relacionado con la condensación de los cromosomas. Según esta hipótesis, dicha fosforilación modula la interacción entre la cola de la H3 con el DNA, permitiendo el acceso de factores condensantes<sup>48, 49</sup>. Células que expresan oncogenes o carentes de genes supresores presentan altos niveles de H1 y H3 fosforiladas, que serían responsables de la estructura menos condensada de la cromatina y la expresión génica aberrante<sup>50</sup>. La proteína quinasa Aurora B que fosforila la S10 de la H3 se sobreexpresa en muchas líneas cancerígenas. La incorporación de ese fosfato se asocia a la inestabilidad cromosómica observada en células malignas<sup>51, 52</sup>.

La fosforilación en S10 de la H3 también se asocia con la inducción de la transcripción en la respuesta al shock térmico<sup>53</sup> y a la activación de c-Fos, c-jun y de los genes regulados por NF- $\kappa$ B<sup>25</sup>. Otra fosforilación, en este caso H3T3, es necesaria para la alineación correcta de los cromosomas metafásicos<sup>54</sup>.

La fosforilación *in vivo* de las histonas H4 y H2A ocurre en el citoplasma poco después de su síntesis<sup>55</sup>. La fosforilación de estas histonas, junto con la acetilación de la histona H4, podría dirigir las selectivamente hacia las chaperonas moleculares involucradas en la formación del nucleosoma y de la horquilla de replicación<sup>56</sup>. La eliminación de la fosforilación de las histonas internas mediante fosfatasa no influye en el espaciado de los nucleosomas<sup>55</sup>, y su función en el ensamblaje de la cromatina continúa sujeta a especulación.

Una de las respuestas asociadas al daño en el DNA es la fosforilación de la variante  $\gamma$  de la H2A en mamíferos. Esta fosforilación se extiende por varias kilobases alrededor del sitio dañado y en levadura recluta el complejo INO80, que presenta actividad remodeladora dependiente de ATP. Las modificaciones H2AS129 y H4S1 se han relacionado con la reparación de las roturas en doble cadena por la vía de la unión de extremos no homólogos. La fosforilación H4S1 también está involucrada en la regulación de la esporulación en levadura y se conserva durante la espermiogénesis de mosca y ratón<sup>57</sup>. En levaduras la

fosforilación H2BS10 también regula la inducción de apoptosis por peróxido, aunque este residuo no está conservado en mamíferos, la fosforilación H2BS14 podría realizar la misma función<sup>58</sup>.

El grado de fosforilación de las histonas del core parece aumentar durante el inicio de la apoptosis, especialmente de la H2A y la H2B. La fosforilación del subtipo H2A.X en S139 parece asociada a la iniciación de la fragmentación del DNA en líneas celulares<sup>59</sup>.

### **Metilación.**

Las histonas internas también se metilan en algunos de sus residuos de lisina y arginina. La posición y el grado de metilación de las histonas se ha relacionado con la expresión génica. Estudios recientes del patrón de modificaciones epigenéticas y su importancia en los procesos celulares, así como en el desarrollo de enfermedades como el cáncer, apuntan a que la metilación de las histonas tendría un papel crucial en estos fenómenos<sup>60</sup>.

La metilación de las lisinas comienza después de la formación de los nucleosomas y es máxima durante la mitosis. Durante la interfase, la metilación de las histonas se da con preferencia en las histonas H3 y H4 que ya hayan sido acetiladas. Es posible que esto ocurra debido a la mayor accesibilidad de las colas N-terminales acetiladas<sup>29, 30</sup>.

La mayoría de las metilaciones de la histona H3 en los vertebrados tienen lugar en las lisinas 9 y 27, que incorporan hasta tres grupos metilo<sup>61</sup>. Estas dos metilaciones están asociadas a la represión transcripcional. La metilación en K9 es importante para el reclutamiento de la proteína de heterocromatina (HP1). Esta proteína se libera en mitosis cuando se fosforila la S10, estableciendo un mecanismo regulador que involucra dos modificaciones postraduccionales adyacentes, una metilación estable y una fosforilación dinámica, denominado "methyl-phos switch"<sup>62, 63</sup>.

La histona H3 también se metila en las lisinas 4, 36 y 79. La metilación de esas tres posiciones es una característica de zonas transcripcionalmente activas de la cromatina<sup>64, 65</sup>. La metilación en la lisina 36 es dependiente de la conformación de la prolina 38. Cuando la prolina 38 se encuentra en *trans*, la metiltransferasa Set-2 cataliza el proceso de metilación. Si la P38 se encuentra en conformación *cis*, proceso catalizado por la prolina isomerasa Frp4, el enzima Set-2 es incapaz de metilar la K36. La conformación *cis/trans* de la P38 también podría tener efecto en la actividad de las desmetilasas<sup>66</sup>.

Los genes transcritos activamente tienen algunas características comunes: hay sitios especialmente ricos en histonas acetiladas en el promotor y en el extremo 5' de la región codificante; dentro del promotor los dos nucleosomas que flanquean el sitio de iniciación de la transcripción están hipoacetilados y enriquecidos en la variante Htz1 de la H2A; el sitio de iniciación no contiene nucleosomas; hay un alto contenido de lisinas trimetiladas en la región codificante, donde H3K4, H3K36 y H3K79 tienen su patrón específico. Aunque las características anteriores se han descrito en levaduras, evidencias en tejido humano y de ratón están a favor de su conservación<sup>25</sup>.

Otra función importante de la metilación de las lisinas es la demarcación de los dominios de eucromatina y heterocromatina. Las fronteras de la heterocromatina se mantienen con la presencia en los dominios adyacentes de eucromatina de las metilaciones H3K4 y H3K9. La heterocromatina se asocia con bajos niveles de acetilación y el patrón de metilación asociado a represión transcripcional H3K9, H3K27 y H4K20. La metilación en H3K27 está implicada en el silenciamiento del gen HOX y el mantenimiento del cromosoma X inactivo. La metilación H3K9 parece importante en el mantenimiento de la heterocromatina pericentromérica<sup>25</sup>.

La histona H4 es metilada en la lisina 20, en la que se han encontrado hasta dos grupos metilo. Esta modificación parece funcionar como una marca en el DNA dañado que es reconocida por los complejos de reparación<sup>67</sup> y está relacionada con la formación de la heterocromatina<sup>25</sup>.

La metilación de las argininas está muy poco caracterizada, puede ocasionar activación y represión de la transcripción. Al parecer se ubica preferentemente en regiones promotoras. En el promotor pS2 regulado por estrógenos la metilación de las argininas es cíclica (aparecen y desaparecen) en el proceso de activación<sup>68</sup>. Este fenómeno se podría derivar de la necesidad de que los promotores regulados por estrógenos respondan rápidamente a los estímulos externos y por tanto la cromatina tenga que permanecer dispuesta para una posible represión transcripcional inmediata<sup>25</sup>.

### **Deiminación.**

Esta modificación consiste en la conversión de la arginina en citrulina. Las argininas de la H3 y la H4 pueden ser convertidas en citrulina por el enzima PADI4. Esta modificación tiene el potencial de impedir la metilación de la arginina ya que la citrulina no puede ser

metilada<sup>69</sup>. La conversión de la citrulina en arginina no se ha descrito, aunque en el promotor pS2 los niveles de citrulina son también cíclicos, lo que apunta a que el proceso de deiminación es reversible<sup>70</sup>.

### **ADP-ribosilación.**

La ADP ribosilación consiste en la adición de unidades de ADP-ribosa a algún residuo aminoacídico de la proteína. Las enzimas que la catalizan son la mono-ADP-ribosil-transferasa y la poli-ADP-ribosil-transferasa, según se incorporen una o más unidades de ADP-ribosa respectivamente. Hay muchos estudios sobre la ADP ribosilación de las histonas, pero hasta el momento solamente se ha caracterizado la H2Bar1, donde el residuo de ADP-ribosa se incorpora en E2<sup>25</sup>.

Algunos resultados sugieren que la ADP-ribosilación de las histonas internas puede dar lugar a la descondesación localizada de la fibra de cromatina. La ADP-ribosilación pudiera tener un papel importante en la reparación del DNA<sup>71</sup>. La síntesis de largas cadenas de ADP-ribosa cargadas negativamente podría facilitar la disrupción de los nucleosomas por intercambio de histonas hacia este polianión competidor.

### **Ubiquitinación.**

La H3, la H2B y especialmente la H2A pueden modificarse por conjugación con ubiquitina (proteína de 76 residuos)<sup>72</sup>. La ubiquitina participa en la regulación de la degradación de las proteínas. Esta proteína se une covalentemente, a través de una reacción dependiente de ATP, a la proteína que debe ser marcada para su proteólisis. Sin embargo, la ubiquitinización de la H2A y la H2B no parece estar relacionada con su degradación. Se ha observado que la histona H2A ubiquitinada se incorpora a los nucleosomas sin cambios importantes aparentes en la organización de la partícula núcleo<sup>73</sup>. Las histonas ubiquitinadas podrían encontrarse en las regiones del DNA activas transcripcionalmente, impidiendo la compactación de la cromatina. En algunos organismos se ha encontrado un aumento de la H2A ubiquitinada durante la espermiogénesis; justo en el momento anterior a la sustitución de la H1 por la protamina. A pesar de estas observaciones no está definido el papel estructural de la ubiquitinación de las histonas internas y su efecto o importancia en la dinámica de la cromatina<sup>74</sup>.



Recientemente se ha caracterizado la ubiquitinación en H2AK119 y H2BK120 (K123 en levaduras). La modificación H2AK119 parece relacionada con la represión de la transcripción, mientras que la H2BK120 produce activación. La ubiquitinación de la H3 y la H4 se ha relacionado con la reparación del DNA inducida por daños producidos por UV, ya que es importante en el reclutamiento de la proteína de reparación XPC. En este mismo proceso está implicada la monoubiquitinación de la H2A<sup>25</sup>.

### **Sumoilación.**

SUMO es una proteína modificadora relacionada con la ubiquitina que tiene aproximadamente 100 aminoácidos y está conservada en organismos eucariotas. La conjugación de la proteína SUMO a las histonas es una modificación voluminosa, como la ubiquitinación, ya que ambas implican la unión de una molécula de gran tamaño. Las cuatro histonas del core experimentan sumoilación. Los sitios de modificación que se han encontrado en levaduras, la K126 en el C-terminal de la H2A, la K6, 7, 16, 17 en el N-terminal de la H2B y las cinco lisinas del N-terminal de la H4, mientras que aún no hay datos concretos sobre la H3. Esta modificación, a diferencia de la acetilación y la ubiquitinación, produce represión transcripcional<sup>25, 75</sup>.

### **Histonas H1.**

La definición más amplia de histonas H1 agrupa a aquellas proteínas básicas y abundantes que interaccionan con el DNA de unión entre los nucleosomas<sup>1</sup>. Otros autores restringen su concepto por motivos estructurales solamente a los metazoos, obviando así a las proteínas con idéntica función de levaduras y protozoos.

Las histonas H1 de metazoos subdividen su estructura en tres dominios: el dominio amino terminal (20-35 aminoácidos), el dominio central de carácter globular (~ 80 aminoácidos), y el dominio carboxilo terminal rico (~ 100 aminoácidos)<sup>76</sup> (Fig. 7).



Figura 7. Dominios de la histona H1.

Los dominios amino y carboxi terminal tienen carácter básico, especialmente el dominio C-terminal que puede contener hasta un 40% de lisinas. La longitud de los dominios terminales es más variable que la del dominio globular. Los dominios amino y carboxi terminal de la H1 son de baja complejidad de secuencia a nivel de nucleótidos y de aminoácidos, mientras que el dominio globular es de alta complejidad de secuencia. La simplicidad de secuencia de los dominios terminales observada en los subtipos de H1 sugiere que estos dominios podrían tener su origen en la amplificación de motivos cortos que posteriormente han evolucionado mediante mutaciones puntuales y mutación por deslizamiento de cadena (slippage)<sup>77</sup>.

La estructura tripartita de la H1 sugiere que cada dominio podría tener funciones específicas en la estructura de la cromatina. El dominio N-terminal podría estabilizar la unión del dominio globular, el dominio globular podría localizar la H1 en el nucleosoma, mientras que el dominio C-terminal es la región que generalmente se asocia con la condensación de la cromatina mediante la unión y la neutralización de las cargas del DNA linker<sup>76</sup>.

La H1 tiene múltiples isoformas o subtipos. En mamíferos encontramos, los subtipos somáticos H1a-e, los subtipos específicos de líneas germinales H1oo y H1t y los subtipos asociados a diferenciación H1° y H1x. El análisis de la evolución de las H1 de vertebrados muestra que la tasa de sustitución de aminoácidos entre los subtipos varía al menos en un orden de magnitud, lo que sugiere la posibilidad de que los subtipos puedan tener funciones específicas<sup>78</sup>.

Las H1 de levadura y de protozoos como *Tetrahymena* no comparten la estructura general descrita anteriormente. En levaduras, la H1 tiene dos dominios globulares con una homología de secuencia significativa con el dominio globular de los metazoos, separados entre sí por un segmento de 42 aminoácidos, rico en lisina, alanina y prolina como el dominio C-terminal de los organismos superiores.

La H1 de *Tetrahymena* carece de dominio globular y consiste en un polipéptido de 163 aminoácidos con una composición similar a la del C-terminal de los metazoos.

## Clases de H1.

Las histonas linker se clasifican en tres clases según el tipo y grado de diferenciación de la célula en que se están expresando<sup>76, 79</sup>.

- Clase I. Se expresan de forma específica en células de la línea germinal. En oocitos encontramos la CS H1 (erizo de mar), H1M y B4 (rana) y la H1oo (ratón y humanos). Algunos de estos genes contienen intrones y elementos reguladores en *cis* propios de oocitos como la Y-box y se encuentran codificados por mRNA poliadenilado. Se expresan en altos niveles, generándose grandes reservas de éstas proteínas para ser usadas durante la embriogénesis temprana<sup>80,81</sup>. En la línea germinal masculina también encontramos una histona específica, la H1t y dos proteínas histone-like llamadas H1T2<sup>82, 83</sup> y HILS1<sup>84</sup>.
- Clase II. Se expresan en células somáticas y tienen un patrón de expresión dependiente de ciclo celular. Carecen de intrones y tienen promotores con características similares, entre las que se encuentran las cajas TATA y CAAT, secuencias ricas en GC y una región específica para genes H1, llamada caja H1, cuya secuencia es 5'AAGAAAACACA 3'. Se comienzan a sintetizar en el embrión a partir de mRNA materno que solamente se libera para la traducción durante la fertilización. Una vez concluido el desarrollo embrionario temprano la síntesis se acopla a la del DNA en la fase S del ciclo celular. A ello contribuyen la caja CAAT y la caja H1<sup>79</sup>. Sus genes están agrupados en una región del genoma conocida como "cluster" de histonas, que incluye también la H1t, ubicada en el cromosoma 6 de humanos<sup>85</sup> y en el 13 de ratón<sup>86</sup>. La H1a, H1c y H1t forman parte de un segmento de 80 kb que contiene las histonas del core; la H1d y H1e pertenecen a un segundo segmento, mientras que la H1b se encuentra 500 kb más lejos en un segmento que contiene 14 genes de histonas del core<sup>87</sup>.
- Clase III. Variantes específicas de diferenciación o también llamadas histonas de reemplazo. En este grupo encontramos la H1<sup>o</sup>, la H5 y la recientemente añadida H1x<sup>88</sup>. Sus promotores son divergentes e incorporan secuencias propias de las histonas internas y secuencias promotoras específicas de tejido<sup>89, 79</sup>. Contienen elementos reguladores particulares que responden según el tipo de tejido y el estadio del desarrollo. La H5 se expresa en eritrocitos de pollo y se incorpora a la cromatina de manera estable. La H1<sup>o</sup>

se acumula en células diferenciadas terminalmente. Su gen se encuentra aislado del resto de los genes de las histonas, en ratón se ubica en el cromosoma 15<sup>90</sup> y en humanos en el cromosoma 22<sup>85</sup>.

### Diversidad y expresión.

En los organismos superiores, las histonas H1 tienen múltiples variantes no alélicas. Los subtipos de la H1 podrían estar diferenciados funcionalmente<sup>91</sup>. En *Xenopus* hay una expresión secuencial de los subtipos de H1: B4 en el estado embrionario, H1 en el adulto y finalmente H1<sup>o</sup>, específica de diferenciación. Los cambios en los subtipos de histonas se asocian a diferencias en la estructura de la cromatina, así como a su capacidad replicativa y transcripcional.

En aves, la acumulación de H5, variante de H1 típica de eritrocitos, se asocia con el establecimiento de una estructura inerte de la cromatina<sup>92</sup>.

En mamíferos se han encontrado al menos 7 variantes de H1. Los subtipos de H1 difieren en la extensión y el momento de la fosforilación, la conservación evolutiva<sup>93,77</sup>, las tasas relativas de síntesis y degradación en células en proliferación o quiescentes<sup>94</sup>, y en su habilidad para condensar el DNA *in vitro*<sup>95</sup>. Las características mencionadas apoyan la hipótesis de la existencia de diferencias funcionales entre los subtipos.

Se han realizado experimentos de sobreexpresión de los subtipos H1<sup>o</sup> y H1c en los que se evaluó el efecto producido en la expresión génica. La sobreexpresión de H1<sup>o</sup> resultó en la detención transitoria en la fase G1 del ciclo y de la progresión a fase S, así como en una disminución de la expresión de los genes de ciclofilina,  $\beta$ -actina y gliceraldehído 3-fosfato deshidrogenasa (GAPDH). De ello se deriva la función propuesta para la H1<sup>o</sup>, consistente en la estabilización de la cromatina inactiva en células diferenciadas. La sobreexpresión de H1c no afectó la progresión del ciclo celular y aumentó 2-3 veces la expresión de los genes de ciclofilina,  $\beta$ -actina y GAPDH. También prepararon proteínas quiméricas realizando un intercambio de dominios entre los subtipos H1<sup>o</sup> y H1c. Con las proteínas obtenidas se encontró que el dominio globular es el responsable de los diferentes efectos de la H1<sup>o</sup> y la H1c en la expresión génica. A pesar de que ambos dominios globulares son similares, existen diferencias que determinan la estructura tridimensional y por tanto en la forma en que interaccionan con los nucleosomas<sup>91</sup>.

Los efectos diferenciales de algunas variantes de H1 en la embriogénesis y la espermatogénesis se han atribuido a diferencias en la longitud y carga de los dominios amino y carboxi terminal<sup>79</sup>.

### **H1 somáticas.**

Los niveles de H1a, H1b y H1d decrecen rápidamente en células terminalmente diferenciadas o en G0, mientras que la H1c y la H1e se continúan expresando. La H1c produce un mRNA poliadenilado en células diferenciadas, además del mRNA fundamental no poliadenilado que termina en una estructura de "stem-loop"<sup>87</sup>.

La H1a se expresa en todos los tejidos en niveles tan bajos que hasta hace poco se consideraba específica de células de bazo, timo y testículos<sup>96</sup>. Su expresión está asociada a la división celular. En el sistema nervioso la H1a se expresa en los neuroblastos, pero no en las neuronas<sup>97</sup>.

En humanos, los niveles de H1c y H1e también se mantienen constantes en el ciclo y tipo celular, expresadas en altos niveles. Las otras histonas sufren variaciones dependientes de la replicación, observándose un aumento en los mRNA de unas 4-15 veces. Estos resultados apoyan la teoría de una función específica para cada subtipo<sup>98</sup>. En neuronas de rata la H1c tiene una tasa de degradación mucho mayor que la de la H1e, lo que origina que en este tipo celular el subtipo más abundante en los adultos sea la H1e<sup>97</sup>.

### **H1°.**

La H1° se acumula durante la diferenciación terminal de las neuronas y las células gliales<sup>97</sup>,<sup>99</sup>,<sup>100</sup> y se activa transcripcionalmente en tipos neuronales específicos. En el córtex adrenal, la tiroides, los testículos y la próstata la pérdida de función por la ausencia de la hormona correspondiente se relaciona con bajos niveles de H1°<sup>101</sup>. En núcleos del arcuato los niveles de esta proteína son regulados por hormonas y la cantidad es sexualmente dimórfica, variando con el ciclo estral<sup>102</sup>. Los niveles de H1° decrecen en tejidos en regeneración, por ejemplo en páncreas<sup>103</sup>, así como en neuronas del núcleo supraóptico como consecuencia de la estimulación osmótica de la transcripción<sup>104</sup>. Estos resultados muestran la capacidad de la histona H1° para responder a señales externas.

En ratón, rata y humanos se observó la presencia de dos subfracciones de H1<sup>o</sup> que se llamaron H1<sup>o</sup>a y H1<sup>o</sup>b. Una caracterización más completa de estas subfracciones demostró que en realidad eran 4 las formas de H1<sup>o</sup> y no dos. La H1<sup>o</sup>a consiste en la H1<sup>o</sup> intacta (H1<sup>o</sup>a-N3) y la forma deamidada de la asparagina en la posición 3 (H1<sup>o</sup>a-D3). La subfracción H1<sup>o</sup>b consiste en la H1<sup>o</sup> acetilada en su N-terminal (H1<sup>o</sup>b-N3) y la forma acetilada y deamidada (H1<sup>o</sup>b-D3). La proporción de cada una de las subfracciones cambia en dependencia del tejido y la edad. Se ha encontrado un aumento progresivo de la forma acetilada y deamidada (H1<sup>o</sup>b-D3) con la edad<sup>105</sup>.

La H1<sup>o</sup> tiene una elevada homología con la H5 de eritrocitos de ave, aunque ésta última se une con mayor fuerza al nucleosoma, permitiendo una cromatina más estable y una disminución de la expresión más drástica<sup>106</sup>.

Las diferencias de expresión de la H1<sup>o</sup> respecto al resto de las histonas somáticas se pueden atribuir a los elementos que presenta en su región promotora<sup>107</sup> (Fig. 8). En este promotor se encuentra la H4-box en lugar de la CAAT-box que está presente en las histonas somáticas. Esta caja solamente se encuentra en promotores asociados a diferenciación celular. El factor de transcripción HBP1, del tipo que se une a las HMG-box y que se ha encontrado unido también a la proteína de retinoblastoma (RB), la cual se expresa en dependencia del grado de diferenciación. Este complejo interacciona directamente con la H4-box y al parecer está involucrado en la activación transcripcional de la H1<sup>o</sup><sup>108</sup>.

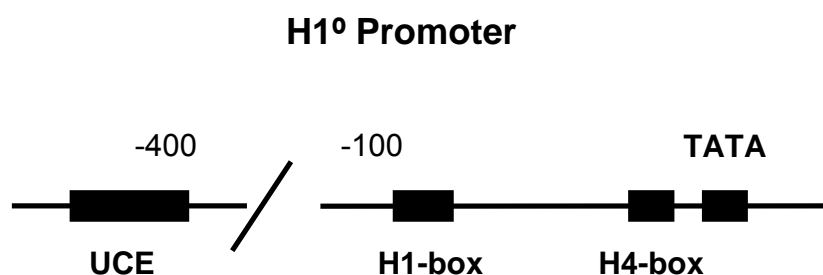


Figura 8. Promotor de la H1<sup>o</sup><sup>107</sup>.

Los otros elementos en *cis* que regulan la expresión de la H1<sup>o</sup> son la TATA-box, la H1-box y el UCE (upstream control element) y están presentes en los promotores del resto de las histonas<sup>107</sup>.

El promotor de la H1° es sensible a la estructura de la cromatina. Se ha observado que la hiperacetilación de la cromatina producida por la inhibición de las histonas deacetilasas induce la expresión de la H1° en células HeLa y en *Xenopus*<sup>107</sup>. La región promotora de la H1° es también sensible a señales externas. El receptor de la hormona tiroidea (TR) y los receptores de ácido retinoico son capaces de unirse al promotor de la H1°, produciendo se activación transcripcional<sup>109</sup>.

La H1x se transcribe también en casi todos los tejidos, aunque la proteína se ha podido detectar únicamente en cultivos celulares<sup>96</sup>. Al igual que la H1° se localiza en regiones de la cromatina que son menos sensibles a la digestión con nucleasa micrococcal y su gen se encuentra en una localización aislada dando lugar a un mRNA poliadenilado. No obstante las semejanzas entre los dos subtipos, la inducibilidad de la expresión, así como la regulación son diferentes<sup>88</sup>.

### **H1t.**

A lo largo de la espermatogénesis en mamíferos las espermatogonias diploides se diferencian en espermatocitos y luego de dos divisiones meióticas sucesivas en espermátidas haploides. Durante la espermiogénesis, las espermátidas atraviesan una fase de elongación, transformándose en espermatozoides maduros. Este proceso conlleva cambios importantes a nivel bioquímico y la reestructuración morfológica de las células de la línea germinal<sup>82</sup> (Fig. 8).

La histona H1t es específica de testículos, se expresa en los espermatocitos durante la profase de la meiosis, específicamente en el paquiteno<sup>110</sup> (Fig. 9) y constituye el 55% de la histona linker presente en este tipo celular<sup>111</sup>. El mRNA de H1t existe en el paquiteno medio y tardío, pero la proteína se mantiene hasta las espermátidas redondas o alargadas<sup>112</sup>. Otros autores han encontrado mRNA mediante microscopía electrónica e hibridización *in situ* en las espermatogonias y una banda difusa de H1t en ratones de 9 días<sup>113</sup>. La histona H1t difiere bastante de las histonas somáticas, pero son muy parecidas entre especies, al menos en su región codificante<sup>114,115</sup>.

## SPERMATOGENESIS

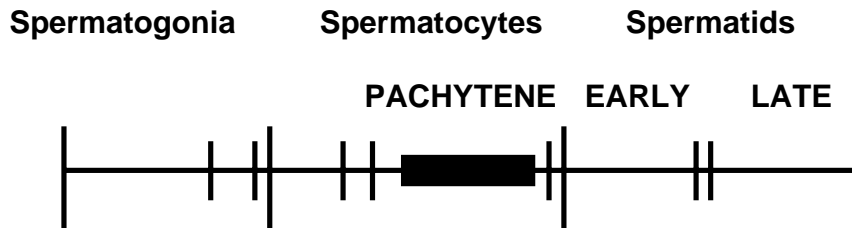


Figura 9. Etapas de la espermatogénesis<sup>116</sup>. Se muestran las tres etapas de la espermatogénesis. Las espermátidas tempranas (estadios 1-8), intermedias (estadios 9-12) y tardías (estadios 13-19). El rectángulo simboliza el período en que se sintetiza la H1t.

El promotor de la H1t (Fig. 10) tiene una gran parecido con el de las histonas somáticas. En su secuencia podemos encontrar las cajas TATA, CCAAT y AC. Existe una región única designada como elemento TE entre las cajas CCAAT y AC, que es esencial para la expresión de la H1t. A este elemento, parecido a la caja X, se une el factor RFX2<sup>117</sup>, que al parecer constituye el principal activador transcripcional de la H1t<sup>116</sup>.

### H1t Promoter

211 pb

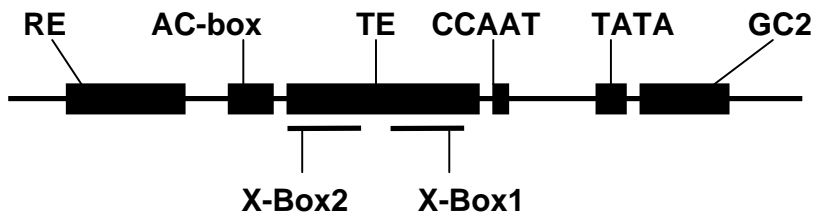


Figura 10. Promotor del gen de la H1t<sup>116</sup>.

La represión de la expresión del gen de la H1t en células no germinales parece estar mediada por una segunda caja GC, ubicada después de la caja TATA. En las células germinales la represión de la transcripción, sin embargo, está relacionada con un elemento llamado RE<sup>118</sup>, ubicado muy cerca de la caja AC (Fig. 10)<sup>116</sup>.

A pesar de la diferente secuencia de aminoácidos y la regulación de la H1t, esta histona no parece ser indispensable en ratones, pues se han realizado experimentos de genesupresión donde la función de la H1t ha sido sustituida por otros subtipos de la histona H1, compensando en su mayoría la ausencia de la H1t<sup>111</sup>. Se ha comprobado que sin H1t los



animales transgénicos realizan la espermatogénesis de forma normal, sin diferencias morfológicas en los testículos, son fértiles y se reproducen como el fenotipo salvaje. También se ha encontrado aumentada la expresión de H1a, H1c y H1e<sup>119</sup>.

### Interacción de la histona H1 con la cromatina y el DNA.

Las diferencias entre los subtipos de la histona H1 se ponen también de manifiesto en la localización nuclear y la afinidad por el DNA y la cromatina. Experimentos de inmunolocalización con fibroblastos humanos, donde no hay prácticamente expresión de H1a y H1<sup>o</sup>, mostraron que los subtipos H1e y H1d son más abundantes en la heterocromatina, mientras que la H1b predomina en los genes activos. Se ha propuesto que la H1c proporciona el nivel basal de condensación de la cromatina que luego es modulado por intercambio con los demás subtipos<sup>120</sup>.

Experimentos de FRAP (Fluorescent Recovery After Photobleaching), en células de neuroblastoma que expresan los diferentes subtipos como proteínas de fusión N-terminales con GFP, mostraron que los subtipos H1<sup>o</sup>, H1a, H1c y H1d tienen preferencia por las regiones de eucromatina, mientras que los subtipos H1e y H1b tienen más afinidad por la heterocromatina. También permiten clasificar los subtipos en dos grupos de afinidad: baja afinidad donde se encuentran la H1a y H1c, mientras que el resto de los subtipos estudiados H1<sup>o</sup>, H1b, H1d y H1e se ubican en el grupo de alta afinidad<sup>121</sup>.

Respecto a la afinidad de los subtipos de la histona H1 por la cromatina, los resultados obtenidos con FRAP son básicamente compatibles con los resultados obtenidos en nuestro laboratorio donde fue posible clasificar los subtipos estudiados en tres grupos de afinidad. Con baja afinidad se encuentra la H1a, con afinidad intermedia la H1b y H1c, mientras que el grupo de alta afinidad incluye la H1d, H1e y H1<sup>o</sup>. La afinidad mayor difiere de la menor en un factor de 19. Estos grupos de afinidad se confirmaron en los experimentos realizados con DNA desnudo, donde las afinidades a pesar de no ser idénticas dentro del mismo grupo conservaban los mismos tres grupos de baja, intermedia y alta afinidad<sup>122</sup>.

Entre las proteínas que se unen al DNA, las hay que poseen especificidad de secuencia. En este grupo encontramos proteínas del tipo de los factores de transcripción, represores y enzimas de restricción. El reconocimiento en estos casos involucra motivos proteicos como

el hélice-giro-hélice, hélice-lazo-hélice, cremallera de leucinas o los dedos de zinc que contactan la doble hélice del DNA fundamentalmente por el surco mayor<sup>123</sup>.

Otras proteínas y algunas drogas se unen al DNA, generalmente por el surco menor, con baja o nula especificidad de secuencia. En este grupo encontramos a las histonas, la DNAsa I, las proteínas tipo HU y muchas otras<sup>123</sup>. Los motivos estructurales propuestos para la interacción con el surco menor del DNA son muy variados e incluyen hélice  $\alpha$ <sup>124, 125</sup>, hoja  $\beta$  antiparalela<sup>126</sup> y el motivo SPKK<sup>127</sup>.

La unión del H1 al DNA es cooperativa en presencia de sal<sup>128, 129</sup>. Probablemente ocurre mediante el reclutamiento de moléculas sucesivas de H1 al complejo H1-DNA.

A pesar de la baja especificidad de secuencia, algunas proteínas de este grupo muestran preferencias por algunos tipos de secuencia como la descrita para la histona H1 por el DNA metilado, las estructuras cruciformes y las secuencias ricas en A/T<sup>123</sup>, entre las que se encuentran las secuencias asociadas al "scaffold", secuencias SAR (scaffold associated regions), que pudieran tener una gran importancia biológica<sup>130</sup>.

La metilación es la modificación más abundante en el DNA, más de un 60% de los motivos CpG están metilados. La presencia de CpG metilados se asocia a la represión transcripcional<sup>131</sup>.

La histona H1 tiene preferencia por el DNA metilado, la cual, además, parece aumentar con el grado de metilación<sup>132, 133</sup>. Esta preferencia tiene lugar en condiciones de unión cooperativa y parece estar determinada por el dominio globular. Las variantes de H1 presentes en núcleos reprimidos transcripcionalmente como la H5 muestran mayor preferencia por este tipo de DNA, sugiriendo que dicho fenómeno podría ser relevante en la estabilización de la superestructura de la cromatina y la represión transcripcional<sup>132</sup>.

En un estudio donde se caracterizó la cromatina activa e inactiva transcripcionalmente en neuronas de ratón se encontró que la histona H1 se superponía en la hibridación *in situ* con las regiones de DNA ribosomal (rDNA) metiladas. Estas regiones de rDNA son inactivas transcripcionalmente y se encuentran en el nucleolo, adyacentes a la heterocromatina perinucleolar. La histona H1 también se encontró unida a la heterocromatina centromérica y al borde perinucleolar<sup>134</sup>.

Es importante señalar que la preferencia de la H1 por el DNA metilado es objeto de controversia pues hay estudios que contradicen estas observaciones<sup>131, 135</sup>.

La histona H1 también se une a las estructuras cruciformes<sup>136, 137</sup>. Este tipo de DNA es un sustrato sintético que se utiliza para estudiar las proteínas que reconocen y resuelven las uniones de Holliday que se forman *in vivo* durante la recombinación. Se ha hipotetizado que las estructuras cruciformes imitan la estructura del DNA en la zona cercana a los sitios de entrada y salida de los nucleosomas<sup>136</sup>. Al parecer la H1 se une al centro de las estructuras cruciformes, en una interacción dependiente del dominio globular. Como en el caso de la unión al DNA metilado, la preferencia por las estructuras cruciformes es mayor en la H5<sup>137</sup>.

### Preferencia por las regiones SAR.

Las secuencias SAR son regiones cortas de DNA, con más de un 70 % de residuos A/T organizados en tramos homopoliméricos, que se unen de forma específica al scaffold nuclear y metafásico<sup>138</sup>. Se supone que estas secuencias constituyen la base estructural de los lazos de la cromatina en las células eucariotas<sup>138</sup> y que pudieran estar involucrados en la dinámica de los cromosomas<sup>139</sup> (Fig. 11).

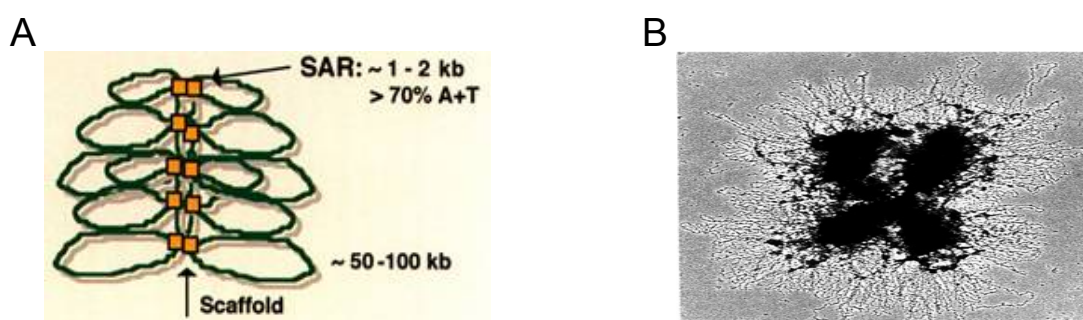


Figura 11. Regiones SAR. A. Representación de las regiones SAR en los cromosomas metafásicos (<http://www.molbio.unige.ch/laemml>). B. Microfotografía de cromosomas humanos descompactados por acción de baja fuerza iónica<sup>140</sup>.

Se han descrito numerosas proteínas con preferencia por las regiones SAR entre las que podemos citar la histona H1<sup>130</sup>, la topoisomerasa II<sup>141</sup>, la HMG I/Y<sup>142</sup>, la SAF-A y la SAF-B<sup>143</sup>.

La organización estructural de los cromosomas en el núcleo es altamente ordenada y se mantiene por la unión al scaffold. Dicha organización permite los movimientos dinámicos de la cromatina durante procesos como la recombinación y la condensación de los cromosomas. Esta organización está determinada supuestamente por elementos en *cis*, que permiten el anclaje de la cromatina al scaffold y actúan como puntos de señalización o

iniciación a partir de los que se extienden los estados de cromatina abierta, cerrada o heterocromatina. Estos elementos también podrían limitar la extensión de dichos estados<sup>144</sup>. Aunque no se puede descartar la posible participación de otros elementos en *cis*, las secuencias SAR se consideran el elemento en *cis* implicado en facilitar los cambios dinámicos entre los diferentes estados de la cromatina y la organización estructural de los cromosomas<sup>144</sup>.

Estudios estructurales han demostrado que las SARs no están localizadas al azar, sino que se encuentran separadas por distancias variables permitiendo la formación de lazos más cortos en las bandas Q de los cromosomas y más largos en las bandas R que contienen la mayoría de los genes expresados constitutivamente (housekeeping genes)<sup>145</sup> (Fig. 12).

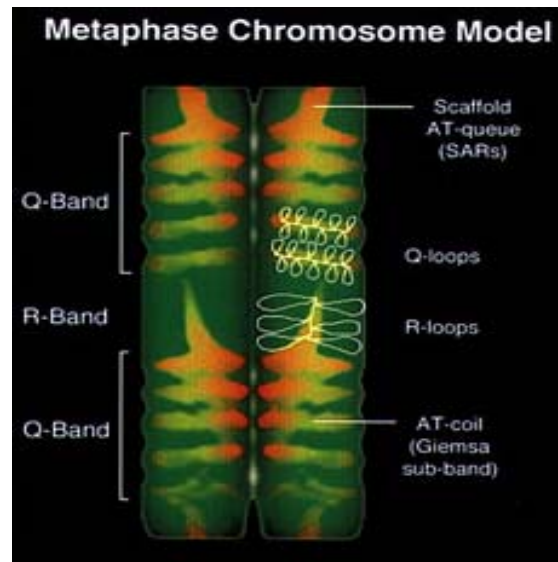


Figura 12. Modelo de un cromosoma metafásico. En amarillo, vemos la subregión interna formada por la yuxtaposición de las regiones SAR, llamada AT-queue, que se extiende de telómero a telómero. Esta subregión está más empaquetada en las bandas Q que en las R<sup>145</sup>.

La longitud relativa de los lazos está de acuerdo con el hecho de que las bandas Q tienen una proporción de DNA SAR mayor que las bandas R<sup>146</sup>.

Además del efecto en el estado de la cromatina, se ha demostrado que casi siempre la presencia de SAR flaqueando un gen reportero estimula su expresión. Esta característica se ha conservado evolutivamente, lo que sugiere que dependa de elementos conservados en la cromatina y no de factores específicos de especie<sup>147</sup>.

El modelo de Laemmli propone que dos tipos de proteínas compiten por la unión a las SARs. Un primer tipo, que incluye la histona H1 y la topoisomerasa II, presenta una unión cooperativa a las SARs ocasionando la compactación de la cromatina, la condensación de los cromosomas y la formación de heterocromatina. El segundo tipo de proteínas, como la HMG I/Y, no se unen de forma cooperativa y pueden desplazar a las proteínas compactantes afectando la cooperatividad entre ellas con el resultado de una mayor accesibilidad de la cromatina<sup>148</sup> (Fig. 13).

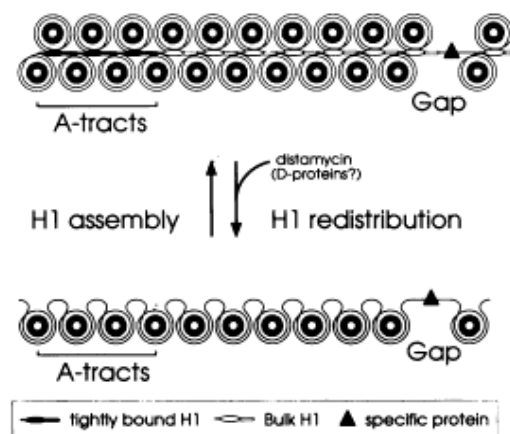


Figura 13. Modelo de la apertura cierre de la cromatina. Arriba observamos los nucleosomas compactados. En negrita está la H1 que se encuentra unida a la secuencia SAR. Abajo observamos la cromatina abierta por la acción de la distamicina o proteínas como la HMG I/Y sobre la H1 unida a la región rica en A/T<sup>148</sup>.

La apretura así originada se detiene por la acción de otro elemento SAR, formando así los dominios. Este modelo está sustentado en experimentos de desplazamiento de la H1 de las SARs por la HMG I/Y y la distamicina<sup>142</sup>.

También hay datos sobre la acción combinada de las SARs y otros elementos como los enhancers, en la activación de la expresión de genes como el de la cadena  $\mu$  de las inmunoglobulinas<sup>149</sup>, lo que pudiera conferir especificidad a la iniciación de la apertura de la cromatina<sup>150</sup>.

La afinidad de las proteínas por las SARs, tanto las de unión cooperativa como la histona H1 o las de unión no cooperativa (tipo HMG I/Y), está determinada por la presencia de los tramos de adenina, A-tracts, y no por una secuencia de bases precisa<sup>148, 151</sup>. Estos tramos de adeninas tienen además la peculiaridad de alterar la estructura de la doble hélice del DNA, la cual tiene un surco menor muy estrecho<sup>152, 153</sup>.

El surco menor de las regiones SAR parece tener especial importancia, ya que la preferencia de las proteínas por dicha región de DNA se pierde en presencia de distamicina<sup>141, 151</sup>. La distamicina A es un antibiótico que se une al DNA en el surco menor y que tiene predilección por las regiones ricas en A/T, lo que sugiere que la interacción ocurre por esta zona, aunque en el caso concreto de la histona H1 no se puede descartar la posibilidad que también se una por el surco mayor<sup>154</sup>.

Algunas causas de la preferencia por un surco menor muy estrecho podrían ser la posibilidad de tener mejores contactos de van der Waals entre las paredes del surco menor y los ligandos o la mayor densidad local de carga negativa que se origina<sup>152</sup>.

En cuanto a qué motivos en las proteínas confieren la preferencia por las regiones SAR, aún no hay una respuesta segura. Las proteínas del grupo HMG tienen un motivo denominado gancho AT (AT-hook), de secuencia RGR que se repite en múltiples copias y que parece asociado a la preferencia por las regiones SAR<sup>155</sup>. Recientemente también se ha descrito un dominio llamado SAF-box que determina la preferencia por las regiones SAR de factores de unión al scaffold<sup>143</sup>. En el caso de la histona H1 se ha señalado que el motivo conocido como SPKK pudiera estar involucrado en la interacción y preferencia por las SAR<sup>123</sup>.

### Motivo SPKK.

El motivo SPKK fue descubierto en la secuencia de las histonas del erizo de mar (*Strongylocentrotus nodus*) determinadas por von Holt<sup>156</sup>, quien señaló que dichas secuencias presentaban repeticiones de un pentapéptido comenzando en prolina. Posteriormente Suzuki postuló que el verdadero motivo repetitivo era S/TPKK y que se estructuraba en forma de giro  $\beta$ <sup>127</sup>.

El giro formado por el motivo SPKK (Fig. 14) está estabilizado por puentes de hidrógeno entre el OH del residuo  $i$ , que puede ser treonina, serina, ácido aspártico o asparagina, y el NH del residuo  $i+2$ . La presencia de este enlace produce un giro conocido como  $\sigma$ . El segundo puente de hidrógeno se establece entre el grupo carbonilo del residuo  $i$  y el NH del residuo  $i+3$ , enlace típico de un giro  $\beta$ . Se ha observado la formación de giros en secuencias del tipo SPXK donde la X es un residuo hidrofílico o alanina<sup>157</sup>.

La estructura del motivo SPKK en forma de giro  $\beta/\sigma$  (Figura 14) ha sido observada por Suzuki et al<sup>157</sup> en estudios de RMN en los péptidos SPRKSPRK y GSPKKSPRK en 90% de DMSO. En nuestro grupo esta estructura se ha observado en un péptido derivado del

dominio C-terminal de la histona H1 en 90% de TFE<sup>158</sup>. En el caso del péptido que contiene la unidad sencilla SPRK se encontró la conectividad correspondiente al giro  $\beta$ , pero no la  $\sigma$ <sup>157</sup>.

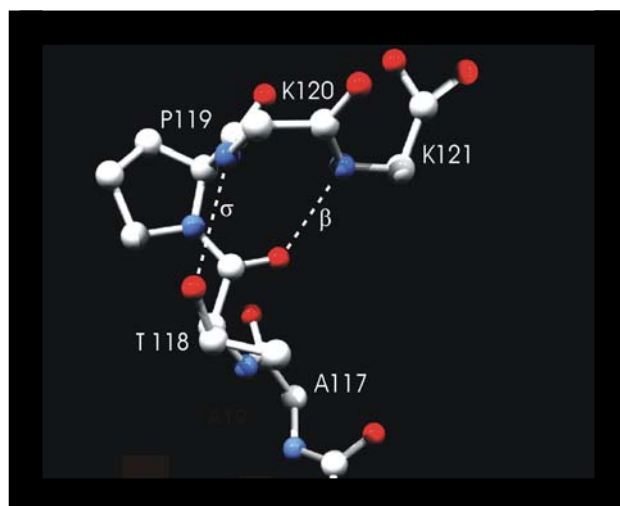


Figura 14. Motivo SPKK con el giro  $\beta/\sigma$  descrito por RMN en el C-terminal de la H1<sup>158</sup>.

La interacción de este motivo con el DNA se supone que ocurre en el surco menor, reemplazando la doble capa de moléculas de agua que estabiliza la forma B del DNA y que se conoce como espina de hidratación<sup>159</sup>. La hipótesis de la unión en el surco estrecho de la hélice de DNA se apoya en la protección observada frente a la metilación de las adeninas, que ocurre en el N3 ubicado en el surco menor, por el motivo SPKK<sup>127</sup> y la protección contra el footprinting por radical hidroxilo de las secuencias ricas en AT, lo que sugiere un modelo de interacción similar al Hoecht<sup>123</sup>. Este tipo de interacción también se ha descrito para drogas de unión al DNA y la netropepsina. Todas estas moléculas poseen grupos amida que se unen a la molécula de DNA por puentes de hidrógeno, reemplazando al agua. El hidrógeno enlazado al N de la serina del motivo SPKK no está involucrado en la formación del giro  $\sigma/\beta$  y pudiera formar dichos puentes de hidrógeno<sup>127</sup>.

De acuerdo con este modelo habría una unidad de SPKK por cada par de bases con dos cargas positivas para compensar las negativas del fosfato.

Se ha descrito un péptido de 16 residuos derivado de la H1e de rata, que contiene dos motivos SPKK y es capaz de condensar el DNA y la cromatina<sup>160</sup>. La estructura formada por dos SPKKs, podría conllevar la interacción entre moléculas de DNA y por tanto facilitar la

condensación del DNA. De forma general este péptido se comporta de manera análoga a la H1 en cuanto a la preferencia por las regiones ricas en A-T, la cooperatividad y la dependencia de la sal de la condensación del DNA.

La prolina del motivo SPKK es un aminoácido que puede sufrir isomerización. Generalmente las prolinas se encuentran en configuración *trans* (enlace de la prolina con el residuo anterior), pero pueden de forma espontánea o enzimática pasar a *cis*<sup>161</sup>. La presencia de la prolina convierte el motivo SPKK en sustrato de los enzimas con actividad prolina isomerasa, que pueden ser dependientes o no de la fosforilación en la serina o treonina adyacente<sup>66</sup>. La conformación de la prolina podría ser importante en la desfosforilación del motivo SPKK, la proteólisis, la estructura y la función del dominio C-terminal<sup>162, 163</sup>.

### **Dominios de la histona H1.**

#### Dominio globular.

El dominio globular tiene aproximadamente 80 residuos y se une al nucleosoma cerca del punto de entrada y salida del DNA, protegiendo 20 pares de bases adicionales de la digestión con nucleasa micrococcal<sup>164</sup>. Es el dominio más conservado evolutivamente hasta el punto que puede ser idéntico entre subtipos o entre especies<sup>78</sup>.

La estructura cristalográfica del dominio globular de la H5<sup>165</sup> (Fig. 15) ha identificado esta proteína como miembro del grupo de proteínas de unión al DNA que contienen el motivo hélice-giro-hélice modificado conocido como “winged helix”, y ha puesto de manifiesto una gran similitud estructural con la proteína activadora por catabolito (catabolite activator protein, CAP) y con un factor de transcripción específico de hígado, el HNF-3 $\gamma$ . La estructura de la H1 de pollo determinada por RMN también pertenece a este grupo. Por la homología existente entre la H5 y la H1<sup>o</sup>, y entre la H1 de pollo y la H1 de mamíferos, es razonable asumir que el mismo patrón estructural existe en los dominios globulares de todas las H1 de mamíferos. Recientemente también se ha resuelto la estructura cristalográfica del dominio globular de la H1<sup>o</sup> de rata<sup>166</sup>.



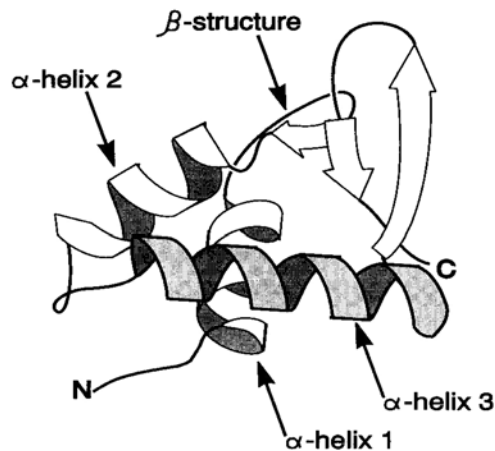


Figura 15. Estructura del dominio globular de la histona H5<sup>13</sup>.

Basándose en estas estructuras y datos bioquímicos se ha postulado la existencia de un sitio primario de unión formado por la hélice III, que se corresponde con el sitio de unión al DNA de CAP y HNF-3 $\gamma$  y por el giro  $\beta$  adyacente. En este sitio se encuentran los aminoácidos K69 y R73 de la hélice III y K85 del lazo  $\beta$ , que se conservan en todas las variantes estudiadas. No obstante, hay una gran divergencia de secuencia en la hélice III, lo que puede afectar el modo y la fuerza de la unión y también determinar algunas preferencias de secuencia. El sitio primario de unión se encuentra en el surco mayor de la doble hélice de DNA, posiblemente cerca del extremo del DNA del cromosoma<sup>165, 167</sup>.

Se piensa que existe un sitio secundario de unión formado por K40 y R42 ubicadas entre las hélices I y II, K52 de la hélice II y R94 del lazo  $\beta$ . Entre K40 y R42, adyacentes en la H5 y la H1 $^{\circ}$  se encuentra una asparagina que es un residuo no cargado, mientras que en la H1c hay un glutámico cuya carga negativa puede disminuir la densidad de carga positiva de la región, reduciendo su afinidad por al DNA<sup>91</sup>. Se ha comprobado que se requieren los dos sitios de unión para que el GH5 se una al nucleosoma<sup>168</sup>. En el caso del sitio de unión secundario, los mutantes simples donde se sustituye uno de los aminoácidos básicos por alanina son capaces de unirse al nucleosoma<sup>169</sup>.

El sitio de unión secundario al parecer interacciona con el surco menor de la doble hélice del DNA sobre el eje binario de simetría<sup>167</sup>. Los aminoácidos del sitio de unión secundario también se conservan en las variantes de las histonas de mamíferos, con la excepción de la H1t<sup>170</sup>. En la secuencia de la H1t la posición equivalente a K52 está ocupada por una

glutamina. Este cambio está presente en la secuencia de la H1t de rata, ratón, cerdo y humanos, lo que hace pensar que sea un cambio específico de este subtipo.

Al realizar el modelo del dominio globular de la H1t, basándose en el GH5, Rao y col<sup>170</sup>. encontraron que la glutamina tiene una orientación diferente de la lisina, y por tanto no está en posición de interactuar con el eje de simetría binario de la partícula nucleosomal. Este cambio tiene como efecto la disminución a la mitad de la afinidad por la estructura cruciforme de DNA y ofrece menos protección frente a la digestión con nucleasa micrococcal. Este cambio se mantiene en las variantes de H1t recientemente descubiertas H1T2p y H1LSp, aunque la significación biológica no está clara todavía.

### Dominio N-terminal.

El dominio N-terminal de la histona H1 tiene alrededor de 30 aminoácidos y contiene dos subregiones<sup>171</sup>. La región distal es rica en alanina, prolina y de carácter hidrofóbico, mientras que la región adyacente al dominio globular tiene carácter básico y podría contribuir a la estabilidad de la interacción del dominio globular<sup>172</sup>.

Este dominio se encuentra desestructurado en disolución pero adquiere una proporción importante de hélice  $\alpha$  en presencia de TFE o DNA<sup>173, 174</sup>.

Estudios de RMN-2D de los péptidos, NH-1 que contiene los primeros 20 aa de la H1° de ratón y NH-2, derivado de la misma proteína pero abarcando los aa 8-29, mostraron que la mitad básica del dominio N-terminal de la H1° adquiriría estructura helicoidal en presencia de 90% de TFE. La espectroscopía de infrarrojo (FTIR) de los complejos con DNA muestra que la estructura del N-terminal unido al DNA es similar a la obtenida en TFE<sup>173</sup>.

El análisis del RMN de los péptidos NH-1 y NH-2 permitió proponer un modelo estructural para el N-terminal de la H1° unido al DNA en el cual de los residuos del 11-24 se encuentran en hélice  $\alpha$ . El límite estructural del dominio N-terminal se extiende hasta la histidina 24, en lugar de la lisina 20, obtenido por digestión trípica (Fig. 16). En esta hélice encontramos dos regiones de tres residuos básicos en caras opuestas de la hélice, lo que permitiría su unión a diferentes regiones de DNA<sup>173</sup>.

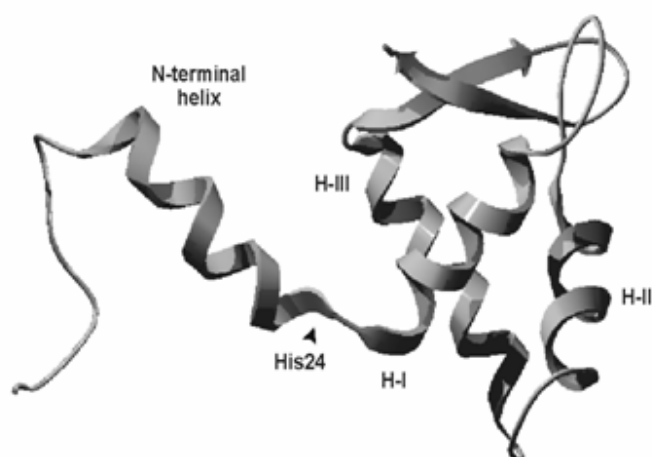


Figura 16. Modelo estructural para el dominio N-terminal de la H1<sup>e</sup> conectado con el dominio globular de la H5<sup>173</sup>.

También se ha estudiado un péptido, correspondiente a la región básica del dominio N-terminal de la H1e, por dicroísmo circular y RMN. Este péptido comprende los residuos 15-36 y en presencia de TFE adquiere un 40% de hélice  $\alpha$ . En estas condiciones los estudios de RMN mostraron que el péptido está estructurado en dos hélices anfipáticas separadas por un motivo Gly-Gly que funciona como una bisagra entre las dos hélices. La primera hélice abarca los residuos 17-27 y la segunda los residuos 29-34, quedando así prácticamente todo el péptido en hélice  $\alpha$ <sup>174</sup> (Fig. 17).

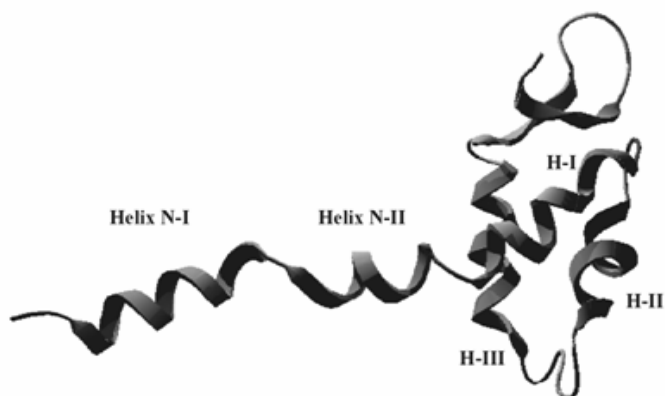


Figura 17. Modelo estructural para el dominio N-terminal de la H1e conectado con el dominio globular de la H5<sup>174</sup>.

### Dominio C-terminal.

El dominio C-terminal de la histona H1 representa alrededor de la mitad de la proteína y tiene un alto contenido (~40%) de aminoácidos básicos, en su mayor parte lisina. El dominio C-terminal también es rico en serina, prolina y alanina. Este dominio es muy variable entre los diferentes subtipos de histonas, aunque la distribución de las cargas positivas es relativamente uniforme a pesar de las diferencias en la secuencia de aminoácidos<sup>175</sup>.

En general se acepta que el dominio C-terminal es responsable de la condensación de la cromatina en la fibra de 30 nm y la agregación intermolecular<sup>176</sup>. El dominio C-terminal también tiene capacidad de condensar el DNA, a juzgar por la inducción de espectro en psi ( $\psi$ ) de dicroísmo circular en el DNA, atribuido a la agregación ordenada de las moléculas de DNA<sup>177</sup>.

Al parecer la condensación del DNA es mediada por la neutralización de las cargas negativas de los fosfatos del DNA linker por las lisinas del C-terminal. En estudios recientes, se ha postulado la presencia de subdominios dentro del C-terminal que serían los responsables de la alteración del DNA linker y la estabilización de estructuras de cromatina condensada<sup>178</sup>.

Dichos experimentos se realizaron con mutantes truncados de la H1<sup>o</sup> de ratón evaluando por ultracentrifugación su capacidad de alterar el DNA linker, el plegamiento y la autoasociación de las fibras de cromatina. Los resultados indican que la mayoría de las funciones pueden ser realizadas por un mutante que contiene los primeros 24 aminoácidos del dominio C-terminal, aunque para el adecuado plegamiento de las fibras de cromatina parecen necesarios los aminoácidos 145-169 del C-terminal<sup>178</sup> (Fig. 18).



Figura 18. Distribución de funciones en la histona H1<sup>o</sup>. L: alteración del DNA linker. F: plegamiento de la fibra de cromatina. SA: Autoasociación de la fibra de cromatina<sup>178</sup>.

En la H1e de rata se ha identificado un fragmento de 34 aminoácidos como la región necesaria para la condensación de la cromatina. Este fragmento comprende los aminoácidos 145-178 y contiene tres de los cuatro motivos SPKK presentes en ese C-terminal<sup>179</sup> (Fig. 19).

KKAASGEAKPKAKKAGAAKAKKPAGAAKKPKKATGT*ATPKKSTKKTP*  
*KKAKKPAAAAGAKKAKSPKKAKAT*TKAKKAPKSPAKARAVKPKAAKPKT  
 SKPKAAKPKKTAACKK

Figura 19. Secuencia del dominio C-terminal de la histona H1e de rata. Subrayado y en cursiva se destaca el fragmento de 34 aa responsable de la condensación del DNA. En rojo los motivos SPKK.

Los motivos SPKK están ubicados en repeticiones imperfectas de un octapéptido que es capaz de mimetizar las propiedades condensantes del dominio C-terminal de la histona H1<sup>160</sup>. Es importante señalar que dos de los tres motivos SPKK presentes en el dominio C-terminal de la H1<sup>o</sup> de ratón se encuentran ubicados en los dos subdominios del C-terminal involucrados en la condensación de la cromatina detectados por Hansen y col<sup>178</sup>.

El dominio C-terminal también es esencial para la unión a la cromatina *in vivo* con alta afinidad<sup>180, 181</sup>. En estudios más recientes, se apunta además a la posibilidad de que esta unión este modulada por fosforilación<sup>182</sup>.

Esta conclusión se extrajo estudiando la H1a humana que tiene solamente dos sitios de fosforilación, S183 en el motivo SPAK y T152 del motivo TPKK. Los experimentos de delección mostraron que si se eliminaba del aminoácido 151 en adelante la proteína resultante se unía con muy poca afinidad a la cromatina. La sustitución por mutagénesis de T152 y S183 por glutámico, para así mimetizar la introducción de una carga negativa por la fosforilación tenía el mismo efecto que la delección a partir de la posición 151, especialmente la mutación en la T152, asociada al motivo SPKK. Estos resultados sugieren que la afinidad de la unión de la H1 a la cromatina puede estar modulada directamente por fosforilación de residuos del C-terminal<sup>182</sup>.

El dominio C-terminal está desestructurado en disolución, pero adquiere un grado sustancial de estructura en hélice  $\alpha$  en presencia de TFE<sup>183, 184</sup>.

La estabilización de hélice  $\alpha$ <sup>183, 184</sup> y giros  $\beta$ <sup>185</sup> en presencia de estabilizantes de estructura secundaria hizo pensar que la interacción con el DNA induciese estructura secundaria en el dominio C-terminal de la histona H1.

El péptido CH-1 cuya secuencia pertenece al dominio C-terminal de la histona H1<sup>o</sup> de ratón (residuos del 99 al 121), adyacente al dominio globular, fue estudiado por DC y RMN-2D. Se observó la inducción en el CH-1 de hasta un 45 % de hélice  $\alpha$  en presencia de TFE por DC. Los estudios de RMN permitieron determinar la estructura del péptido en 90% de TFE, que incluye una región de hélice  $3_{10}$  en los residuos 100-103, una hélice  $\alpha$  de marcado carácter anfipático desde el 104 hasta 117, con 8 residuos básicos en una cara de la hélice y 6 hidrofóbicos en la otra. Esta región está seguida por un giro  $\sigma/\beta$  (118-121), que se corresponde con el motivo SPKK<sup>158</sup> (Fig. 20).

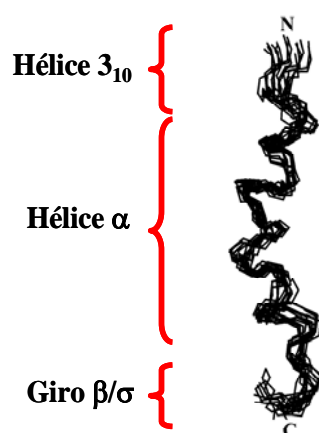


Figura 20. Estructura del péptido CH-1 (residuos 99-121 de la H1<sup>o</sup>) en presencia de TFE por RMN<sup>158</sup>.

La interacción del péptido CH-1 con el DNA se estudió por espectroscopía de infrarrojo (FTIR). La interacción con el DNA induce la estructuración del péptido. La componente mayoritaria es de hélice  $\alpha$ , con un 43%, distribuida en dos picos ( $1647, 1657 \text{ cm}^{-1}$ ), un 20% de hélice  $3_{10}$  a  $1637 \text{ cm}^{-1}$  y un 20% de giros a  $1670 \text{ cm}^{-1}$ , atribuido al SPKK<sup>186</sup>.

Sobre la base de estudios de modelado molecular, Rao y col.<sup>187</sup> han propuesto que el segmento de 34 aa del C-terminal de la H1e, con tres motivos SPKK y gran capacidad condensante, forma un dominio del tipo HMG-box con tres hélices  $\alpha$  cortas, igual al encontrado en las proteínas Lef-1 y SRY<sup>188</sup> (Fig. 21A). A partir de esta predicción Rao y col.<sup>187</sup> han elaborado un modelo de cromatosoma que incluye la histona H1 entera (Fig. 21B).

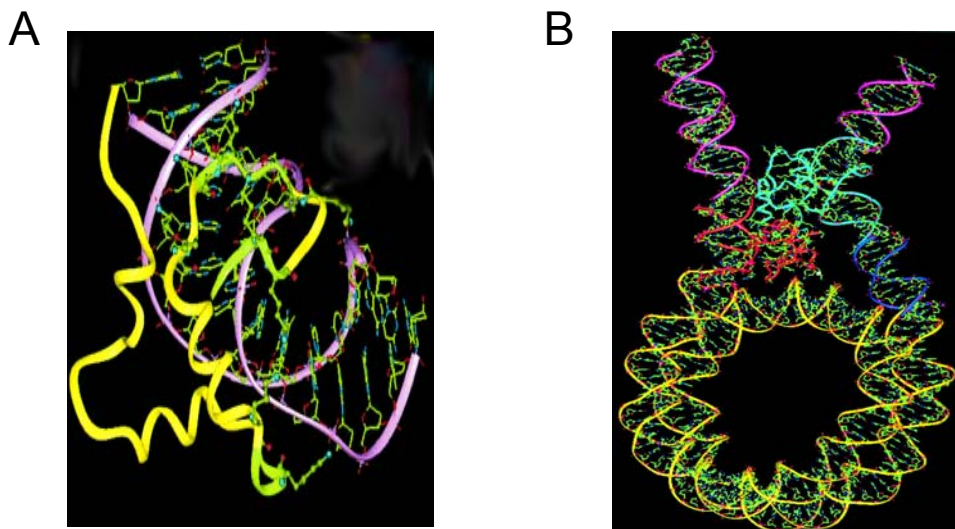


Figura 21. Interacción de la H1e de rata y el DNA. A. Dominio HMG-box de la H1e de rata (amarillo) unido al DNA (lila). En verde aparecen señalados los 3 motivos SPKK. B. Modelo del cromatosoma con la H1e de rata. El dominio globular en rojo y el C-terminal en azul claro. Los residuos de la doble hélice del DNA en contacto con cada una se muestran en el mismo color que los dominios de la histona. El segmento en azul indica el DNA adicional que se necesita para acomodar el dominio C-terminal en esta posición<sup>187</sup>.

Recientemente se ha relacionado el dominio C-terminal de la H1 con la activación de una de las nucleasas apoptóticas, el factor de fragmentación nuclear DFF40/CAD<sup>189</sup>. Una de las características de las etapas finales de la apoptosis es la fragmentación del DNA nuclear en mononucleosomas que luego son fagocitados. Este proceso es importante ya que elimina la transformación potencial de otras células por oncogenes<sup>190</sup>. Defectos en la apoptosis pueden causar predisposición a enfermedades autoinmunes<sup>191</sup>.

El DFF40 es inactivo en condiciones normales ya que está unido a una molécula inhibidora, formando un heterodímero. El estímulo para la apoptosis activa la caspasa-3, que corta la subunidad inhibidora dejando activo el DFF40<sup>192</sup>. La histona H1 se une al DFF40 formando un complejo con alta afinidad por el DNA. Esto produce un aumento marcado de la degradación de DNA desnudo e internucleosomal por el DFF40<sup>193</sup>. La unión de la H1 al DFF40 dirige el complejo hacia el DNA linker de la cromatina, donde el DFF40 corta la cadena de DNA<sup>194</sup>. La activación se produce con una molécula de H1 por cada 1-2 nucleosomas. La unión del DFF40 a la histona H1, permite el funcionamiento del DFF40 a concentraciones salinas fisiológicas (150 mM NaCl)<sup>189</sup>.

Al analizar el efecto en la activación de los mutantes de delección del C-terminal<sup>178</sup> se detectó que es necesario el dominio C-terminal intacto para la máxima activación del DFF40. También se encontró que los seis subtipos de H1 somáticas de ratón eran capaces de producir igual activación. Esta falta de especificidad asegura la apoptosis independientemente de la composición específica de isoformas de un tipo celular<sup>189</sup>.

Cuando se induce la apoptosis se produce la desfosforilación de las histonas<sup>195</sup>. Esto puede ser consecuencia, al menos en parte, de la parada del ciclo celular<sup>196</sup>. La presencia de las histonas hipofosforiladas en el momento de la interacción de la H1 con DFF40, hizo pensar la posible relación entre el estado de fosforilación y la modulación de esta proteína. Los experimentos realizados en líneas celulares tumorales humanas demostraron que el efecto de la H1 en la endonucleasa DFF40 era independiente del estado de fosforilación de la histona H1<sup>59</sup>.

### **Modificaciones postraduccionales de la histona H1.**

La histona H1, al igual que las histonas internas, experimenta varios tipos de modificación postraduccionales: acetilación, metilación, poliADP-ribosilación, formilación, deimidación, ubiquitinación y fosforilación<sup>197, 198, 199</sup>. La acetilación y metilación de las histonas H1 ocurre fundamentalmente en células en rápida proliferación, como en las del bazo<sup>96</sup>. La fosforilación se considera la modificación fundamental en la H1 ya que ocurre en múltiples sitios en la molécula y está asociada al ciclo celular. La diversidad de efectos estructurales y funcionales que podría provocar la convierten en un tema muy importante a estudiar para entender el papel de la H1 en la cromatina.

#### **Acetilación.**

La histona H1 se modifica mediante dos tipos diferentes de acetilación, la que ocurre en el N-terminal y la que tiene lugar en el grupo  $\epsilon$ -amino de las lisinas.

La mayoría de las proteínas eucariotas pierden la metionina N-terminal y el aminoácido que queda en su lugar al inicio de la cadena polipéptidica es acetilado. Esta modificación ocurre co-traduccionalmente y no tiene una secuencia específica, aunque el 95% de los residuos acetilados son: serina, alanina, metionina, glicina y treonina. La acetilación del N-terminal *in vivo* se ha caracterizado en la mayoría de los subtipos de H1 por espectrometría de masas a partir de cultivos celulares y en tejidos de ratón<sup>96, 105</sup>. Esta modificación había sido



documentada previamente en experimentos de microsecuenciación<sup>197</sup>. En los estudios realizados en tejidos y líneas celulares se ha encontrado también una pequeña proporción del N-terminal de las histonas H1 procesado, pero no acetilado.

La H5 y la H1° se encuentran acetiladas y no acetiladas<sup>200, 201</sup>. Mientras que en la H5 no varía el grado de acetilación, en la H1° la acetilación del N-terminal aumenta con el envejecimiento en ratas. En hígado de ratón se ha encontrado la metionina inicial de la H1° acetilada<sup>96</sup>.

La acetilación del grupo ε-amino de la lisina es muy común en las histonas internas. En la histona H1 se encontró inicialmente *in vivo* la presencia de un sitio de acetilación interno en el dominio globular (K64) que es compartido por los subtipos c, d, e<sup>199</sup>. En estudios más recientes se han encontrado múltiples sitios de acetilación internos (Tabla 1)<sup>96</sup>.

**Tabla 1. Sitios de acetilación internos de la histona H1<sup>96</sup>.**

Dominio	Cultivos celulares		Tejidos de ratón	
	Subtipo	Posición	Subtipo	Posición
N-terminal	H1b, c, d, e	K17	H1°	K12
	H1a	K22	H1a, c	K16
			H1e	K17
			H1a	K22
Globular	H1a	K64, K85, K90, K97	H1a	K75, K87, K90
	H1b	K49, K88	H1b	K85
	H1c, d, e	K34, K46, K52, K64, K90	H1c, d	K34, K46, K52, K75, K85, K90
			H1e	K46, K52, K75, K85, K90
C-terminal	H1b	K168	H1a	K121
	H1c, d, e	K169	H1b	K158, K165
	H1c	K192	H1c	K160, K168, K178
			H1d	K168
			H1e	K160, K168

La numeración de las posiciones se realizó usando la secuencia de la H1c.

Los tejidos analizados: cerebro, corazón, músculo, pulmón, bazo, riñón y vesícula.

La mayoría de los sitios de acetilación están localizados en el dominio globular (Tabla 1). En particular, la acetilación afecta residuos que se consideran directamente involucrados en la unión al DNA como son K52, K64, K85 y K97<sup>168</sup>. La presencia de residuos acetilados

puede interferir en la unión al DNA confiriendo a algunos subtipos propiedades específicas en la cromatina. En los dominios N y C-terminal se encuentran también residuos modificados (Tabla 1)<sup>96</sup>. El papel de esta modificación en la estructura y función de la histona H1 aún está por estudiar.

### Metilación.

La lisina y la arginina pueden metilarse en sus cadenas laterales. Inicialmente se describieron dos sitios metilados en la H1<sup>62, 227</sup>. Una de las modificaciones se ha encontrado en la H1e de células HeLa en la lisina 25, adyacente a la serina 26 fosforilada. Estas modificaciones al parecer forman un "methyl-phos switch" donde mientras la lisina está metilada se une la proteína HP1 y al fosforilarse la serina 26 se bloquea dicha unión<sup>202</sup>. La modificación de la lisina 25 al parecer es realizada por la histona-lisina metiltransferasa Ezh2, la cual se ha asociado a la represión transcripcional<sup>203</sup>.

También se ha encontrado H1b metilada en la lisina 26 en fibroblastos de ratón. En este caso la metilación es indispensable para la interacción de este subtipo con las variantes de la HP1 $\beta$  y 1 $\gamma$ . Como en el caso de la interacción con HP1 $\alpha$ , la fosforilación de la H1 actúa como señal para la eliminación de la interacción<sup>204</sup>.

En estudios más recientes se han descrito múltiples sitios de metilación (Tabla2), algunos acoplados a acetilación o a metilación de residuos adyacentes<sup>96</sup>.

**Tabla 2. Sitios de metilación de la histona H1<sup>96</sup>.**

Cultivos celulares		Tejidos de ratón	
Subtipo	Posición	Subtipo	Posición
H1b	aK168-mK169	H1a	mK22-mK23, md-K63, dK92
H1d, e	aK169-mK170, a-mK169	H1b	mK34, mK75
		H1c	mK26, mK34, mdK46, mdK63, mdK75, mK106
		H1d	mK17, mK34, mdK46, mdK63, mdK75, mK106
		H1e	dK17, mR33-mK34, mK34, mdK63, mdK75, mK106

a: acetilado, m: metilado, d: dimetilado.

Las acetilaciones se señalan cuando la metilación aparece acoplada a esta modificación.

La numeración de las posiciones se realizó usando la secuencia de la H1c.

Los tejidos analizados son; corazón, músculo, pulmón, bazo, riñón y vesícula.

### ADP-ribosilación.

La ADP-ribosilación de la H1 puede involucrar la unión covalente de la ADP-ribosa, con 8-10 ADP-ribosas por H1 o la unión no covalente de polímeros ramificados de ADP-ribosa con hasta 100-200 unidades por proteína. La unión no covalente de ADP-ribosa es resistente a la acción de ácidos fuertes, detergentes, agentes caotrópicos y condiciones de elevada fuerza iónica<sup>205</sup>. Es evidente que la presencia de polímeros de ADP-ribosa en la H1 cambia su carga y se espera que afecte la interacción H1-DNA y H1-H1. *In vitro* la poliADP-ribosilación conlleva a una relajación significativa de la cromatina, pero no desplaza la H1 del DNA linker<sup>206, 207</sup>. *In vivo* esta modificación podría tener un papel regulador en la protección del patrón de metilación del DNA, ya que se ha encontrado necesaria para el mantenimiento del estado no metilado en las islas CpG del promotor del gen Htf9 de fibroblastos<sup>208, 209</sup>.

### Formilación.

Esta modificación novedosa se encontró en algunos residuos de lisina solamente en tejidos de ratón, pero no en líneas celulares (Tabla3)<sup>96</sup>. Su origen y función potencial son desconocidos. La formilación puede ser catalizada a partir del formiltetrafolato o a partir del formaldehído liberado en la dimetilación de la lisina mediante un enzima específico, o la LSD1 dimetilasa en la reacción inversa<sup>210</sup>.

**Tabla 3. Sitios formilados en la histona H1<sup>96</sup>.**

Tejidos de ratón	
Subtipo	Posición
H1a	K63
H1c	K34, K52, K63, K85, K90, K106
H1d	K34, K52, K63, K85
H1e	K52, K63, K85

La numeración de las posiciones se realizó usando la secuencia de la H1c.

Los tejidos analizados son; cerebro, corazón, músculo, pulmón, bazo, riñón y vesícula.

### Deimidación.

La deimidación de las proteínas es un proceso generalmente no enzimático. Ocurre en residuos de asparagina y glutamina teniendo como resultado la formación de ácido aspártico y glutámico, respectivamente<sup>211, 212</sup>. Esta modificación implica la incorporación de una

carga negativa a la secuencia de aminoácidos. Algunos estudios sugieren que la deamidación aumenta con la edad en proteínas con poco o ningún recambio durante el ciclo celular. Este aumento se debe probablemente a la disminución de la tasa de degradación de las proteínas deamidadas. En la histona H1<sup>o</sup> ocurre la deamidación de la asparagina en la posición 3 y la presencia de dicha modificación aumenta con la edad<sup>105</sup>.

### Ubiquitinación.

La ubiquitinación *in vitro* de la H1 se describió por primera vez en *Drosophila*, aunque sin identificar el sitio concreto<sup>213</sup>. Actualmente se han identificado dos lisinas ubiquitinadas, la K46 en la H1c, d, e y la K116 en la H1a<sup>96</sup>. En las histonas del core, esta modificación se asocia con el aumento de la expresión<sup>214</sup>, pero su presencia en las H1 tiene efectos aún indeterminados.

### Fosforilación.

La fosforilación de proteínas es un mecanismo de regulación fundamental en las células eucariotas. Se estima que al menos un tercio de las proteínas son fosforiladas de forma reversible<sup>215</sup>. La fosforilación modula la actividad de numerosas proteínas involucradas en la transducción de señales, regula la afinidad de unión de factores de transcripción a sus coactivadores y al DNA, alterando la expresión génica, el crecimiento celular y la diferenciación<sup>216</sup>.

El estudio de la fosforilación de las proteínas como mecanismo fundamental de control de la progresión del ciclo celular ha llevado a la identificación de proteínas quinasa dependientes de ciclinas (CDKs) que son necesarias para el avance del ciclo celular (Tabla 4)<sup>217</sup>.

**Tabla 4. Quinasas dependientes de ciclinas relacionadas con el control del ciclo celular<sup>217</sup>.**

Etapa del ciclo	CDK	Ciclina	Inhibidores
G1	CDK4/6	D	p21, p27, p15, p16
G1/S	CDK2	E	p21, p27
S	CDK2	A	p21
G2/M	cdc2	A	p21
M	cdc2	A, B	

Las CDKs están compuestas por una subunidad catalítica y otra reguladora. Su actividad se regula mediante modificaciones postraduccionales y la asociación a inhibidores<sup>218</sup>.

La fosforilación de la histona H1 es dependiente del ciclo celular, alcanzando su grado máximo durante la mitosis y mínimo en la fase G1<sup>198</sup>. Progresiva desde un fosfato por molécula al final de G1 a tres en la fase S y G2, llegando a 5 fosfatos, en algunos subtipos, en mitosis<sup>218, 198</sup> (Tabla 5). Todos los grupos fosfato se han perdido al final de la telofase, de forma que al principio del nuevo ciclo la H1 no tiene fosfatos.

**Tabla 5. Número de fosfatos encontrados en los subtipos de H1 en fibroblastos de ratón en las etapas del ciclo celular**<sup>198, 219</sup>.

Subtipo	G1 tardío	S tardía	Mitosis
H1a	1	1	4
H1b	2	3 (3)	5 (5)
H1c	1	1 (1)	4
H1d	1	2 (1)	5
H1e	1	3 (2)	4
H1°	1	1	3

Entre paréntesis se muestran los sitios fosforilados *in vivo* en linfoblastos humanos.

Los estudios realizados indican que la fosforilación de la H1 está más directamente relacionada con la descondensación de la cromatina que con su condensación. Green y col.<sup>220</sup> mostraron que la H1 hipofosforilada se une con mayor fuerza al DNA que la H1 hiperfosforilada. La fosforilación disminuye la afinidad de la H1 por el DNA, lo que permitiría la descondensación necesaria en las diferentes etapas del ciclo celular<sup>219</sup>.

La fosforilación moderada durante la interfase podría propiciar la descondensación necesaria en la cromatina para la replicación del DNA y la transcripción<sup>221</sup>. También se ha descrito el papel de la fosforilación de la H1 en la expresión de genes específicos<sup>222</sup>.

Durante la mitosis la hiperfosforilación de la H1, permitiría la unión de factores específicos de condensación de la cromatina<sup>219</sup>. Así sería posible relacionar el aumento en los niveles de fosforilación de la H1 con el aumento de la condensación de la cromatina durante G2, hasta su momento de máxima compactación en los cromosomas metafásicos. El tratamiento con estaurosporina impide la fosforilación de la H1 y la entrada en mitosis de la célula. El efecto inverso se puede obtener añadiendo la droga a células detenidas en mitosis<sup>223, 224</sup>. Se acepta que el estado global de fosforilación de la H1 es el resultado del balance entre las actividades de la proteína fosfatasa 1 y de las quinasas CDK2/cdc2<sup>219</sup>.

La histona H1 se fosforila por quinasas dependientes de ciclinas en el motivo (S/T)-P-X-Z, donde X es cualquier aminoácido, mientras que Z es un aminoácido básico<sup>199, 96</sup>. Esta secuencia consenso incluiría los motivos SPKK, que están involucrados en aspectos estructurales y funcionales de la H1 (Fig. 22).

H1a_m	SETAPVAQAA	STATEKPAAA	KKT.KKPAKA	AAP.RKKPAG	PSVSELIVQA	VSSSKERSGV
H1b_m	SETAPAETAA	PAPVEK <b>SPAK</b>	KKTTKK...	AGAAKRKATG	PPVSELITKA	VSASKERGGV
H1c_m	SEAAPAAPAA	APPAEKAPAK	KKAACK.PAG	V...RRKASG	PPVSELITKA	VAASKERSGV
H1d_m	SETAAPAPAA	PAPVEK <b>TPVK</b>	KKA.KK.T.G	AAAGKRKASG	PPVSELITKA	VAASKERSGV
H1e_m	SETAAPAPAA	PAPAEK <b>TPVK</b>	KKA.RKAA.G	G..AKRKTSG	PPVSELITKA	VAASKERSGV
H1a_h	SETVPPAPAA	SAAPEKPLAG	KKA.KKPAKA	AAASKKKPAG	PSVSELIVQA	ASSSKERGGV
H1b_h	SETAPAETAT	PAPVEK <b>SPAK</b>	KKATKAAA.G	AGAAKRKATG	PPVSELITKA	VAASKERNGL
H1c_h	SETAAPAPAA	APPAEKAPVK	KKAACK.AGG	<b>TP...RKASG</b>	PPVSELITKA	VAASKERSGV
H1d_h	SETAPLAPTI	PAPAEK <b>TPVK</b>	KKA.KK.A.G	ATAGKRKASG	PPVSELITKA	VAASKERSGV
H1e_h	SETAAPAPAA	PAPAEK <b>TPVK</b>	KKA.RKSA.G	A..AKRKASG	PPVSELITKA	VAASKERSGV
	** *	**	** *	* *	* * * * * *	* * * * *
H1a_m	SLAALKKSLA	AAGYDVEKNN	SRIKLGLKSL	VNKGTLVQTK	GTGAAGSFKL	NKKA...ESK
H1b_m	SLPALKKALA	AGGYDVEKNN	SRIKLGLKSL	VSKGTLVQTK	GTGASGSFKL	NKKAASGEAK
H1c_m	SLAALKKALA	AAGYDVEKNN	SRIKLGLK <b>SL</b>	VSKGILVQTK	GTGASGSFKL	NKKAASGEAK
H1d_m	SLAALKKALA	AAGYDVEKNN	SRIKLGLK <b>SL</b>	VSKGTLVQTK	GTGASGSFKL	NKKAASGEAK
H1e_m	SLAALKKALA	AAGYDVEKNN	SRIKLGLK <b>SL</b>	VSKGTLVQTK	GTGASGSFKL	NKKAASGEAK
H1a_h	SLAALKKALA	AAGYDVEKNN	SRIKLGIKSL	VSKGTLVQTK	GTGASGSFKL	NKKASSVETK
H1b_h	SLAALKKALA	AGGYDVEKNN	SRIKLGLKSL	VSKGTLVQTK	GTGASGSFKL	NKKAASGEAK
H1c_h	SLAALKKALA	AAGYDVEKNN	SRIKLGLKSL	VSKGTLVQTK	GTGASGSFKL	NKKAASGEAK
H1d_h	SLAALKKALA	AAGYDVEKNN	SRIKLGLKSL	VSKGTLVQTK	GTGASGSFKL	NKKAASGEAK
H1e_h	SLAALKKALA	AAGYDVEKNN	SRIKLGLKSL	VSKGTLVQTK	GTGASGSFKL	NKKAASGEAK
	** * * * *	** * * * * * *	* * * * * * *	* * * * * * *	* * * * * *	* * * * *
H1a_m	AITTKVSVKA	KA...SGA.A	KKPKKTAGAA	A.KKTV.K <b>TP</b>	<b>KKPKKPAV...SKKTSKSPK</b>	
H1b_m	PKA.KKTGAA	KAKKPAGAT <b>TP</b>	<b>KKPKKTAGA...KKTVKKTP</b>	<b>KKAKKPAAA</b>	<b>GVKKVAKSPK</b>	
H1c_m	PQA.KKAGAA	KAKKPAGA.A	KKPKKATGAA	<b>TPKKAACKTP</b>	<b>KKAKKPAAAA</b>	<b>VTKKVAKSPK</b>
H1d_m	PKA.KKAGAA	KAKKPAGA.A	KKPKKATGAA	<b>TPKKTAKKTP</b>	<b>KKAKKPAAAA</b>	<b>GAKKVS</b> <b>SKSPK</b>
H1e_m	PKA.KRAGAA	KAKKPAGA.A	KKPKKAAGTA	TAKKSTK <b>TP</b>	<b>KKAKKPAAAA</b>	<b>GAKK.AKSPK</b>
H1a_h	PGASKVATKT	KA...TGA.S	KKLKKATGA.	S.KKSV.K <b>TP</b>	<b>KKAKKPAA...TRKSSKNPK</b>	
H1b_h	PKA.KKAGAA	KAKKPAGAT <b>TP</b>	<b>KKAKKAAGA...KKAVK<b>TP</b></b>	<b>KKAKKPAAA</b>	<b>GVKKVAKSPK</b>	
H1c_h	PKV.KKAGGT	KPKKPVGA.A	KKPKKAAGGA	<b>TPKKS<b>AKKTP</b></b>	<b>KKAKKPAAAT</b>	<b>VTKKVAK<b>SPK</b></b>
H1d_h	PKA.KKAGAA	KPRKPAGA.A	KKPKKVAGAA	<b>TPKKS<b>IKKTP</b></b>	<b>KKVKKPATAA</b>	<b>GTKK<b>VAKSAK</b></b>
H1e_h	PKA.KKAGAA	KAKKPAGA.A	KKPKKATGAA	<b>TPKKS<b>AKKTP</b></b>	<b>KKAKKPAAAA</b>	<b>GAKK.AK<b>SPK</b></b>
	*	* **	** * * *	** **	** * * * *	* * *
H1a_m	KPKVV.KAKK	VAK <b>SPAKAKA</b>	VKPKASKAKV	TKPK <b>TPAKP</b> ...	...KKAAPK	KK
H1b_m	KAKAAAKPKK	AAK <b>SPAKPKA</b>	VKSKASKPKV	TKPKT.AKPK	AAKAKKAVSK	KK
H1c_m	KAKVT. <b>KPKK</b>	V.KSAS...KA	VKPKAAKPKV	AKAK.....	...KVAAPK	KK
H1d_m	KVKAA.KPKK	AAK <b>SPAKAKA</b>	PKAKASKPKA	SKPKAT...	...KAKKAAPR	KK
H1e_m	KAKAT.KAKK	APK <b>SPAKAKT</b>	VKPKAAKPKT	SKPKA.AKP...	...KKTAAK	KK
H1a_h	KPKTV.KPKK	VAK <b>SPAKAKA</b>	VKPKAAKARV	TKPKT.AKP...	...KKAAPK	KK
H1b_h	KAKAAAKPKK	ATK <b>SPAKPKA</b>	VKPKAAKPKA	AKPKA.AKPK	AAKAKKAAAK	KK
H1c_h	KAKVA.KPKK	AAKSAA...KA	VKPKAAKPKV	VKPK.....	...KAAPK	KK
H1d_h	KVK <b>TP</b> .QPKK	AAK <b>SPAKAKA</b>	PKPKAAKPKS	GKPKVT...	...KAKKAAPK	KK
H1e_h	KAKAA.KPKK	APK <b>SPAKAKA</b>	VKPKAAKPKT	AKPKA.AKP...	...KKAAPK	KK
	**	** **	* * * *	**	*	**

Figura 22. Comparación entre los subtipos H1a, b, c, d, e de la histona H1 de ratón (m) y humano (h). En rojo se muestran los sitios de fosforilación *in vivo* encontrados en humanos. Asterisco (\*), aminoácido conservado en todas las secuencias analizadas. En negrita los motivos del tipo (S/T)-P-X-Z<sup>96, 199, 200</sup>.

Se encontraron 7 sitios de fosforilación que no comparten esta secuencia consenso<sup>96, 199, 200</sup>, dos de los cuales podrían ser fosforilados por la proteína quinasa A, ya que son secuencias del tipo R-X-X-S y R-X-S. El péptido KARKSAGAAKR del dominio N-terminal de la H1e se encontró fosforilado en la serina 26 y metilado en la lisina 29, lo que abre la discusión sobre la posibilidad de la existencia de algún tipo de señal combinada entre las dos modificaciones<sup>199</sup>.

Las histonas H1<sup>o</sup> y H1t difieren bastante de las H1 somáticas. Las secuencias de H1<sup>o</sup> de humano y ratón son muy parecidas y se conservan los motivos de fosforilación (S/T)-P-X-Z (Fig. 23A). En la H1t la homología de secuencia entre las dos especies es menor e incluso se pierde uno los motivos de fosforilación en humanos (Fig. 23B).

### A.

```

H1o_m TENSTAPAA  KPKRAKASKK  STDHPKYS DM  IVAAIQAEKN  RAGSSRQSIQ  KYIKSHYKVG
H1o_h TENSTAPAA  KPKRAKASKK  STDHPKYS DM  IVAAIQAEKN  RAGSSRQSIQ  KYIKSHYKVG
*****

H1o_m ENADSQIKLS  IKRLVTTGVL  KQTKGVGASG  SFRLAKGDEP  KRSVAFKTKK  KEVKKVATPK
H1o_h ENADSQIKLS  IKRLVTTGVL  KQTKGVGASG  SFRLAKSDEP  KRSVAFKTKK  KEIKKVATPK
*****

H1o_m KAAKPKKAAS  KAPSKKPKAT  PVKKAKKKPA  ATPKKAKKPK  VVKVKPVKAS  KPKKAKTVKP
H1o_h KASKPKKAAS  KAPTKKPKAT  PVKKAKKKLA  ATPKKAKKPK  TVKAKPVKAS  KPKKAKPVKP
** *****

H1o_m KAKSSAKRAS  KKK
H1o_h KAKSSAKRAG  KKK
*****

```

### B.

```

H1t_m SETAPAASST  LVPAPVEEKP  SSKRRGKKPG  .LAPARKPR  GFSVSKLIPE  ALSTSQERAG
H1t_h SETVPAASAS  AGVAAMEKLP  .TKKRGRKPA  GLISASRKVP  NLSVSKLITE  ALSVSQERVG
*** *****

H1t_m MSLAALKKAL  AAAGYDVEKN  NSRIKLALKR  LVNKGVLVQT  KGTGASGSFK  LSKKAASGND
H1t_h MSLVALKKAL  AAAGYDVEKN  NSRIKLSLSK  LVNKGILVQT  RGTGASGSFK  LSKKVIPKST
*** *****

H1t_m KGKGKSASA  KAKKMGLPRA  SRSPKSSKT  .KAVKKPKATP  TKASGSGRKT  KGAKGVQQRK
H1t_h RSKAKKSVSA  KTKKLVLSRD  SKSPKTAKTN  KRAKKPRATT  PKTVRSGRKA  KGAKGKQQK
* * * * *

H1t_m SPAKARAANP  NSGKAKMVMQ  .KTDLRKAAG  RK
H1t_h SPVKARAS  . . . KSKLTQH  HEVNVRKATS  KK
** * * *

```

Figura 23. Comparación de las secuencias de H1<sup>o</sup> y H1t de humanos y ratón. A, H1<sup>o</sup>. B, H1t. En rojo se muestran los sitios de fosforilación *in vivo* encontrados en humanos. Asterisco (\*), aminoácido conservado en todas las secuencias analizadas. En negrita los motivos del tipo (S/T)-P-X-Z<sup>96</sup>

De estos subtipos casi no hay información sobre la fosforilación *in vivo*. Hasta ahora se la encontrado la H1<sup>o</sup> fosforilada en S5, diana no fosforilada por la CDKs. También se ha encontrado fosforilada la H1x en S31<sup>96</sup>.

El enzima responsable del alto grado de fosforilación de la H1 en mitosis fue descrita por Lake y Salzman<sup>225</sup>. Esta enzima se denominó H1-quinasa asociada al crecimiento y se encontró en muchos tipos de células en división. El enzima era independiente de cAMP y tenía homología de secuencia con la primera quinasa dependiente de ciclina que se identificó, p34<sup>CDK2</sup>, en levaduras. La fosforilación específica de mitosis está localizada en la secuencia SET del N-terminal, distinta de la secuencia consenso que fosforilan las CDKs, pero que al parecer también es fosforilada por las quinasas dependientes de ciclinas. Se fosforilan los residuos de serina y de treonina. La fosforilación de la secuencia SET es la última que ocurre en la H1 durante el ciclo celular. Se ha apuntado que dicha modificación podría tener un papel activo en la mitosis o en la modulación de otras proteínas o factores involucrados en la condensación de la cromatina<sup>226</sup>. De ser así, la modificación de estos dos sitios en otra etapa del ciclo celular podría ser responsable de la catástrofe mitótica observada en condiciones que propician la hiperfosforilación de la H1 durante la interfase<sup>223</sup>. La histona H1 es sustrato tanto de la CDK2 como de la cdc2. Estudios *in vitro* han mostrado que tanto la CDK2 como la cdc2 son capaces de fosforilar todos los sitios del tipo (S/T)-P-X-Z y el sitio específico de mitosis. Esto demuestra que el nivel de fosforilación durante el ciclo celular no se debe a la especificidad de la quinasa. Una posible hipótesis apunta a que la extensión de la fosforilación en cada etapa está asociada a la exposición de los sitios de fosforilación. Ello conlleva que la estructura de la H1 en la cromatina difiera durante las etapas del ciclo celular y que esté involucrada en la determinación de los sitios de fosforilación. No obstante, la accesibilidad de los sitios de fosforilación podría estar también relacionada con otros factores dependientes del ciclo celular, la modificación postraduccional de las histonas del core o de otras proteínas no histonas<sup>226</sup>.

Recientemente, Lindner y col.<sup>219</sup> han encontrado que la fosforilación durante la interfase solamente ocurre en los residuos de serina, mientras que la treonina se fosforila durante la mitosis. Un estudio más exhaustivo sobre la H1b mostró que la primera fosforilación era casi siempre en S17, ubicada en el N-terminal, seguida de S172 y finalizaba con la fosforilación de la S188. Durante la mitosis este subtipo llega a estar pentafosforilado,



supuestamente por la fosforilación de la treonina en los dos sitios consenso restantes. Sin embargo, las formas pentafoforiladas encontradas contenían fosfatos en T10, sitio de secuencia no consenso (Fig. 22) y uno de los dos sitios consenso restantes T137 y T154. La fosforilación de la T10 recuerda el caso de la fosforilación de S10 de la H3 discutido anteriormente, pero cualquier otro tipo de relación está por demostrar.

La histona H1b y su fosforilación podrían estar involucradas en el mantenimiento de la heterocromatina<sup>227</sup>. La proteína de heterocromatina 1 $\alpha$  (HP1 $\alpha$ ) se asocia de forma específica con la heterocromatina. Se piensa que el reclutamiento de HP1 $\alpha$  y el establecimiento de dominios silenciados está determinado por la interacción del dominio N-terminal de la HP1 $\alpha$  con la H3 metilada en la lisina 9 y de la región de bisagra con el DNA y la H1 no fosforilada. La fosforilación de la H1 provocaría una ruptura de la interacción, permitiendo la relajación necesaria de la cromatina para el proceso de replicación<sup>228</sup>.

En la H1 también encontramos fosforilación dependiente de cAMP<sup>10</sup> en sitios cercanos al dominio globular o en el inicio del dominio globular. Se ha encontrado en el hígado en regeneración, pero está ausente en la interfase y en células bloqueadas en mitosis<sup>226</sup>.

La fosforilación de la histona H1 está presente en organismos sencillos como *Tetrahymena thermophila*<sup>229, 230, 231</sup>. Este protozoo tiene dos núcleos, el macronúcleo que se divide de forma amitótica y el micronúcleo que se divide por mitosis.

En el micronúcleo encontramos la proteína quinasa A (PKA), que es dependiente de cAMP. En cada uno de los polipéptidos correspondientes a la H1:  $\alpha$ ,  $\beta$ ,  $\gamma$ , y  $\delta$ , encontramos dianas para la PKA y se han detectado dos sitios fosforilados *in vivo* en el polipéptido  $\delta$ , RKNS y KSSV. En ninguno de estos polipéptidos se encuentran dianas para la quinasa cdc2<sup>229</sup>.

El macronúcleo, además de dividirse de forma amitótica, es transcripcionalmente activo. La H1 del macronúcleo tiene composición aminoacídica parecida a la de metazoos, pero carece de dominio globular. También presenta solubilidad en ácidos y su gen es de copia única. La histona H1 se encuentra altamente fosforilada en la interfase al contrario de muchos otros organismos. El macronúcleo se divide sin condensación de los cromosomas, lo que sugiere que la fosforilación de la H1 esté relacionada solamente con la regulación transcripcional. En un fragmento de 20 aminoácidos se han encontrado 5 sitios de fosforilación, tres de ellos correspondientes a la diana consenso de la cdc2, dos sitios con secuencia TPVK y un motivo TPTK. Los dos sitios de fosforilación restantes, encontrados en la secuencia

AASASTTPVK, donde se pueden fosforilar las dos serinas, no guardan parecido con ninguna secuencia consenso de las quinasas conocidas y su fosforilación solamente ocurre cuando el sitio consenso adyacente ha sido fosforilado por cdc2. A pesar de la importancia que pudiera tener la fosforilación de la histona H1, la mutación de los 5 sitios no afecta la viabilidad de *Tetrahymena thermophila*<sup>230</sup>.

La fosforilación de la histona H1 de macronúcleo juega un papel importante en la regulación de la expresión génica durante la transición entre células en crecimiento y en condiciones limitantes de nutrientes<sup>231</sup> (Fig. 24).

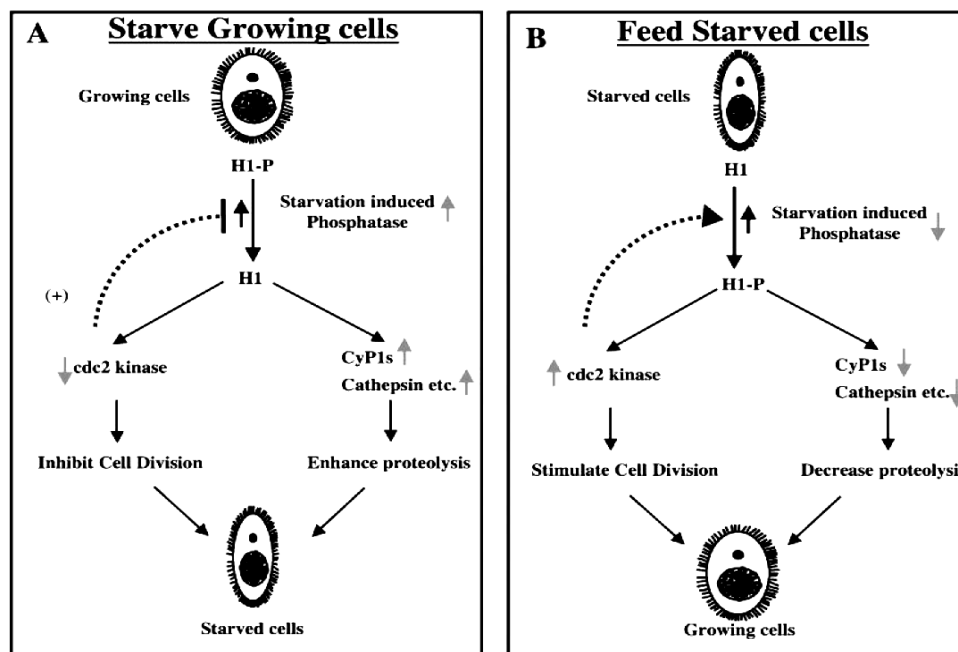


Figura 24. Papel de la fosforilación de la H1 en la transición entre células en crecimiento y células en condiciones limitantes de nutrientes<sup>231</sup>.

Cuando a células en crecimiento se les suprimen los nutrientes la respuesta inicial es la inducción de la actividad H1-fosfatasa que desfosforila la H1. La desfosforilación de la H1 tiene dos efectos. Primero, inhibe la expresión del gen de cdc2, que acelera la transición a condiciones de no crecimiento, lo que conlleva a la detención del ciclo celular. El segundo efecto consiste en la inducción de la expresión de genes específicos, entre los que se encuentran proteasas y otros genes necesarios para mantener la viabilidad y/o preparar la

conjugación (Fig. 24A). Cuando las células vuelven a disponer de nutrientes, una reducción de la actividad fosfatasa revertiría los efectos descritos (Fig. 24B).

### **La histona H1 en el núcleo celular.**

Las histonas H1 tienen gran variedad de funciones en la célula, entre las que se encuentran, la formación de la superestructura de la cromatina<sup>2,232</sup>, la estabilización de hasta dos vueltas de DNA nucleosomal alrededor del octámero por la interacción con el DNA en la entrada y la salida del nucleosoma<sup>233</sup>, la regulación transcripcional y la participación en otros procesos celulares. Sin embargo, su posición exacta en el cromosoma y su papel en la regulación transcripcional de la expresión génica, todavía no se han establecido.

Inicialmente se pensaba que las H1 eran represores generales de la transcripción<sup>91</sup>. Los estudios *in vitro* apuntaban a que H1 era un represor transcripcional y que algunos factores de transcripción actuaban como antirepresores. En otros estudios se ha descrito una acción más específica de la H1 sobre determinados genes que puede implicar tanto la represión como la activación de la transcripción, lo que sugiere un papel mucho más complejo para la H1.

La modificación de la estructura de orden superior de la cromatina ocasionada por la unión preferencial de la H1 a regiones SAR y su participación en el posicionamiento de los nucleosomas, podrían intervenir en la regulación de la transcripción<sup>234</sup>. La modulación de la expresión por la H1 puede deberse también a la modificación de la accesibilidad de secuencias promotoras o de secuencias potenciadoras a los factores de transcripción.

La estructura de la cromatina y su actividad transcripcional están estrechamente relacionadas. La compactación del DNA en la fibra de cromatina presenta una barrera para la maquinaria transcripcional. En muchos promotores o enhancers los sitios de unión para los factores de transcripción se encuentran ocultos en la estructura nucleosomal. A pesar de que algunos factores son capaces de unirse al DNA enrollado en los nucleosomas, no parece ser la situación más común<sup>235</sup>. Es posible que el nucleosoma se elimine para permitir la iniciación de la transcripción, pero en la mayoría de los casos el nucleosoma permanece en su sitio, aunque alterado, permitiendo el acceso de los factores de transcripción<sup>236</sup>.

Un mecanismo sencillo para permitir el acceso a la maquinaria de transcripción consiste en el desenrollamiento parcial del DNA alrededor del core nucleosomal. Este fenómeno parece

ser espontáneo en la cromatina sin H1<sup>237</sup>. Por tanto la unión de la H1 puede verse como una puerta que estaría cerrada cuando la proteína está unida y abierta a la invasión del nucleosoma por otros factores proteicos cuando la H1 es desplazada. Dependiendo de la naturaleza del factor proteico que se una podríamos tener la activación o la represión de la transcripción<sup>236</sup>.

Hay proteínas que se unen directamente a la H1 como es el caso de la nucleolina que induce la descondensación de la cromatina al unirse a la H1<sup>238</sup> o la protimosina  $\alpha$  que es una proteína ácida nuclear al parecer involucrada en la proliferación celular<sup>239</sup>.

La proteína MeCP2 es capaz de desplazar la H1 del DNA con motivos CpG metilados ocasionando una represión de la transcripción. Esta proteína, además, recluta una desacetilasa que desacetila las histonas internas, induciendo la condensación de la cromatina asociada a la metilación<sup>240, 241</sup>. También, como comentamos anteriormente, la sobreexpresión de la H1<sup>o</sup> en fibroblastos de ratón tiene un efecto represor en los genes de ciclofilina, GAPDH y actina, todos transcritos por la RNA polimerasa II<sup>91</sup>.

El desplazamiento de la H1 de sitios concretos en el DNA puede resultar en la activación transcripcional. El factor de transcripción HNF-3 de hígado desplaza la H1 y estimula la transcripción de la albúmina<sup>242</sup>. Lo mismo ocurre con la unión del UBF (Upstream binding factor) que desplaza la H1 de un potenciador de r-DNA, ocasionando la activación de la transcripción<sup>243</sup>. La sobreexpresión de la H1c en fibroblastos de ratón también aumenta los niveles de transcrito de los genes de ciclofilina, GAPDH y actina<sup>91</sup>.

Las modificaciones postraduccionales de las histonas internas y de la propia H1 están activamente implicadas en la modificación de la afinidad de la H1 por la cromatina y la regulación de la transcripción<sup>236</sup>.

Se ha mencionado anteriormente que las regiones con las histonas internas acetiladas son pobres en H1. Algunos experimentos sugieren que la modificación de las colas de las histonas internas podría afectar la afinidad de la H1 facilitando la salida de la H1 adyacente al nucleosoma modificado<sup>244</sup>. La importancia de la acetilación en la regulación de la transcripción se evidencia en la presencia de acetilasas y desacetilasas asociadas a la maquinaria transcripcional<sup>245, 246</sup>.

En cambio, la poliADP-ribosilación de la H1 parece estar involucrada en mantener sin metilar las islas CpG en los promotores de los genes transcritos constitutivamente (housekeeping genes)<sup>208, 209</sup>.

La fosforilación de la H1 se ha relacionado directamente con la expresión. Solamente la H1 fosforilada puede ser desplazada del promotor MMTV (mouse mammary tumor virus) por la acción del receptor de glucocorticoides. La exposición prolongada a glucocorticoides provoca la desfosforilación de la H1 y lleva el promotor en cuestión a un estado refractario<sup>247</sup>.

El promotor MMTV es un sistema conveniente para estudiar los cambios de expresión asociados a la remodelación de la cromatina porque el posicionamiento de los nucleosomas está estudiado y además es modulable mediante la adición de hormonas<sup>248</sup>. El MMTV, está organizado en 6 nucleosomas regularmente posicionados. En la región ocupada por el segundo nucleosoma se encuentra el elemento de respuesta a hormonas (HRE) y sitios de unión para el factor de transcripción NF1 y el factor de transcripción de octámero 1 (octamer transcription factor, OTF1)<sup>249</sup>.

Un posible modelo para la inducción del MMTV sugiere que la unión de la hormona al HRE permite el reclutamiento de la CDK2 que fosforila la H1, lo que origina su salida de la región del promotor. La pérdida de la H1, permitiría la entrada del complejo de remodelación de la cromatina hSWI/SNF y de los factores de transcripción<sup>250</sup>.

La regulación de la actividad del promotor MMTV, al menos en respuesta a la dexametasona en células de ratón, parece estar relacionada con un subtipo concreto de histona H1, la H1d<sup>248</sup>.

A pesar del efecto inductor de la expresión asociado a la fosforilación de la H1, durante la mitosis, la acción coordinada de los altos niveles de fosforilación de las histonas H1 y H3 promueve la formación de cromatina altamente condensada donde no es posible la transcripción del MMTV<sup>251</sup>.

Por tanto, a partir de los datos anteriores se infiere la posición de la H1 en una red complicada de interacciones moleculares que modifican la expresión génica (Fig. 25).

Las distintas modificaciones postraduccionales de las histonas internas actúan de forma secuencial o en combinación para regular la estructura de la cromatina, activando o reprimiendo la transcripción. Este conjunto de señales se conoce como el código de las

histonas. El papel de las modificaciones postraduccionales en la H1 y su acción coordinada con las histonas internas hace pensar que la H1 también podría incluirse en este mecanismo de regulación epigenética.

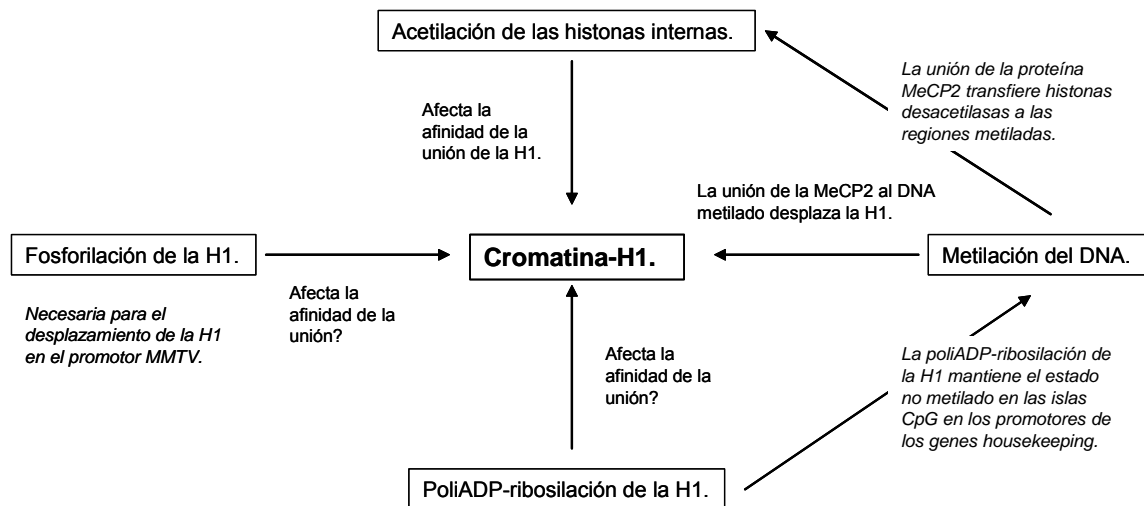


Figura 25. Interacciones que afectan la unión de la H1 a la cromatina. Adaptada de Zlatanova y col<sup>236</sup>.

### Otros efectos de la fosforilación de la histona H1.

La fosforilación de la histona H1 también se ha relacionado con algunos procesos nucleares como la remodelación de la cromatina dependiente de ATP<sup>252</sup>, la ligación<sup>253</sup>, la diferenciación celular<sup>254</sup> y, posiblemente, la apoptosis<sup>255</sup>.

La H1 fosforilada afecta la actividad de los complejos de remodelación de la cromatina dependientes de ATP. Estos complejos juegan un papel fundamental en la regulación de la transcripción, el desarrollo, la reparación del DNA y la regulación del ciclo celular. La incorporación de la histona H1 a un polinucleosoma formado por 11 repeticiones de la secuencia posicionante del rDNA 5S, ocasiona la inhibición de varios complejos de remodelación de la cromatina dependientes de ATP: ySWI/SNF, hSWI/SNF, xMi-2 y xACF. La fosforilación de la histona H1 por cdc2/ciclina B restaura la actividad de los complejos de remodelación, sugiriendo un control global de la histona H1 sobre la remodelación de la cromatina y la necesidad del acoplamiento entre actividad de las quinasas de H1 y los enzimas de remodelado dependientes de ATP<sup>252</sup>. Esta acción coordinada ya se ha comentado anteriormente en el MMTV<sup>250</sup>.

La histona H1 fosforilada participa en la ligación de extremos de doble cadena no homólogos. Este tipo de ligación *in vivo* requiere de la presencia de DNA-PK (proteína quinasa dependiente de DNA) y LX (DNA ligasa). La actividad de LX se reduce sustancialmente *in vitro* por la adición de la histona H1. Cuando se añade la DNA-PK, que cataliza la fosforilación de la H1, la actividad de LX se recupera. Esto sugiere que *in vivo* la DNA-PK pudiera fosforilar la histona H1 en la vecindad de los sitios de corte en el DNA, ocasionando la salida de la H1 y permitiendo la ligación de los extremos de doble cadena por LX<sup>253</sup>.

La diferenciación terminal de la mayoría de los tipos celulares está asociada a cambios estructurales significativos en la conformación de la cromatina, aunque no está claro si esos cambios son la causa o el efecto de los múltiples procesos que contribuyen al establecimiento de un fenotipo diferenciado. La desfosforilación de la histona H1 se puede relacionar con la diferenciación celular<sup>254</sup>.

Las células de leucemia eritrocítica de ratón (MEL) constituyen un sistema modelo para el estudio de la diferenciación celular terminal<sup>256</sup>. En las células MEL, al avanzar el proceso de diferenciación hay una pérdida progresiva de la actividad CDK, que conlleva la acumulación de H1 desfosforilada y la restricción de la flexibilidad de la cromatina.

La expresión de la H1<sup>o</sup> cambia considerablemente durante la diferenciación en las células MEL. Unido a ello hay un cambio en sus propiedades de unión a la cromatina entre las 48-60 h después de la inducción de la diferenciación<sup>254</sup>. Este cambio ocurre una vez las células están comprometidas en el proceso de diferenciación, pero antes de la parada del ciclo celular en G<sub>0</sub> o la expresión abierta de los genes específicos de eritrocito, lo que sugiere que este proceso podría contribuir a la finalización del proceso de diferenciación. Esta conclusión se apoya en que tanto el bloqueo de la pérdida de la actividad CDK<sup>257, 258</sup>, como la adición de pequeñas cantidades de H1<sup>o</sup> fosforilada bloquean la diferenciación<sup>254</sup>, lo que implica un efecto dominante de la H1<sup>o</sup> fosforilada en relación con la no fosforilada. Al parecer la fosforilación reduce la afinidad de la H1 por la cromatina, pero no su liberación inmediata. Por lo tanto, el efecto de la fosforilación podría estar relacionado con cambios estructurales globales de la cromatina, mientras la H1 permanece unida a ella<sup>254</sup>.

Se ha encontrado que durante las fases S y M del ciclo celular, la histona H1 fosforilada es translocada al citoplasma<sup>259</sup>. En enfermedades neurodegenerativas como el Alzheimer y el

Parkinson se ha encontrado una activación de la CDK5, una quinasa dependiente de ciclinas, parecida a la cdc2. Esta proteína es capaz de fosforilar la histona H1 en el cerebro<sup>260</sup>. El aumento de la fosforilación de la H1 mediado por la CDK5 podría explicar el incremento de la H1 citoplasmática que se ha encontrado en neuronas y astrocitos afectados por Alzheimer y Parkinson<sup>261</sup>.

La histona H1 citoplasmática interacciona preferentemente con las proteínas  $\alpha\beta$ -amiloide y con la  $\alpha$ -synucleína en forma amiloide. Estas proteínas son las principales responsables del Alzheimer y el Parkinson respectivamente. La H1 se ha encontrado asociada *in vivo* a las placas amiloides y a las fibras amiloides intracelulares. La fosforilación anormal de la H1 y el cambio en el patrón de localización intracelular podrían explicar la entrada en ciclo celular de las neuronas en enfermos de Alzheimer y Parkinson, así como la activación de la apoptosis. La inhibición de la interacción entre la H1 y las proteínas amiloides podría impedir estos eventos y por tanto podría constituir una diana terapéutica en el tratamiento de estas enfermedades neurodegenerativas<sup>261</sup>.

### **La histona H1 y el sistema inmune.**

A pesar del papel esencial en el núcleo, la histona H1 también tiene funciones fuera de la envoltura nuclear. En la última década se han acumulado evidencias de este hecho, que sitúa a las histonas como componentes del sistema inmune innato y en el punto de mira de terapias contra el cáncer, infección microbiana o enfermedades autoinmunes<sup>262</sup> (Fig 26).

La H1c juega un papel esencial en la inducción de la apoptosis al migrar del núcleo a la mitocondria. Esta migración es dependiente de p53 y forma parte del complejo Bak que promueve la liberación de citocromo c. Hay otros dos subtipos capaces de mediar esta función, la H1a y la H1b, aunque de forma menos efectiva<sup>263, 262</sup>. Una característica importante es que al parecer las histonas H1 liberadas durante la apoptosis tienen reducido su grado de fosforilación a mono fosforiladas o no fosforiladas<sup>195</sup>.

Durante la apoptosis aumenta la exposición en la superficie celular de histonas e incluso nucleosomas. La liberación de contenido nuclear da pie a la hipótesis de algunos autores que sostienen que el fallo de la limpieza de los restos de las células muertas o la ruptura de las estructuras expuestas en la superficie celular pueden llevar al desarrollo de enfermedades autoinmunes. Algunos resultados sustentan esta hipótesis, como es el hecho de que los anticuerpos antinucleares aparezcan junto con los anticuerpos antifosfolípidos, años antes



del desarrollo del lupus eritematoso (SLE) y mucho antes de la aparición de los anticuerpos anti-DNA<sup>262</sup>.

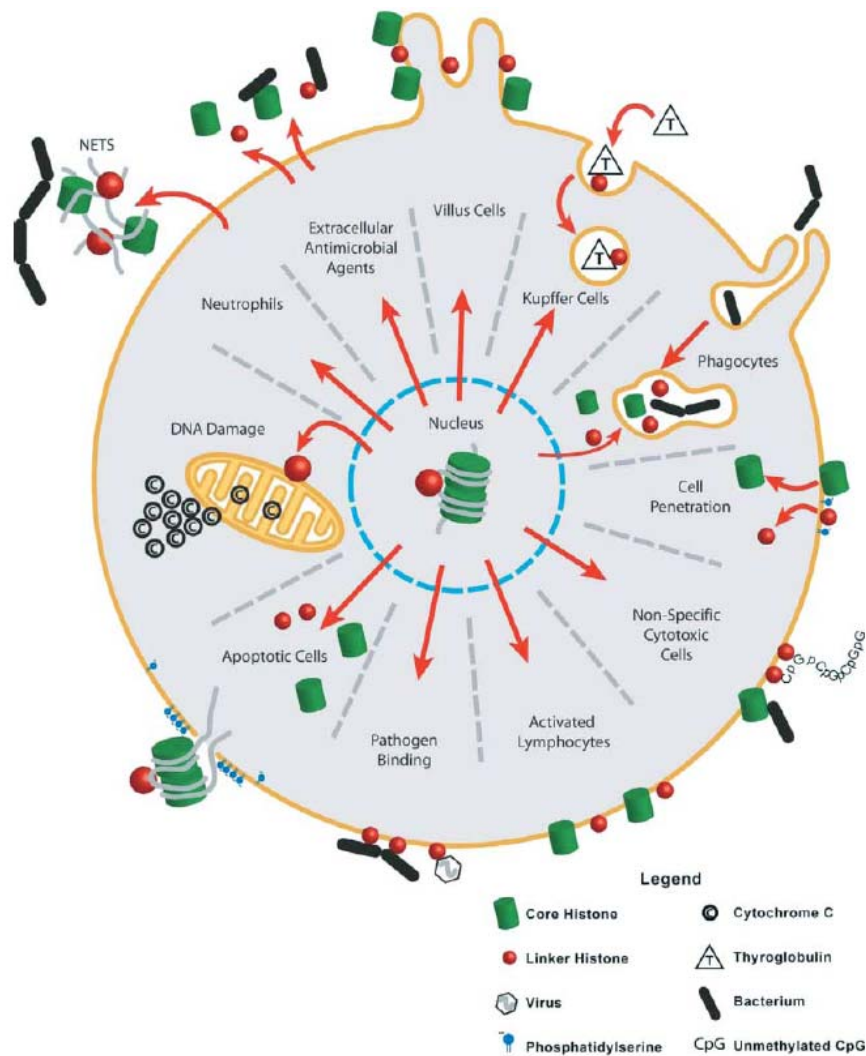


Figura 26. Papel de la histona H1 en el sistema inmune<sup>262</sup>.

En los pacientes de SLE es muy frecuente la presencia de anticuerpos contra el C-terminal de la histona H1<sup>264</sup>. Estos anticuerpos son considerados un marcador altamente específico del avance de la enfermedad<sup>265</sup>. En un estudio realizado con 9 pacientes de SLE se encontró que los subtipos de la H1 no eran reconocidos todos por igual. De los 9 sueros analizados, todos reconocían la H1c, principal subtipo involucrado en la inducción de la apoptosis, y 7 reconocían la H1b<sup>266</sup>.

El desarrollo de anticuerpos contra el C-terminal de la histona H1 también se observa en aproximadamente el 90% de los pacientes infectados por el HIV<sup>267</sup>. Es de destacar que al menos 3 de los anticuerpos neutralizantes del HIV tienen reactividad cruzada con la histona H1. Esto podría explicar por qué los pacientes con SLE casi nunca contraen el HIV<sup>268</sup>. Los anticuerpos contra la H1 que se encuentran en los pacientes pueden ser aquellos que reconocen un epitopo que no es crítico para la supervivencia o que tal vez no está ni siquiera presente en el virus. Estudios con una línea de células T-CD4 resistentes al HIV demostraron que al aislar el factor de resistencia, éste coelúa con la H1b monoubiquitinada. El silenciamiento de este subtipo con RNA de interferencia conlleva la pérdida de la resistencia al virus<sup>269</sup>.

Otro ligando de la histona H1 es la tiroglobulina. Esta proteína es secretada por las células epiteliales de la tiroides y se elimina del organismo mediante la acción de los macrófagos hepáticos (Kupffer cells). En las vesículas endocíticas del macrófago se pudo observar la colocalización de la tiroglobulina y la H1, lo que sugiere la implicación de esta última en la internalización de la tiroglobulina *in vivo*<sup>270</sup>.

Una de las funciones efectoras de los macrófagos es la fagocitosis de microbios y su destrucción mediante la liberación intracelular de agentes antimicrobianos almacenados en gránulos. Al identificar tres de estas proteínas antimicrobianas de ratón (MUMP) se encontró que dos de ellas eran H1. Estas proteínas mostraban actividad antimicrobiana contra gran variedad de microorganismos, incluyendo *Salmonella typhimurium*, *Escherichia coli* y *Staphylococcus aureus*<sup>271</sup>.

La presencia de la histona H1 en las células del sistema inmune no se limita a los macrófagos. Las células NK (Natural Killer Cells) reconocen y eliminan bacterias, protozoos, células transformadas por virus e incluso células tumorales. Al identificar una de las proteínas antimicrobianas típicas de este tipo celular en peces se encontró que era la H1x, cuya función hasta ahora no ha sido determinada. También se ha confirmado la presencia de H1x en macrófagos de ratón<sup>272</sup>. La histona H1 también está presente en las redes de DNA e histona liberadas por los neutrófilos, que actúan como barrera física para atrapar patógenos bacterianos y que además tienen actividad bactericida<sup>273</sup>.

La actividad antimicrobiana de la histona H1 se manifiesta en el intestino delgado de mamíferos<sup>274, 275</sup>; en hígado, intestino, estómago y piel del salmón atlántico<sup>276</sup>; en bazo,

agallas y piel de la trucha y la perca<sup>277</sup>. La acción antimicrobiana puede estar mediada por la molécula entera de H1 o por fragmentos discretos y es eficaz contra bacterias Gram(+) y Gram(-); también actúa contra *Saccharomyces cerevisiae* y parásitos como el *Amyloodinium ocellatum* e incluso agentes virales como el *Norovirus*<sup>262</sup>.

La implicación de la histona H1 en el sistema inmune ha traído como consecuencia su estudio como agente terapéutico<sup>262</sup>. La histona H1 se ha empleado como antígeno en una vacuna contra *Leishmania*<sup>278</sup>. Se utilizó un fragmento de 60 aminoácidos de la H1, obteniéndose una protección parcial contra la enfermedad y una reducción importante de la severidad de las lesiones cutáneas<sup>279</sup>. También es la molécula central de la terapia de la zona necrótica del tumor (Tumor Necrosis Therapy). Esta terapia emplea anticuerpos anti-H1 acoplados a alguna molécula efectora como el I<sup>131</sup>. Los anticuerpos son así dirigidos a la zona necrótica del tumor sólido y la molécula efectora elimina las células adyacentes, atacando el tumor de dentro hacia fuera. Actualmente esta terapia se emplea en el tratamiento de tumores cerebrales<sup>280</sup> y de pulmón<sup>281</sup>.

Por último, los anticuerpos contra la histona H1 son candidatos a ser utilizados en el tratamiento para evitar el rechazo de trasplantes<sup>282</sup>. Los anticuerpos anti-H1 son el componente principal de los anticuerpos autoreactivos que aparecen de forma transitoria después de un trasplante, a las dos o tres semanas, y que contribuyen a la inhibición de la proliferación de las células T que causan el rechazo.



## **Objetivos.**

La histona H1 tiene un papel esencial en la estructura de la cromatina y participa en la regulación génica. La H1 está formada por tres dominios: el dominio N-terminal con 20-35 residuos, el dominio globular central con 80 residuos aproximadamente y el dominio C-terminal que contiene alrededor de 100 residuos.

El dominio C-terminal es muy rico en lisina, prolina y alanina, y es una región muy variable entre los diferentes subtipos de la histona H1. Su función se relaciona con la condensación de la cromatina. En este dominio también se ubican la mayoría de los sitios de fosforilación, que constituye la modificación postraduccional más importante en la H1. En este trabajo, hemos estudiado la preferencia por el DNA SAR y las propiedades estructurales del dominio C-terminal en disolución y unido al DNA, así como los efectos estructurales de la fosforilación de los sitios CDK2. Los objetivos concretos han consistido en:

1. Estudiar la contribución de los dominios de la histona H1 a la preferencia por el DNA SAR.
2. Determinar la estructura secundaria del dominio C-terminal de la histona H1 por infrarrojo y dicroísmo circular en disolución acuosa, en presencia de TFE y en los complejos con DNA.
3. Determinar el efecto de la aglomeración (crowding) macromolecular en la estructura del dominio C-terminal de la histona H1.
4. Determinar el efecto de la fosforilación de los motivos (S/T)-P-X-(K/R) en la estructura secundaria del dominio C-terminal.



## **Capítulo 1**

*La preferencia de la histona H1 por el DNA SAR está determinada por el dominio C-terminal.*





## Introduction

H1 linker histones are thought to be primarily responsible for the condensation of the nucleosome chain in the thick chromatin fibre. It is currently accepted that histone H1 could have a regulatory role in transcription through the modulation of chromatin higher-order structure. H1 has been described as a general transcriptional repressor because it contributes to chromatin condensation, which limits the access of the transcriptional machinery to DNA. Other studies indicate that H1 may regulate transcription at a more specific level, participating in complexes that either activate or repress specific genes<sup>1, 89, 222, 247, 283, 284, 285, 286</sup>. Preferential binding to SARs (for scaffold-associated regions; also termed matrix-associated regions, MARs)<sup>130</sup> and participation in nucleosome positioning<sup>287</sup> are other mechanisms by which H1 could contribute to transcriptional regulation.

H1 has multiple isoforms. In mammals, six somatic subtypes, designated H1a-e and H1<sup>o</sup>, a male germ line-specific subtype, H1t, and an oocyte specific subtype, H1oo, have been identified<sup>80, 93, 288, 289</sup>. The subtypes differ in their timing of expression<sup>79</sup>, extent of phosphorylation<sup>290</sup> and turnover rate<sup>97, 291</sup>. Analysis of the evolution of vertebrate H1 subtypes has shown that amino acid substitution rates differ among subtypes by almost one order of magnitude, suggesting that each subtype has acquired a unique function<sup>78</sup>. Differences in DNA condensing capacity have also been demonstrated for some subtypes<sup>292, 293, 294</sup>.

Histone H1 has been characterized as an SAR binding protein<sup>130</sup>. SARs were identified as DNA sequences rich in AT base pairs (>70%) and in homopolymeric oligo(dA).oligo(dT) tracts (A-tracts)<sup>138</sup> that were specifically bound by the nuclear and the metaphase scaffold. SARs have been proposed as DNA elements that would define the bases of chromatin loops in eukaryotic cells<sup>138</sup>, and that could also be involved in chromosome dynamics<sup>139</sup>. In addition to H1, several SAR binding proteins have been identified, including topoisomerase II, lamin B1, nucleoline, HMG I/Y, SAT B1, SAF-A and SAF-B<sup>141, 142, 143, 295, 296, 297</sup>. It has been proposed that regulated H1 dissociation or assembly with SARs is implicated in the regional opening or closing of chromatin loops and, consequently, contributes to transcriptional regulation<sup>130</sup>. The high affinity cooperative binding of H1 and other SAR-binding proteins to SARs is determined by the presence of the A-tracts, rather than a precise

base sequence<sup>141, 142, 143, 151, 295, 296, 297</sup>. A-tracts have peculiar structural properties, including a narrower minor groove<sup>152, 153</sup>. The importance of the conformational features of the DNA in H1 binding is stressed by the preferential affinity of H1 for supercoiled DNA, cruciforms and DNA fragments with intrinsic curvature<sup>298, 299, 300</sup>.

H1 linker histones present a tripartite structure consisting of a central globular domain flanked by highly basic amino- and carboxy-terminal tail-like domains<sup>301</sup>. The N- and C-terminal domains are very different in length, the C-terminal tail comprising almost 50 % of the protein. The distribution of charge in the C-terminal domain is extremely uniform in spite of the variation in sequence of the different H1 subtypes<sup>175</sup>. The structure of the globular domain has been described by X-ray diffraction<sup>165</sup> and nuclear magnetic resonance<sup>302, 303</sup>. It contains a three-helix bundle, which resembles the winged-helix motif found in some sequence-specific DNA-binding proteins, and a C-terminal  $\beta$ -sheet. The N- and C-terminal domains are largely unstructured in solution. However, they acquire a substantial amount of secondary structure upon interaction with DNA. Helix and turn elements, inducible upon interaction with the DNA, have been described in both the N- and C-terminal domains<sup>158, 173, 174, 186</sup>. The distinct structure of the H1 domains suggests that they could play specific functions in chromatin structure. The N-terminal domain could contribute to the binding stability of the globular domain; the globular domain is very likely to localise the molecule in the nucleosome; while the C-terminal domain is the main region of the molecule involved in chromatin condensation through binding and neutralization of the charge of the linker DNA.

The tripartite structure of histone H1 raises the question of the involvement of the individual H1 domains in the preferential binding of the protein to SARs. Here, we identify the C-terminal domain as the main determinant of the SAR binding properties of histone H1. We also show that salmine, a typical protamine that could have its evolutionary origin in histone H1<sup>304</sup>, also shows preferential binding to SARs. The possible significance of these findings in relation to H1 and SAR functions is discussed.

---

## **Materials and methods**

### Peptides and protamine

The peptides Ac-EKTPVKKKARKAAGGAKRKTSG-NH<sub>2</sub> (NE-1) and Ac-TENSTSAPAAKPKRAKASKK-NH<sub>2</sub> (NH-1) were synthesized by standard methods (NE-1 by Neosystem Laboratoire, Strasbourg, France, and NH-1 by DiverDrugs, Barcelona, Spain. Peptide homogeneity was determined by HPLC on Kromasil C8. The peptide composition was confirmed by amino acid analysis and the molecular mass was checked by mass spectrometry. The sequence of NE-1 corresponds to residues 15 to 36 at the N-terminus of mouse H1e and that of NH-1 to residues 1 to 20 of mouse H1°.

Protamine (salmine) was from salmon (Sigma P4005).

### Separation of histone H1 subtypes

Nuclei were isolated from the brain of adult mice<sup>305, 306</sup>. Histone H1 was extracted with 0.35M NaCl, following the method of García-Ramírez et al.<sup>307</sup>. The mixture of subtypes was digested with alkaline phosphatase to eliminate small amounts of phosphorylated forms that could be present. H1° was purified by gel-filtration chromatography, according to Böhm et al.<sup>308</sup>. The subtypes H1a-e were separated by reverse phase HPLC according to Brown et al.<sup>309</sup>. All subtypes were obtained as homogeneous peaks, except H1d and H1e, which largely overlapped. The latter subtypes were separated by acetic acid/urea gel electrophoresis<sup>99</sup> and recovered by electroelution using a Biotrap camera (Schleider & Schuell). H1t was purified from mouse testes according to Khadake et al.<sup>293</sup>. Before being used in binding experiments, H1 subtypes were subjected to a cycle of denaturation/renaturation by stepwise dialysis from 6 M urea into, successively, 3.0 M, 1.5 M, 0.7 M, 0.3 M and 0.0 M urea, in 0.2 M NaCl, 0.01 M phosphate buffer, pH 7. Finally, the proteins were dialysed against 0.14 M NaCl, 0.01 M phosphate buffer, pH 7.0. The concentration of protein was estimated by amino acid analysis.

### Cloning, expression and purification of the globular and C-terminal domains of H1 histone subtypes

The sequences encoding the globular domains of histones H1° and H5 and the C-terminal domains of histones H1e, H1° and H1t were cloned and expressed. All gene fragments were

amplified from mouse genomic DNA by PCR. The primers were 5'GGCCGCCCATATGTCCACGGACCACCCCAAG 3' and 5'CTTGGATCCCTACGACCTCTTGGGCTC 3' for the globular domain of H1<sup>o</sup>; 5'GGCATCGCATATGTCGGCATCGCACCCACC 3' and 5'GCCGGATCCTTAGGACCTCTTGGCC 3' for the globular domain of H5; 5'CCACCATGGATGAGCCTAAAAGGTC 3' and 5'GGAGATCTCTTCTTCTTGCTGGCCCTCT 3' for the C-terminal domain of H1<sup>o</sup>; 5'AAACCATGGCTGCTTCCGGTGAGGCTAA 3' and 5'ACAGATCTCTTTTCTTGGCTGCGGTTTT 3' for the C-terminal domain of H1e; and 5'GTACCATGGCGGCTTCAGGGAACGAC 3' and 5'ACGGATCCCTCCTCCCTGCTGCCTTCT 3' for the C-terminal domain of H1t. The amplification products of H1<sup>o</sup> and H5 globular domains were cloned in the pET11b vector (Novagen), using NdeI and BamHI restriction sites to yield the expression vectors pGH1<sup>o</sup> and pGH5, respectively. The C-terminal domains were cloned in the pQE-60 vector (Qiagen) using the NcoI and BglII restriction sites to yield the expression vectors pCTH1<sup>o</sup>, pCTH1e and pCTH1t.

The recombinant plasmids pGH1<sup>o</sup> and pGH5 were transformed into *E. coli* BL21(DE3). Cells were grown to an OD<sub>600</sub> of 0.8 and then induced with 1 mM IPTG, allowing expression to proceed for 4 hours at 37°C. Cells were then harvested and stored at -80 °C. The protein was purified according to the protocol described for the globular domain of chicken histone H5<sup>168</sup>.

The expression vectors pCTH1<sup>o</sup>, pCTH1e and pCTH1t were transformed into *E. coli* M15 (Qiagen). Cells were grown and induced as previously described for pGH1<sup>o</sup> and pGH5. Cells were lysed in lysis buffer (0.05 M NaH<sub>2</sub>PO<sub>4</sub>, 0.75 M NaCl, 0.02 M imidazol) plus 4 M guanidine hydrochloride, pH 8.0 for 15 min at room temperature. Guanidine hydrochloride was found to be necessary to avoid degradation and aggregation of the expressed protein. The extract was centrifuged at 20,000 × g for 25 minutes. The supernatants were loaded on a HiTrap chelating HP column (Amersham Biosciences) equilibrated with lysis buffer. The column was then washed in three steps with lysis buffer containing increasing amounts of imidazol: 40, 60 and 80 mM. Finally, the proteins were eluted with 250 mM imidazol in lysis buffer and desalted by gel filtration through Sephadex G-25 (Amersham Biosciences).

## Preparation of DNA fragments

An SAR fragment of 657 bp from the histone cluster of *Drosophila melanogaster* was obtained by digestion of p1314<sup>130</sup> with KpnI and BamHI. Another DNA fragment, of 587 bp, was excised from pUC19 by digestion with HaeIII. Both inserts were separated on agarose gels and electroeluted using a Biotrap camera (Schleider & Schuell). The 587 bp pUC19 fragment was extended from both ends to obtain a fragment of 763 bp. The long fragment was obtained by PCR on pUC19 using the primers 5'GCGGTTAGCTCCTTCGGTCCTC 3' and 5' CACCCGCTGACGCGCCCTGACG 3'. An AT-rich sequence (75% AT) was prepared by polymerisation of the 5'-phosphorylated oligomer 5'CTATGATATATAGATAGTTAATGTAATATGATATAGATATAGGGATCC 3', annealed with a complementary sequence that left five overhanging nucleotides. The annealed DNA was ligated overnight at 16°C with T4-ligase (Roche). The products of the ligation reaction were separated on agarose gels. Fragments ranging in size from about 500 bp to 5000 bp were electroeluted using a Biotrap camera (Schleider & Schuell).

## Binding assays

Binding experiments were performed by mixing equivalent amounts of SAR and pUC19 fragments with the proteins at different ratios. Binding conditions were 0.01 M phosphate buffer, pH 7.0, 5% glycerol and either 0.07 or 0.14 M NaCl. After 1 h of incubation at 37 °C, the mixtures were centrifuged at 14,000 × g for 10 min. The pellets and the supernatants were digested with proteinase K at 37°C overnight and the DNA was purified by phenol-chloroform extraction. The proportions of the SAR and pUC19 fragments in the complex and the supernatant were analysed by electrophoresis on 2% agarose (Metaphor) gels.

In some experiments, histone H1 C-terminal domain and protamine binding was performed in the presence of the DNA-binding drug distamycin A (Sigma). The DNA was preincubated with the drug for 30 min at 37°C before adding the proteins. The mixture was then incubated for 90 min at 37°C, as described.

## Results

### All histone subtypes are SAR binding proteins

The specific interaction of H1 with SARs was originally established for the SAR of the *Drosophila* histone-gene cluster, using total H1 from rat liver and DNA fragments from pBR322 as competitors<sup>130</sup>. The H1 complement from rat liver is dominated by the subtypes H1e and H1c, while H1a, b, d and H1° are either present in low amounts or absent. In order to determine whether the six mammalian somatic subtypes, H1a-e and H1° and the more divergent male germ line-specific subtype, H1t, are all SAR-binding proteins, we performed binding experiments with the minimal 657 bp SAR (75% AT) derived from the *Drosophila* histone-gene cluster and each of the purified subtypes. An HaeIII/HaeIII fragment of similar length from pUC19 (587 bp, 56% AT) was used as a competitor. The experiments consisted of mixing approximately equal amounts of the SAR and the pUC19 fragments with a limited amount of one of the subtypes. The experiments were performed at physiological salt concentration (0.14 M NaCl), where at subsaturating protein concentrations fast-sedimenting fully-complexed DNA molecules co-exist with free DNA<sup>310</sup>. Analysis of the DNA in the soluble and insoluble fractions showed that under conditions of limited protein, H1 bound first to the SAR whilst the pUC19 fragment remained in the supernatant (Fig. 27, lanes 2-15).

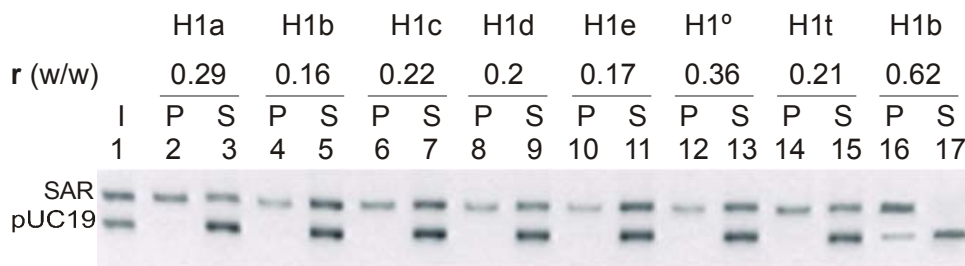


Figure 27. Preferential binding of histone H1 subtypes to the *Drosophila* SAR. The subtypes H1a-e, H1° and H1t (lanes 2-15) show preferential binding to the SAR fragment compared with the pUC19 fragment. Once the SAR fragment is saturated, H1 binds to pUC19 (lanes 16 and 17). The subtype is indicated above. The protein/DNA ratio (w/w) is indicated. I, input mixture of the SAR and pUC19 fragments; P, pellet; S, supernatant.

The pUC19 appeared in the pellet only when the SAR was saturated, and was thus absent from the supernatants, which occurred, starting with a 1:1 mixture of SAR and pUC19, at a protein/(SAR+pUC19) weight ratio higher than 0.5 (Fig. 27, lanes 16 and 17). Above a

weight ratio of 1.0, all DNA was saturated and further added protein was found as free protein in the supernatants. As shown in Figure 27, all subtypes, including H1<sup>o</sup> and H1t, showed strong cooperative binding to the SAR fragment.

The preferential binding of histone H1 is determined by its C-terminal domain

To establish the contribution of the C-terminal domain to the SAR binding character of the entire H1 molecule, we performed experiments similar to those described for the entire H1 molecule, using the purified recombinant C-terminal domains of subtypes H1e, H1<sup>o</sup> and H1t in physiological salt. As with the entire molecule, in competition experiments the C-terminal domains bound first to the SAR fragment, and only when it was saturated did they bind to the pUC19 fragment, which occurred, starting with a 1:1 mixture of SAR and pUC19, at a protein/(SAR+pUC19) weight ratio of about 0.3 (Fig. 28).

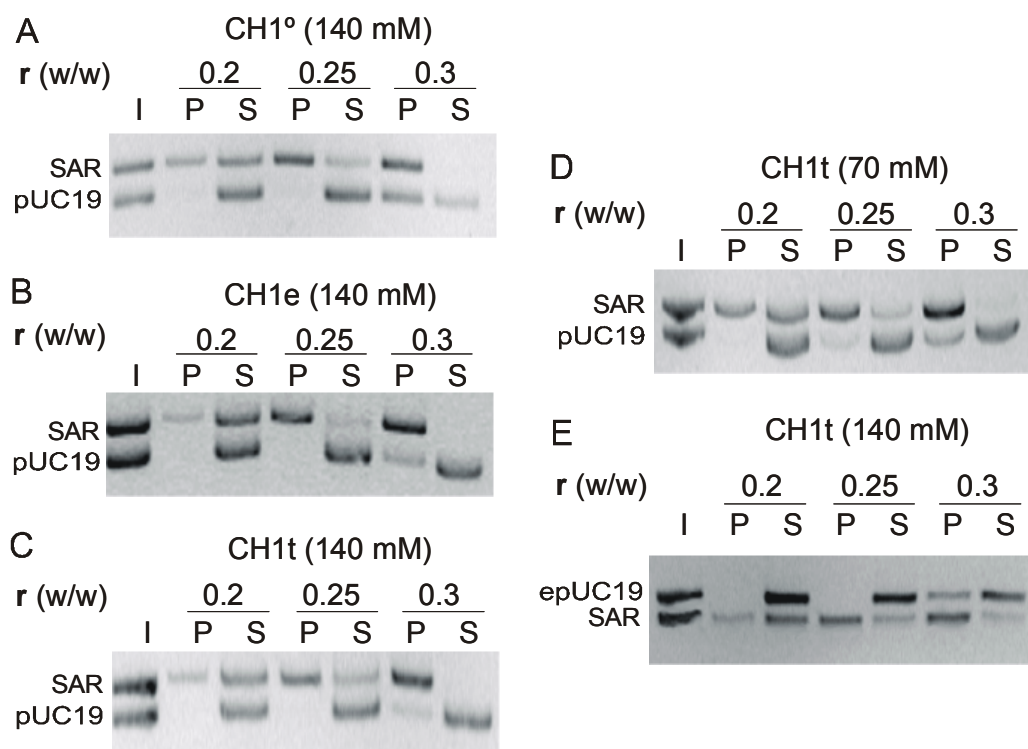


Figure 28. Preferential binding of the histone H1 C-terminal domain to the SAR fragment. (A) C-terminal domain of H1<sup>o</sup> (CH1<sup>o</sup>); (B) C-terminal domain of H1e (CH1e); (C), (D) and (E) C-terminal domain of H1t (CH1t). DNA fragments are indicated: SAR (SAR fragment of 657 bp), pUC19 (pUC19 fragment of 587 bp), epUC19 (extended pUC19 fragment of 763 bp). The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the fragments; P, pellet; S, supernatant.

The preferential binding of the C-terminal domain to the SAR was maintained in 70 mM NaCl, although under these conditions the binding was slightly less cooperative, as indicated by the presence of a faint pUC19 band in the pellet, coexisting with a similarly faint band of SAR in the supernatant (Fig. 28D). In spite of the large sequence divergence between the C-terminal domains of H1e, H1<sup>o</sup> and H1t, all showed the same strong preference for the SAR sequences (Fig. 29).

**A**

```

H1a  KKA---ESKAITTKVSVKAKA---SGA-AKKPKKTAGAAA-KKTV-KTPK
H1b  KKAASGEAKPKA-KKTGAAKAKKPAGATPKPKKKTAGA---KKTVKKTPK
H1c  KKAASGEAKPQA-KKAGAAKAKKPAGA-AKKPKKATGAATPKKAAKKTPK
H1d  KKAASGEAKPKA-KKAGAAKAKKPAGA-AKKPKKATGAATPKKTAKKTPK
H1e  KKAASGEAKPKA-KRAGAAKAKKPAGA-AKKPKKAAGTATAKKSTKKTTPK
H1o  KGDEPKRSVAF--KKTKEVKKVATPKKAAKPKKAASKAPSKKPKATPVK
H1t  KKAASGNDKGGK-KKSASAKAKK-MGLP-----RASR---SPKSSK--TK
      *                *                               *      *

```

```

H1a  KPKK-PAV---SKKTSKSPKKPKVV-KAKKVAKSPAKAKAVKPKASKAKV
H1b  KAKK-PAAAG-VKKVAKSPKKAKAAAKPKKAAKSPAKPKAVKSKASKPKV
H1c  KAKK-PAAAAVTKKVAKSPKKAKVT-KPKKV-KSAS--KAVKPKAAKPKV
H1d  KAKK-PAAAAGAKKVS KSPKKVKAA-KPKKAAKSPAKAKAPKAKASKPKA
H1e  KAKK-PAAAAGAKK-AKSPKKAKAT-KAKKAPKSPAKAKTVKPKAAKPKT
H1o  KAKKKPAATP--KK-AKKPKVVKV--KPKVASKPK-KAKTVKPKAKSSAK
H1t  AVKK-PKATP--TKASGSRRKTK-GAKGVQQRKSPAKARAANP---NSGK
      ** *          *          *      *      *

```

```

H1a  TKPKTPAKP-----KKAAPKKK
H1b  TKPKT-AKPKA AKKAVSKKK
H1c  AKAK-----KVAAKKK
H1d  SKPKAT-----KAKKAAPRKK
H1e  SKPKA-AKP-----KKTAAKKK
H1o  R-----ASKKK
H1t  AKMVMQKTDL-----RKAAGRK
                        *

```

**B**

	H1b	H1c	H1d	H1e	H1 <sup>o</sup>	H1t
H1a	63.1	58.2	61.5	63.1	40.2	32.0
H1b		67.2	74.6	72.1	36.9	34.4
H1c			77.0	77.0	45.9	36.1
H1d				78.7	41.0	36.1
H1e					44.3	37.7
H1 <sup>o</sup>						27.9

Figure 29. Sequence identity between the C-terminal domains of mouse histone H1 subtypes. (A) Alignment of the C-terminal domain sequences of H1a-e, H1<sup>o</sup> and H1t. Multiple alignment was performed using the ClustalW resource under DNASTAR. The conserved positions are indicated by an asterisk (\*). (B) Percentage similarity table.



To make sure that the slightly longer size of the SAR fragment over the pUC19 fragment had no influence on the preference of the C-terminal domain for the SAR, we performed a competition experiment with an extended pUC19 fragment of 763 bp that included the 587 bp pUC19 sequence. It can be seen in Figure 28E that the preference of the H1 C-terminal domain for the SAR remained unaffected.

It has been reported that the preferential binding of histone H1 to SARs is not determined by its high AT content, but by the presence of abundant A-tracts. As a control of the preferential binding of the H1 C-terminal domain, we have included a competition between the SAR and a mixture of multimers of a 50 bp sequence containing 75% AT, but lacking A-tracts (see Materials and Methods). The size distribution of the multimers spanned from about 500 bp to 5000 bp (Fig. 30A). As shown in Figure 30B, the C-terminal domain bound with extremely high preference to the SAR in spite of the excess and longer size of a large fraction of the polymerised sequence.

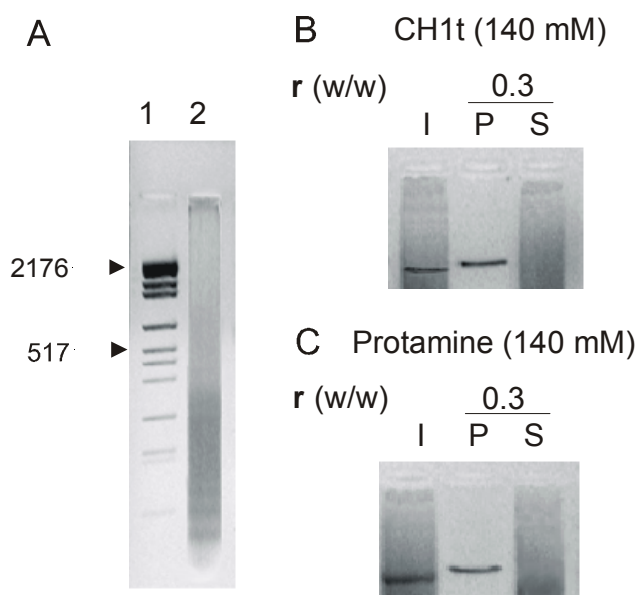


Figure 30. Preferential binding of the histone H1 C-terminal domain and protamine to the SAR in the presence of AT-rich DNA. The AT-rich DNA was obtained by polymerization of 50 bp sequence containing 75% AT and lacking A-tracts as described in Materials and Methods. (A) Analysis of the ligated DNA by 2% agarose gel electrophoresis; 1, DNA molecular size markers; 2, ligation reaction. The arrows indicate the limits of the size distribution of oligomers used in the binding experiments (from 500 bp to 5000 bp). (B) Preferential binding of the C-terminal domain of H1t to the SAR. (C) Preferential binding of protamine to the SAR. The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the DNA fragments, the discrete band corresponds to the SAR and the smear to the AT-rich fragments; P, pellet; S, supernatant.

In the study of the interaction of the N-terminal domain with the SAR, we used peptides that had been characterized previously<sup>173, 174</sup>. One peptide corresponded to the N-terminal domain of H1<sup>o</sup> (residues 1-20). Another peptide comprised the basic region of the N-terminal domain of H1e (residues 15-36). The affinity of these peptides for the DNA was much lower than that of the C-terminal domain, as expected from the lower number of positive charges involved in the interaction: 6-9 charges in the N-terminal domain compared with 33-46 in the C-terminal domain, depending on the subtype. Both N-terminal peptides precipitated the DNA, but a large excess of peptide was necessary. The N-terminal peptide of H1e had a moderate preference for the SAR in physiological salt (Fig. 31A). The preference was most apparent at protein/DNA ratios that precipitated a small amount of DNA, and was lost at higher protein/DNA ratios that precipitated most of the DNA. In 70 mM NaCl, the preference for the SAR fragment was completely lost (Fig. 31B). In 140 mM NaCl, the preference for the SAR fragment was completely lost (Fig. 31C). The N-terminal peptide of H1<sup>o</sup> did not show any preference for the SAR, either in 140 mM or in 70 mM NaCl (Fig. 31C, D).

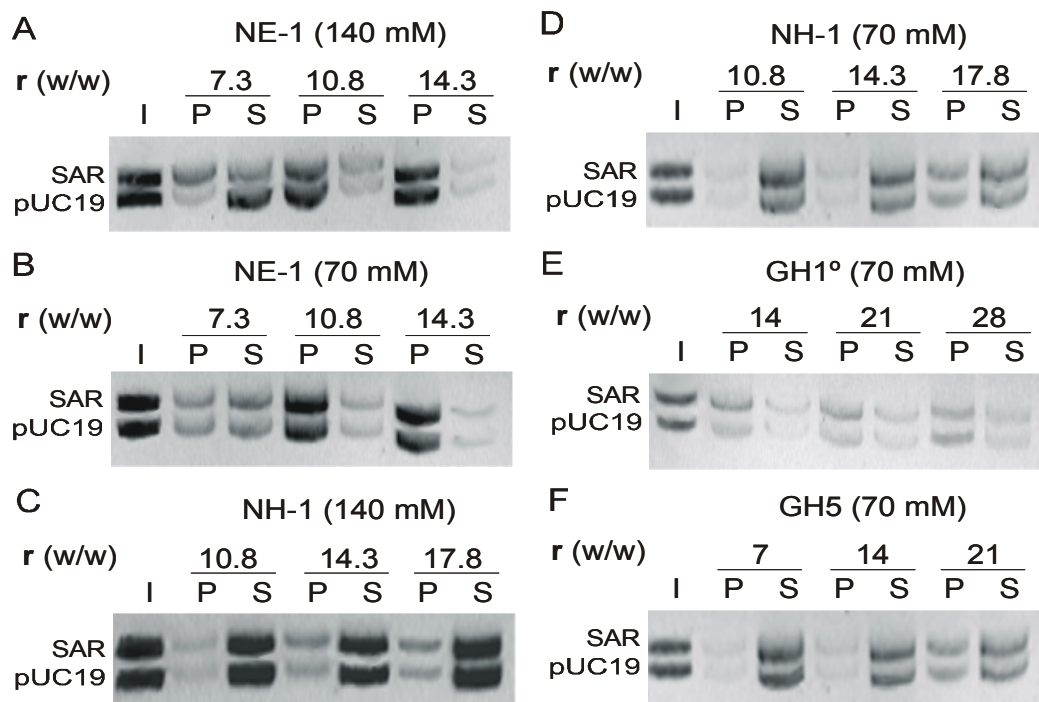


Figure 31. Binding of the N-terminal peptides and the globular domain of histone H1 to the SAR and pUC19 fragments. (A) and (B) N-terminal peptide of H1e (NE-1); (C) and (D) N-terminal peptide of H1<sup>o</sup> (NH-1); (E) globular domain of H1<sup>o</sup> (GH1<sup>o</sup>); and (F) globular domain of H5 (GH5). The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the SAR and pUC19 fragments; P pellet; S, supernatant.

With the globular domain, precipitation experiments could not be performed in 140 mM NaCl because the interaction with the DNA was too weak in this salt concentration. Binding experiments were carried out instead in 70 mM NaCl, where with a high enough protein/DNA ratio all the DNA could be precipitated. The globular domain of H1<sup>o</sup> showed no preference for the SAR fragment, not even at protein/DNA ratios that precipitated a small proportion of the DNA (Fig. 31E). The globular domain of H5, an avian subtype closely related to H1<sup>o</sup>, also failed to bind preferentially to the SAR (Fig. 31F).

### Protamines also bind preferentially to SARs

Typical protamines are sperm basic proteins with an arginine content of 60% or higher. We used salmine in the binding experiments, a typical protamine from salmon sperm. It has 32 residues, 21 of which are arginine. The Arg residues are mostly present in clusters of up to 6 residues. Figure 32 shows that protamine also bound with high preference to the SAR fragment either in physiological salt or in 70 mM NaCl.

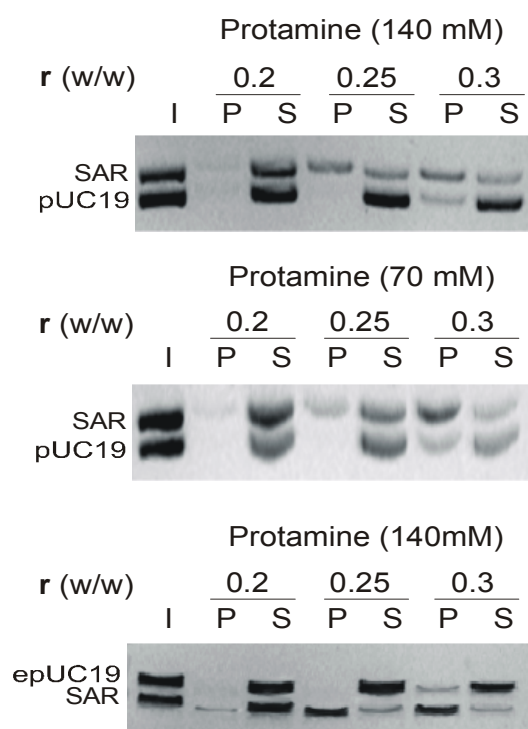


Figure 32. Preferential binding of protamine to the SAR fragment. DNA fragments are indicated: SAR (SAR fragment of 657 bp), pUC19 (pUC19 fragment of 587 bp), epUC19 (extended pUC19 fragment of 763 bp). The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the fragments; P, pellet; S, supernatant.

The preference is, however, not so high as displayed by the C-terminal domain, as shown by the presence of a small amount of competitor pUC19 in the pellets coexisting with a small amount of SAR in the supernatants. The preference for the SAR was maintained when the longer pUC19 fragment (763 bp) was used in the experiments (Fig. 32). As in the case of the H1 C-terminal domain, protamine also bound preferentially to the polymerised sequence lacking A-tracts (Fig. 30C).

### Distamycin abolishes the SAR binding character of the C-terminal domain of H1 and protamine

Distamycin is an antibiotic drug with high selectivity for A-tract DNA<sup>311, 312</sup>. Binding to A-tracts by distamycin abolishes the specific interaction of SARs with the nuclear scaffold and SAR binding proteins. In particular, binding of distamycin results in the suppression of preferential binding of histone H1 to SAR-containing DNA, leading to a redistribution of histone H1 between SAR and non-SAR DNA<sup>151</sup>. Distamycin has been used to confirm the SAR binding properties of several proteins. We examined the effect of distamycin on the binding of the C-terminal domain of H1 and protamine, which were shown to bind highly preferentially to the SAR fragment. In both cases, distamycin abolished the preferential binding of the proteins to the SAR (Fig. 33). The effect was so clear as to reverse the binding preferences, binding to pUC19 being stronger than to the SAR in the presence of distamycin.

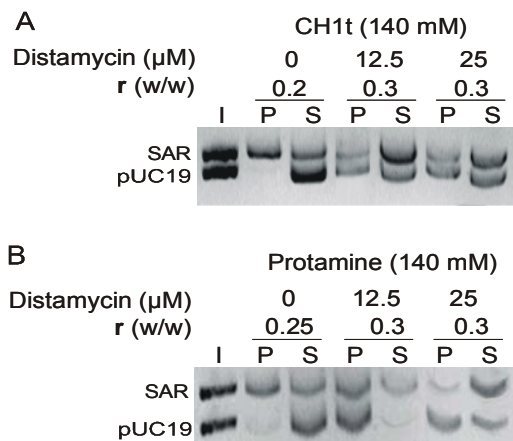


Figure 33. Distamycin inhibits the preferential binding of the histone H1 C-terminal domain and protamine to the SAR fragment. (A) C-terminal domain of H1t (CH1t); (B) salmon protamine. The concentration of NaCl is indicated in parenthesis. The concentration of distamycin and the protein /DNA ratio (w/w) are indicated. I, input mixture of the SAR and pUC19 fragments; P, pellet; S, supernatant.

## **Discussion**

Histone H1 has been generically described as a SAR-binding protein. Here, we have shown that the highly conserved mammalian subtypes, H1a-e, and the highly divergent H1<sup>o</sup> and H1t are all SAR binding proteins. The C-terminal domain of histone H1 contributes to a large extent to the binding free energy of the entire molecule to DNA, providing the localized charge neutralization of the DNA necessary for the chromatin folded state<sup>178, 182, 313</sup>. It was thus of interest to investigate whether the SAR binding character of H1 subtypes was a property of the C-terminal domain. We studied the C-terminal domains of subtypes H1e, the most abundant somatic subtype, H1<sup>o</sup>, a subtype that accumulates in quiescent cells, and H1t, the male germ line-specific subtype. In spite of their large sequence divergence (44.3% sequence identity between H1e and H1<sup>o</sup> and 27.9% between H1<sup>o</sup> and H1t; Fig. 29), all three C-terminal domains appear to have very strong preference for the SAR. H1e and H1<sup>o</sup> contain SPKK motifs, presumably contributing to SAR binding specificity; however, this motif is not present in H1t, which indicates that its presence is not essential to determine the SAR binding character of histone H1 subtypes. As shown previously for the entire molecule<sup>151</sup>, distamycin, a minor groove binder with high affinity for A-tracts, abolishes the preferential binding of the C-terminal domain to the SAR fragment. This suggests involvement of the DNA minor groove in the interaction with the C-terminal domain, although it does not exclude the possibility that the C-terminal domain also binds to the wide groove<sup>154</sup>. As shown by X-ray crystallography, distamycin occupies the minor groove with preference for narrow minor grooves<sup>152</sup>. However, distamycin may also cause subtle changes in the structural parameters and mecanoelastic properties of the DNA that could contribute to the suppression of the preferential binding of SAR-binding proteins<sup>312</sup>.

In agreement with its reduced number of positive charges compared with the C-terminal domain, the N-terminal domain exhibits much weaker DNA binding, and a large excess of protein is necessary to precipitate the DNA. The N-terminal peptide of H1<sup>o</sup> showed no preference for the SAR fragment. The N-terminal domain of H1e had a moderate preference for the SAR fragment in physiological salt, but in 70 mM salt the binding preference was completely lost. With the globular domain, binding experiments had to be performed at 70 mM NaCl, as in 140 mM the binding was too weak to observe a precipitate even with a large excess of protein. Under these conditions, the globular domain did not show any preference

for the SAR sequences over the pUC19 fragments. Given the low selectivity and weak binding of the N-terminal peptides and the globular domain, it is unlikely that they make a substantial contribution to the SAR binding character of the whole protein. It thus appears that the preferential binding of histone H1 to SARs is basically determined by its C-terminal domain.

Phosphorylation is the main post-translational modification undergone by histone H1. Most phosphorylation sites are located in the C-terminal domain. Phosphorylation weakens the binding of H1 to DNA, and could thus facilitate SAR chromatin opening through cooperative H1 dissociation<sup>147, 314</sup>. The C-terminal domain could also be responsible for the targeting of H1 molecules to SARs. At substoichiometric concentrations, the higher affinity of H1 for SARs would guarantee the saturation of SAR sequences with H1 in preference to other sequences.

Evolutionary evidence has recently been obtained for chordate protamines having originated from histone H1 through general substitution of Lys residues by Arg<sup>304</sup>. In view of the preferential binding of the C-terminal domain of H1 to SARs, we investigated whether such a property was conserved in protamines. The results show that protamine also binds to the SAR fragment with high specificity. Moreover, the binding preference is also determined by the oligo-A tracts as shown by the suppression of preferential binding to the SAR by distamycin. The effect is so strong as to reverse the binding preferences of protamine. This is indicative of minor groove recognition by protamine. The role, if any, of the preferential binding of protamines to SARs cannot be ascertained at this stage; however, it is possible that this property could lend some spatial and temporal order to the substitution of core histones by protamines in spermatogenesis, the SARs providing nucleation sites for the substitution process.

The features that confer SAR binding specificity to a protein or protein motif have yet to be established. Mammalian HMGA non-histone proteins contain multiple copies of a DNA binding motif called “AT hook” that preferentially binds to the narrow minor groove of AT sequences<sup>155</sup>; the core of the AT hook is the sequence Arg-Gly-Arg. The interaction of the RGR element with the minor groove has been characterized by NMR spectroscopy<sup>315</sup>. No RGA motifs are found in either histone H1 or protamine. While histone H1 is lysine-rich, the core motif of AT hooks and protamine contain Arg. The C-terminal domain is an even

better SAR binding protein than protamine. The choice between Arg and Lys does not therefore appear essential in itself in determining the SAR-binding character. Presumably, the interaction should as a rule involve minor groove recognition, as all SAR binding proteins so far described are competed by distamycin. The narrower minor groove associated with A-tracts would be preferred because it would give better van der Waals contacts between the walls of the minor groove and the ligands<sup>152</sup>. In the case of highly cationic ligands, such as the C-terminal domain of histone H1 and protamine, potentially interacting with the DNA phosphates, a narrower minor groove would also offer a more intense electrostatic potential along the minor groove path that could also contribute to the preferential binding.





## **Capítulo 2**

*Estructura secundaria del dominio C-terminal de la histona H1 inducida por el DNA .*



### **Introduction**

H1 linker histones are thought to be primarily responsible for the condensation of the thick chromatin fibre. It is currently accepted that histone H1 could have a regulatory role in transcription through the modulation of chromatin higher-order structure. H1 has been described as a general transcriptional repressor because it contributes to chromatin condensation, which limits the access of the transcriptional machinery to DNA. However, H1 may regulate transcription at more specific level, participating in complexes that either activate or repress specific genes<sup>1, 89, 222, 247, 283, 284, 285, 286</sup>. Binding to SARs and participation in nucleosome positioning are other mechanisms by which H1 could contribute to transcriptional regulation<sup>130, 287</sup>.

H1 has multiple isoforms. In mammals, six somatic subtypes, designated H1a-e and H1<sup>o</sup>, a male germline-specific subtype, H1t, and an oocyte-specific subtype, H1oo, have been identified<sup>80, 93, 288, 289</sup>. The subtypes differ in timing of expression<sup>79</sup>, extend of phosphorylation<sup>290</sup>, turnover rate<sup>97, 291</sup>, binding affinity<sup>121</sup> and evolutionary stability<sup>78</sup>. Differences in DNA condensing capacity have also been demonstrated for some subtypes<sup>292, 293, 294</sup>.

Linker histones contain three distinct domains: a short N-terminal domain (20-35 amino acids), a central globular domain (~80 amino acids), consisting of a helix bundle and a  $\beta$ -hairpin, and a long C-terminal domain (~100 amino acids)<sup>301</sup>. The N- and C-terminal domains are highly basic. The distribution of charge in the C-terminal domain is extremely uniform in spite of the variation in sequence in the different subtypes<sup>175</sup>.

The C-terminal domain is the primary determinant of histone H1 binding to chromatin *in vivo*<sup>121, 182</sup>. Several studies indicate that the ability of linker histones to stabilize chromatin folding resides in the C-terminal domain of the protein<sup>172, 178, 316</sup>. The preferential binding of histone H1 to scaffold-associated regions (SARs) appears to be determined by the C-terminal domain<sup>317</sup>. The C-terminal domain is also responsible for activation of apoptotic nuclease<sup>189</sup>. Knowledge of the structure of the C-terminal domain of H1 once bound to DNA is thus important to our understanding of H1 function.

The N- and C-terminal domains have little defined structure in solution, although they may contain a considerable amount of turn like conformations in rapid equilibrium with unfolded states. Molecular modelling<sup>188</sup> and induction of  $\alpha$ -helix in the presence of secondary

structure stabilizers<sup>183</sup>, such as TFE and HClO<sub>4</sub>, suggest that the C-termini would acquire a folded conformation on binding to DNA. We have shown previously that interaction with DNA induces stable helix and turn structures in a 23 residue peptide from the H1° C-terminus (CH-1, residues 99-121)<sup>158, 186</sup>. This region, adjacent to the globular domain, was shown to behave as a specific sub-domain in the stabilization of folded chromatin structures by histone H1<sup>178</sup>. Inducible helical elements were also characterized in the N-terminal domains of H1° and H1e<sup>173, 174</sup>.

In the present study we have used infrared spectroscopy to study the secondary structure of the C-terminal domains of H1° and H1t in aqueous solution and in the complexes with DNA. The C-terminal domains have little structure in solution, but become fully structured upon interaction with DNA in physiological salt (140 mM NaCl). The secondary structure elements present in the complexes include  $\alpha$ -helix,  $\beta$ -structure, turns and open loops. The study of the entire C-terminal domain has provided new information, not readily accessible to studies with peptides, on the variety of structural motifs present in the DNA-bound domain, the salt dependence and thermal stability of the folded structure and the conservation of the protein structure among divergent subtypes. It has also shown that the property of coupled binding and folding, already demonstrated for the CH-1 peptide<sup>186</sup>, is shared by the entire domain.

### **Materials and methods**

#### Cloning, expression and purification of the carboxy-terminal domains of histone H1 subtypes

The sequences encoding the C-terminal domains of histones H1° and H1t were cloned and expressed. All gene fragments were amplified from mouse genomic DNA by PCR. The primers were 5'GGAGATCTTTACTTCTTCTTGCTGGCCCTCT 3' and 5'CCACCATGGATGAGCCTAAAAGGTC 3' for the C-terminal domain of H1° and 5'GTACCATGGCGGCTTCAGGGAACGAC 3' and 5'ACGGATCCTTACTTCTCCCTGCTGCCTTCT 3' for the C-terminal domain of H1t. The amplification products were cloned in the pQE60 vector (Qiagen) using the NcoI and BglII restriction sites to yield the expression vectors pCTH1° and pCTH1t. The recombinant plasmids were transformed into *E. coli* M15 (Qiagen). Cells were grown to an OD<sub>600</sub> of 0.8 and then induced with 1 mM IPTG, allowing

the expression to proceed for 4 h at 37°C. Cells were lysed in 10 mM potassium phosphate plus 4 M guanidinium hydrochloride, for 15 min at room temperature. The extract was centrifuged at 20,000×g for 25 min. The supernatants were loaded on a CHT-II cartridge filled with ceramic hydroxyapatite type II (Bio-Rad), equilibrated with lysis buffer. The bound fraction, corresponding to the recombinant proteins, was eluted with 200 mM potassium phosphate, 4 M guanidinium hydrochloride, pH 7.0. Finally, the proteins were desalted by gel filtration through Sephadex G-25 (Amersham Biosciences).

### Circular dichroism spectroscopy

Samples for CD spectroscopy were  $2 \times 10^{-5}$  M of C-terminal domain in 10 mM phosphate buffer, pH 7.0. Samples in aqueous solution and in 60% 2,2,2- trifluoroethanol (TFE) (v/v) were prepared. Spectra were obtained on a Jasco J-715 spectrometer in 1 mm cells at 20°C. The results were analyzed with Standard Analysis software (JACSO) and expressed as mean residue molar ellipticity  $[\theta]$ . The helical content was estimated from the ellipticity value at 222 nm ( $\theta_{222}$ ), according to the empirical equation of Chen et al.<sup>318</sup>: % helical content =  $100 [\theta_{222}/-39,500 \times (1-2.57/n)]$ , where  $n$  is the number of peptide bonds.

### Infrared spectroscopy measurements

The C-terminal domains of histone H1 were measured at 5 mg/ml in 10 mM HEPES plus 10 mM or 140 mM NaCl. DNA-protein complexes contained 7.0 mg/ml mouse DNA and the appropriate amount of protein.

Measurements were performed on a Nicolet Magna II 550 spectrometer equipped with a MCT detector, using a demountable liquid cell with calcium fluoride windows and 50  $\mu\text{m}$  spacers for D<sub>2</sub>O medium and 6  $\mu\text{m}$  spacers for H<sub>2</sub>O medium measurements. Typically, 1000 scans for each background and sample were collected and the spectra were obtained with a nominal resolution of 2  $\text{cm}^{-1}$ , at 22°C. Data treatment and band decomposition have been described previously<sup>319</sup>. The DNA contribution to the spectra of the complexes with the C-terminal domain was subtracted using a DNA sample of the same concentration; the DNA spectrum was weighted so as to cancel the symmetric component of the phosphate vibration at 1087  $\text{cm}^{-1}$  in the difference spectra as described in Vila et al.<sup>186</sup>. In addition, spectra of complexes of different protein to DNA ratio were recorded (0.4 and 0.7, w/w). The resulting

protein spectra were independent of the protein/DNA ratio inside the statistical error, indicating that the amide I region was not significantly affected by DNA spectral changes induced by interaction with the protein.

For the analysis of the thermal stability of the DNA-C-H1t complexes, a tungsten-copper thermocouple was placed directly onto the window and the cell placed on a thermostatted cell mount. Thermal analyses were performed by heating continuously from 20 to 80°C at a rate of 1°C/min. Spectra were recorded using a Rapid Scan software running under OMNIC (Nicolet). For each degree of temperature interval, 305 interferograms were averaged, Fourier-transformed and ratioed against background, obtaining the spectra with a nominal resolution better than 2 cm<sup>-1</sup> <sup>320</sup>. The solvent contribution was subtracted as described earlier<sup>321</sup>. Thermal analyses were performed only in D<sub>2</sub>O. The DNA contribution to the spectra at each temperature was subtracted using a DNA sample at the same temperature and concentration. The subtraction was weighted using the phosphate symmetric vibration at 1087 cm<sup>-1</sup> as described above. In the amide I' regions, the spectra of native and denatured mixed-sequence calf-thymus DNA were similar enough not to affect the protein difference spectrum.

## Results

### CD and IR spectroscopy analysis of the carboxy-terminal domain of histone H1 in aqueous solution and in TFE

We have used IR spectroscopy to examine the secondary structure of the C-terminal domains of the histone H1 subtypes H1° (C-H1°) and H1t (C-H1t) (Figure 34) in aqueous (D<sub>2</sub>O) and in TFE solutions (Figure 35). Values corresponding to band position and percentage area are given in Table 6.

```

H1°  KGDEPKRSVAF-KKTKKEVKKVATPKKAAKPKKAASKAPSKKPKATPVKKAKKKPAATPKK-AKKP
H1t  KKAASGNDKGGKKSASAKA-----KKMGLP--RASR--SPKSSKT--KAVKK-PKATPTKASGSG
      *           **             ** *   **   * *   * *   * *   ** * * * * *
H1°  KVVK-VKPVKASK-PKKAKTVKPKAKSSAK-----RASKKK
H1t  RKTGKAGGVQQRKSPAKARAANP---NSGKAKMVMQKTDLRKAAGRK
      * * *   * * * *   *   * *           *   *

```

Figure 34. Alignment of the C-terminal domain sequences of mouse histone H1 subtypes H1° and H1t. The conserved positions are indicated by an asterisk (\*).

In aqueous solution, the amide I' of both C-H1° and C-H1t was dominated by the random coil at 1643  $\text{cm}^{-1}$ , with about 50% of total amide I' intensity. The other main component at 1661-1662  $\text{cm}^{-1}$ , with 35% of the total area in C-H1° and 29% in C-H1t, was assigned to turns. In previous studies with H1 peptides, the component at 1661  $\text{cm}^{-1}$  was related to the nascent helix: turns in rapid equilibrium with the unfolded state that in the presence of secondary structure stabilizers are transformed into helical structure<sup>173, 186, 322</sup>. Small amounts of turns at 1670  $\text{cm}^{-1}$  (6-9%) and extended structure (8-9%) were also present.

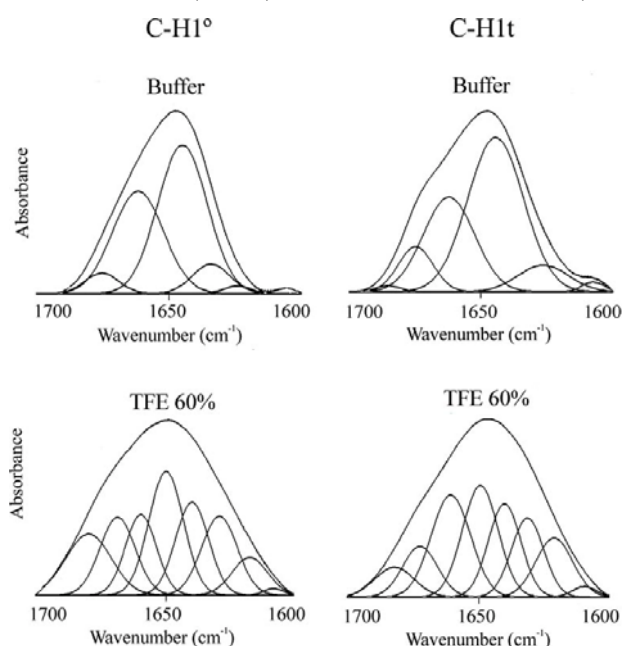


Figure 35. Amide I decomposition of the carboxy-terminal domain of histone H1. The left column shows the spectra of the C-terminal domain of histone H1° (C-H1°) and the right column the spectra of the C-terminal domain of histone H1t (C-H1t). The spectra were recorded in water ( $\text{D}_2\text{O}$ ) and in 60 % TFE solution at 20°C. The buffer was 10 mM HEPES plus 10 mM NaCl, pD 7.0. The protein concentration was 5 mg/ml.

In 60% TFE, C-H1° and C-H1t were extensively folded. Turns (40-38%),  $\alpha$ -helix (23-22%) and  $\beta$ -structure (21-25%) were identified, together with a component of random coil/open loops (16%), which are not resolved in  $\text{D}_2\text{O}$ . The fraction of turns at 1659-1660  $\text{cm}^{-1}$  represented 13% in C-H1° and 21% in C-H1t. In addition to turns, this component could contain a variable proportion of  $3_{10}$  helix<sup>323, 324, 325</sup>, in which case, the proportion of helical structure would increase considerably.

TABLE 6				
Percentages of secondary structure of the carboxy-terminal domain of histones H1 <sup>o</sup> and H1t in aqueous and TFE solutions.				
Assignment	C-H1 <sup>o</sup>		C-H1t	
	Band	%	Band	%
	(cm <sup>-1</sup> )		(cm <sup>-1</sup> )	
IR spectroscopy <sup>a</sup>				
Buffer				
Turns			1685	1
Turns	1675	6	1674	9
Turns	1661	35	1662	29
Random coil/open loops	1644	50	1643	53
β-sheet	1634	7		
Low frequency β-sheet	1624	2	1626	8
60% TFE				
Turns	1679	13	1682	7
Turns	1668	14	1672	10
Turns/3 <sub>10</sub> helix	1659	13	1660	21
α-helix	1650	23	1649	22
Random coil/open loops	1639	16	1640	16
β-sheet	1629	14	1631	14
Low frequency β-sheet	1617	7	1621	11
CD <sup>b</sup>				
Buffer, α-helix	2		2	
60% TFE, α-helix	23		24	
<sup>a</sup> Band position (cm <sup>-1</sup> ), percentage area (%) and assignment of the components obtained after curve fitting of the amide I band of the C-H1 <sup>o</sup> and C-H1t in aqueous (D2O) and TFE solution. The values were rounded off to the nearest integer				
<sup>b</sup> Percentages of α helical structure estimated by circular dichroism from the mean residue molar ellipticity at 222 nm as described under "Materials and Methods".				

The values of α-helix estimated by CD from the ellipticity value at 222 nm<sup>318</sup> were similar to those estimated by IR for both C-H1<sup>o</sup> and C-H1t (Table 6, Figure 36).

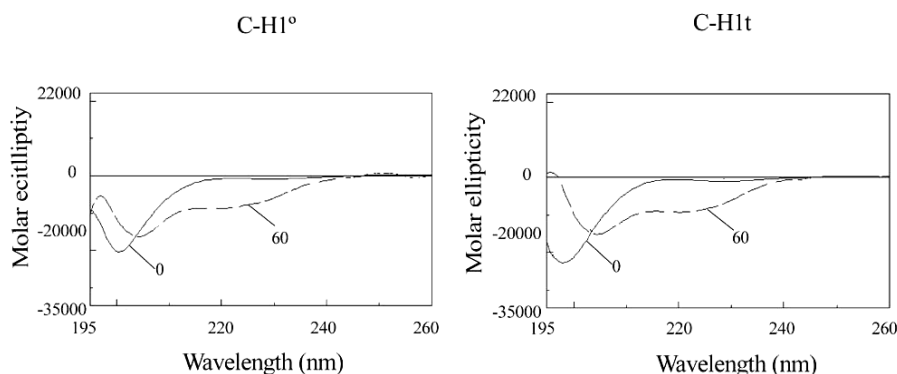


Figure 36. TFE-dependent conformational transition of the carboxy-terminal domain of histone H1 measured by CD. Far-UV CD spectra in water and 60% TFE in 10 mM NaCl, both in phosphate buffer 10 mM, pH 7.0 at 20°C. The numbers refer to the TFE concentration in percentage by volume. C-H1<sup>o</sup>, C-terminal domain of histone H1<sup>o</sup>; C-H1t, C-terminal domain of histone H1t.

### IR spectroscopy of the carboxy-terminal domain of histone H1 bound to DNA

We have examined the secondary structure of the C-terminal domain of histone H1 bound to DNA. The spectra of the complexes were recorded both in H<sub>2</sub>O and D<sub>2</sub>O to facilitate the assignment of the amide I components. Deuteration has a large effect on the position of



some components of the amide I, while others are shifted slightly or not at all<sup>319</sup>. The vibrations of loops and the  $\alpha$ -helix are only slightly affected by isotopic substitution and are thus found at similar positions in H<sub>2</sub>O and D<sub>2</sub>O; the  $\alpha$ -helix at about 1652 cm<sup>-1</sup> and loops at about 1643 cm<sup>-1</sup>. In contrast, deuteration has a large effect on the position of the random coil, so that in H<sub>2</sub>O it overlaps with the  $\alpha$ -helix, while in D<sub>2</sub>O it overlaps with loops. Therefore, the  $\alpha$ -helix could be estimated directly in D<sub>2</sub>O, while loops could be estimated directly in H<sub>2</sub>O. When both  $\alpha$ -helix and loops are present in the protein structure, the percentage of random coil can be obtained from either the difference of the components around 1652 cm<sup>-1</sup> in H<sub>2</sub>O and D<sub>2</sub>O or from the difference of the components around 1643 cm<sup>-1</sup> in D<sub>2</sub>O and H<sub>2</sub>O, as indicated by the expression:

$$cb_{\alpha\text{-helix} + \text{RC}}^{\text{H}_2\text{O}} (\sim 1652 \text{ cm}^{-1}) - cb_{\alpha\text{-helix}}^{\text{D}_2\text{O}} (\sim 1652 \text{ cm}^{-1}) = cb_{\text{loops} + \text{RC}}^{\text{D}_2\text{O}} (\sim 1643 \text{ cm}^{-1}) - cb_{\text{loops}}^{\text{H}_2\text{O}} (\sim 1643 \text{ cm}^{-1}) = cb_{\text{RC}}$$

where  $cb$  is the percentage of the assigned amide I component band in either H<sub>2</sub>O or D<sub>2</sub>O and  $cb_{\text{RC}}$  the percentage of random coil.

Considering that the percentages of the different secondary structure elements in H<sub>2</sub>O and D<sub>2</sub>O should be comparable, the similarity of the percentages of random coil estimated either from the differences of the components around 1643 cm<sup>-1</sup> in D<sub>2</sub>O and H<sub>2</sub>O and around 1652 cm<sup>-1</sup> in H<sub>2</sub>O and D<sub>2</sub>O warrants the proper assignment of the amide I components containing contributions of  $\alpha$ -helix, loops and random coil.  $\beta$ -structure components appear approximately at the same frequency in D<sub>2</sub>O and H<sub>2</sub>O, while in D<sub>2</sub>O turn components are slightly shifted to lower frequencies.

The complexes of the C-terminal domain with DNA were examined in 10 mM and 140 mM NaCl. In 10 mM NaCl, the analysis of the complexes of C-H1° with DNA in D<sub>2</sub>O and H<sub>2</sub>O yielded 27% of turns, 19% of  $\alpha$ -helix, 25% of open loops, 22% of  $\beta$ -structure and 7% of random coil. The amount of random coil was calculated from the difference of the components around 1654 cm<sup>-1</sup> in H<sub>2</sub>O and D<sub>2</sub>O and from the difference of the components around 1643 cm<sup>-1</sup> in D<sub>2</sub>O and H<sub>2</sub>O. The values obtained were equal within 1% (Figures 37 and 38, Table 7).

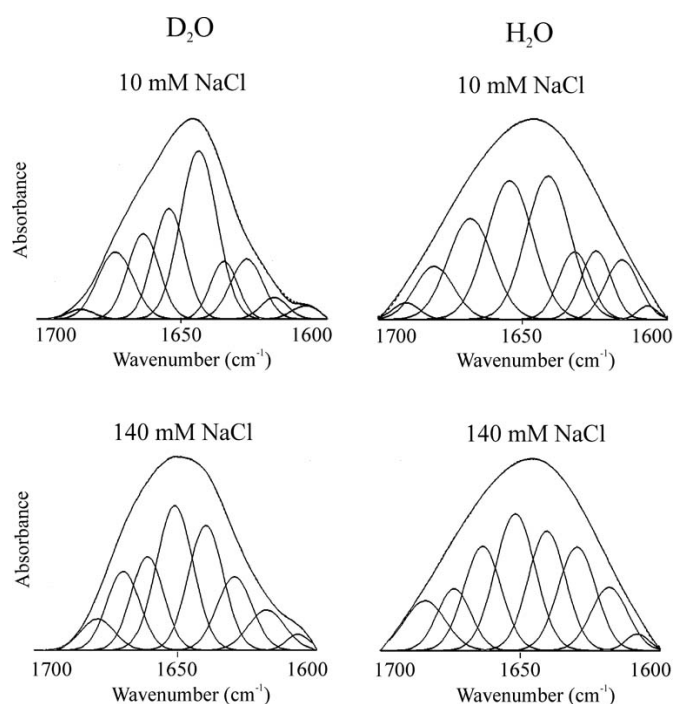


Figure 37. Amide I decomposition of the spectra of the carboxy-terminal domain of histone H1° (C-H1°) bound to DNA. The left column shows the spectra measured in D<sub>2</sub>O and the right column those measured in H<sub>2</sub>O. The salt concentration, 10 mM or 140 mM NaCl, is indicated. The reconstitution of the absorption spectra is indicated by dashed lines. The DNA contribution to the spectra of the complexes was subtracted as described in Materials and Methods.  $r$  (protein/DNA ratio = 0.7 (w/w)).

The representation of the different types of secondary structure in C-H1t bound to DNA was similar to that of C-H1°, although some small differences were observed; namely, a higher amount of random coil (12%), a lower amount of  $\alpha$ -helix (15%) and open loops (14%) and a higher amount of  $\beta$ -structure (25-27%) and turns (32-34%) (Figures 37 and 38, Table 7).

When the salt concentration was increased up to 140 mM NaCl, the DNA-bound C-terminal domains of both H1° and H1t underwent significant structural changes. In the first place, the protein became fully structured, as judged by the total absence of a random coil component. In C-H1°, the disappearance of the random coil was correlated with an increase of  $\alpha$ -helix,  $\beta$ -structure and turns, which reached values of 24%, 24-25% and 33-34%, respectively. A decrease in the percentage of open loops from 25% in 10 mM to 18% in 140 mM was also observed. The increase of turns was most apparent for C-H1t, reaching 40-42%; the  $\alpha$ -helix also increased up to 21%; the contributions of the remaining structural motifs did not change significantly (Figures 37 and 38, Table 7). As in TFE solution, the component at  $\sim 1660$  cm<sup>-1</sup> could contain a variable proportion of  $3_{10}$  helix.

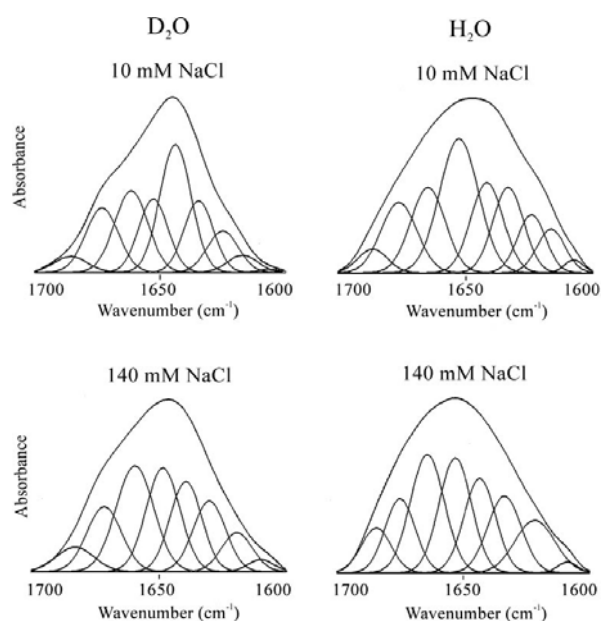


Figure 38. Amide I decomposition of the spectra of the carboxy-terminal domain of histone H1t (C-H1t) bound to DNA. The left column shows the spectra measured in D<sub>2</sub>O and the right column those measured in H<sub>2</sub>O. The salt concentration, 10 mM or 140 mM NaCl, is indicated. The reconstitution of the absorption spectra is indicated by dashed lines. The DNA contribution to the spectra of the complexes was subtracted as described in Materials and Methods.  $r$  (protein/DNA ratio = 0.7 (w/w)).

TABLE 7								
Percentages of secondary structure of the carboxy-terminal domain of histones H1 <sup>o</sup> and H1t bound to DNA.								
Band position (cm <sup>-1</sup> ), percentage area (%) and assignment of the components obtained after curve fitting of the amide I band of the C-H1 <sup>o</sup> and C-H1t bound to DNA in H <sub>2</sub> O and D <sub>2</sub> O and in 10 mM and 140 mM NaCl. The values were rounded off to the nearest integer.								
Assignment	Complexes with DNA							
	10 mM NaCl				140 mM NaCl			
	D <sub>2</sub> O		H <sub>2</sub> O		D <sub>2</sub> O		H <sub>2</sub> O	
	Band	%	Band	%	Band	%	Band	%
	(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )	
C-H1 <sup>o</sup>								
Turns	1684	1	1688	2	1677	6	1682	8
Turns	1673	12	1679	8	1668	13	1673	9
Turns/3 <sub>10</sub> helix	1663	14	1667	17	1659	15	1663	16
α-helix	1654	19			1650	24		
α-helix/random coil			1655	26			1652	24
Open loops			1641	25			1640	18
Random coil/open loops	1643	32			1639	18		
Random coil <sup>a</sup>		7				0		
Random coil <sup>b</sup>				7				0
β-sheet	1634	8	1632	8	1630	16	1630	16
Low frequency β-sheet	1627	10	1625	7				
Low frequency β-sheet	1617	4	1616	7	1617	8	1618	9
C-H1t								
Turns	1686	4	1688	4	1680	7	1685	6
Turns	1673	13	1676	13	1671	13	1676	12
Turns/3 <sub>10</sub> helix	1662	17	1665	15	1659	22	1664	22
α-helix	1653	15			1649	21		
α-helix/random coil			1653	27			1653	21
Open loops			1642	14			1644	15
Random coil/open loops	1644	26			1640	15		
Random coil <sup>a</sup>		12				0		
Random coil <sup>b</sup>				12				0
β-sheet	1635	13	1633	13	1631	13	1634	13
Low frequency β-sheet	1626	8	1624	8	1621	9	1622	11
Low frequency β-sheet	1617	4	1616	6				

<sup>a</sup> The value corresponds to the difference between the components at ~ 1642 cm<sup>-1</sup> in D<sub>2</sub>O and H<sub>2</sub>O.

<sup>b</sup> The value corresponds to the difference between the components at ~ 1652 cm<sup>-1</sup> H<sub>2</sub>O and D<sub>2</sub>O.

## Melting experiments

We monitored the changes in the amide I' components of the DNA-bound C-H1t upon heating in the 20-80°C interval. The DNA contribution to the spectra at each temperature was subtracted using a DNA sample at the same temperature and concentration; the subtraction was adjusted using the phosphate symmetric vibration at 1087 cm<sup>-1</sup> as described. Although free DNA melted before complexed DNA, the spectra of native and denatured DNA were similar enough not to affect the amide I' region of the difference spectra. Melting experiments were performed in either 10 mM or 140 mM NaCl at a protein/DNA ratio (w/w) of 0.4. In 10 mM NaCl, the  $\alpha$ -helix and  $\beta$ -sheet components remained essentially unaltered up to 78°C (Figure 39A).

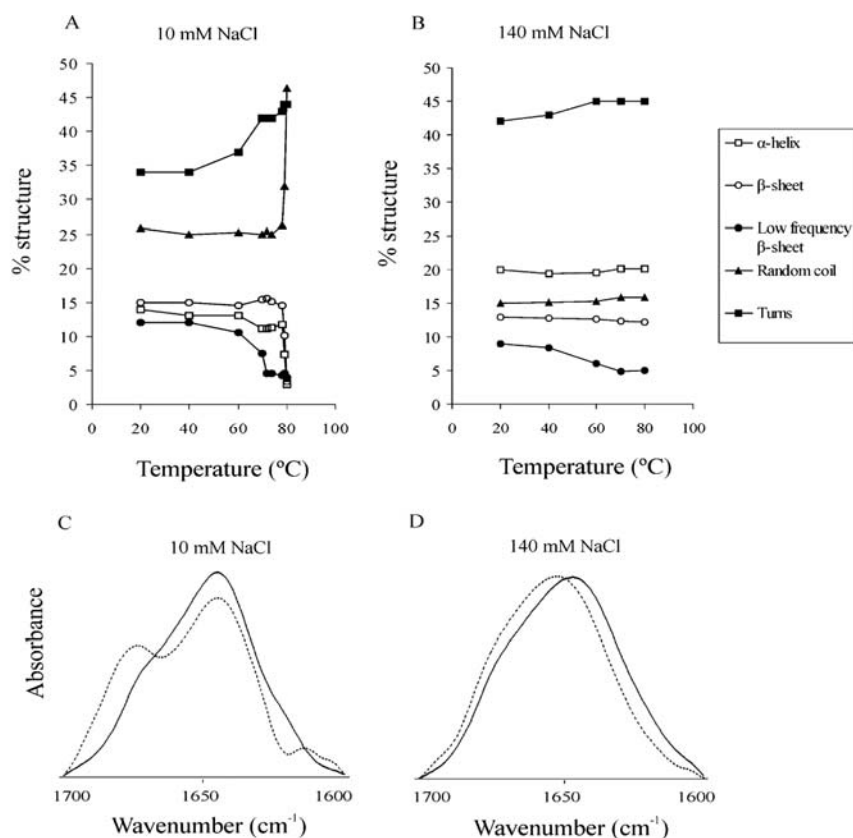


Figure 39. Thermal profiles of the amide I band of the secondary structure motifs of the carboxy-terminal domain of histone H1t (C-H1t) bound to DNA. The symbols of the different structural motifs are shown in the inset. Panel A: melting in 10 mM NaCl. Panel B: melting in 140 mM NaCl. The amide I bands of C-H1t at 20°C (continuous line) and 80°C (dotted line) are shown superimposed, both in 10 mM (panel C) and 140 mM NaCl (panel D). The DNA contribution to the spectra of the complexes was subtracted as described in Materials and Methods.

Between 78°C and 80°C, they decreased sharply down to a residual 2-3%, which may disappear at higher temperatures. Concomitantly with the decrease of the  $\alpha$ -helix and the  $\beta$ -sheet, the random coil component increased from 25% to 45%. The low frequency  $\beta$ -sheet ( $\sim 1620\text{ cm}^{-1}$ ) melted along a wide temperature interval with a transition midpoint of  $\sim 60^\circ\text{C}$ ; an equivalent increase of turns was observed coinciding with the melting of the low frequency  $\beta$ -sheet (Figure 39A,C).

In 140 mM NaCl, the structure of the DNA-bound protein was even more stable, the proportions of  $\alpha$ -helix,  $\beta$ -sheet and random coil remaining essentially constant up to 80°C. A small decrease of the low frequency  $\beta$ -sheet and a concomitant increase of the turn component was observed (Figure 39B,D).

### **Discussion**

We have examined the secondary structure of the C-terminal domain of linker histone subtypes H1<sup>o</sup> and H1t, free in solution and bound to DNA. The use of the entire C-terminal domain is important in the study of properties that are not localized to any specific subdomain<sup>178</sup>, but rather determined by the whole domain, such as differential affinity<sup>121</sup> or apoptotic nuclease stimulation<sup>189</sup>.

The IR spectra of the C-terminal domain in aqueous solution were dominated by the random coil ( $1643\text{-}1644\text{ cm}^{-1}$ ) with 50% of the total amide I intensity. However, even in this most highly unfolded state, an intense band (35% in CH1<sup>o</sup> and 29% in CH1t) at  $1662\text{ cm}^{-1}$  was observed, suggesting the presence of short-lived turn-like conformations in rapid equilibrium with extended forms<sup>173, 186</sup>. Small amounts of local structure consisting of turns at  $1670\text{ cm}^{-1}$  (6-10%) and  $\beta$ -structure (8-9%) and were also present in the free protein.

Upon interaction with DNA, the C-terminal domain acquired an extensively folded conformation. The structure of the bound domain showed a significant dependence on salt concentration. In 10 mM NaCl the DNA-bound C-H1<sup>o</sup> contained  $\alpha$ -helix (19%),  $\beta$ -structure (22%), turns (27%) and open loops (25%), together with a residual 7% of random coil. In physiological salt (140 mM), the C-terminus became fully structured, as indicated by the complete absence of random coil. Under these conditions, CH1<sup>o</sup> contained 24% of  $\alpha$ -helix, 25% of  $\beta$ -structure, 18% of open loops and 33-34% of turns. The component at  $\sim 1660\text{ cm}^{-1}$  in TFE and in the complexes with DNA could arise totally or in part from  $3_{10}$  helix. This

position has been assigned to  $3_{10}$  helix in several studies<sup>323, 324, 325</sup>. The presence of  $3_{10}$  helix was demonstrated in a peptide from the H1° C-terminal domain (residues 99-121)<sup>158, 186</sup>. The amount of  $\alpha$ -helix should thus be considered as a lower limit of the amount of helical structure. All the structural motifs found in the C-terminal domain have the potential to interact with the major or minor groove DNA phosphates<sup>326</sup>. The dependence of the secondary structure of the DNA-bound C-terminal domain on salt concentration, although moderate, is an indication of conformational versatility, which could have a role in chromatin structure and dynamics.

In spite of the large sequence divergence between C-H1° and C-H1t (30% sequence identity) (Figure 1), the two appear to have common secondary structure components, both in 10 mM and 140 mM NaCl. This suggests that selective constraints may act to preserve the secondary structure of the DNA-bound domain in different subtypes.

The induced secondary structure of the DNA-bound domain appears to be extremely stable. In 10 mM NaCl, the  $\alpha$ -helix and the  $\beta$ -sheet ( $1630\text{ cm}^{-1}$ ) begin to melt at  $78^{\circ}\text{C}$ ; only the low frequency  $\beta$ -sheet ( $1620\text{ cm}^{-1}$ ) melts earlier with a transition midpoint of  $\sim 65^{\circ}\text{C}$ . In 140 mM, the  $\alpha$ -helix and the  $\beta$ -sheet are even more stable and no changes are observed in their proportions up to the experimental temperature limit of  $80^{\circ}\text{C}$ . The low frequency  $\beta$ -sheet melts in a temperature interval similar to that observed at low salt, with a concomitant increase in the proportion of turns. The high thermal stability of the DNA-bound form may result from increased hydrophobicity following extensive compensation of the positive charge of the protein by the DNA phosphates. Considering the high proportion of Lys residues in the C-terminus ( $\sim 40\%$ ), charge neutralization should have a large effect on the mean complex hydrophobicity. Evidence for extensive reciprocal charge compensation in C-H1/DNA complexes is given by the +/- ratio of  $\sim 1.0$  in saturated complexes<sup>310, 317</sup>. In addition, CH1/DNA complexes form the kind of toroidal complexes predicted by counterion condensation theory when the DNA charge is fully neutralized ( $>90\%$  charge compensation)<sup>310, 327, 328</sup>.

Folding could contribute to hydrophobicity in a more specific way. In the CH-1 peptide, helical folding (from Glu99 to Ala117) creates a hydrophobic patch due to the marked amphipathic character of the helix, with all positively charged residues on one face of the helix and all the hydrophobic residues on the other face<sup>158</sup>.

The C-terminal domain appears to belong to the group of “disordered proteins”, also known as natively unfolded or intrinsically unstructured<sup>329, 330, 331</sup>. In free solution, these proteins lack specific secondary or tertiary structure and are composed of an ensemble of conformations. Many disordered proteins fold into stable secondary or tertiary structures upon binding to their targets, such as other proteins, nucleic acids or membranes. Intrinsic disorder is characterised by a low sequence complexity, a low content of hydrophobic amino acids and a high content of polar and charged amino acids. All these features are present in the C-terminal domains and also in the N-terminal domains of H1<sup>77</sup>. In addition, disordered regions are often present in proteins regulated by phosphorylation, which is also the case with H1<sup>215</sup>.

The H1 globular domain is believed to be involved in specific binding site recognition in chromatin. A stably folded C-terminal domain with its high charge density could interfere with binding site recognition. Given that folding of the C-terminus is associated with charge neutralization, binding to unspecific sites, such as nucleosomal DNA that is already involved in interaction with core histones, may lead to partial folding and lower affinity of the C-terminus. One of the possible advantages of intrinsic disorder would thus be to kinetically favour the recognition of specific H1 binding sites in chromatin.





### **Capítulo 3**

*Estructura secundaria del dominio C-terminal de la histona H1 inducida por agentes aglomerantes*



## Introduction

H1 linker histones are involved in chromatin structure and gene regulation. It is currently accepted that histone H1 could have a regulatory role in transcription through the modulation of chromatin condensation. H1 may mediate transcription on a more specific level, participating in complexes that either activate or repress specific genes<sup>89, 130, 222, 236, 247, 284, 332</sup>. Binding to scaffold-associated regions<sup>130</sup> and participation in nucleosome positioning<sup>287</sup> are other mechanisms by which H1 could contribute to transcriptional regulation.

H1 contains three distinct domains: a short amino-terminal domain (20-35 amino acids), a central globular domain (~80 amino acids), consisting of a three-helix bundle and a  $\beta$ -hairpin, and a long carboxy-terminal domain (~100 amino acids)<sup>301</sup>. The amino acid composition of the C-terminus is dominated by Lys (~40%), Ala (~17%) and Pro (~12%). The C-terminal domain is the primary determinant of H1 binding to chromatin *in vivo*<sup>121, 182</sup>. Several studies indicate that the ability of linker histones to stabilize chromatin folding resides in the C-terminal domain of the molecule<sup>172, 178</sup>. Preferential binding of histone H1 to scaffold-associated regions and activation of apoptotic nuclease also appear to be determined by the C-terminal domain<sup>189, 317</sup>.

In aqueous solution, the C-terminal domain is dominated by the random coil and turn-like conformations in rapid equilibrium with the unfolded state, but upon interaction with DNA it folds cooperatively<sup>333</sup>. The DNA-bound structure is extremely stable and includes  $\alpha$ -helix,  $\beta$ -sheet, turns and open loops. The H1 carboxyl-terminal domain thus appears to belong to the so-called intrinsically disordered proteins undergoing coupled binding and folding<sup>329, 331, 333, 334, 335, 336</sup>. Like most natively disordered proteins, the C-terminus is of low hydrophobicity, large net charge and low sequence complexity<sup>77, 337, 338</sup>.

Proteins are usually studied in dilute solution; however, in the cellular environment macromolecules and small molecule solutes are present at high concentrations so that a significant fraction of the intracellular space is not available to other macromolecules. The excluded volume effects are predicted to favor the adoption of compact as opposed to expanded macromolecular conformations, resulting in a reduction of the total excluded volume. However, experimental evidence has shown that although crowding agents in

general favor refolding<sup>339, 340, 341, 342</sup>, in some cases they appear to be ineffective<sup>343, 344</sup>. A role of molecular crowding in accelerating  $\alpha$ -synuclein fibrillation has been described<sup>345, 346</sup>. In the present study we have used IR spectroscopy to estimate the proportions of secondary structure motifs of the C-terminus of H1<sup>o</sup> (C-H1<sup>o</sup>) and H1t (C-H1t) in the presence of Ficoll 70 and Polyethyleneglycol (PEG) 6000 as macromolecular crowding agents. Our results show that crowding is highly effective in promoting folding of the C-terminal domain. In the presence of 30% PEG or Ficoll the proportions of secondary structure motifs appear to be similar to those of the DNA-bound domain. In these conditions, Kratky plots of the X-ray scattering indicate that the compactness of the C-terminus is that of a globular state. However, binding of 1-anilinonaphthalene-8-sulfonate (ANS), thermal melting and differential scanning calorimetry (DSC) show that in a crowded environment the C-terminus is not cooperatively folded. The properties of the C-terminus in crowding agents are thus those of a molten globule state.

### **Materials and methods**

#### Cloning, expression and purification of the C-terminal domains of histone H1 subtypes

The C-terminal domains of histones H1<sup>o</sup> and H1t were obtained from recombinant *Escherichia coli* (M15) as described previously<sup>333</sup>.

#### Circular dichroism spectroscopy

The C-terminal domains were at 0.3 mg/ml in 10 mM phosphate buffer, pH 7.0, plus 140 mM NaCl and in the presence of PEG 6000 or Ficoll 70. Spectra were obtained on a Jasco J-717 spectrometer in 1 mm cells at 20°C. The results were analyzed with Standard analysis software (JACSO) and expressed as mean residue molar ellipticity  $[\theta]$ . The helical content was estimated from the ellipticity value at 222 nm ( $\theta_{222}$ ), according to the empirical equation of Chen *et al.*<sup>318</sup>: % helical content =  $100[\theta_{222}/-39,500 \times (1-2.57/n)]$ , where n is the number of peptide bonds.

## Infrared spectroscopy measurements

The C-terminal domains were at 5 mg/ml in 10 mM HEPES, pH 7.0, plus 140 mM NaCl and in the presence of PEG 6000 or Ficoll 70 at 300 g/L. Data was collected on an FT600 Bio-Rad spectrometer equipped with an MCT detector. Typically, 1000 scans for each background and sample were collected and the spectra were obtained with a nominal resolution of  $2\text{ cm}^{-1}$  at  $20^\circ\text{C}$ . Secondary structure content was determined by curve fitting to the original spectrum using the component band positions identified by Fourier self-deconvolution as previously described<sup>186, 319</sup>. The baseline contributed by the solvent (10 mM HEPES, 140 mM NaCl, pH 7.0 in  $\text{H}_2\text{O}$  or  $\text{D}_2\text{O}$  and with or without PEG 6000 or Ficoll 70) was removed before fitting. Spectra were recorded both in  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$  to distinguish between the overlapping contributions in  $\text{D}_2\text{O}$  of random coil and open loops. The vibrations of loops and the  $\alpha$ -helix are found at similar positions in  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$ , the  $\alpha$ -helix at about  $1652\text{ cm}^{-1}$  and loops at about  $1643\text{ cm}^{-1}$ . In contrast, deuteration has a major effect on the position of the random coil so that in  $\text{H}_2\text{O}$  it overlaps with the  $\alpha$ -helix, whereas in  $\text{D}_2\text{O}$  it overlaps with loops. The  $\alpha$ -helix was, therefore, estimated directly in  $\text{D}_2\text{O}$ , whereas loops were estimated directly in  $\text{H}_2\text{O}$ . When both  $\alpha$ -helix and loops are present in the protein structure, the percentage of random coil can be obtained either from the difference of the components around  $1552\text{ cm}^{-1}$  in  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$  or from the difference of the components around  $1643\text{ cm}^{-1}$  in  $\text{D}_2\text{O}$  and  $\text{H}_2\text{O}$ , as previously described<sup>333</sup>.

## Small-angle X-ray scattering

Measurements were performed in a MBraun instrument equipped with a Siemens Kristalloflex 760 (K-760) generator, producing a radiation with a wavelength of  $1.54\text{ \AA}$  ( $\text{CuK}\alpha$ ) operating at 50 kV and 40 mA, and a Kratky Hecus camera. The collimator was a slit window and the scattering was detected with a linear position sensitive detector OED-50M. The sample to detector distance was 268 mm. Samples were measured at  $25^\circ\text{C}$ . All samples were placed in glass capillaries of 1 mm diameter and  $10\text{ }\mu\text{m}$  wall thickness. The background was subtracted by measuring blanks without protein. Both the samples and blanks were measured for 4 hours, unless otherwise stated. The concentration of protein was 4-10 mg/ml.

The scattering intensity  $I(Q)$  is represented as a function of the momentum transfer  $Q$ ;  $Q=4\pi\sin\theta/\lambda$ , where  $\lambda$  is the wavelength of X-rays and  $2\theta$  the scattering angle. The inner part of the scattering profile can be described by the Guinier approximation<sup>347</sup>,  $I(Q) = I_0 e^{-Q^2 R_g^2/3}$ . The scattering at zero angle,  $I_0$ , is proportional to the molecular mass of the particle and the square of the contrast between the particle and the solvent. The radius of gyration,  $R_g$ , which is the root-mean-square of the distances of all volume elements to the center-of-mass of the electronic volume of the particle, was obtained from the slope of a Guinier plot of  $\ln I(Q)$  vs.  $Q^2$ . The  $R_g$  can also be measured with the Debye equation<sup>348</sup>:  $P_D(x) = I(Q)/I_0 = 2(x-1+e^{-x})$ , where  $x = Q^2 R_g^2$ , using the approximation given by Calmettes et al.<sup>349</sup>:  $[P_D(x)]^{-1} = 1 + 0.359 x^{1.103}$ . The  $R_g$  is given by  $R_g = (a/0.359b)^{0.453}$ , where  $a$  is the slope and  $b$  the intercept of the straight line.

### Thermal melting

For the analysis of the thermal stability of C-H1t in 30% Ficoll 70 the temperature of the sample was adjusted using a cover jack connected to a circulating thermostatic bath and monitored with a fitted external probe. Thermal analyses were performed by heating from 20 to 85°C at a rate of 1°C/min. For each degree of temperature interval, 400 interferograms were averaged, Fourier-transformed and ratioed against background. Thermal analyses were performed in D<sub>2</sub>O. The Ficoll contribution at each temperature was subtracted using a Ficoll sample at the same temperature and concentration.

### ANS fluorescence assays

1-anilinonaphthalene-8-sulfonate (ANS) binding assays were performed in a Cary Eclipse spectrofluorometer. The spectra were measured in 140 mM NaCl, 10 mM phosphate buffer, pH 7.0, at 20°C. The concentration of ANS was 0.3 mM and the concentration of protein 10 mM. The ANS emission was scanned between 400 and 650 nm with an excitation wavelength of 380 nm.

### Differential scanning calorimetry

DSC measurements were carried out on a Microcal MC-2 calorimeter supplied with an Origin software package for data analysis and curve fitting. Protein and background scans

were performed at a scan rate of 1.5°C /min between 20 and 100°C. We used RNase A as a control at a concentration of 150  $\mu$ M. C-H1t was analyzed at a concentration of 100  $\mu$ M in buffer and in buffer plus 30% Ficoll 70. The buffer was 10 mM phosphate, pH 7.0, plus 140 mM NaCl.

## Results

### CD and IR spectroscopy analysis of the C-terminal domain of histone H1 in the presence of crowding agents

We used CD to explore the effects of different concentrations of Ficoll 70 and PEG 6000 (10, 20 and 30%) on the secondary structure of the C-terminal domains of the histone H1 subtypes H1<sup>o</sup> and H1t in physiological salt (0.14 M NaCl) (Fig. 40). We found that in 30% of both crowding agents C-H1<sup>o</sup> and C-H1t became significantly folded. In 30% PEG, the estimated values of  $\alpha$ -helix were 16% for C-H1<sup>o</sup> and 13% for C-H1t. In 30% Ficoll, the estimated helical content was 16% for both C-terminal domains.

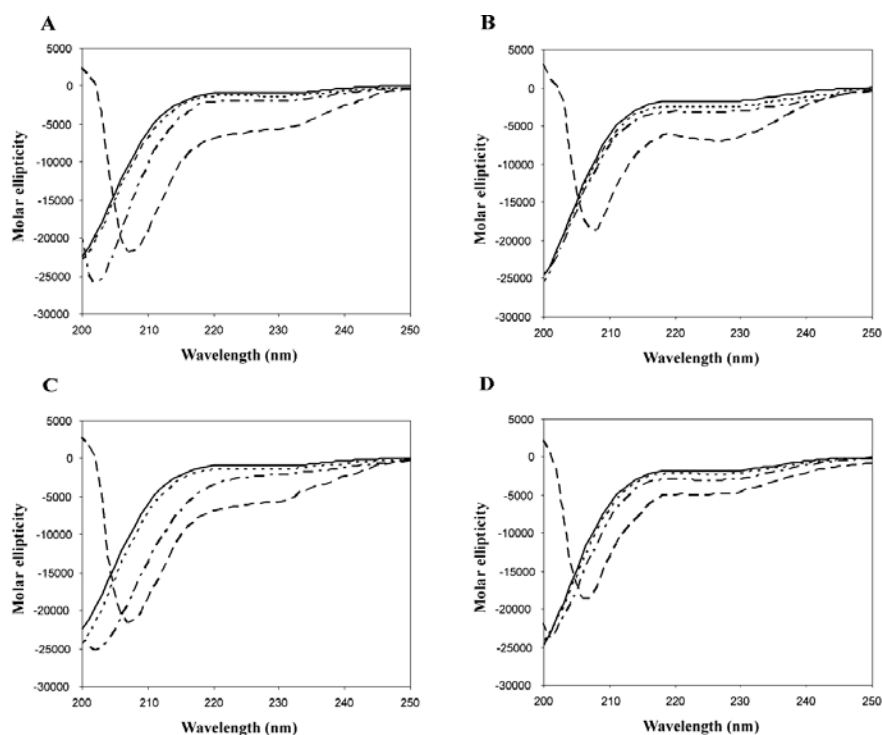


Figure 40. Conformational transition of the C-terminal domain of histone H1 measured by CD. C-H1<sup>o</sup> in Ficoll (A) and PEG (C). C-H1t in Ficoll (B) and PEG (D). The buffer was 10 mM phosphate, pH 7.0, plus 140 mM NaCl. Continuous line: C-terminal domains in buffer; dotted line: buffer plus 10% PEG or Ficoll; dashed-dotted line: buffer plus 20% PEG or Ficoll; dashed line: buffer plus 30% PEG or Ficoll.

IR spectroscopy was used to further characterize the secondary structure of the C-terminal domains in crowding agents. In the absence of crowding agents, the amide I (D<sub>2</sub>O) of both C-H1° and C-H1t was dominated by the unordered structure (at 1642 cm<sup>-1</sup> in D<sub>2</sub>O) with ~50% of the total intensity. The other main component (at 1663 cm<sup>-1</sup>), with 28-31 % of the total area in both C-H1° and C-H1t, was assigned to turns in rapid equilibrium with the unfolded state. Well defined structure was represented by a 9-12% of  $\beta$ -sheet (at 1626-1630 cm<sup>-1</sup>) (Fig. 41 and Table 8).

<b>Table 8</b>										
<b>Percentages (%) of secondary structure of the carboxy-terminal domain of histones H1° (C-H1°) and H1t (C-H1t) in presence of crowding agents</b>										
Assignment	Buffer*		Ficoll <sup>†</sup>				PEG <sup>‡</sup>			
	D <sub>2</sub> O		D <sub>2</sub> O		H <sub>2</sub> O		D <sub>2</sub> O		H <sub>2</sub> O	
	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%
<b>C-H1°</b>										
turns	1676	3	1685	1	1685	3	1680	1	1680	7
turns	1661	31	1667	26	1668	21	1668	15	1668	10
$\alpha$ -helix			1655	15			1656	19		
$\alpha$ -helix/random coil					1655	15			1656	19
open loops					1643	23			1642	31
random coil/open loops	1642	52	1644	23			1645	31		
random coil <sup>‡</sup>				0				0		
random coil <sup>¶</sup>						0				0
$\beta$ -sheet	1630	9	1635	14	1634	14	1635	18	1633	12
$\beta$ -sheet			1626	13	1625	13	1626	10	1623	11
Low frequency $\beta$ -sheet	1617	5	1616	8	1616	11	1616	6	1615	10
<b>C-H1t</b>										
turns	1685	2	1684	1	1681	8	1681	1	1681	8
turns	1674	4	1668	29	1668	14	1666	23	1666	16
turns	1663	28								
$\alpha$ -helix			1656	15			1655	14		
$\alpha$ -helix/random coil					1656	15			1656	14
open loops					1645	23			1643	27
random coil/open loops	1643	53	1645	23			1645	27		
random coil <sup>‡</sup>				0				0		
random coil <sup>¶</sup>						0				0
$\beta$ -sheet			1637	16	1635	15	1634	17	1634	13
$\beta$ -sheet	1626	12	1628	11	1625	16	1625	13	1624	11
Low frequency $\beta$ -sheet	1616	1	1616	5	1615	9	1615	5	1615	11

\* Buffer: 10mM HEPES, 140 mM NaCl, pH 7.0.  
<sup>†</sup> The concentration of Ficoll 70 and PEG 6000 was 30%.  
<sup>‡</sup> The value corresponds to the difference between the components at ~ 1643 in D<sub>2</sub>O and H<sub>2</sub>O  
<sup>¶</sup> The value corresponds to the difference between the components at ~ 1655 in H<sub>2</sub>O and D<sub>2</sub>O



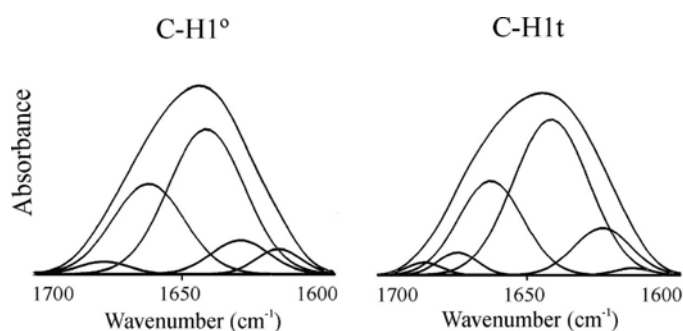


Figure 41. Amide I decomposition of the spectra of the C-terminal domain of histone H1. C-H1°, carboxy-terminal domain of histone H1°. C-H1t, carboxy-terminal domain of histone H1t. The spectra were recorded in D<sub>2</sub>O at 20°C. The buffer was 10 mM HEPES plus 140 mM NaCl, pH 7.0. The protein concentration was 5 mg/ml.

Addition of either PEG or Ficoll forced the C-terminus to fold. In 30% PEG 6000, deconvolution of the amide I of C-H1° gave 19%  $\alpha$ -helix (1656 cm<sup>-1</sup>), 28%  $\beta$ -sheet (1635, 1626 cm<sup>-1</sup>), 16% turns (1680, 1668 cm<sup>-1</sup>) and 31% open loops (1642 cm<sup>-1</sup>) (Fig. 42 and Table 8).

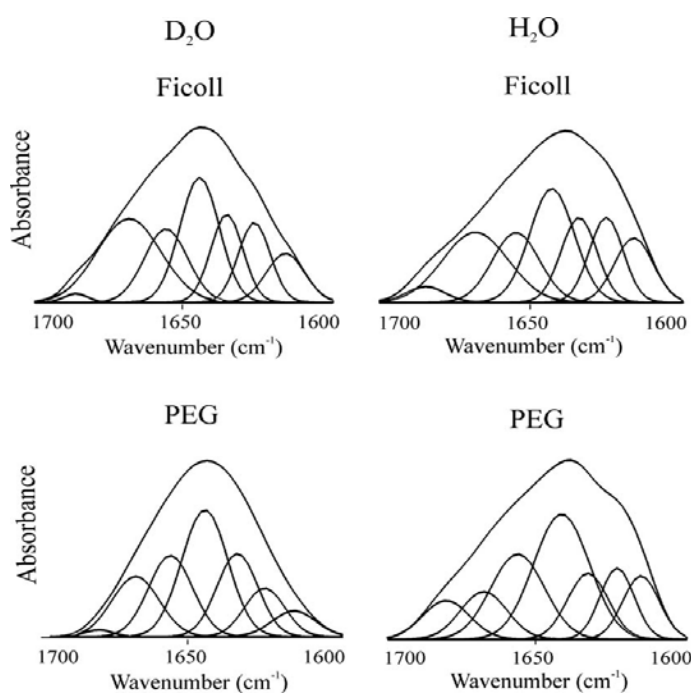


Figure 42. Amide I decomposition of the spectra of C-H1° in the presence of macromolecular crowding agents. The left column shows the spectra measured in D<sub>2</sub>O and the right column shows those measured in H<sub>2</sub>O. The buffer was 10 mM HEPES plus 140 mM NaCl, pH 7.0, at 20°C. The protein concentration was 5 mg/ml. The concentration of Ficoll 70 and PEG 6000 was 30%.

Spectra were recorded both in H<sub>2</sub>O and D<sub>2</sub>O to distinguish between the overlapping contributions in D<sub>2</sub>O of random coil and open loops as described under “Materials and

Methods”. The proportions of the different secondary structure motifs in C-H1t in the presence of PEG were similar to those found in C-H1°, although slight differences were present: a lower proportion of  $\alpha$ -helix (14%) and open loops (27%) and a higher proportion of turns (24%); the proportion of  $\beta$ -sheet was very similar (30%) (Fig. 42 and 43). The amount of random coil was less than 1% in both C-H1° and C-H1t. The results in Ficoll 70 were similar to those found in PEG (Fig. 42 and 43, Table 8).

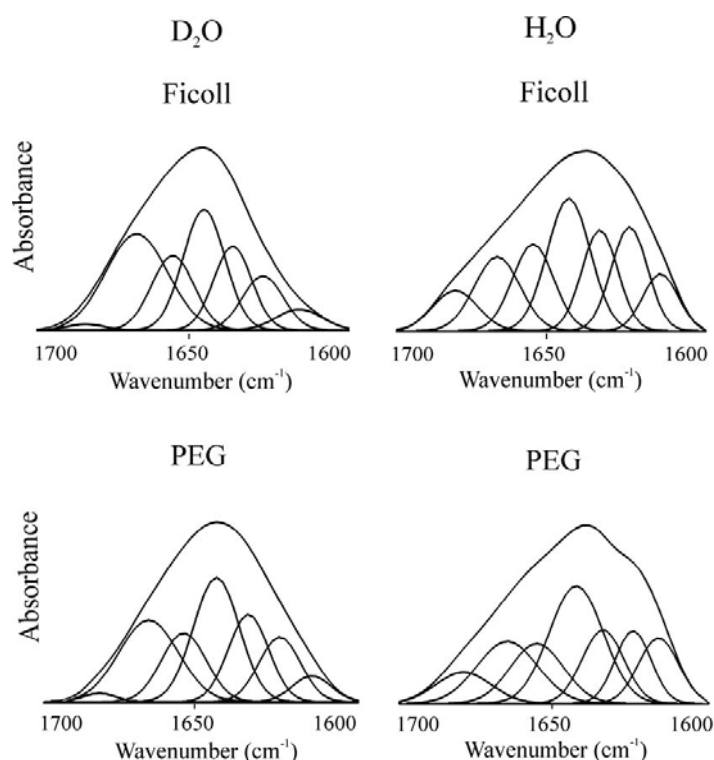


Figure 43. Amide I decomposition of the spectra of C-H1t in the presence of macromolecular crowding agents. The left column shows the spectra measured in D<sub>2</sub>O and the right column shows those measured in H<sub>2</sub>O. The buffer was 10 mM HEPES plus 140 mM NaCl, pH 7.0, at 20°C. The protein concentration was 5 mg/ml. The concentration of Ficoll 70 and PEG 6000 was 30%.

### Small-angle X-ray scattering

The conformation of the C-terminus was examined by small-angle X-ray scattering (SAXS). The changes in compaction can be qualitatively analyzed by means of Kratky plots ( $Q^2I(Q)$  vs.  $Q$ ). The scattering curve in Kratky plots has a characteristic bell-shape when the protein is in the native state or in the molten globule state. Such a maximum is absent in completely unfolded proteins or in premolten globule states<sup>334, 350, 351</sup>. In 30% Ficoll, the C-terminal domain of H1t showed the characteristic maximum of a well developed globular structure (Fig. 44A). The position of the maximum ( $0.1 \text{ \AA}^{-1}$ ) is that expected for the molecular mass

of the C-terminus ( $M_r = 11082$ )<sup>351</sup>. The Guinier plot gave an  $R_g$  of  $14.8(\pm 0.7)$  Å (Fig. 44B). In the absence of crowding agents the Kratky plot lacked the maximum of globular conformations and exhibited the monotonic increase indicative of an expanded coil-like conformation (Fig. 44A).

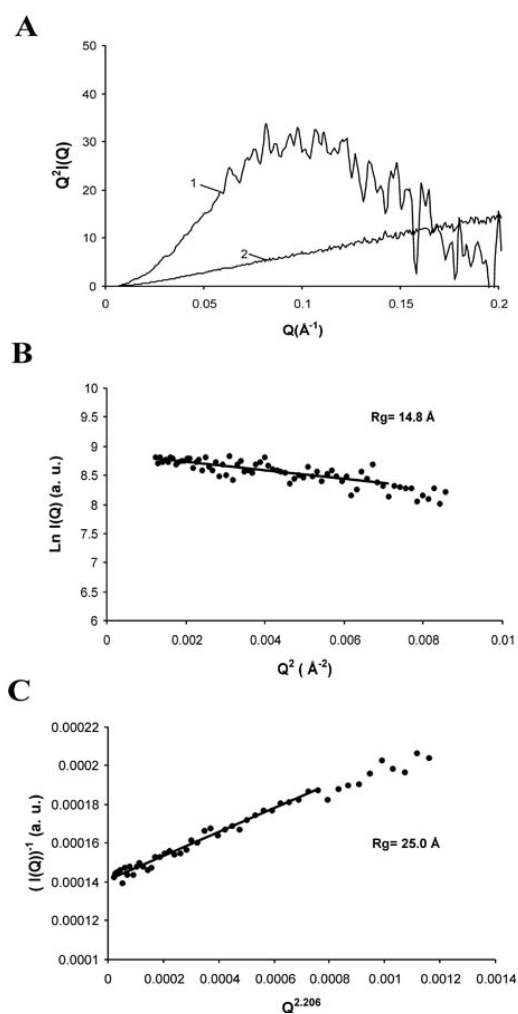


Figure 44. Small-angle X-ray scattering of the C-terminal domain of histone H1t. (A) Kratky plots ( $Q^2 I(Q)$ ) as a function of  $Q$  in 30% Ficoll (1) and in buffer solution (2). (B) Guinier plot ( $\text{Ln } I(Q)$ ) as a function of  $Q^2$  of the protein in 30% Ficoll. (C) Representation of the X-ray scattering of the protein in buffer solution using the approximation to the Debye function ( $(I(Q))^{-1}$  as a function of  $Q^{2.206}$ );  $Q$  is in Å<sup>-1</sup>. The  $Q$  ranges used for the fits were  $0.035 < Q < 0.084$  for the protein in Ficoll and  $0.0074 < Q < 0.0385$  for the protein in buffer solution.

The  $R_g$  in buffer solution was estimated with the Debye equation, which is more appropriate than the Guinier approximation in the case of unfolded proteins<sup>349, 352</sup>. In these conditions, the  $R_g$  was  $25.0(\pm 0.2)$  Å (Fig. 44C). Although this value is considerably higher than the  $R_g$  in crowded conditions, it is smaller than the value of  $\sim 31.5$  Å that can be estimated for

random coils with the expression:  $R_g^U = (2.07 \pm 0.18)N^{(0.585 \pm 0.018)}$ , where N is the number of amino acid residues<sup>337, 353</sup>. This indicates that the C-terminus in crowded conditions is more compact than an unperturbed random coil polypeptide chain. This is consistent with the IR results, showing the presence of a residual proportion of turns and  $\beta$ -sheet in buffer solution. Turns may not contribute to a tightly folded core, but may reduce the dimensions of the conformational ensemble.

### ANS binding

The C-terminal domains were tested for 1-anilinonaphthalene-8-sulfonate (ANS) binding in either the absence or presence of PEG and Ficoll (Fig. 45). This hydrophobic dye has been used as a probe for the formation of collapsed, partially folded states such as the premolten and molten globules<sup>354, 355</sup>. In the absence of crowding agents the emission of ANS was not affected by the presence of the C-terminus. In contrast, in Ficoll, the  $\lambda_{\max}$  of fluorescence of protein-bound ANS shifted from 525 to 485 nm and the fluorescence intensity increased ~11-fold. In PEG, a shift from 525 to 501 nm and a ~7-fold increase in fluorescence intensity were observed. (Fig. 45A, B)

### Thermal melting and DSC

To better understand the features of the folded states of the C-terminus in crowding agents, we analyzed the thermal melting by IR spectroscopy and the change in partial heat capacity (Cp) with temperature by DSC of C-H1t in 30% Ficoll. The changes in the amide I (D<sub>2</sub>O) components were monitored in the 20-85°C interval. The  $\alpha$ -helix and the  $\beta$ -sheet were lost fairly linearly as temperature was increased; all defined structure was lost by 85°C. Concomitantly with the decrease in  $\alpha$ -helix and  $\beta$ -sheet, the random coil component increased from 23 to 56% (Fig. 45C).

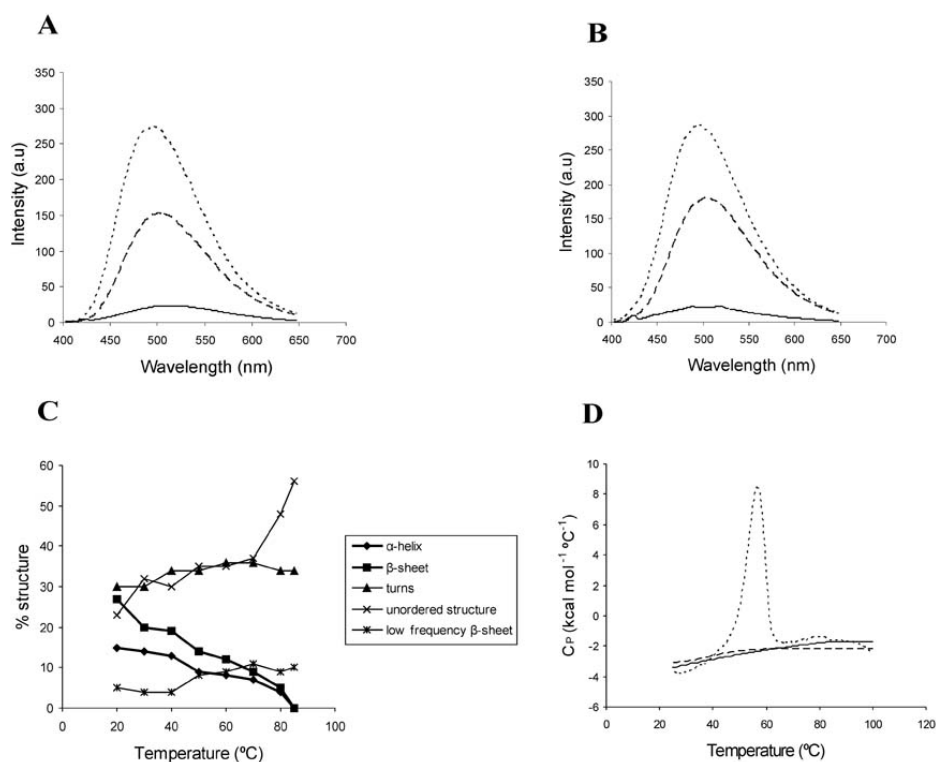


Figure 45. Effects of macromolecular crowding on the C-terminal domain of histone H1. (A) Fluorescence emission spectra of 0.3 mM ANS in the presence of C-H1<sup>o</sup> without crowding agent (continuous line) and with 30% PEG 6000 (dashed line) or 30% Ficoll 70 (dotted line). (B) Fluorescence emission spectra of 0.3 mM ANS in the presence of C-H1t without crowding agent (continuous line) and with 30% PEG 6000 (dashed line) or 30% Ficoll 70 (dotted line). (C) Thermal profiles of the secondary structure motifs obtained from amide I (D<sub>2</sub>O) decomposition in 140 mM NaCl, 30% Ficoll. (D) Temperature dependence of the partial heat capacity (C<sub>p</sub>) in the absence (dashed line) and in the presence of 30% Ficoll 70 (continuous line). The DSC curve of ribonuclease A is shown for comparison (dotted line). Binding, thermal denaturation and DSC experiments were performed in 10 mM phosphate buffer, pH 7.0, plus 140 mM NaCl.

In parallel with the dissipation of secondary structure, the C<sub>p</sub> increased almost linearly between 25 and 85°C, where it leveled off coinciding with the complete dissipation of the α-helix and β-sheet (Fig. 45D). Cooperatively folded proteins have significant enthalpy difference between their native and unfolded states which lead to large heat capacity peaks corresponding to heat absorption upon denaturation<sup>356</sup>. Thermal melting and C<sub>p</sub> measurements indicate that despite globular compaction and native-like secondary structure the C-terminus is not cooperatively folded and, instead, has features characteristic of the molten globule state<sup>357, 358, 359, 360</sup>, which possesses an extensive secondary structure but does not have a fixed tertiary structure.

## Discussion

The C-terminal domain of histone H1 is mostly unstructured in dilute solution, but extensively folds upon interaction with DNA. The C-terminal domain of histone H1 can thus be considered an intrinsically disordered protein<sup>329, 331</sup>. Some intrinsically disordered proteins may adopt compact conformations in crowded environments<sup>344</sup>. The C-terminal domain appears to belong to this kind of intrinsically disordered proteins. In the presence of Ficoll 70 or PEG 6000, the C-terminal domain becomes extensively structured with 14-19%  $\alpha$ -helix, 22-31%  $\beta$ -sheet, 16-26% turns and 20-30% loops, depending on the subtype and crowding agent. Folding is associated with the differentiation of a limited hydrophobic core as shown by the increased ANS binding by the C-terminus in crowded conditions. Interestingly, the high positive charge of the C-terminus (about 40% Lys) does not impair folding in crowded conditions.

The proportions of the secondary structure motifs of the C-terminus in crowded conditions are comparable to those of the DNA-bound domain<sup>333</sup>, although the  $\alpha$ -helix was proportionally more abundant in the DNA complexes and the  $\beta$ -sheet more abundant in crowded conditions (Table 9).

<b>Table 9</b>										
<b>Percentages of secondary structure of the C-terminal domain in crowded conditions and bound to DNA</b>										
	<b>Complexes with DNA*</b>				<b>Ficoll</b>		<b>PEG</b>			
	<b>C-H1°</b>		<b>C-H1t</b>		<b>C-H1°</b>	<b>C-H1t</b>	<b>C-H1°</b>	<b>C-H1t</b>		
Turns/ $3_{10}$ helix	33-34		40-42		24-27		22-30	16-17	24	
$\alpha$ -helix	24		21		15		15	19	14	
Open loops	18		15		23		23	31	27	
$\beta$ -sheet	16		13		27		27-31	22-28	24-30	
Random coil	<1		<1		<1		<1	<1	<1	

\* This data was taken from reference<sup>333</sup>.

Despite the low sequence identity between C-H1° and C-H1t (~30%), the proportions of their secondary structure motifs appear to be similar as also occurs in the DNA-bound domains<sup>333</sup>. This suggests that selective constraints may act to preserve the secondary structure of the C-terminus in the different subtypes.

The compactness of the C-terminus in the presence and in the absence of crowding agents was examined using SAXS. Global shape changes were analyzed by means of Kratky plots ( $Q^2I(Q)$  vs.  $Q$ ). The Kratky representation has a characteristic maximum in globular states (native and molten globule). This maximum is absent in non globular states (random coil, premolten globule)<sup>334</sup>. In the presence of crowding agents the Katky plot of the X-ray

scattering by the C-terminus of H1t in 30% Ficoll exhibited the bell-shaped appearance typical of globular states. In the absence of crowding agents the Kratky plot increased monotonously as observed for many unfolded proteins. The decrease of the  $R_g$  from 25.0 Å in buffer solution to 14.8 Å in crowding agents, confirms the condensation of the C-terminus in crowded conditions.

Analysis of the folded state of the C-terminus in crowding agents by thermal melting and DSC reveals the progressive dissipation of secondary structure in the temperature interval 20-85°C, which is paralleled by a lineal increase in partial heat capacity, lacking the heat absorption peak typical of cooperatively folded proteins. These features indicate that the C-terminus in crowded conditions is not cooperatively folded despite native-like secondary structure and compaction, and has, instead, the characteristics of the molten globule state<sup>357, 358, 359, 360, 361</sup>. We have previously shown that upon interaction with DNA, the secondary structure of the protein becomes extremely stable and melts cooperatively<sup>333</sup>. The transition from the molten-globule state to the DNA-bound cooperatively folded state is probably driven by increased hydrophobicity, following charge compensation of the positive charge of the abundant Lys residues of the C-terminus by the DNA phosphates.

The H1 globular domain is thought to be responsible for the location of H1 by interacting with a specific site in the nucleosome. The dependence of the tightening of the C-terminus molten globule structure on charge compensation may afford a mechanism to avoid the interference of the strong electrostatic binding of the C-terminus with binding site recognition by the globular domain, since binding of the C-terminus to unspecific sites, such as nucleosomal DNA, already partially charge-compensated by core histones, may lead to partial folding and lower affinity.

The similarity of the proportions of secondary structure motifs of the domain in crowded conditions and bound to DNA suggests that in the presence of crowding agents the structure approximates that of the DNA-bound domain. An advantage related to chromatin dynamics of the condensation of the chain into a native-like structure could be an increased rate of the transition toward the native DNA-bound state<sup>361</sup>. The folding of the C-terminus in crowded conditions might also facilitate diffusion of H1 inside cell nuclei. This effect would compensate for the slowing down of protein diffusion by crowding<sup>362, 363</sup>. Improved

diffusion could have functional consequences since H1 molecules exchange *in vivo* between chromatin binding sites through a soluble intermediate<sup>180</sup>.

In summary, the molten globule state of the C-terminus and the tightening of the structure upon interaction with DNA could kinetically favor H1 binding and exchange.



## **Capítulo 4**

*Modulación de la estructura secundaria del dominio C-terminal de la histona H1<sup>o</sup> por fosforilación.*



## **Introduction**

H1 linker histones are involved in chromatin structure and gene regulation. H1 may mediate transcription on a more specific level, participating in complexes that either activate or repress specific genes<sup>89, 130, 222, 236, 247, 284</sup>.

H1 contains three distinct domains: a short amino-terminal domain (20-35 amino acids), a central globular domain (80 amino acids) and a long carboxy-terminal domain (100 amino acids)<sup>301</sup>. The C-terminus contributes with a very large proportion to the binding free energy of the molecule<sup>178, 182</sup>. H1 has multiple isoforms. In mammals, six somatic subtypes (designated H1a-e and H1<sup>o</sup>), a male germ line-specific subtype (H1t), and an oocyte-specific subtype (H1oo) have been identified<sup>80, 93, 288, 289</sup>. The subtypes can differ in timing of expression<sup>79</sup>, extent of phosphorylation<sup>290</sup>, turnover rate<sup>97, 291</sup>, binding affinity<sup>121</sup> and evolutionary stability<sup>78</sup>. Differences in DNA condensing capacity have also been reported for some subtypes<sup>292, 293, 294</sup>. Several studies indicate that the ability of linker histones to stabilize chromatin folding resides in the C-terminal domain of the molecule<sup>172, 178, 316</sup>. The C-terminal domain determines the preferential binding to DNA scaffold associated regions (SARs)<sup>317</sup> it is also responsible of the activation of the apoptotic nuclease<sup>189</sup> and the interaction with HP1 $\alpha$ <sup>227</sup>.

In aqueous solution, the C-terminus is dominated by the random coil and turn-like conformation in rapid equilibrium with the unfolded state, but upon interaction with the DNA it folds cooperatively. The DNA bound structure is extremely stable and includes  $\alpha$ -helix,  $\beta$ -sheet, turns and open loops. The H1 carboxy-terminal domain thus appears to belong to the so-called intrinsically disordered proteins undergoing coupled binding and folding<sup>333</sup>. Like most natively unfolded proteins, the C-terminus is of low hydrophobicity, large net charge and low sequence complexity<sup>77</sup>. In the presence of macromolecular crowding agents the C-terminus acquires the properties of a molten globule with native-like secondary structure and compaction<sup>364</sup>.

The phenotypic roles of histone H1 may be determined by complementary and overlapping effects of stoichiometry, subtype composition and post-translational modifications. Phosphorylation of the consensus sequences of cyclin-dependent kinases, (S/T)-P-X-(K/R) is the main post-translational modification affecting histone H1<sup>314</sup>. Most of these sequences are localized to the C-terminal domain. The maximal number of phosphate groups often

corresponds to the number of (S/T)-P-X-(K/R) sites of the molecule. Histone H1 is phosphorylated in a cell cycle-dependent manner<sup>198</sup>. The highest number of phosphorylated sites is found in mitosis, when chromatin is maximally condensed<sup>198, 218</sup>. It is not clear how H1 phosphorylation affects chromatin condensation at interphase and mitosis.

A number of studies indicate that interphase phosphorylation is involved in chromatin relaxation<sup>219, 220, 221</sup>; still, in metaphase chromosomes H1 is highly phosphorylated. In considering these apparently conflicting effects of H1 phosphorylation, it should be noted that hyperphosphorylated H1 is restricted to metaphase and specific properties could be associated to full phosphorylation. In interphase, H1 subtypes are present as a mixture of unphosphorylated and low-phosphorylated species with a proportion of 35-75% of unphosphorylated forms, according to the subtype and cell line and the moment in the cell cycle<sup>198, 219, 365</sup>.

In the present study we have used infrared spectroscopy to describe the effect of phosphorylation on the secondary structure of the C-terminal domain of histone H1<sup>o</sup> in aqueous solution and bound to DNA.

### **Materials and methods**

Cloning, expression and purification of the C-terminal domains.

The C-terminal domains of histones H1<sup>o</sup> and H1t were obtained from recombinant *Escherichia coli* (M15) as described previously<sup>333</sup>.

*In vitro* phosphorylation assay.

The C-terminal domain of histone H1<sup>o</sup> was phosphorylated *in vitro* with CDK2-cyclin A kinase (New England Biolabs). Phosphorylation reactions were carried out in 50 mM Tris-HCl, 10 mM MgCl<sub>2</sub>, 1 mM EGTA, 2 mM dithiothreitol, pH 7.5 plus 200 μM ATP and 1 U of CDK2-cyclin A per 5 μg of C-terminal domain. The mixture was incubated at 30 °C for one hour and the reaction buffer was eliminated by gel filtration on a Hi-Trap desalting column (GE Healthcare). The *in vitro* phosphorylation assay was occasionally performed with <sup>32</sup>P (GE Healthcare). The product of the reaction was then analyzed by SDS-PAGE and autoradiography. The extent of phosphorylation was evaluated by MALDI-TOF mass spectrometry.

### Construction of mutants in CDK2 sites.

The mutant clones for the CDK2 phosphorylation sites for the C-terminal domain of histone H1<sup>o</sup> were obtained by PCR using the QuickChange<sup>®</sup> Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's instructions. The primers used for each phosphorylation site were mutated so that the threonine residue was changed to alanine. For threonine in the position 118 the primers were 5'TCAAGAAAGTGGCAGCTCCAAAGAAGGCA3' and 5'TGCCTTCTTTGGAGCTGCCACTTTCTTGA3'. For the position T140 the primers were 5'AGAAACCCAAAGCCGCCCTGTCAAGAAG3' and 5'CTTCTTGACAGGGGCGGCTTTGGGTTTCT3'. For the position T152 the primers were 5'GCTGCCGCGCCCAAGAAAGCCAAAAAGCC3' and 5'GGCTTTTTGGCTTTCTTGGGCGCGGCAGC3'. Double and triple mutants were obtained in successive rounds of PCR. The introduction of the correct mutation was evaluated by DNA sequencing of the recombinant clones. The recombinant plasmids were named by the position mutated in each case. These clones were expressed and purified as described for the wild-type C-terminal domain<sup>333</sup>.

### MALDI-TOF mass spectrometry.

The MALDI-TOF mass spectrometry was performed in the Servei de Proteòmica i de Bioinformàtica, Universidad Autònoma de Barcelona. The incorporation of each phosphate group was associated with an increase of 80 Da in the molecular mass of the recombinant protein. MALDI-mass spectra were recorded in the positive ion mode on a Bruker Daltonics Ultraflex MALDI-TOF spectrometer.

### IR spectroscopy.

The C-terminal domains of histone H1 were measured at 5 mg/ml in 10 mM HEPES pH 7.0 plus 140 mM NaCl. DNA-protein complexes contained the appropriate amount of salmon DNA for each protein/DNA ratio (w/w). The protein samples at pH 10.2, 11 and 12 did not contain NaCl.

Measurements were performed on a Nicolet Magna II 550 spectrometer equipped with a MCT detector, using a demountable liquid cell with calcium fluoride windows and 50 µm spacers for D<sub>2</sub>O medium and 6 µm spacers for H<sub>2</sub>O medium measurements. Typically, 1000 scans for each background and sample were collected and the spectra were obtained with a

nominal resolution of  $2 \text{ cm}^{-1}$ , at  $22^\circ\text{C}$ . The protein concentration for the  $\text{D}_2\text{O}$  measurements was of  $5 \text{ mg/mL}$  while for the  $\text{H}_2\text{O}$  measurements was  $20 \text{ mg/mL}$ . Data treatment and band decomposition have been described previously<sup>319</sup>. The DNA contribution to the spectra of the complexes with the C-terminal domain was subtracted using a DNA sample of the same concentration; the DNA spectrum was weighted so as to cancel the symmetric component of the phosphate vibration at  $1087 \text{ cm}^{-1}$  in the difference spectra as described in Vila *et al.*<sup>186</sup>.

## Results

### *In vitro* phosphorylation of CDK2 sites.

The C-terminal domain of the histone  $\text{H1}^\circ$  and its mutants were phosphorylated *in vitro* with CDK2. C- $\text{H1}^\circ$  contains three consensus CDK2 sites (Fig. 46A). Phosphorylation was evaluated by the increase in  $80 \text{ Da}$  corresponding to the incorporation of each phosphate group by MALDI-TOF mass spectrometry. The difference between the unphosphorylated (C- $\text{H1}^\circ$ ) and the completely phosphorylated C- $\text{H1}^\circ$  (C- $\text{H1}^\circ\text{abcP}$ ) was  $240 \pm 2 \text{ Da}$ , corresponding to the incorporation of three phosphate groups (Fig. 46B).

A.

DEPKRSVAFKKTKEVKKVA**TPKK**AAKPKKAASKAPSKKP  
 KA**TPVK**KAKKKPAA**TPKK**AKKPKVVKVPKASKPKKAK  
 TVKPKAKSSAKRASKKK

B.

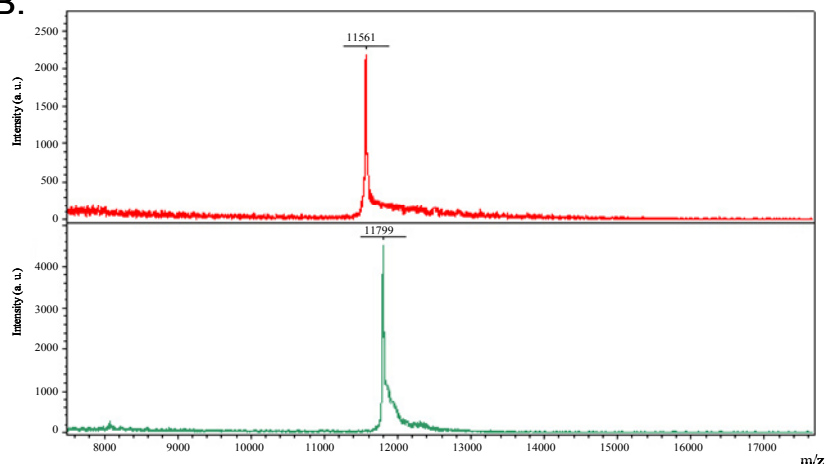


Figure 46. *In vitro* phosphorylation of the C-terminal domain of histone  $\text{H1}^\circ$ . A. Sequence of the C- $\text{H1}^\circ$ , in red the three phosphorylation sites for CDK2. These positions will be referred as *a* for T118, *b* for T140 and *c* for T152. The residues are numbered according to their position in the entire protein. B. Mass spectrometry of C- $\text{H1}^\circ$  (red peak) and triphosphorylated C- $\text{H1}^\circ$  (green peak) (C- $\text{H1}^\circ\text{abcP}$ ).

We obtained T→A mutants of all the CDK2 sites in the C-terminal domain. These mutants were phosphorylated and analyzed by mass spectrometry. The increase in molecular mass of the single mutants was of 160 Da while it was of 80 Da in the double mutants. We also performed a phosphorylation assay with  $^{32}\text{P}$ . The products of the phosphorylation reaction were analyzed by SDS-PAGE and autoradiography (Fig. 47). No signal of  $^{32}\text{P}$  corresponding to the triple mutant with all CDK2 sites mutated was observed. This result confirmed that the CDK2 phosphorylated the predicted sites and that no unspecific phosphorylation occurred.

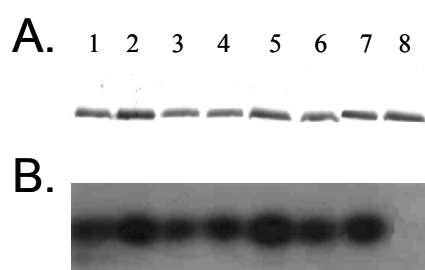


Figure 47. Incorporation of  $^{32}\text{P}$  in the C-H1° mutants of the CDK2 sites after the phosphorylation reaction. A. SDS-PAGE stained with Coomassie blue. B. Autoradiography. C-H1°*abc*P (lane 1), single and double mutants (lanes 2-7) and triple mutant (lane 8).

### Effects of hyperphosphorylation on the structure of the free C-terminal domain of histone H1.

In free solution triphosphorylated C-H1° (C-H1°*abc*P) was mostly unstructured with a 57% of random coil/open loops ( $1642\text{ cm}^{-1}$  in  $\text{D}_2\text{O}$ ) (Fig. 48, Table 10). The main difference with unphosphorylated C-H1° was the increase in the low frequency  $\beta$ -sheet in C-H1°*abc*P from 5 to 11%.

Macromolecular crowding induces a molten globule state in the C-terminal domain of histone H1° with proportions of secondary structure motifs similar to those of the DNA bound domain<sup>364</sup>. Phosphorylation did not alter the proportions of secondary structure motifs in crowded conditions (Ficoll 30%) with only a slight decrease in  $\beta$ -structure from 34% to 28% (Fig 48, Table 10).

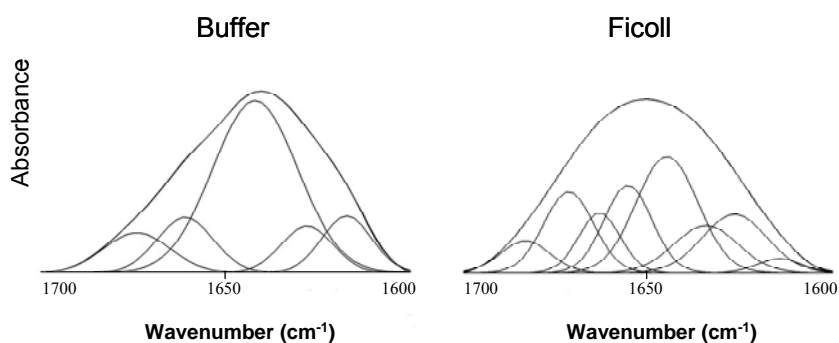


Figure 48. Amide I decomposition of the spectra of hyperphosphorylated C-H1° (C-H1°abcP) in buffer and in Ficoll 70. The spectra were measured in D<sub>2</sub>O. The buffer was 10 mM HEPES plus 140 mM NaCl, pH 7.0, at 20°C. The protein concentration was 5 mg/ml. The concentration of Ficoll 70 was 30%.

TABLE 10								
Percentages (%) of secondary structure by FTIR of the unphosphorylated (C-H1°) and triphosphorylated (C-H1°abc P) C-terminal domain of histone H1° in aqueous solution and in the presence of Ficoll 70.								
Assignment	Buffer*				Ficoll†			
	C-H1°‡		C-H1°abc P		C-H1°‡		C-H1°abc P	
	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%
Turns							1682	6
Turns	1676	3	1674	10	1685	1	1671	16
Turns	1661	31	1661	12	1667	26	1663	9
α-helix					1655	15	1655	15
random coil/open loops	1642	52	1642	57	1644	23	1645	26
β-sheet					1635	14	1635	12
β-sheet	1630	9	1628	10	1626	13	1627	14
Low frequency β-sheet	1617	5	1617	11	1616	8	1616	2

\* Buffer: 10mM HEPES, 140 mM NaCl, pH 7.0, in D<sub>2</sub>O.  
† The concentration of Ficoll 70 was 30%.  
‡ This values were taken form Roque et al. <sup>365</sup>

Effects of hyperphosphorylation on the secondary structure of the C-terminal domain of histone H1 bound to DNA.

We have examined the structural effects of the phosphorylation at the CDK2 sites on the DNA-bound C-terminal domain of H1°. Phosphorylation of C-H1° simultaneously at all three sites brought about a large structural change, consisting in the decrease of the proportion of the α-helix and the increase of β-structure (Fig. 49, Table 11).

When complexes at different protein to DNA ratio (r, w/w) were examined, it appeared that the extent of the structural change was critically dependent on r. At r=0.25, the lowest ratio examined, there was a drastic change in the proportion of α-helix, which decreased form 24% in the unphosphorylated C-terminus to 8%, while the β-structure increased from 24% to 35-37%. There was also a significant 13% of random coil, which was absent in the



unphosphorylated domain<sup>333</sup>. At  $r=0.5$  the  $\alpha$ -helix decreased to 6% and the  $\beta$ -structure increased up to 45-47% (at this ratio the random coil was almost non-existent). At  $r=0.7$ , which is approximately the saturation ratio, the C-terminus eventually became an all- $\beta$  protein, with no  $\alpha$ -helix, 6% of open loops, 18-22% of turns and 72-76% of  $\beta$ -structure.

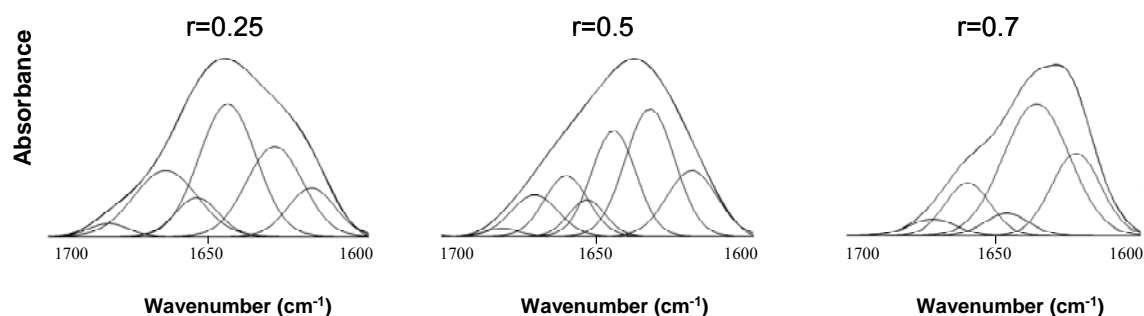


Figure 49. Amide I decomposition of the spectra of C-H1° triphosphorylated (C-H1°*abcP*) in different protein/DNA ratio ( $r$ ) (w/w). The spectra were measured in D<sub>2</sub>O. The buffer was 10 mM HEPES plus 140 mM NaCl, pH 7.0, at 20°C. The protein concentration was 5 mg/ml.

<b>TABLE 11</b>												
<b>Percentages (%) of secondary structure of the C-H1° triphosphorylated (C-H1°<i>abcP</i>) at different protein/DNA ratio (<math>r</math>)</b>												
	<b><math>r=0.25</math></b>				<b><math>r=0.5</math></b>				<b><math>r=0.7</math></b>			
	<b>D<sub>2</sub>O</b>		<b>H<sub>2</sub>O</b>		<b>D<sub>2</sub>O</b>		<b>H<sub>2</sub>O</b>		<b>D<sub>2</sub>O</b>		<b>H<sub>2</sub>O</b>	
<b>Assignment</b>	<b>Band</b>	<b>%</b>	<b>Band</b>	<b>%</b>	<b>Band</b>	<b>%</b>	<b>Band</b>	<b>%</b>	<b>Band</b>	<b>%</b>	<b>Band</b>	<b>%</b>
	(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )	
Turns					1681	1						
Turns	1681	2	1679	6	1670	10	1680	8	1671	4	1671	9
Turns	1663	20	1661	14	1660	13	1663	19	1659	14	1661	13
$\alpha$ -helix	1654	8			1653	6						
$\alpha$ -helix/ random coil			1655	21			1654	7				
open loops			1641	22			1646	21			1646	6
Random coil/open loops	1644	35			1645	23			1646	6		
Random coil <sup>a</sup>		13				2						
Random coil <sup>b</sup>				13				1				
$\beta$ -sheet	1630	24	1632	28	1633	31	1633	26	1636	52	1637	45
Low frequency $\beta$ -sheet									1623	24	1624	27
Low frequency $\beta$ -sheet	1618	11	1616	9	1620	16	1619	19				

<sup>a</sup> The value corresponds to the difference between the components at 1641-46 cm<sup>-1</sup> in D<sub>2</sub>O and H<sub>2</sub>O.

<sup>b</sup> The value corresponds to the difference between the components at 1653-55 cm<sup>-1</sup> H<sub>2</sub>O in and D<sub>2</sub>O.

Effects of partial phosphorylation on the secondary structure of the C-terminal domain of histone H1° bound to DNA.

In order to examine the effects of phosphorylating one or two of the three TPKK sites present in the C-terminus of H1°, we prepared single and double T→A mutants (Table 12).

None of these T/A substitutions altered significantly the structure of the DNA-bound domain (supplementary Table 1).

**Table 12. Mutations in the C-terminal domain of histone H1°.**

Mutated positions	Phosphorylated positions	Protein name
-	T118, T140, T152	C-H1° <i>abcP</i>
T118A ( <i>a</i> )	T140, T152	C-H1° <i>bcP</i>
T140A ( <i>b</i> )	T118, T152	C-H1° <i>acP</i>
T152A ( <i>c</i> )	T118, T140	C-H1° <i>abP</i>
T118A, T140A	T152	C-H1° <i>cP</i>
T118A, T152A	T140	C-H1° <i>bP</i>
T140A, T152A	T118	C-H1° <i>aP</i>

We examined the secondary structure of the complexes of the mutants phosphorylated in one or two positions with DNA (Fig. 50, Table 13). The structural changes in the DNA-protein complexes using mono and diphosphorylated species were much less apparent than those observed in the C-H1°*abcP* and did not depend on the protein/DNA ratio (supplementary Table 2).

In the complexes with the partially phosphorylated domains the amount of  $\beta$ -structure varied between 23% and 32%, and was similar to that of the unphosphorylated domain (24%)<sup>333</sup>. Of the three phosphorylation positions, T118 (C-H1°*aP*) affected the structure the most, leading to a decrease of the  $\alpha$ -helical content from 24% to 10% and to the appearance of 18-20% of random coil (supplementary Table 3), which is absent in the unphosphorylated domain<sup>333</sup>. Phosphorylation of T118 lead thus to significant loss of defined structure.

Phosphorylation of T140 (C-H1°*bP*) also increased the amount of random coil (9-10%), but the proportion of the  $\alpha$ -helix decreased to a lesser extent (19%). The structure when T152 was phosphorylated (C-H1°*cP*) was very similar to that of the unphosphorylated domain with only a slightly higher amount of  $\beta$ -structure. The central role of phosphorylation of the position T118 in  $\alpha$ -helix destabilization is further supported by the doble-phosphorylated domains, C-H1°*abP* and C-H1°*acP*, where the content of  $\alpha$ -helix was 9-10% while in C-H1°*bcP* was 17%.

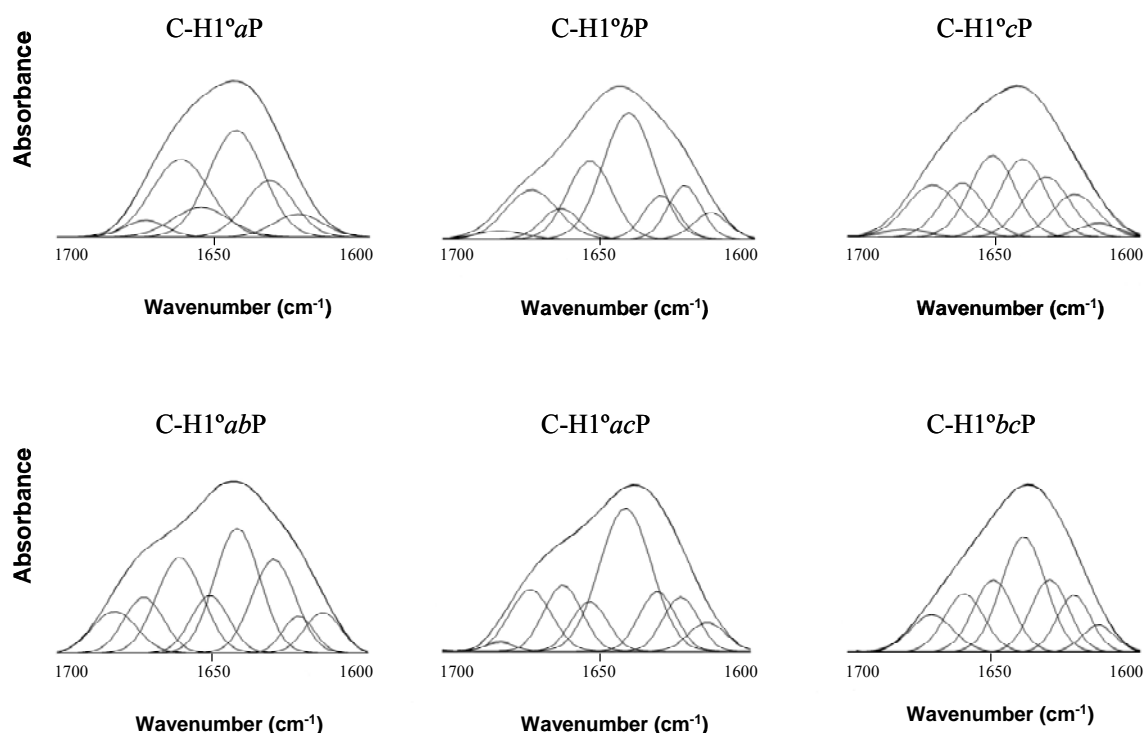


Figure 50. Amide I decomposition of the spectra of the complexes DNA/C-H1° phosphorylated in one or two positions. The spectra were measured in D<sub>2</sub>O. The buffer was 10 mM HEPES plus 140 mM NaCl, pH 7.0, at 20°C. The protein concentration was 5 mg/ml. The DNA/protein ratio was  $r=0.5$  (w/w).

<b>TABLE 13</b>														
<b>Percentages (%) of secondary structure of the DNA-bound carboxy-terminal domain of histone H1° phosphorylated on one, two or three positions.</b>														
Assignment	C-H1°a P		C-H1°b P		C-H1°c P		C-H1°ab P		C-H1°ac P		C-H1°bc P		C-H1°abc P	
	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%
Turns			1682	2	1680	1	1682	8	1682	2			1681	1
Turns	1672	4	1672	14	1671	15	1672	10	1672	13	1671	9	1670	10
Turns	1661	27	1662	6	1661	13	1661	19	1661	13	1660	13	1660	13
$\alpha$ -helix	1654	10	1653	19	1650	21	1650	10	1653	9	1650	17	1653	6
Random coil/open loops	1643	35	1640	36	1640	20	1642	24	1641	37	1640	29	1645	23
$\beta$ -sheet	1632	17	1630	8	1632	16	1630	18	1630	11	1631	16	1633	31
Low frequency $\beta$ -sheet	1623	7	1622	10	1622	11	1622	5	1623	10	1622	11	1620	16
Low frequency $\beta$ -sheet			1614	5	1614	3	1615	6	1614	5	1615	5		

The protein:DNA was 0.5 (w/w). Spectra were obtained in D<sub>2</sub>O

Effects of charge neutralization on the folding of the C-terminus of H1°.

We previously proposed that charge neutralization provided by DNA phosphate groups could be an important factor in the folding of the C-terminal domain of histone H1 when

bound to DNA<sup>333</sup>. Alkaline pH has been previously used efficiently for charge suppression in histone H1<sup>183</sup>. We examined the effect of charge neutralization by increasing pH in the induction of secondary structure in the unphosphorylated and the triphosphorylated domain of histone H1°.

The increase of the pH from 7 to 10.2, 11 and 12 results in a neutralization of ~46, ~65 and ~97% of the lysine charge, respectively. These values were calculated with the Henderson-Hasselbach equation assuming a pK<sub>a</sub> of for the lysine ε-amine group of 10.54<sup>366</sup>.

At pH 10.2 the unphosphorylated domain (Table 14) had 34% of β-structure, 30% of random coil/open loops and 36% of turns. At pH 11 there was only a slight change in the secondary structure consisting in the appearance of 5% of α-helix. However, at pH 12, when 97 % of the lysines were neutralized, the α-helix increased to 34%, the β-structure was 44% and the component of random coil/open loops disappeared.

**TABLE 14**  
**Percentages (%) of secondary structure by FTIR of the unphosphorylated (C-H1°) and the triphosphorylated (C-H1°*abc* P) C-terminal domains of histone H1° at different pH.**

Assignment	C-H1°						C-H1° <i>abc</i> P					
	pH 10.2		pH 11		pH 12		pH 10.2		pH 11		pH 12	
	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%	Band (cm <sup>-1</sup> )	%
Turns	1678	5	1677	2	1680	4	1681	7			1681	2
Turns	1667	4	1667	14	1669	12	1670	3	1671	9	1672	7
Turns	1659	27	1660	7	1659	6	1659	18	1659	13	1660	15
α-helix			1654	5	1650	34					1650	10
Random coil/open loops	1645	30	1645	38			1646	25	1646	26		
β-sheet	1633	19	1630	31	1634	36	1637	19	1633	27	1636	47
Low frequency β-sheet	1622	15					1627	13	1625	15	1625	6
Low frequency β-sheet			1617	3	1619	8	1618	15	1616	10	1617	13

Spectra were obtained in D<sub>2</sub>O

The triphosphorylated domain (Figure 51, Table 14) at pH 10.2 had 47% of β-structure, 25% of random coil/open loops and 28% of turns. At pH 11 the β-structure increased to 53%. At pH 12 C-H1°*abc*P became fully structured, the spectrum was composed of 66% of β-structure, 9% of α-helix and 25% of turns. The component of random coil/open loops also disappeared in the triphosphorylated domain at pH 12.

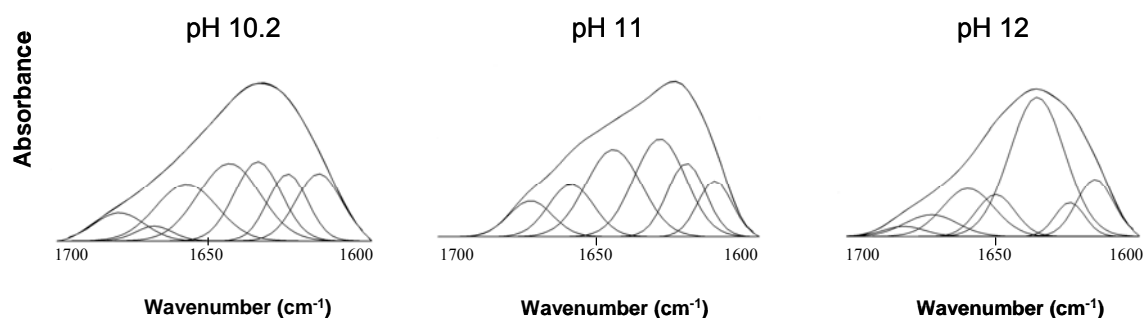


Figure 51. Amide I decomposition of the spectra of the triphosphorylated C-H1°*abcP* at different pH. The spectra were measured in D<sub>2</sub>O. The buffer was 10 mM HEPES plus NaOH to achieve the desired pH, at 20°C. The protein concentration was 5 mg/ml.

### Discussion

Protein phosphorylation represents an important regulatory mechanism in eukaryotic cells. At least one-third of all eukaryotic proteins are estimated to undergo reversible phosphorylation<sup>367</sup>. Phosphorylation modulates the activity of numerous proteins involved in signal transduction and regulates the binding affinity of transcription factors to their coactivators and DNA targets, thereby altering gene expression, cell growth and differentiation<sup>215</sup>. Phosphorylation sites frequently cluster within functionally important protein domains. Regarding the structural consequences of phosphorylation, both disorder to order and order to disorder transitions have been observed following phosphorylation<sup>368</sup>. Conformational changes upon phosphorylation often affect protein function<sup>369, 370</sup>.

Histone H1 undergoes reversible phosphorylation in a cell cycle dependent manner<sup>198</sup>. H1 phosphorylation is catalyzed by the cyclin-dependent kinases and the phosphorylation sites are mainly located in the C-terminal domain<sup>314</sup>.

In the presence of crowding agents, as Ficoll 70, the proportion of secondary structure motifs in the triphosphorylated domain were similar to those of the unphosphorylated domain. This fact suggests that the molten globule could give rise to different cooperatively-DNA bound structures depending on the phosphorylation state of the molecule.

Phosphorylation alters the secondary structure of the C-terminal domain of histone H1° when bound to DNA. The structural changes in the protein-DNA complexes with the mono and diphosphorylated species were dependent on the position that was phosphorylated. The secondary structure when T152 was phosphorylated is closer to the unphosphorylated

domain, with changes in the proportions of the structural motifs not exceeding 6%. Phosphorylation of T140 led to a decrease in  $\alpha$ -helix from 24% to 19% and the appearance of 9-10% of random coil. Phosphorylation of T118 causes a significant decrease in the helical component in the mono, di, and triphosphorylated domain. This effect in the triphosphorylated C-terminus would not be affected by the protein/DNA ratio, since the decrease in  $\alpha$ -helix is already observed at  $r=0.25$ .

A strong  $\alpha$ -helical tendency of the region immediately upstream the T118 phosphorylation site was previously observed by NMR and FTIR in a peptide belonging to the C-terminal domain of histone H1<sup>o</sup>, so it is likely that this is one of the regions in  $\alpha$ -helix conformation in the DNA-bound C-terminal domain<sup>158, 186</sup>. It appears that phosphorylation of position T118 interferes with the formation of the  $\alpha$ -helix in this region. The effect of the phosphorylation in T118 could also have functional consequences in chromatin, as this position is located in the subdomain of histone H1<sup>o</sup> involved in linker DNA alteration and folding and self-association of the chromatin fiber<sup>178</sup>.

Complete phosphorylation of C-H1<sup>o</sup> is associated to a major structural change. This structural rearrangement implies the loss of almost all the  $\alpha$ -helix probably following phosphorylation at T118, and a large increase in  $\beta$ -structure. The extent of the conformational change appears to be dependent on triphosphorylation and the protein/DNA ratio. The final state of the structural change consists in an all- $\beta$  protein at ratios near saturation. The dependence on the amount of protein in the complexes coupled to the concomitant increase in the  $\beta$ -structure, especially of the low frequency  $\beta$ -sheet, which has been proposed that could arise from intermolecular interactions of extended strands<sup>371</sup>, suggest that protein-protein interactions could play an important role in the structural changes of the C-terminal domain of histone H1 associated to complete phosphorylation.

The increase of the  $\beta$ -structure content and the loss of  $\alpha$ -helical structure following phosphorylation is a kind of structural conversion similar to that observed in amyloidogenic proteins that in prion encephalopathies has been associated with infectivity. In the case of prion proteins, the analogy can be pushed further as it has been shown that interaction with DNA converts the  $\alpha$ -helical cellular isoform into a soluble  $\beta$ -isoform similar to that found in the fibrillar state<sup>372</sup>. Furthermore, H1 has been found associated to disease associated amyloid-like fibrils<sup>261</sup>.

---

We have used alkaline pH to mimic the effect of protein charge neutralization by the DNA phosphates. At pH 12, where ~97% of the lysine positive charge is neutralized, the C-terminal domain becomes fully structured. The effect is sensitive to the phosphorylation state: the unphosphorylated domain has 36% of  $\alpha$ -helix and 34%  $\beta$ -structure while the triphosphorylated domain has 10% of  $\alpha$ -helix and 66% of  $\beta$ -structure. The helix is thus more abundant in the unphosphorylated domain while the  $\beta$ -structure is more abundant in the triphosphorylated domain as observed in the complexes with DNA. Therefore, charge neutralization upon interaction with DNA contributes to the cooperative folding of the C-terminus. The DNA-bound final structure appears to be critically dependent of the phosphorylation state.

The differences in the secondary structure of the unphosphorylated, partially phosphorylated and hyperphosphorylated C-terminal domain give the structural framework to understand the effects of site specificity and level of phosphorylation in interphase chromatin relaxation<sup>221, 222</sup> and mitotic chromosome condensation<sup>223</sup>.

**Supplementary material.**

<b>Supplementary Table 1</b>														
<b>Percentages (%) of secondary structure of the DNA-bound carboxy-terminal domain of histone H1<sup>o</sup> mutated on one, two or three positions.</b>														
Assignment	Mutated positions													
	<i>a</i>		<i>b</i>		<i>c</i>		<i>ab</i>		<i>ac</i>		<i>bc</i>		<i>abc</i>	
	Band	%	Band	%	Band	%	Band	%	Band	%	Band	%	Band	%
	(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )	
Turns	1685	3					1678	12	1685	2	1685	1	1678	2
Turns	1674	4	1673	15	1674	10	1667	14	1674	9	1674	8	1668	12
Turns	1664	17	1660	19	1661	23	1660	5	1663	16	1659	19	1660	14
$\alpha$ -helix	1653	23	1653	27	1650	22	1651	26	1653	20	1652	28	1650	24
Random coil/open loops	1643	29	1644	20	1640	18	1641	16	1642	25	1644	20	1639	23
$\beta$ -sheet	1631	13	1632	15	1630	15	1630	18	1633	11	1632	16	1629	16
Low frequency $\beta$ -sheet			1624	3			1620	6	1624	11	1620	8		
Low frequency $\beta$ -sheet	1619	11	1615	1	1619	12	1613	3	1615	6			1615	10

The protein:DNA was 0.5 (w/w). Spectra were obtained in D<sub>2</sub>O

<b>Supplementary Table 2</b>														
<b>Percentages (%) of secondary structure of the DNA-bound carboxy-terminal domain of histone H1<sup>o</sup> phosphorylated on one, two or three positions.</b>														
Assignment	C-H1 <sup>a</sup> P		C-H1 <sup>b</sup> P		C-H1 <sup>c</sup> P		C-H1 <sup>ab</sup> P		C-H1 <sup>ac</sup> P		C-H1 <sup>bc</sup> P		C-H1 <sup>abc</sup> P	
	Band	%	Band	%	Band	%	Band	%	Band	%	Band	%	Band	%
	(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )	
Turns			1682	2	1681	3	1682	8	1682	3			1679	12
Turns	1670	9	1672	11	1671	5	1672	8	1672	11	1675	13		
Turns	1660	20	1662	11	1662	19	1661	22	1661	17	1662	17	1648	9
$\alpha$ -helix	1653	11	1654	22	1650	19	1650	11	1654	9	1652	17		
Random coil/open loops	1643	31	1640	32	1643	25	1642	25	1643	33	1641	24	1636	28
$\beta$ -sheet	1632	21	1630	10	1631	22	1630	19	1631	17	1630	17	1626	29
Low frequency $\beta$ -sheet	1620	7	1622	9	1624	4	1622	3	1623	5	1622	6	1616	19
Low frequency $\beta$ -sheet			1614	3	1614	3	1614	4	1615	5	1615	6	1608	3

The protein:DNA was 0.7 (w/w). Spectra were obtained in D<sub>2</sub>O

<b>Supplementary Table 3</b>														
<b>Percentages (%) of secondary structure of the DNA-bound carboxy-terminal domain of histone H1<sup>o</sup> phosphorylated on one, two or three positions.</b>														
Assignment	C-H1 <sup>a</sup> P		C-H1 <sup>b</sup> P		C-H1 <sup>c</sup> P		C-H1 <sup>ab</sup> P		C-H1 <sup>ac</sup> P		C-H1 <sup>bc</sup> P		C-H1 <sup>abc</sup> P	
	Band	%	Band	%	Band	%	Band	%	Band	%	Band	%	Band	%
	(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )		(cm <sup>-1</sup> )	
Turns			1683	2	1681	4	1681	3						
Turns	1675	16	1674	7	1672	7	1672	6	1680	5	1680	8	1680	8
Turns	1665	13	1661	11	1660	14	1660	25	1667	18	1668	8	1663	19
$\alpha$ -helix/random coil	1656	30	1655	29	1656	22	1656	11	1656	17	1657	20	1654	7
Open loops	1643	17	1641	27	1644	22	1644	26	1643	31	1644	27	1646	21
$\beta$ -sheet	1635	18	1629	15	1634	17	1633	17	1633	12	1632	19	1633	26
Low frequency $\beta$ -sheet	1624	7	1622	3	1622	12	1621	11	1624	10	1622	13	1619	19
Low frequency $\beta$ -sheet			1614	5	1614	2	1616	1	1616	7	1615	5		

The protein:DNA was 0.5 (w/w). Spectra were obtained in H<sub>2</sub>O



## Discusión general.

### ***El dominio C-terminal determina la preferencia de la histona H1 por las regiones SAR.***

La histona H1 se ha descrito de forma genérica como una proteína de unión a las regiones SAR<sup>130</sup>. Se supone que las regiones SAR constituyen la base estructural de los lazos de cromatina en las células eucariotas<sup>138</sup>. La preferencia de la histona H1 por las SAR podría implicar a ésta histona en los cambios dinámicos de la cromatina y la organización estructural de los cromosomas<sup>144</sup>.

En mamíferos la histona H1 tiene múltiples subtipos. Los subtipos H1a-e están bastante conservados, mientras que otros subtipos como la H1<sup>o</sup> y la H1t son bastante divergentes. Utilizando histonas H1 purificadas de ratón, H1a-e, H1<sup>o</sup> y H1t, pudimos comprobar que la preferencia por las regiones SAR se mantenía a pesar de las diferencias de secuencia entre los subtipos.

La histona H1 tiene tres dominios: N-terminal, globular y C-terminal<sup>301</sup>. Cada dominio pudiera tener funciones diferentes en la proteína. Para determinar si la preferencia por las regiones SAR estaba determinada especialmente por alguno de sus dominios estudiamos dos péptidos derivados del dominio N-terminal, dos dominios globulares y tres dominios C-terminales.

Al examinar la preferencia por las regiones SAR de los dominios globulares y N-terminales encontramos una baja afinidad por el DNA. Ni los dominios globulares estudiados, correspondientes a la H1<sup>o</sup> y la H5, ni el péptido N-terminal de la H1<sup>o</sup> mostraron preferencia por las SAR. El péptido derivado del dominio N-terminal de la H1e mostró una preferencia moderada por las SAR en condiciones fisiológicas que se pierde en 70 mM de NaCl. Es importante señalar que mientras el N-terminal de la H1<sup>o</sup> tiene todos los residuos básicos concentrados en el extremo más cercano al dominio globular, el péptido derivado del N-terminal de la H1e contiene una parte del dominio N-terminal que tiene una distribución homogénea de las cargas positivas a lo largo de la molécula. Ello podría influir en la preferencia moderada por las regiones SAR observada en condiciones fisiológicas.

Teniendo en cuenta la poca selectividad y la debilidad de la unión de los dominios N-terminal y globular a las regiones SAR, es poco probable que tengan una contribución importante en la preferencia por las SAR mostrada por la H1 entera.

Estudiamos la preferencia por las SAR de los dominios C-terminales de la H1e, el subtipo más abundante de la H1<sup>o</sup>, propia de células quiescentes, y de la H1t, propia de la línea germinal masculina. La conservación de secuencia en los dominios C-terminales de estos tres subtipos es menor del 50%. A pesar de las diferencias de secuencia entre los subtipos analizados, todos ellos mostraron una alta preferencia por las secuencias SAR. La preferencia por las SAR del dominio C-terminal se mantuvo en condiciones de fuerza iónica inferiores a la fisiológica, 70 mM, y cuando se usó un DNA competidor de mayor tamaño que el DNA SAR. También se confirmó el papel de los tramos homopoliméricos en la preferencia por las SAR ya que no se encontró preferencia por un fragmento de DNA con un 75% de pares de bases A/T, dispuestos de forma alternada.

Por lo tanto, de los resultados obtenidos podemos concluir que la unión preferencial a las SAR de la histona H1 está básicamente determinada por su dominio C-terminal. El dominio C-terminal podría contribuir también a la ubicación de la H1, pues en condiciones subestequiométricas la preferencia de la H1 por la SAR podría garantizar la saturación de las SAR con H1 en detrimento de otras secuencias.

Existen evidencias que apuntan a que las protaminas podrían haberse originado mediante sustituciones de los residuos de lisina por arginina en el dominio C-terminal de la histona H1<sup>304</sup>. Los resultados mostraron que la protamina también presenta una alta especificidad por las SAR. No sabemos qué importancia puede tener la preferencia por las SAR en la función de las protaminas, es posible que contribuya al orden espacial y temporal de la sustitución de las histonas por las protaminas durante la espermatogénesis, aportando los sitios de nucleación de dicho proceso.

Una característica común de las proteínas que tienen preferencia por las regiones SAR es la pérdida de dicha preferencia en presencia de distamicina. Esto se había demostrado para la H1 entera<sup>151</sup> y nosotros comprobamos que también ocurría con el dominio C-terminal y con la protamina. La distamicina se une al surco menor del DNA con preferencia por el surco menor más estrecho de los tramos homopoliméricos de adenina<sup>152</sup>. El hecho de que la distamicina anule la preferencia del dominio C-terminal y de la protamina por las SAR

apunta a que la interacción entre dichas proteínas y el DNA involucra al surco menor, aunque no excluye la posibilidad de que también se unan al surco mayor. Aunque la distamicina tiene preferencia por el surco menor del DNA, también puede causar cambios sutiles en los parámetros estructurales y las propiedades mecanoelásticas del DNA que podrían contribuir a la supresión de la preferencia por las SAR<sup>312</sup>.

Las características que determinan que una proteína tenga preferencia por las regiones SAR no están establecidas. La proteína HMG de mamíferos contiene un motivo repetido llamado gancho AT (AT-hook) de secuencia RGR, que se une preferentemente al surco menor de las secuencias SAR<sup>155</sup>. También se ha descrito otro motivo llamado SAF-box<sup>143</sup>. Estos motivos no los encontramos ni en el dominio C-terminal de la histona H1 ni en la protamina. La histona H1 es rica en lisinas, mientras que en la HMG y la protamina predominan las argininas (la protamina puede tener hasta un 70% de argininas). Al parecer el tipo de aminoácido básico no es esencial en determinar la preferencia por las SAR. Presumiblemente la interacción involucra el reconocimiento del surco menor del DNA, cuyo campo electrostático es más intenso en los tramos homopoliméricos. Teniendo en cuenta que entre las proteínas con preferencia por las SAR es común una elevada densidad de carga positiva, la preferencia por las SAR podría ser un fenómeno básicamente electrostático.

### **Estructura del dominio C-terminal de la histona H1.**

El dominio C-terminal de la histona H1 es una proteína rica en aminoácidos básicos, especialmente lisinas, pero en su secuencia también encontramos en abundancia prolina, alanina y serina. Todos ellos son considerados aminoácidos que promueven desorden<sup>373</sup>. El dominio C-terminal tiene una elevada carga neta, baja complejidad de secuencia, baja hidrofobicidad y una movilidad electroforética anómala<sup>77</sup>. En disolución acuosa el dominio C-terminal está desestructurado, pero al unirse al DNA adquiere un elevado porcentaje de estructura secundaria, lo que permite clasificar a este dominio en el grupo de las proteínas intrínsecamente desordenadas (Intrinsically disordered proteins, IDP) que funcionan mediante el reconocimiento molecular<sup>329, 330, 331</sup>.

El dominio C-terminal en disolución acuosa se encuentra poco estructurado. El espectro de infrarrojo está dominado por la componente de ovillo estadístico (random coil)/lazos abiertos (open loops) ( $1643\text{-}1644\text{ cm}^{-1}$ ) que representa aproximadamente el 50% de la amida I. La segunda componente en importancia se encuentra a  $1662\text{ cm}^{-1}$ , y representa un 35% de la

amida I del CH1° y un 29% del CH1t. Esta componente la hemos asignado a giros de corta vida media que se encuentran en equilibrio con formas más extendidas<sup>173, 186</sup>. También encontramos otra componente de giros a 1670 cm<sup>-1</sup> (6-10%) y estructura β (8-9%). El espectro de dicroísmo circular del dominio C-terminal es el característico del ovillo estadístico, con un mínimo cercano a 200 nm y muy poca elipticidad a 220 nm, lo que permite incluir al dominio C-terminal en el grupo de IDP consideradas coil-like<sup>335</sup>.

La interacción con el DNA induce un alto grado de estructura secundaria en los dominios C-terminales de las histonas H1° y H1t. La estructura del complejo C-terminal-DNA es dependiente de la concentración salina. En 10 mM de NaCl el dominio C-terminal de la H1° está formado por un 19% de hélice α, un 22% de estructura β, un 27% de giros, un 25% de open loops y un 7% de ovillo estadístico. En concentraciones de sal fisiológicas (140 mM NaCl) el dominio CH1° se estructura completamente, ya que el ovillo estadístico desaparece. En estas condiciones encontramos un 24% de hélice α, un 25% de estructura β, un 33-34% de giros y un 18% de lazos abiertos. En el dominio C-terminal de la H1t también se observa la dependencia de la estructura de la concentración de sal; en concentraciones fisiológicas aumentan las proporciones de estructura secundaria definida, hélice y estructura β, junto con la desaparición del ovillo estadístico. También es importante señalar que en los dos dominios la estructura secundaria encontrada en presencia de 60% de TFE es casi idéntica a la encontrada cuando las proteínas están unidas al DNA.

En los espectros de C-H1° y C-H1t en 60% de TFE y en 140 mM de NaCl unidos a DNA encontramos una componente ubicada alrededor de 1660 cm<sup>-1</sup>. Esta componente pudiera tener una contribución de hélice 3<sub>10</sub> ya que dicha posición se ha asignado a esa estructura en algunos estudios<sup>323, 325</sup> y en el péptido derivado del C-terminal de la H1° (residuos 99-121) también se encontró hélice 3<sub>10</sub> en estudios de RMN<sup>158</sup>. Si hubiese una contribución en esta componente proveniente de hélice 3<sub>10</sub>, la cantidad de la hélice α solamente representaría el límite inferior de la cantidad total de estructura helicoidal presente en los dominios C-terminales, la cual podría aumentar hasta casi un 40%.

La dependencia de la estructura secundaria de los dominios C-terminales de la concentración de sal constituye una indicación de la versatilidad conformacional de este dominio, que podría ser importante en la dinámica de la cromatina.

A pesar de las grandes diferencias en las secuencias aminoacídicas de los dominios C-terminales de la H1<sup>o</sup> y la H1t, que tienen sólo un 27,7% de identidad de secuencia, existe una gran similitud en la estructura secundaria inducida por el DNA. Esto sugiere la presencia de presiones selectivas que ayudan a preservar la estructura secundaria de los diferentes subtipos unidos al DNA.

La estructura inducida en el dominio C-terminal por la unión al DNA es muy estable. Realizamos el análisis de la pérdida de la estructura secundaria con la temperatura por IR. El C-terminal en los complejos en 10 mM de sal mantiene casi toda la estructura hasta una temperatura de 78°C, a partir de la cual se desnaturaliza cooperativamente. En 140 mM de NaCl la estructura del C-terminal se mantiene prácticamente inalterada en el rango de temperaturas estudiado (20-80°C). La componente más sensible a la temperatura en ambos casos es la hoja  $\beta$  de baja frecuencia que comienza a disminuir alrededor de los 65°C, ocasionando un aumento de la componente de giros.

La elevada estabilidad térmica de los complejos C-terminal-DNA puede resultar del aumento de la hidrofobicidad asociado a la extensiva compensación de las cargas positivas de la proteína por los fosfatos del DNA. Si tenemos en cuenta el elevado contenido de lisinas en los dominios C-terminales (~40%), el efecto de la neutralización de las cargas en la hidrofobicidad puede ser importante. La neutralización de cargas se manifiesta en la relación +/- de ~1 que se encuentra en los complejos saturados y en la formación de estructuras toroidales por los complejos con el DNA, que sólo se forman cuando las cargas del DNA están compensadas en más del 90%<sup>310, 327, 328</sup>.

El proceso de estructuración como tal también puede contribuir a la hidrofobicidad. En el péptido derivado del C-terminal de la H1<sup>o</sup>, la hélice  $\alpha$  formada por los residuos del 99 al 117 tiene carácter anfipático. En una cara de la hélice se agrupan los residuos básicos, mientras que en la otra se crea una zona hidrofóbica por la presencia de residuos con estas características<sup>158</sup>.

Las IDP se asocian a procesos celulares como el control del ciclo celular, la regulación transcripcional y traduccional, el transporte en las membranas y la transducción de señales. Un elevado número de proteínas de señalización celular y asociadas al cáncer contienen regiones desordenadas, sugiriendo la importancia del desorden intrínseco en la señalización y

la regulación. Estas regiones desordenadas, muchas veces se asocian al reconocimiento molecular y a las modificaciones proteicas, incluida la fosforilación<sup>215</sup>.

Los sitios de fosforilación son frecuentes en regiones de desorden intrínseco. Como consecuencias estructurales de la fosforilación se han encontrado transiciones del orden al desorden y viceversa<sup>368</sup>. Los cambios conformacionales producto de la fosforilación pueden afectar la función de la proteína<sup>369, 370</sup>.

La fosforilación es la modificación post-traducciona l más importante de la histona H1. Es un fenómeno dependiente del ciclo celular y su grado máximo se encuentra en los cromosomas metafásicos<sup>198, 218</sup>. El número de fosfatos varía con el subtipo y la mayoría de los sitios diana se encuentran en el dominio C-terminal<sup>198</sup>.

La estructura del dominio C-terminal de la histona H1<sup>o</sup> está modulada por fosforilación. Los estados intermedios de fosforilación tienen cantidades variables de estructura secundaria. La fosforilación en T152 es la que afecta menos la estructura, con cambios en la proporción de los motivos estructurales que no exceden el 6%. La fosforilación en T140 ocasiona la disminución de la hélice  $\alpha$  del 24% al 19% y la aparición de un 9-10% de ovillo estadístico. De las tres posiciones de fosforilación analizadas en el dominio C-terminal de la H1<sup>o</sup>, la fosforilación en la posición T118 parece afectar directamente la cantidad de hélice de los complejos con DNA. La tendencia helicoidal de la región inmediatamente anterior a la T118 ya se observó en los estudios estructurales por RMN y FTIR de un péptido que contiene esta zona<sup>158, 186</sup>. Esta disminución marcada en la componente helicoidal parece indicar que la región adyacente a la T118 se encuentra en hélice  $\alpha$  en el dominio C-terminal y que la fosforilación disminuye la estabilidad de la hélice. Tanto la fosforilación en T118, como las modificaciones estructurales que ocasiona, podrían tener un efecto significativo en la cromatina, ya que esta posición se ubica en un subdominio que interviene en la asociación con el DNA internucleosomal, en el plegamiento y en la autoasociación de la fibra de cromatina<sup>178</sup>.

Cuando el dominio C-terminal está completamente fosforilado el complejo formado con el DNA tiene características estructurales diferentes. En este caso la hélice  $\alpha$  también disminuye, posiblemente debido a la fosforilación de T118, y hay un aumento muy importante de la estructura  $\beta$ , que depende de la relación proteína/DNA y del estado trifosforilado de la molécula. A medida que aumenta la relación proteína/DNA la

componente de estructura  $\beta$  aumenta y se hace máxima (72-76%) a relaciones proteína/DNA cercanas a la saturación ( $r=0,7$ ). En esta relación proteína/DNA el C-H1<sup>o</sup> trifosforilado adopta una estructura todo- $\beta$ . El aumento de la estructura  $\beta$  con la relación proteína/DNA podría indicar la interacción proteína-proteína<sup>371</sup> o entre los complejos DNA-proteína.

El aumento de la estructura  $\beta$  y la disminución de la hélice  $\alpha$  como consecuencia de la fosforilación se ha descrito en proteínas amiloides. En las encefalopatías priónicas esta transición está asociada con la infectividad. En el caso de los priones la analogía es aún mayor ya que se ha encontrado que el DNA convierte la estructura celular rica en hélice  $\alpha$  en una isoforma  $\beta$  soluble, similar a la del estado fibrilar<sup>372</sup>. También es importante señalar que se ha encontrado histona H1 asociada a fibras amiloides<sup>261</sup>.

La neutralización de las cargas positivas de las lisinas a pH alcalino favorece el plegamiento del dominio C-terminal de la H1<sup>o</sup>. A pH 12, con el ~97% de las cargas positivas neutralizadas, el C-terminal no fosforilado y trifosforilado están estructurados significativamente. En ambos casos la estructura secundaria inducida por pH es similar a la encontrada con el DNA, donde al parecer la neutralización de las cargas positivas por los grupos fosfato constituye uno de los mecanismos moleculares que hacen posible el plegamiento del C-terminal unido al DNA. También la estructura secundaria inducida a pH 12 pone de manifiesto las diferentes tendencias de plegamiento dependientes de la fosforilación.

Las diferencias en la estructura del dominio C-terminal unido al DNA dependientes de fosforilación podrían constituir la base estructural de la pérdida de afinidad por el DNA<sup>220</sup>, de la relajación de la cromatina en la interfase<sup>221, 222</sup> y del papel de la hiperfosforilación de la histona H1 en los cromosomas metafásicos<sup>223</sup>. Los cambios estructurales asociados a la fosforilación también podrían afectar la interacción con otras proteínas como ocurre con la HP1 $\alpha$ , que a su vez podría influir en el grado de condensación de la cromatina<sup>227</sup>.

A pesar de que en la mayoría de los casos las proteínas se estudian en disolución diluida, el ambiente intracelular contiene macromoléculas y solutos pequeños que se encuentran en altas concentraciones que hacen que una fracción importante del volumen intracelular no esté disponible para otras macromoléculas. Se acepta generalmente que los efectos del volumen excluido favorecen que las moléculas adopten una conformación más compactada, disminuyendo así su propio volumen<sup>339, 340, 341, 342</sup>. En el caso de las IDP se ha encontrado

que en algunos casos la aglomeración (crowding) macromolecular induce la formación de conformaciones compactas<sup>344</sup>. El dominio C-terminal de la histona H1 pertenece a este tipo de IDP.

En presencia de Ficoll 70 y PEG 6000 el dominio C-terminal se estructura significativamente, adquiriendo un 14-19% de hélice  $\alpha$ , un 22-31% de hoja  $\beta$ , un 16-26% de giros y un 20-30% de lazos abiertos, según el subtipo y el agente aglomerante utilizado. La estructuración está asociada con la aparición de un núcleo hidrofóbico, como se puede constatar por el aumento de la unión de ANS en condiciones de aglomeración. Es interesante destacar que el elevado contenido de aminoácidos básicos presentes en el dominio C-terminal no impide el plegamiento en condiciones de aglomeración. Las proporciones de los motivos de estructura secundaria en condiciones de aglomeración son comparables a las encontradas en los complejos formados con el DNA, aunque la hélice  $\alpha$  es algo más abundante en los complejos con DNA y la hoja  $\beta$  en condiciones de aglomeración<sup>333</sup>.

La compactación del C-terminal en presencia y ausencia de agentes aglomerantes se estudió por SAXS. Los cambios en la forma global fueron analizados mediante el plot de Kratky (Q2I(Q) vs. Q). La representación de Kratky tiene un máximo característico en los estados nativos y de glóbulo fundido (molten globule), que no se observa en los estados no globulares como son el ovillo estadístico o el pre-glóbulo fundido<sup>334</sup>. En presencia de agentes aglomerantes el plot de Kratky de la dispersión de rayos X del C-terminal de la H1t en 30% de Ficoll tiene una forma de campana típica de los estados globulares. En ausencia de agentes aglomerantes en el plot de Kratky encontramos un aumento gradual característico de muchas proteínas desestructuradas. El radio de giro calculado en presencia de tampón, o sea en condiciones no estructurantes, fue de 25,0 Å. Este valor disminuye en condiciones de aglomeración hasta 14,8 Å, confirmando el plegamiento del dominio C-terminal en un estado más compacto.

Se analizó la desnaturalización térmica del C-terminal en presencia de agentes aglomerantes mediante IR y DSC (Differential scanning calorimetry). El efecto de la temperatura analizado por IR muestra una disminución progresiva de la estructura secundaria entre 20 y 85 °C. En este mismo intervalo de temperaturas en el DSC encontramos un aumento lineal de la capacidad calorífica, sin el pico de absorción de calor típico de las proteínas plegadas cooperativamente<sup>356</sup>.



La combinación de los resultados obtenidos mediante SAXS, IR y DSC indica que el dominio C-terminal en presencia de agentes de aglomerantes no está plegado cooperativamente a pesar de que la estructuración y compactación se aproximan a los del estado nativo, considerando como estado nativo la estructura del dominio C-terminal unido a DNA. Estas características indican que el dominio C-terminal en presencia de agentes aglomerantes adopta un estado de glóbulo fundido<sup>357, 358, 359, 360, 361</sup>. El glóbulo fundido es un estado intermedio en el plegamiento de las proteínas que se define como un estado dinámico y compacto de la cadena polipeptídica caracterizado por un contenido de estructura secundaria muy parecido al estado nativo, pero que carece de los contactos fijos de la estructura terciaria nativa y que por tanto no experimenta desnaturalización cooperativa<sup>360</sup>. También se caracteriza por una mayor sensibilidad a las proteasas, mayor afinidad por el ANS y un aumento del radio hidrodinámico del 15% que corresponde a un aumento de volumen del 50%<sup>335</sup>. El plegamiento inducido por la unión al DNA en el dominio C-terminal de la histona H1 sí es cooperativo y muy estable<sup>333</sup>. La transición del estado de glóbulo fundido al estado nativo podría estar favorecida por el aumento de hidrofobicidad causado por la compensación de las cargas positivas del C-terminal por los fosfatos del DNA.

La similitud de las proporciones de los diferentes motivos de estructura secundaria encontrados en presencia de aglomeración y en los complejos con el DNA, sugieren que en presencia de agentes aglomerantes la estructura se aproxima a la del complejo con DNA. Ello podría ser ventajoso en la dinámica estructural de la cromatina, ya que el paso del glóbulo fundido al estado nativo (unido al DNA) podría ser mucho más rápido<sup>361</sup>. El plegamiento del C-terminal en condiciones de aglomeración también podría facilitar la difusión en el núcleo celular. Este efecto compensaría la disminución de la difusión asociada a la aglomeración<sup>362, 363</sup>. El incremento de la difusión podría tener consecuencias funcionales ya que el intercambio de las moléculas de H1 entre los sitios de la cromatina ocurre mediante un intermediario soluble<sup>180</sup>.

También estudiamos el efecto de los agentes aglomerantes en el dominio C-terminal completamente fosforilado por la CDK2 y no encontramos cambios en la estructura secundaria, por lo que al parecer el efecto inducido por la aglomeración macromolecular no está afectado por la fosforilación. A partir del mismo glóbulo fundido podrían originarse las diferentes estructuras encontradas en los complejos con DNA, dependientes del estado de

fosforilación del dominio C-terminal. Estos efectos observados en la estructura del dominio C-terminal reflejan la versatilidad estructural que se ha encontrado en algunas IDP<sup>374</sup>.

Dunker y col. postularon el Modelo de la Trinidad Proteica<sup>375</sup> (Protein Trinity Model) (Fig. 52) que establece que una proteína intracelular o sus regiones funcionales pueden existir en tres estados termodinámicos, ordenado, glóbulo fundido y desordenado u ovillo estadístico. La función de dicha proteína puede encontrarse en cualquiera de las tres conformaciones o en las transiciones entre sí.

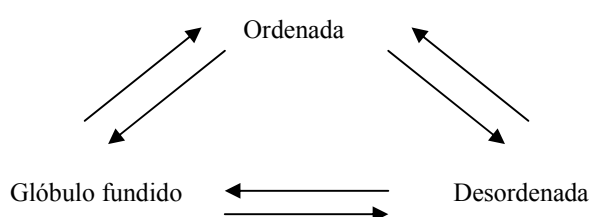


Figura 52. Modelo de la Trinidad Proteica<sup>375</sup>.

En este trabajo hemos caracterizado desde el punto de vista estructural estos tres estados para el dominio C-terminal de la histona H1. El estado desestructurado se encuentra fundamentalmente en disolución acuosa, mientras que las condiciones de aglomeración intracelular inducen el glóbulo fundido con las ventajas cinéticas antes mencionadas. El plegamiento cooperativo resulta de la interacción con el DNA y su estructura secundaria está modulada por la fosforilación de los motivos (S/T)-P-X-(K/R) por la CDK2.

El dominio C-terminal de la histona H1 se ha descrito usualmente como la cola C-terminal, haciendo alusión a su estado desestructurado en disolución acuosa<sup>183, 184</sup>. Sin embargo, la caracterización estructural del C-terminal unido al DNA y del glóbulo fundido inducido en condiciones de aglomeración macromolecular demuestra que al parecer en el interior de la célula, ya sea en estado libre o unido al DNA, el dominio C-terminal tiene una estructura secundaria definida y compacta.

## Conclusiones.

1. La preferencia por las secuencias SAR de la histona H1 no es específica de subtipo. Los subtipos de la histona H1 de mamíferos, H1a-e, H1<sup>o</sup> y H1t, tienen preferencia por las SAR.
2. La preferencia de la histona H1 por las SAR está determinada por el dominio C-terminal. La protamina, relacionada evolutivamente con la histona H1, también tiene preferencia por las SAR.
3. La preferencia por las SAR del dominio C-terminal y de la protamina es independiente de la longitud y del contenido de AT. La distamicina es capaz de eliminar dicha preferencia, lo que indica que la interacción involucra el surco menor del DNA.
4. Los dominios C-terminales de las histonas H1<sup>o</sup> y H1t se encuentran desestructurados en disolución acuosa. La interacción con el DNA induce la estructuración completa de dichos dominios, caracterizada por la presencia de hélice  $\alpha$ , estructura  $\beta$ , giros y lazos abiertos. El TFE también es capaz de inducir estructura secundaria en los dominios C-terminales con características similares a las encontradas en los complejos con DNA.
5. La estructuración inducida por el DNA en los dominios C-terminales es dependiente de la fuerza iónica, alcanzándose la estructuración completa en condiciones de sal fisiológicas (140 mM NaCl).
6. La estructura de los dominios C-terminales en los complejos con DNA se mantiene estable a altas temperaturas. Los complejos en 10 mM de NaCl se desnaturalizan cooperativamente a 78°C, mientras que los complejos en 140 mM de NaCl se mantienen estables hasta el límite experimental de 80°C.
7. El acoplamiento entre la unión al DNA y la inducción de estructura secundaria en el C-terminal de la histona H1 permite incluir este dominio en el grupo de las proteínas intrínsecamente desordenadas que funcionan mediante el reconocimiento molecular.
8. Los dominios C-terminales de las histonas H1<sup>o</sup> y H1t adquieren estructura secundaria en presencia Ficoll 70 (30%) y PEG 6000 (30%). La estructura secundaria inducida es similar a la encontrada con el dominio C-terminal unido a DNA y no está afectada por la fosforilación de los motivos (S/T)-P-X-(K/R).

9. En presencia de agentes aglomerantes la representación de Kratky de la dispersión de rayos X a bajo ángulo tiene la forma de campana típica de los estados compactos (glóbulo fundido, estado nativo), que está de acuerdo con la disminución del radio de giro de 25,0 Å en tampón a 14,8 Å en condiciones de aglomeración. La compactación va acompañada de la aparición de un núcleo hidrofóbico como indica el aumento de la unión de ANS.
10. El dominio C-terminal de la histona H1 en presencia de agentes aglomerantes no está estructurado cooperativamente, como indican la desnaturalización térmica por IR y el DSC que muestran una pérdida gradual de la estructura secundaria.
11. Una proporción de estructura secundaria semejante a la nativa junto con una compactación propia de los estados globulares en ausencia de plegamiento cooperativo indica que el el dominio C-terminal de la histona H1 en presencia de agentes aglomerantes se encuentra en un estado de glóbulo fundido.
12. La fosforilación de los tres motivos TPKK del dominio C-terminal de la histona H1 induce un cambio estructural caracterizado por la disminución de la proporción de la hélice  $\alpha$  y un aumento de la estructura  $\beta$ . La magnitud del cambio depende de la relación proteína/DNA. A relaciones cercanas a la saturación ( $r=0,7$ ) la conformación de la proteína es del tipo todo- $\beta$  (72-76%).
13. La fosforilación de uno o dos de los tres sitios TPKK presentes en el C-terminal de la H1° tiene diferentes efectos estructurales. La fosforilación en T152 es la que menos afecta la estructura con una variación en la proporción de los diferentes motivos estructurales que no supera el 6%. La fosforilación de T140 disminuye la proporción de hélice  $\alpha$  (del 24% al 19%) y origina un 9-10% de ovillo estadístico. La fosforilación en T118 es la que afecta más profundamente la estructura del C-H1°, con una disminución significativa de la hélice  $\alpha$  (del 24% al 10%), que va acompañada de la aparición de un 18-20% de ovillo estadístico. El efecto desestabilizador de la fosforilación en T118 sobre la hélice  $\alpha$  se confirma en los dominios doblemente fosforilados.
14. La neutralización de las cargas positivas de las lisinas por los grupos fosfato puede ser una de las causas que permiten la estructuración del dominio C-terminal en los complejos con el DNA. A pH 12, con un ~97% de las lisinas neutralizadas, se induce

una estructuración extensiva del dominio C-H1° no fosforilado y trifosforilado. En la estructura secundaria inducida en el dominio C-H1° no fosforilado predomina la hélice  $\alpha$  y en el C-H1° trifosforilado predomina la estructura  $\beta$ . La estructura en ambos casos se aproxima a la encontrada en los complejos con el DNA.



---

## Bibliografía.

1. Wolffe AP, Khochbin S, Dimitrov S. What do linker histones do in chromatin? *BioEssays*. 1997;12(3):249-255.
2. Thoma F, Koller T and Klug A. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol*. 1979;83:403-27.
3. Woodcock CL, Frado LL and Rattner JB. The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol*. 1984;99:42-52.
4. Dorigo B, Schalch T, Kulangara A, Duda S, Schroeder RR, Richmond TJ. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science*. 2004;306(5701):1571-3.
5. Dabán JR and Bermúdez A. Interdigitated solenoid model for compact chromatin fibers. *Biochemistry*. 1998;37(13):4299-304.
6. Alberts B, Bray D, Johnson A, Lewis J, Raaf M, Roberts K, Walter P. *Essential Cell Biology. An Introduction to the Molecular Biology of the Cell*. Garland Publishing, Inc. New York. 1998.
7. Laemmli UK, Kas E, Poljak L, Adachi Y. Scaffold associated regions: cis actin determinants of chromatin structural loops and functional domains. *Curr Opin Genet Dev*. 1992; 2: 275-285.
8. Wolffe AP. New insights into chromatin function in transcriptional control. *FASEB J*. 1992; 6:3354-61.
9. Maison C, Almouzni G. HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol*. 2004;5(4):296-304.
10. van Holde KE. *Chromatin*. 1988. Springer-Verlag, New York.
11. De Lange RJ, Farnbrough DM, Smith EL and Bonner J. Calf and pea histone IV: complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone. *J. Biol. Chem*. 1969;244:5669-5679.
12. Arents G, Burlingame RW, Wang BW, Love WE and Moudrianakis EN. The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl Acad. Sci*. 1991;88:10148-10152.
13. Wolffe AP. Packaging principle: how DNA methylation and histone acetylation control the transcriptional activity of chromatin. *J Exp Zool*. 1998;282(1-2):239-44.
14. Kuo MH, Brownell JE, Sobal RE, Ranalli TA, Cook RG, Edmondson DG, Roth SY and Allis CD. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*. 1996;383:269-272.
15. Hecht A, Laroche T, Strahl-Bolsinger S, Gasser SM, Grunstein M. Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell*. 1995;80(4):583-92.
16. Edmondson DG, Smith MM and Roth SY. Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev*. 1996;10:1247-1259.
17. Prevelige PE Jr. and Fasman GD. Structural studies of acetylated and control inner core histones. *Biochemistry*. 1987;26:2944-2955.
18. Wang X, Moore SC, Laszczak M and Ausió J. Acetylation increases the alpha-helical content of the histone tails of the nucleosome. *J Biol Chem*. 2000;275:35013-35020.
19. Arents G and Moudrianakis EN. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc. Natl Acad. Sci*. 1993; 90:10489-10493.
20. Ichimura S, Mita K and Zama M. Essential role of arginine residues in the folding of deoxyribonucleic acid into nucleosome cores. *Biochemistry*. 1982;21:5329-5334.
21. Hayes JJ, Clark DJ and Wolffe AP. Histone contribution to the structure of DNA in a nucleosome. *Proc. Natl Acad. Sci*. 1991;88:6829-6833.
22. Ausio J, Dong F, and van Holde KE. Use of selectively trypsinized nucleosome core particles to analyze the role of the histone tails in the stabilization of the nucleosome. *J. Mol. Biol*. 1989;206:451-463.
23. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997;389(6648):251-60.
24. Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell*. 2007;128(4):707-19.
25. Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128(4):693-705.
26. Jenuwein T, Allis CD. Translating the histone code. *Science*. 2001;293(5532):1074-80.
27. Turner BM. Cellular memory and the histone code. *Cell*. 2002;111(3):285-91.

28. Allis CD, Jenuwein T and Reinberg D. Epigenetics. 2007. Cold Spring Harbor.
29. Zhang K, Williams KE, Huang L, Yau P, Siino JS, Bradbury EM, Jones PR, Minch MJ, Burlingame AL. Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Mol Cell Proteomics*. 2002;1(7):500-8.
30. Zhang K, Tang H, Huang L, Blankenship JW, Jones PR, Xiang F, Yau PM, Burlingame AL. Identification of acetylation and methylation sites of histone H3 from chicken erythrocytes by high-accuracy matrix-assisted laser desorption ionization-time-of-flight, matrix-assisted laser desorption ionization-postsource decay, and nanoelectrospray ionization tandem mass spectrometry. *Anal Biochem*. 2002;306(2):259-69.
31. Zhang L, Eugeni EE, Parthun MR, Freitas MA. Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma*. 2003;112(2):77-86.
32. Turner BM. Histone acetylation and control of gene expression. *J. Cell. Sci.* 1991;99:13-20.
33. Oliva R, Bazett-Jones DP, Locklear L and Dixon G. Histone hyperacetylation can induce unfolding of the nucleosome core particle. *Nucleic. Acids Res.* 1990;18:2739-2747.
34. Lee DY, Hayes JJ, Pruss D and Wolffe AP. A positive role for histone acetylation in transcription factor binding to nucleosomal DNA. *Cell*. 1993;72:73-84.
35. Vettesse-Dadey M, Grant PA, Hebbes RT, Crane-Robinson C, Allis CD, and Workman JL. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *EMBO J*. 1996;15:2508-2518.
36. Ura K, Kurumizaka H, Dimitrov S, Almouzni G, and Wolffe AP. Histone acetylation: influence on transcription by RNA polymerase, nucleosome mobility and positioning, and linker histone dependent transcriptional repression. *EMBO J*. 1997;16:2096-2107.
37. Annunziato AT, Frado LL, Seale RL and Woodcock CL. Treatment with sodium butyrate inhibits the complete condensation of interphase chromatin. *Chromosoma*. 1988;96:132-138.
38. Kan PY, Lu X, Hansen JC, Hayes JJ. The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. *Mol Cell Biol*. 2007;27(6):2084-91.
39. Mathis DJ, Oudet P, Waslyk B and Chambon P. Effect of histone acetylation on structure and in vitro transcription of chromatin. *Nucleic. Acids Res.* 1978;5:3523-3547.
40. Wang X, Moore SC, Laszczak M and Ausió J. Acetylation increases the alpha-helical content of the histone tails of the nucleosome. *J Biol Chem*. 2000;275:35013-35020.
41. Strahl BD and Allis CD. The language of covalent histone modifications. *Nature*. 2000;403:41-45.
42. Maas NL, Miller KM, DeFazio LG, Toczyski DP. Cell cycle and checkpoint regulation of histone H3 K56 acetylation by Hst3 and Hst4. *Mol Cell*. 2006;23(1):109-19.
43. Qin S, Parthun MR. Recruitment of the type B histone acetyltransferase Hat1p to chromatin is linked to DNA double-strand breaks. *Mol Cell Biol*. 2006;26(9):3649-58.
44. Iizuka M, Matsui T, Takisawa H, Smith MM. Regulation of replication licensing by acetyltransferase Hbo1. *Mol Cell Biol*. 2006;26(3):1098-108.
45. Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science*. 2006;311(5762):844-7.
46. Vaquero A, Scher MB, Lee DH, Sutton A, Cheng HL, Alt FW, Serrano L, Sternglanz R, Reinberg D. SirT2 is a histone deacetylase with preference for histone H4 Lys 16 during mitosis. *Genes Dev*. 2006;20(10):1256-61.
47. Mahadevan LC, Willis AC and Barrah MJ. Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid and protein synthesis inhibitors. *Cell*. 1991;65:775-783.
48. Wei Y, Mizzen CA, Cook RG, Gorovsky MA, Allis CD. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena. *Proc Natl Acad Sci U S A*. 1998;95(13):7480-4.
49. Murnion ME, Adams RR, Callister DM, Allis CD, Earnshaw WC, Swedlow JR. Chromatin-associated protein phosphatase 1 regulates aurora-B and histone H3 phosphorylation. *J Biol Chem*. 2001 ;276(28):26656-65.
50. Chadee DN, Peltier CP, Davie JR. Histone H1(S)-3 phosphorylation in Ha-ras oncogene-transformed mouse fibroblasts. *Oncogene*. 2002;21(55):8397-403.
51. Ota T, Suto S, Katayama H, Han ZB, Suzuki F, Maeda M, Tanino M, Terada Y, Tatsuka M. Increased mitotic phosphorylation of histone H3 attributable to AIM-1/Aurora-B overexpression contributes to chromosome number instability. *Cancer Res*. 2002;62(18):5168-77.



52. Katayama H, Brinkley WR, Sen S. The Aurora kinases: role in cell transformation and tumorigenesis. *Cancer Metastasis Rev.* 2003;22(4):451-64.
53. Nowak SJ, Corces VG. Phosphorylation of histone H3 correlates with transcriptionally active loci. *Genes Dev.* 2000;14(23):3003-13.
54. Dai J, Sultan S, Taylor SS, Higgins JM. The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev.* 2005;19(4):472-88.
55. Dimitrov S, Dasso MC and Wolffe AP. Remodeling sperm chromatin in *Xenopus laevis* egg extracts: the role of core histone phosphorylation and linker histone B4 in chromatin assembly. *J. Cell Biol.* 1994;126:591-601.
56. Wade PA, Pruss D and Wolffe AP. Histone acetylation: chromatin in action. *Trends Biochem. Sci.* 1997;22:128-132.
57. Krishnamoorthy T, Chen X, Govin J, Cheung WL, Dorsey J, Schindler K, Winter E, Allis CD, Guacci V, Khochbin S, Fuller MT, Berger SL. Phosphorylation of histone H4 Ser1 regulates sporulation in yeast and is conserved in fly and mouse spermatogenesis. *Genes Dev.* 2006;20(18):2580-92.
58. Ahn SH, Cheung WL, Hsu JY, Diaz RL, Smith MM, Allis CD. Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in *S. cerevisiae*. *Cell.* 2005;120(1):25-36.
59. Goebel W, Obermeyer N, Bleicher N, Kratzmeier M, Eibl HJ, Doenecke D, Albig W. Apoptotic DNA fragmentation is not related to the phosphorylation state of histone H1. *Biol Chem.* 2007;388(2):197-206.
60. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet.* 2007;8(4):286-98.
61. Annunziato AT, Eason MB and Perry CA. Relationship between methylation and acetylation of arginine-rich histones in cycling and arrested HeLa cells. *Biochemistry.* 1995;34 (9):2916-24.
62. Fischle W, Wang Y and Allis CD. Binary switches and modification cassettes in histone biology and beyond. *Nature.* 2003;425(6957):475-9.
63. Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature.* 2005;438(7071):1116-22.
64. Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N, Morillon A, Weise C, Schreiber SL, Mellor J, Kouzarides T. Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Mol Cell.* 2003;12(5):1325-32.
65. Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J Biol Chem.* 2005;280(18):17732-6.
66. Nelson CJ, Santos-Rosa H, Kouzarides T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell.* 2006;126(5):905-16.
67. Sanders SL, Portoso M, Mata J, Bahler J, Allshire RC, Kouzarides T. Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell.* 2004;119(5):603-14.
68. Metivier R, Penot G, Hubner MR, Reid G, Brand H, Kos M, Gannon F. Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell.* 2003;115(6):751-63.
69. Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ, Kouzarides T. Histone deimination antagonizes arginine methylation. *Cell.* 2004;118(5):545-53.
70. Bannister AJ, Kouzarides T. Reversing histone methylation. *Nature.* 2005;436(7054):1103-6.
71. Mathis G and Althaus FR. Uncoupling of DNA excision repair and nucleosomal unfolding in poly (ADP-ribose) depleted mammalian cells. *Carcinogenesis.* 1990;11:1237-1239.
72. West MHP and Bonner WM. Histone H2B can be modified by the attachment of ubiquitin. *Nucleic Acids Res.* 1980;8:4671-4680.
73. Kleinschmidt AM and Martison HG. Structure of nucleosome core particles containing uH2A. *Nucleic Acids Res.* 1981;27:565-574.
74. Moore SC, Jason L and Ausió J. The elusive structural role of ubiquitinated histones. *Biochem Cell Biol.* 2002;80(3):311-9.
75. Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Johnson ES, Berger SL. Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. *Genes Dev.* 2006;20(8):966-76.

76. Wolffe AP. Histone H1. *Int J Biochem Cell Biol.* 1997;29(12):1463-6.
77. Ponte I, Vila R, Suau P. Sequence complexity of histone H1 subtypes. *Mol Biol Evol.* 2003;20(3):371-80.
78. Ponte I, Vidal-Taboada JM, Suau P. Evolution of the vertebrate H1 histone class: evidence for the functional differentiation of the subtypes. *Mol Biol Evol.* 1998; 15(6):702-8.
79. Khochbin S. Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene.* 2001; 271(1):1-12.
80. Tanaka M, Hennebold JD, Macfarlane J, Adashi EY. A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog. *Development.* 2001;128(5):655-64.
81. Tanaka Y, Kato S, Tanaka M, Kuji N, Yoshimura Y. Structure and expression of the human oocyte-specific histone H1 gene elucidated by direct RT-nested PCR of a single oocyte. *Biochem Biophys Res Commun.* 2003;304(2):351-7.
82. Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, Sassone-Corsi P, Davidson I. Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. *Proc Natl Acad Sci U S A.* 2005;102(8):2808-13.
83. Tanaka H, Iguchi N, Isotani A, Kitamura K, Toyama Y, Matsuoka Y, Onishi M, Masai K, Maekawa M, Toshimori K, Okabe M, Nishimune Y. HANP1/H1T2, a novel histone H1-like protein involved in nuclear formation and sperm fertility. *Mol Cell Biol.* 2005;25(16):7107-19.
84. Yan W, Ma L, Burns KH, Matzuk MM. HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. *Proc Natl Acad Sci U S A.* 2003;100(18):10546-51.
85. Albig W, Drabent B, Kunz J, Kalff-Suske M, Grzeschik KH, Doenecke D. All known human H1 histone genes except the H1(0) gene are clustered on chromosome 6. *Genomics.* 1993;16(3):649-54.
86. Wang ZF, Krasikov T, Frey MR, Wang J, Matera AG, Marzluff WF. Characterization of the mouse histone gene cluster on chromosome 13: 45 histone genes in three patches spread over 1Mb. *Genome Res.* 1996;6(8):688-701.
87. Wang ZF, Sirotkin AM, Buchold GM, Skoultchi AI, Marzluff WF. The mouse histone H1 genes: gene organization and differential regulation. *J Mol Biol* 1997;271:124-38.
88. Happel N, Schulze E, Doenecke D. Characterisation of human histone H1x. *Biol Chem.* 2005;386(6):541-51.
89. Khochbin S, Wolffe AP. Developmentally regulated expression of linker-histone variants in vertebrates. *Eur. J. Biochem.* 1994;225:501-510.
90. Brannan CI, Gilbert DJ, Ceci JD, Matsuda Y, Chapman VM, Mercer JA, Eisen H, Johnston LA, Copeland NG, Jenkins NA. An interspecific linkage map of mouse chromosome 15 positioned with respect to the centromere. *Genomics.* 1992;13(4):1075-81.
91. Brown DT, Gunjan A, Alexander BT, Sitman DB. Differential effect of H1 variant overproduction on gene expression is due to differences in the central globular domain. *Nucl. Acid. Res.* 1997;25(24):5003-5009.
92. Sun JM, Wiaderkeiwicz R, Ruiz-Carrillo A. Histone H5 in the control of DNA synthesis and cell proliferation. *Science.* 1990;245:68-71.
93. Lennox RW. Differences in evolutionary stability among mammalian H1 subtypes. Implications for the roles of H1 subtypes in chromatin. *J Biol Chem.* 1984; 259(1):669-72.
94. Lennox RW, Cohen LH. The histone H1 complements of dividing and nondividing cells of the mouse. *J Biol Chem.* 1983; 258(1):262-8.
95. Welsh SL, Cole RD. Differences between subfractions of H1 histone in their interaction with DNA. *J. Biol. Chem.* 1979; 254: 662-665.
96. Wisniewski JR, Zougman A, Krueger S, Mann M. Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations and phosphorylation as well as differences between cell culture and tissue. *Mol Cell Proteomics.* 2007;6(1):72-87.
97. Dominguez V, Pina B, Suau P. Histone H1 subtype synthesis in neurons and neuroblasts. *Development.* 1992;115(1):181-5.
98. Meergans T, Albig W, Doenecke D. Varied expression patterns of human H1 histone genes in different cell lines. *DNA Cell Biol* 1997;16:1041-9.
99. Franke K, Drabent B, Doenecke D. Expression of murine H1 histone genes during postnatal development. *Biochem Biophys Acta* 1998;1398:232-42.

100. Piña B, Martínez P, Suau P. Changes in H1 complement in differentiating rat-brain cortical neurons. *Eur J Biochem.* 1987; 164(1):71-6.
101. Gjerset R, Gorka C, Hasthorpe S, Lawrence JJ, Eisen H. Developmental and hormonal regulation of protein H1 degrees in rodents. *Proc Natl Acad Sci U S A.* 1982; 79(7):2333-7.
102. Garcia-Segura LM, Luquin S, Martínez P, Casas MT, Suau P. Differential expression and gonadal hormone regulation of histone H1(0) in the developing and adult rat brain. *Brain Res Dev Brain Res.* 1993; 73(1):63-70.
103. Marsh WH, Fitzgerald PJ. Pancreas acinar cell regeneration. Histone synthesis and modification. *Fed Proc.* 1973; 32(11):2119-25.
104. Lafarga M, García-Segura LM, Rodríguez JR, Suau P. Expression of histone H1 (zero) in transcriptionally activated supraoptic neurons. *Brain Res Mol Brain Res.* 1995; 29(2):317-24.
105. Lindner H, Sarg B, Hoertnagl B, Helliger W. The microheterogeneity of the mammalian H1(0) histone. Evidence for an age-dependent deamidation. *J Biol Chem.* 1998;273(21):13324-30.
106. Bergman M, Ringertz N. Gene expression pattern of chicken erythrocyte nuclei in heterokaryons. *J Cell Sci.* 1990; 97 (Pt 1):167-75.
107. Khochbin S, Lawrence JJ. Molecular basis of the activation of basal histone H1(0) gene expression. *Nucleic Acids Res.* 1994;22(15):2887-93.
108. Lemerrier C, Duncliffe K, Boibessot I, Zhang H, Verdel A, Angelov D, Khochbin S. Involvement of retinoblastoma protein and HBp1 in histone H1(0) gene expression. *Mol Cell Biol.* 2000;20(18):6627-37.
109. Bouterfa HL, Piedrafita FJ, Doenecke D, Pfahl M. Regulation of H1(0) gene expression by nuclear receptors through an unusual response element: implications for regulation of cell proliferation. *DNA Cell Biol.* 1995;14(11):909-19.
110. Grimes SR, Wolfe SA, Anderson JV, Stein GS, Stein JL. Structural and functional analysis of the rat testis-specific histone H1t gene. *J Cell Biochem* 1990;44(1):1-17.
111. Lin Q, Sirotkin A, Skoultchi AI. Normal spermatogenesis in mice lacking the testis-specific linker histone H1t. *Mol Cell Biol* 2000;20(6):2122-8.
112. Drabent B, Bode C, Bramlage B, Doenecke D. Expression of the mouse testicular histone gene H1t during spermatogenesis. *Histochem Cell Biol* 1996;106:247-51.
113. Drabent B, Bode C, Miosge N, Herken R, Doenecke D. Expression of the mouse histone gene H1t begins at premeiotic stages of spermatogenesis. *Cell Tissue Res* 1998;291:127-32.
114. Drabent B, Bode C, Doenecke D. Structure and expression of the mouse testicular H1 histone gene (H1t). *Biochim Biophys Acta* 1993;1216:311-3.
115. vanWert JM, Wolfe SA, Grimes SR. Binding of nuclear proteins to a conserved histone H1t promoter element suggests an important role in testis-specific transcription. *J Cell Biochem* 1996;60(3):348-62.
116. Grimes SR. Testis-specific transcriptional control. *Gene.* 2004;343(1):11-22.
117. Wolfe SA, van Wert J, Grimes SR. Transcription factor RFX2 is abundant in rat testis and enriched in nuclei of primary spermatocytes where it appears to be required for transcription of the testis-specific histone H1t gene. *J Cell Biochem.* 2006;99(3):735-46.
118. Wolfe SA, Grimes SR. Transcriptional repression of the testis-specific histone H1t gene mediated by an element upstream of the H1/AC box. *Gene.* 2003;308:129-38.
119. Drabent B, Saftig P, Bode C, Doenecke D. Spermatogenesis proceeds normally in mice without linker histone H1t. *Histochem Cell Biol* 2000;113:433-42.
120. Parseghian MH, Hamkalo BA. A compendium of the histone H1 family of somatic subtypes: an elusive cast of characters and their characteristics. *Biochem Cell Biol.* 2001;79(3):289-304.
121. Th'ng JP, Sung R, Ye M, Hendzel MJ. H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. *J Biol Chem.* 2005;280(30):27809-14.
122. Orrego M, Ponte I, Roque A, Buschati N, Mora X, Suau P. Differential affinity of mammalian histone H1 somatic subtypes for DNA and chromatin. *BMC Biol.* 2007;5(1):22.
123. Churchill ME and Suzuki M. 'SPKK' motifs prefer to bind to DNA at A/T-rich sites. *EMBO J.* 1989;8(13):4189-95.
124. Warrant RW, Kim SH. Alpha-helix-double helix interaction shown in the structure of a protamine-transfer RNA complex and a nucleoprotamine model. *Nature.* 1978;271(5641):130-5.
125. Turnell WG, Satchwell SC, Travers AA. A decapeptide motif for binding to the minor groove of DNA. A proposal. *FEBS Lett.* 1988;232(2):263-8.

126. Kim SH. In Mizobuchi, K., Watanabe, I. and Watson, J.D. (eds). 1983. *Nucleic Acids Research: Future Development*. Academic Press, Orlando, FL. pp. 165-177.
127. Suzuki, M. SPKK, a new nucleic acid-binding unit of protein found in histone. *EMBO J.* 1989;8(3):797-804.
128. Clark DJ, Thomas JO. Salt-dependent co-operative interaction of histone H1 with linear DNA. *J Mol Biol.* 1986;187(4):569-80.
129. Clark DJ, Thomas JO. Differences in the binding of H1 variants to DNA. Cooperativity and linker-length related distribution. *Eur J Biochem.* 1988;178(1):225-33.
130. Izaurralde E, Kas E, Laemmli UK. Highly preferential nucleation of histone H1 assembly on scaffold-associated regions. *J Mol Biol.* 1989;210(3):573-85.
131. Campoy FJ, Meehan RR, McKay S, Nixon J, Bird A. Binding of histone H1 to DNA is indifferent to methylation at CpG sequences. *J Biol Chem.* 1995;270(44):26473-81.
132. McArthur M, Thomas JO. A preference of histone H1 for methylated DNA. *EMBO J.* 1996;15(7):1705-14.
133. Jost JP, Hofsteenge J. The repressor MDBP-2 is a member of the histone H1 family that binds preferentially in vitro and in vivo to methylated nonspecific DNA sequences. *Proc Natl Acad Sci U S A.* 1992;89(20):9499-503.
134. Akhmanova A, Verkerk T, Langeveld A, Grosveld F, Galjart N. Characterisation of transcriptionally active and inactive chromatin domains in neurons. *J Cell Sci.* 2000;113 Pt 24:4463-74.
135. Nightingale K, Wolffe AP. Methylation at CpG sequences does not influence histone H1 binding to a nucleosome including a *Xenopus borealis* 5 S rRNA gene. *J Biol Chem.* 1995;270(9):4197-200.
136. Hill DA, Reeves R. Competition between HMG-I(Y), HMG-1 and histone H1 on four-way junction DNA. *Nucleic Acids Res.* 1997;25(17):3523-31.
137. Varga-Weisz P, Zlatanova J, Leuba SH, Schroth GP, van Holde K. Binding of histones H1 and H5 and their globular domains to four-way junction DNA. *Proc Natl Acad Sci U S A.* 1994;91(9):3525-9.
138. Mirkovitch J, Mirault ME, Laemmli UK. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell.* 1984;39(1):223-32.
139. Strick R, Laemmli UK. SARs are cis DNA elements of chromosome dynamics: synthesis of a SAR repressor protein. *Cell.* 1995;83(7):1137-48.
140. Earnshaw WC and Laemmli UK. Architecture of metaphase chromosomes and chromosome scaffolds. *J Cell Biol.* 1983;96(1):84-93.
141. Adachi Y, Kas E, Laemmli UK. Preferential, cooperative binding of DNA topoisomerase II to scaffold-associated regions. *EMBO J.* 1989;8(13):3997-4006.
142. Zhao K, Kas E, Gonzalez E, Laemmli UK. SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J.* 1993;12(8):3237-47.
143. Kipp M, Gohring F, Ostendorp T, van Drunen CM, van Driel R, Przybylski M, Fackelmayer FO. SAF-Box, a conserved protein domain that specifically recognizes scaffold attachment region DNA. *Mol Cell Biol.* 2000;20(20):7480-9.
144. Hart CM and Laemmli UK. Facilitation of chromatin dynamics by SARs. *Curr Opin Genet Dev.* 1998;8(5):519-25.
145. Saitoh Y, Laemmli UK. Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell.* 1994;76(4):609-22.
146. Craig JM, Boyle S, Perry P, Bickmore WA. Scaffold attachments within the human genome. *J Cell Sci.* 1997;110 ( Pt 21):2673-82.
147. Laemmli UK, Kas E, Poljak L, Adachi Y. Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. *Curr Opin Genet Dev.* 1992;2(2):275-85.
148. Kas E, Poljak L, Adachi Y, Laemmli UK. A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin. *EMBO J.* 1993;12(1):115-26.
149. Jenuwein T, Forrester WC, Fernandez-Herrero LA, Laible G, Dull M, Grosschedl R. Extension of chromatin accessibility by nuclear matrix attachment regions. *Nature.* 1997;385(6613):269-72.
150. Gasser SM, Laemmli UK. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell.* 1986;46(4):521-30.
151. Kas E, Izaurralde E, Laemmli UK. Specific inhibition of DNA binding to nuclear scaffolds and histone H1 by distamycin. The role of oligo(dA).oligo(dT) tracts. *J Mol Biol.* 1989;210(3):587-99.

152. Coll M, Frederick CA, Wang AH, Rich A. A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc Natl Acad Sci U S A*. 1987;84(23):8385-9.
153. Nelson HC, Finch JT, Luisi BF, Klug A. The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature*. 1987;330(6145):221-6.
154. Mamoon NM, Song Y, Wellman SE. Histone h1(0) and its carboxyl-terminal domain bind in the major groove of DNA. *Biochemistry*. 2002;41(29):9222-8.
155. Reeves R, Nissen MS. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem*. 1990;265(15):8573-82.
156. von Holt C, de Groot P, Schwager S and Brandt WF. In Stein G S, Stein JL and Marzluff WF (eds), *Histone genes*. 1984. Wiley-Interscience, New York, pp 65-105.
157. Suzuki M, Gerstein M and Johnson T. An NMR study on the DNA-binding SPKK motif and a model for its interaction with DNA. *Protein Eng*. 1993;6(6):565-74.
158. Vila R, Ponte I, Jimenez MA, Rico M, Suau P. A helix-turn motif in the C-terminal domain of histone H1. *Protein Sci*. 2000;9(4):627-36.
159. Kopka ML, Fratini AV, Drew HR, Dickerson RE. Ordered water structure around a B-DNA dodecamer. A quantitative study. *J Mol Biol*. 1983;163(1):129-46.
160. Khadake JR, Rao MR. Condensation of DNA and chromatin by an SPKK-containing octapeptide repeat motif present in the C-terminus of histone H1. *Biochemistry*. 1997;36(5):1041-51.
161. Fischer G, Aumuller T. Regulation of peptide bond cis/trans isomerization by enzyme catalysis and its implication in physiological processes. *Rev Physiol Biochem Pharmacol*. 2003;148:105-50.
162. Wulf G, Finn G, Suizu F, Lu KP. Phosphorylation-specific prolyl isomerization: is there an underlying theme? *Nat Cell Biol*. 2005;7(5):435-41.
163. Luders T, Birkemo GA, Nissen-Meyer J, Andersen O, Nes IF. Proline conformation-dependent antimicrobial activity of a proline-rich histone h1 N-terminal Peptide fragment isolated from the skin mucus of Atlantic salmon. *Antimicrob Agents Chemother*. 2005;49(6):2399-406.
164. Bucle RS, Mapam JD, Allan J. Site-directed mutagenesis studies on the binding of the globular domain of linker histone H5 to the nucleosome. *J Mol Biol*. 1992; 223(3):651-9.
165. Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*. 1993;362(6417):219-23.
166. Matsui T, Nishimura S, Kijima K, Kawasaki Y, Tashiro F, Miura S. Crystallization and preliminary X-ray diffraction data of the rat histone H1(0) globular domain. *Acta Crystallograph Sect F Struct Biol Cryst Commun*. 2007;63(Pt 5):403-5.
167. Zhou YB, Gerchman SE, Ramakrishnan V, Travers A, Muyldermans S. Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature*. 1998;395(6700):402-5.
168. Goytisolo FA, Gerchman SE, Yu X, Rees C, Graziano V, Ramakrishnan V, Thomas JO. Identification of two DNA-binding sites on the globular domain of histone H5. *EMBO J*. 1996;15(13):3421-9.
169. Duggan MM, Thomas JO. Two DNA-binding sites on the globular domain of histone H5 are required for binding to both bulk and 5 S reconstituted nucleosomes. *J Mol Biol*. 2000;304(1):21-33.
170. Ramesh S, Bharath MM, Chandra NR, Rao MR. A K52Q substitution in the globular domain of histone H1t modulates its nucleosome binding properties. *FEBS Lett*. 2006;580(25):5999-6006.
171. Bohm L, Mitchell TC. Sequence conservation in the N-terminal domain of histone H1. *FEBS Lett*. 1985;193(1):1-4.
172. Allan J, Mitchell T, Harborne N, Bohm L, Crane-Robinson C. Roles of H1 domains in determining higher order chromatin structure and H1 location. *J Mol Biol*. 1986;187(4):591-601.
173. Vila R, Ponte I, Collado M, Arrondo JL, Jimenez MA, Rico M, Suau P. DNA-induced alpha-helical structure in the NH2-terminal domain of histone H1. *J Biol Chem*. 2001;276(49):46429-35.
174. Vila R, Ponte I, Jimenez MA, Rico M, Suau P. An inducible helix-Gly-Gly-helix motif in the N-terminal domain of histone H1e: a CD and NMR study. *Protein Sci*. 2002;11(2):214-20.
175. Subirana JA. Analysis of the charge distribution in the C-terminal region of histone H1 as related to its interaction with DNA. *Biopolymers*. 1990;29(10-11):1351-7.
176. Allan J, Mitchell T, Harborne N, Bohm L, Crane-Robinson C. Roles of H1 domains in determining higher order chromatin structure and H1 location. *J Mol Biol*. 1986;187(4):591-601.
177. Moran F, Montero F, Azorin F, Suau P. Condensation of DNA by the C-terminal domain of histone H1. A circular dichroism study. *Biophys Chem*;22(1-2):125-9.

178. Lu X and Hansen JC. Identification of specific functional subdomains within the linker histone H1O C-terminal domain. *J Biol Chem.* 2004;279(10):8701-7.
179. Bharath MM, Ramesh S, Chandra NR, Rao MR. Identification of a 34 amino acid stretch within the C-terminus of histone H1 as the DNA-condensing domain by site-directed mutagenesis. *Biochemistry.* 2002;41(24):7617-27.
180. Lever MA, Th'ng JP, Sun X, Hendzel MJ. Rapid exchange of histone H1.1 on chromatin in living human cells. *Nature.* 2000;408(6814):873-6.
181. Misteli T, Gunjan A, Hock R, Bustin M, Brown DT. Dynamic binding of histone H1 to chromatin in living cells. *Nature.* 2000;408(6814):877-81.
182. Hendzel MJ, Lever MA, Crawford E, Th'ng JP. The C-terminal domain is the primary determinant of histone H1 binding to chromatin in vivo. *J Biol Chem.* 2004;279(19):20028-34.
183. Clark DJ, Hill CS, Martin SR, Thomas JO. Alpha-helix in the carboxy-terminal domains of histones H1 and H5. *EMBO J.* 1988;7(1):69-75.
184. Hill CS, Martin SR, Thomas JO. A stable alpha-helical element in the carboxy-terminal domain of free and chromatin-bound histone H1 from sea urchin sperm. *EMBO J.* 1989;8(9):2591-9.
185. Erard M, Lakhdar-Ghazal F, Amalric F. Repeat peptide motifs which contain beta-turns and modulate DNA condensation in chromatin. *Eur J Biochem.* 1990;191(1):19-26.
186. Vila R, Ponte I, Collado M, Arrondo JL, Suau P. Induction of secondary structure in a COOH-terminal peptide of histone H1 by interaction with the DNA: an infrared spectroscopy study. *J Biol Chem.* 2001;276(33):30898-903.
187. Bharath MM, Chandra NR, Rao MR. Molecular modeling of the chromatosome particle. *Nucleic Acids Res.* 2003;31(14):4264-74.
188. Bharath MM, Chandra NR, Rao MR. Prediction of an HMG-box fold in the C-terminal domain of histone H1: insights into its role in DNA condensation. *Proteins.* 2002;49(1):71-81.
189. Widlak P, Kalinowska M, Parseghian MH, Lu X, Hansen JC, Garrard WT. The histone H1 C-terminal domain binds to the apoptotic nuclease, DNA fragmentation factor (DFF40/CAD) and stimulates DNA cleavage. *Biochemistry.* 2005;44(21):7871-8.
190. Zhang J and Xu M. Apoptotic DNA fragmentation and tissue homeostasis. *Trends Cell Biol.* 2002;12(2):84-9.
191. Yasutomo K, Horiuchi T, Kagami S, Tsukamoto H, Hashimura C, Urushihara M, Kuroda Y. Mutation of DNASE1 in people with systemic lupus erythematosus. *Nat Genet.* 2001;28(4):313-4.
192. Woo EJ, Kim YG, Kim MS, Han WD, Shin S, Robinson H, Park SY, Oh BH. Structural mechanism for inactivation and activation of CAD/DFF40 in the apoptotic pathway. *Mol Cell.* 2004;14(4):531-9.
193. Liu X, Zou H, Widlak P, Garrard W, Wang X. Activation of the apoptotic endonuclease DFF40 (caspase-activated DNase or nuclease). Oligomerization and direct interaction with histone H1. *J Biol Chem.* 1999;274(20):13836-40.
194. Widlak P, Li P, Wang X, Garrard WT. Cleavage preferences of the apoptotic endonuclease DFF40 (caspase-activated DNase or nuclease) on naked DNA and chromatin substrates. *J Biol Chem.* 2000;275(11):8226-32.
195. Kratzmeier M, Albig W, Hanecke K, Doenecke D. Rapid dephosphorylation of H1 histones after apoptosis induction. *J Biol Chem.* 2000;275(39):30478-86.
196. Happel N, Sommer A, Hanecke K, Albig W, Doenecke D. Topoisomerase inhibitor induced dephosphorylation of H1 and H3 histones as a consequence of cell cycle arrest. *J Cell Biochem.* 2005;95(6):1235-47.
197. Rall SC, Cole RD. Amino acid sequence and sequence variability of the amino-terminal regions of lysine-rich histones. *J Biol Chem.* 1971;246(23):7175-90.
198. Talasz H, Helliger W, Puschendorf B, Lindner H. In vivo phosphorylation of histone H1 variants during the cell cycle. *Biochemistry.* 1996;35(6):1761-7.
199. Garcia BA, Busby SA, Barber CM, Shabanowitz J, Allis CD, Hunt DF. Characterization of phosphorylation sites on histone H1 isoforms by tandem mass spectrometry. *J Proteome Res.* 2004;3(6):1219-27.
200. Sarg B, Helliger W, Hoertnagl B, Puschendorf B, Lindner H. The N-terminally acetylated form of mammalian histone H1(0), but not that of avian histone H5, increases with age. *Arch Biochem Biophys.* 1999;372(2):333-9.

201. Lindner H, Sarg B, Hoertnagl B, Helliger W. The microheterogeneity of the mammalian H1(0) histone. Evidence for an age-dependent deamidation. *J Biol Chem.* 1998;273(21):13324-30.
202. Daujat S, Zeissler U, Waldmann T, Happel N, Schneider R. HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. *J Biol Chem.* 2005;280(45):38090-5.
203. Kuzmichev A, Jenuwein T, Tempst P, Reinberg D. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol Cell.* 2004;14(2):183-93.
204. Lake RS, Salzman NP. Occurrence and properties of a chromatin-associated F1-histone phosphokinase in mitotic Chinese hamster cells. *Biochemistry.* 1972;11(25):4817-26.
205. Panzeter PL, Realini CA, Althaus FR. Noncovalent interactions of poly(adenosine diphosphate ribose) with histones. *Biochemistry.* 1992;31(5):1379-85.
206. Poirier GG, de Murcia G, Jongstra-Bilen J, Niedergang C, Mandel P. Poly(ADP-ribosylation) of polynucleosomes causes relaxation of chromatin structure. *Proc Natl Acad Sci U S A.* 1982;79(11):3423-7.
207. D'erme M, Zardo G, Reale A, Caiafa P. Co-operative interactions of oligonucleosomal DNA with the H1e histone variant and its poly(ADP-ribosylated) isoform. *Biochem J.* 1996;316:475-80.
208. Zardo G, Marenzi S, Caiafa P. H1 histone as a trans-acting factor involved in protecting genomic DNA from full methylation. *Biol Chem.* 1998;379(6):647-54.
209. Zardo G, Caiafa P. The unmethylated state of CpG islands in mouse fibroblasts depends on the poly(ADP-ribosylation) process. *J Biol Chem.* 1998;273(26):16517-20.
210. Shi Y, Lan F, Matson C, Mulligan P, Whetstone JR, Cole PA, Casero RA, Shi Y. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell.* 2004;119(7):941-53.
211. Clarke S. The role of aspartic acid and asparagine residues in the aging of erythrocyte proteins: cellular metabolism of racemized and isomerized forms by methylation reactions. *Prog Clin Biol Res.* 1985;195:91-107.
212. Wright HT. Sequence and structure determinants of the nonenzymatic deamidation of asparagine and glutamine residues in proteins. *Protein Eng.* 1991;4(3):283-94.
213. Pham AD, Sauer F. Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in *Drosophila*. *Science.* 2000;289(5488):2357-60.
214. Moore SC, Jason L, Ausio J. The elusive structural role of ubiquitinated histones. *Biochem Cell Biol.* 2002;80(3):311-9.
215. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 2004;32(3):1037-49.
216. Zor T, Mayr BM, Dyson HJ, Montminy MR, Wright PE. Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. *J Biol Chem.* 2002;277(44):42241-8.
217. Morgan DO. Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu Rev Cell Dev Biol.* 1997;13:261-91.
218. Baatout S, Derradji H. About histone H1 phosphorylation during mitosis. *Cell Biochem Funct.* 2006;24(2):93-4.
219. Sarg B, Helliger W, Talasz H, Forg B, Lindner HH. Histone H1 phosphorylation occurs site-specifically during interphase and mitosis: identification of a novel phosphorylation site on histone H1. *J Biol Chem.* 2006;281(10):6573-80.
220. Green GR, Lee HJ, Poccia DL. Phosphorylation weakens DNA binding by peptides containing multiple "SPKK" sequences. *J Biol Chem.* 1993;268(15):11247-55.
221. Halmer L and Gruss C. Effects of cell cycle dependent histone H1 phosphorylation on chromatin structure and chromatin replication. *Nucleic Acids Res.* 1996;24(8):1420-7.
222. Dou Y, Mizzen CA, Abrams M, Allis CD, Gorovsky MA. Phosphorylation of linker histone H1 regulates gene expression in vivo by mimicking H1 removal. *Mol Cell.* 1999;4(4):641-7.
223. Th'ng JP, Guo XW, Swank RA, Crissman HA, Bradbury EM. Inhibition of histone phosphorylation by staurosporine leads to chromosome decondensation. *J Biol Chem.* 1994;269(13):9568-73.
224. Gurley LR, Valdez JG, Buchanan JS. Characterization of the mitotic specific phosphorylation site of histone H1. Absence of a consensus sequence for the p34cdc2/cyclin B kinase. *J Biol Chem.* 1995;270(46):27653-60.
225. Lake RS, Salzman NP. Occurrence and properties of a chromatin-associated F1-histone phosphokinase in mitotic Chinese hamster cells. *Biochemistry.* 1972;11(25):4817-26.

226. Swank RA, Th'ng JP, Guo XW, Valdez J, Bradbury EM, Gurley LR. Four distinct cyclin-dependent kinases phosphorylate histone H1 at all of its growth-related phosphorylation sites. *Biochemistry*. 1997;36(45):13761-8.
227. Hale TK, Contreras A, Morrison AJ, Herrera RE. Phosphorylation of the linker histone H1 by CDK regulates its binding to HP1alpha. *Mol Cell*. 2006;22(5):693-9.
228. Alexandrow MG, Hamlin JL. Chromatin decondensation in S-phase involves recruitment of Cdk2 by Cdc45 and histone H1 phosphorylation. *J Cell Biol*. 2005;168(6):875-86.
229. Sweet MT, Carlson G, Cook RG, Nelson D, Allis CD. Phosphorylation of linker histones by a protein kinase A-like activity in mitotic nuclei. *J Biol Chem*. 1997;272(2):916-23.
230. Mizzen CA, Dou Y, Liu Y, Cook RG, Gorovsky MA, Allis CD. Identification and mutation of phosphorylation sites in a linker histone. Phosphorylation of macronuclear H1 is not essential for viability in tetrahymena. *J Biol Chem*. 1999;274(21):14533-6.
231. Dou Y, Song X, Liu Y, Gorovsky MA. The H1 phosphorylation state regulates expression of CDC2 and other genes in response to starvation in *Tetrahymena thermophila*. *Mol Cell Biol*. 2005;25(10):3914-22.
232. Graziano V, Gerchman SE, Schneider DK, Ramakrishnan V. Histone H1 is located in the interior of the chromatin 30-nm filament. *Nature*. 1994; 368(6469):351-4.
233. Staynov DZ, Crane-Robinson C. Footprinting of linker histones H5 and H1 on the nucleosome. *EMBO J*. 1988;7(12):3685-91.
234. Izaurralde E, Kas E, Laemmli UK. Highly preferential nucleation of histone H1 assembly on scaffold-associated regions. *J Mol Biol*. 1989; 210(3):573-85.
235. Adams CC, Workman JL. Nucleosome displacement in transcription. *Cell*. 1993;72(3):305-8.
236. Zlatanova J, Caiafa P, Van Holde K. Linker histone binding and displacement: versatile mechanism for transcriptional regulation. *FASEB J*. 2000;14(12):1697-704.
237. Yang G, Leuba SH, Bustamante C, Zlatanova J, van Holde K. Role of linker histones in extended chromatin fibre structure. *Nat Struct Biol*. 1994;1(11):761-3.
238. Erard MS, Belenguer P, Caizergues-Ferrer M, Pantalón A and Almaric F. A major nucleolar protein, nucleolin, induces chromatin decondensation by binding to histone H1. *Eur. J. Biochem*. 1999;175:525-530.
239. Karetsou Z, Sandaltzopoulos R, Frangou-Lazaridis M, Lai CY, Tsolas O, Becker PB, Papamarcaki T. Prothymosin alpha modulates the interaction of histone H1 with chromatin. *Nucleic Acids Res*. 1998;26(13):3111-8.
240. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*. 1998;393(6683):386-9.
241. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet*. 1998;19(2):187-91.
242. Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, Shim EY, Clark KL, Burley SK, Zaret KS. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J*. 1998;17(1):244-54.
243. Kermekchiev M, Workman JL, Pikaard CS. Nucleosome binding by the polymerase I transactivator upstream binding factor displaces linker histone H1. *Mol Cell Biol*. 1997;17(10):5833-42.
244. An W, Zlatanova J, Leuba SH, van Holde K. The site of binding of linker histone to the nucleosome does not depend upon the amino termini of core histones. *Biochimie*. 1999;81(7):727-32.
245. Kuo MH, Allis CD. Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays*. 1998;20(8):615-26.
246. Davie JR, Chadee DN. Regulation and regulatory parameters of histone modifications. *J Cell Biochem Suppl*. 1998;30-31:203-13.
247. Lee HL, Archer TK. Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. *EMBO J*. 1998;17(5):1454-66.
248. Banks GC, Deterding LJ, Tomer KB, Archer TK. Hormone-mediated dephosphorylation of specific histone H1 isoforms. *J Biol Chem*. 2001;276(39):36467-73.
249. Hsiao PW, Deroo BJ, Archer TK. Chromatin remodeling and tissue-selective responses of nuclear hormone receptors. *Biochem Cell Biol*. 2002;80(3):343-51.
250. Stavreva DA, McNally JG. Role of H1 phosphorylation in rapid GR exchange and function at the MMTV promoter. *Histochem Cell Biol*. 2006;125(1-2):83-9.



251. Bhattacharjee RN, Archer TK. Transcriptional silencing of the mouse mammary tumor virus promoter through chromatin remodeling is concomitant with histone H1 phosphorylation and histone H3 hyperphosphorylation at M phase. *Virology*. 2006;346(1):1-6.
252. Horn PJ, Carruthers LM, Logie C, Hill DA, Solomon MJ, Wade PA, Imbalzano AN, Hansen JC, Peterson CL. Phosphorylation of linker histones regulates ATP-dependent chromatin remodeling enzymes. *Nat Struct Biol*. 2002;9(4):263-7.
253. Kysela B, Chovanec M, Jeggo PA. Phosphorylation of linker histones by DNA-dependent protein kinase is required for DNA ligase IV-dependent ligation in the presence of histone H1. *Proc Natl Acad Sci U S A*. 2005;102(6):1877-82.
254. Yellajoshyula D, Brown DT. Global modulation of chromatin dynamics mediated by dephosphorylation of linker histone H1 is necessary for erythroid differentiation. *Proc Natl Acad Sci U S A*. 2006;103(49):18568-73.
255. Mishra S, Saleh A, Espino PS, Davie JR, Murphy LJ. Phosphorylation of histones by tissue transglutaminase. *J Biol Chem*. 2006;281(9):5532-8.
256. Tsiftoglou AS, Pappas IS, Vizirianakis IS. Mechanisms involved in the induced differentiation of leukemia cells. *Pharmacol Ther*. 2003;100(3):257-90.
257. Matushansky I, Radparvar F, Skoultchi AI. Reprogramming leukemic cells to terminal differentiation by inhibiting specific cyclin-dependent kinases in G1. *Proc Natl Acad Sci U S A*. 2000;97(26):14317-22.
258. Matushansky I, Radparvar F, Skoultchi AI. CDK6 blocks differentiation: coupling cell proliferation to the block to differentiation in leukemic cells. *Oncogene*. 2003;22(27):4143-9.
259. Bleher R, Martin R. Nucleo-cytoplasmic translocation of histone H1 during the HeLa cell cycle. *Chromosoma*. 1999;108(5):308-16.
260. Tsai LH, Delalle I, Caviness VS Jr, Chae T, Harlow E. p35 is a neural-specific regulatory subunit of cyclin-dependent kinase 5. *Nature*. 1994;371(6496):419-23.
261. Duce JA, Smith DP, Blake RE, Crouch PJ, Li QX, Masters CL, Trounce IA. Linker histone H1 binds to disease associated amyloid-like fibrils. *J Mol Biol*. 2006;361(3):493-505.
262. Parseghian MH, Luhrs KA. Beyond the walls of the nucleus: the role of histones in cellular signaling and innate immunity. *Biochem Cell Biol*. 2006;84(4):589-604.
263. Konishi A, Shimizu S, Hirota J, Takao T, Fan Y, Matsuoka Y, Zhang L, Yoneda Y, Fujii Y, Skoultchi AI, Tsujimoto Y. Involvement of histone H1.2 in apoptosis induced by DNA double-strand breaks. *Cell*. 2003;114(6):673-88.
264. Costa O, Tchouatcha-Tchouassom JC, Roux B, Monier JC. Anti-H1 histone antibodies in systemic lupus erythematosus: epitope localization after immunoblotting of chymotrypsin-digested H1. *Clin Exp Immunol*. 1986;63(3):608-13.
265. Schett G, Smole J, Zimmermann C, Hiesberger H, Hoefler E, Fournel S, Muller S, Rubin RL, Steiner G. The autoimmune response to chromatin antigens in systemic lupus erythematosus: autoantibodies against histone H1 are a highly specific marker for SLE associated with increased disease activity. *Lupus*. 2002;11(11):704-15.
266. Wesierska-Gadek J, Penner E, Lindner H, Hitchman E, Sauermann G. Autoantibodies against different histone H1 subtypes in systemic lupus erythematosus sera. *Arthritis Rheum*. 1990;33(8):1273-8.
267. Muller S, Richalet P, Laurent-Crawford A, Barakat S, Riviere Y, Porrot F, Chamaret S, Briand JP, Montagnier L, Hovanessian A. Autoantibodies typical of non-organ-specific autoimmune diseases in HIV-seropositive patients. *AIDS*. 1992;6(9):933-42.
268. Haynes BF, Fleming J, St Clair EW, Katinger H, Stiegler G, Kunert R, Robinson J, Searce RM, Plonk K, Staats HF, Ortel TL, Liao HX, Alam SM. Cardiophilin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science*. 2005;308(5730):1906-8.
269. Lesner A, Kartvelishvili A, Lesniak J, Nikolov D, Kartvelishvili M, Trillo-Pazos G, Zablorna E, Simm M. Monoubiquitinated histone H1B is required for antiviral protection in CD4(+)T cells resistant to HIV-1. *Biochemistry*. 2004;43(51):16203-11.
270. Brix K, Summa W, Lottspeich F, Herzog V. Extracellularly occurring histone H1 mediates the binding of thyroglobulin to the cell surface of mouse macrophages. *J Clin Invest*. 1998;102(2):283-93.
271. Hiemstra PS, Eisenhauer PB, Harwig SS, van den Barselaar MT, van Furth R, Lehrer RI. Antimicrobial proteins of murine macrophages. *Infect Immun*. 1993;61(7):3038-46.

272. Evans DL, Kaur H, Leary J 3rd, Praveen K, Jaso-Friedmann L. Molecular characterization of a novel pattern recognition protein from nonspecific cytotoxic cells: sequence analysis, phylogenetic comparisons and anti-microbial activity of a recombinant homologue. *Dev Comp Immunol.* 2005;29(12):1049-64.
273. Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A. Neutrophil extracellular traps kill bacteria. *Science.* 2004;303(5663):1532-5.
274. Rose FR, Bailey K, Keyte JW, Chan WC, Greenwood D, Mahida YR. Potential role of epithelial cell-derived histone H1 proteins in innate antimicrobial defense in the human gastrointestinal tract. *Infect Immun.* 1998;66(7):3255-63.
275. Zhu G, Chen H, Choi BK, Del Piero F, Schifferli DM. Histone H1 proteins act as receptors for the 987P fimbriae of enterotoxigenic *Escherichia coli*. *J Biol Chem.* 2005;280(24):23057-65.
276. Richards RC, O'Neil DB, Thibault P, Ewart KV. Histone H1: an antimicrobial protein of Atlantic salmon (*Salmo salar*). *Biochem Biophys Res Commun.* 2001;284(3):549-55.
277. Fernandes JM, Molle G, Kemp GD, Smith VJ. Isolation and characterisation of oncorhynchin II, a histone H1-derived antimicrobial peptide from skin secretions of rainbow trout, *Oncorhynchus mykiss*. *Dev Comp Immunol.* 2004;28(2):127-38.
278. Solioz N, Blum-Tirouvanziam U, Jacquet R, Rafati S, Corradin G, Mauel J, Fasel N. The protective capacities of histone H1 against experimental murine cutaneous leishmaniasis. *Vaccine.* 1999;18(9-10):850-9.
279. Masina S, M Gicheru M, Demotz SO, Fasel NJ. Protection against cutaneous leishmaniasis in outbred vervet monkeys using a recombinant histone H1 antigen. *J Infect Dis.* 2003;188(8):1250-7.
280. Patel SJ, Shapiro WR, Laske DW, Jensen RL, Asher AL, Wessels BW, Carpenter SP, Shan JS. Safety and feasibility of convection-enhanced delivery of Cotara for the treatment of malignant glioma: initial experience in 51 patients. *Neurosurgery.* 2005;56(6):1243-53.
281. Chen S, Yu L, Jiang C, Zhao Y, Sun D, Li S, Liao G, Chen Y, Fu Q, Tao Q, Ye D, Hu P, Khawli LA, Taylor CR, Epstein AL, Ju DW. Pivotal study of iodine-131-labeled chimeric tumor necrosis treatment radioimmunotherapy in patients with advanced lung cancer. *J Clin Oncol.* 2005;23(7):1538-47.
282. Nakano T, Kawamoto S, Lai CY, Sasaki T, Aki T, Shigeta S, Goto T, Sato S, Goto S, Chen CL, Ono K. Liver transplantation-induced antihistone H1 autoantibodies suppress mixed lymphocyte reaction. *Transplantation.* 2004;77(10):1595-603.
283. Zlatanova J, Van Holde K. Histone H1 and transcription: still an enigma? *J Cell Sci.* 1992;103 ( Pt 4):889-95.
284. Bouvet P, Dimitrov S, Wolffe AP. Specific regulation of *Xenopus* chromosomal 5S rRNA gene transcription in vivo by histone H1. *Genes Dev.* 1994;8(10):1147-59.
285. Shen X, Gorovsky MA. Linker histone H1 regulates specific gene expression but not global transcription in vivo. *Cell.* 1996;86(3):475-83.
286. Vermaak D, Steinbach OC, Dimitrov S, Rupp RA, Wolffe AP. The globular domain of histone H1 is sufficient to direct specific gene repression in early *Xenopus* embryos. *Curr Biol.* 1998;8(9):533-6.
287. Pennings S, Meersseman G, Bradbury EM. Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. *Proc Natl Acad Sci U S A.* 1994;91(22):10275-9.
288. Panyim S, Chalkley R. A new histone found only in mammalian tissues with little cell division. *Biochem Biophys Res Commun.* 1969;37(6):1042-9.
289. Bucci LR, Brock WA, Meistrich ML. Distribution and synthesis of histone 1 subfractions during spermatogenesis in the rat. *Exp Cell Res.* 1982;140(1):111-8.
290. Lennox RW, Oshima RG, Cohen LH. The H1 histones and their interphase phosphorylated states in differentiated and undifferentiated cell lines derived from murine teratocarcinomas. *J Biol Chem.* 1982;257(9):5183-9.
291. Hall JM, Cole RD. Modulation in proportions of H1 histone subfractions by differential changes in synthesis and turnover during butyrate treatment of neuroblastoma cells. *Biochemistry.* 1985;24(26):7765-71.
292. Liao LW, Cole RD. Differences among H1 histone subfractions in binding to linear and superhelical DNA. Sedimentation velocity studies. *J Biol Chem.* 1981;256(21):11145-50.
293. Khadake JR, Rao MR. DNA- and chromatin-condensing properties of rat testes H1a and H1t compared to those of rat liver H1bdec; H1t is a poor condenser of chromatin. *Biochemistry.* 1995;34(48):15792-801.
294. Talasz H, Sapojnikova N, Helliger W, Lindner H, Puschendorf B. In vitro binding of H1 histone subtypes to nucleosomal organized mouse mammary tumor virus long terminal repeat promoter. *J Biol Chem.* 1998;273(48):32236-43.

295. Luderus ME, de Graaf A, Mattia E, den Blaauwen JL, Grande MA, de Jong L, van Driel R. Binding of matrix attachment regions to lamin B1. *Cell*. 1992;70(6):949-59.
296. Dickinson LA, Kohwi-Shigematsu T. Nucleolin is a matrix attachment region DNA-binding protein that specifically recognizes a region with high base-unpairing potential. *Mol Cell Biol*. 1995;15(1):456-65.
297. Dickinson LA, Joh T, Kohwi Y, Kohwi-Shigematsu T. A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell*. 1992;70(4):631-45.
298. Varga-Weisz P, van Holde K, Zlatanova J. Preferential binding of histone H1 to four-way helical junction DNA. *J Biol Chem*. 1993;268(28):20699-700.
299. Yaneva J, Schroth GP, van Holde KE, Zlatanova J. High-affinity binding sites for histone H1 in plasmid DNA. *Proc Natl Acad Sci U S A*. 1995;92(15):7060-4.
300. Ivanchenko M, Zlatanova J, van Holde K. Histone H1 preferentially binds to superhelical DNA molecules of higher compaction. *Biophys J*. 1997;72(3):1388-95.
301. Hartman PG, Chapman GE, Moss T, Bradbury EM. Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryote chromatin. The three structural regions of the histone H1 molecule. *Eur J Biochem*. 1977;77(1):45-51.
302. Clore GM, Gronenborn AM, Nilges M, Sukumaran DK, Zarbock J. The polypeptide fold of the globular domain of histone H5 in solution. A study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *EMBO J*. 1987;6(6):1833-42.
303. Cerf C, Lippens G, Ramakrishnan V, Muyldermans S, Segers A, Wyns L, Wodak SJ, Hallenga K. Homo- and heteronuclear two-dimensional NMR studies of the globular domain of histone H1: full assignment, tertiary structure, and comparison with the globular domain of histone H5. *Biochemistry*. 1994;33(37):11079-86.
304. Lewis JD, Saperas N, Song Y, Zamora MJ, Chiva M, Ausio J. Histone H1 and the origin of protamines. *Proc Natl Acad Sci U S A*. 2004;101(12):4148-52.
305. Thompson RJ. Studies on RNA synthesis in two populations of nuclei from the mammalian cerebral cortex. *J Neurochem*. 1973;21(1):19-40.
306. Piña B, Martínez P, Simon L, Suau P. Differential kinetics of histone H1(0) accumulation in neuronal and glial cells from rat cerebral cortex during postnatal development. *Biochem Biophys Res Commun*. 1984;123(2):697-702.
307. Garcia-Ramirez M, Leuba SH, Ausio J. One-step fractionation method for isolating H1 histones from chromatin under nondenaturing conditions. *Protein Expr Purif*. 1990;1(1):40-4.
308. Bohm EL, Strickland WN, Strickland M, Thwaites BH, van der Westhuizen DR, von Holt C. Purification of the five main calf thymus histone fractions by gel exclusion chromatography. *FEBS Lett*. 1973;34(2):217-21.
309. Brown DT, Alexander BT, Sittman DB. Differential effect of H1 variant overexpression on cell cycle progression and gene expression. *Nucleic Acids Res*. 1996;24(3):486-93.
310. Rodriguez AT, Perez L, Moran F, Montero F, Suau P. Cooperative interaction of the C-terminal domain of histone H1 with DNA. *Biophys Chem*. 1991;39(2):145-52.
311. Van Dyke MW, Hertzberg RP, Dervan PB. Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA.Fe(II). *Proc Natl Acad Sci U S A*. 1982;79(18):5470-4.
312. Fox KR, Waring MJ. Kinetic evidence for redistribution of actinomycin molecules between potential DNA-binding sites. *Eur J Biochem*. 1984;145(3):579-86.
313. Lu X, Hansen JC. Revisiting the structure and functions of the linker histone C-terminal tail domain. *Biochem Cell Biol*. 2003;81(3):173-6.
314. Hill CS, Rimmer JM, Green BN, Finch JT, Thomas JO. Histone-DNA interactions and their modulation by phosphorylation of -Ser-Pro-X-Lys/Arg- motifs. *EMBO J*. 1991;10(7):1939-48.
315. Huth JR, Bewley CA, Nissen MS, Evans JN, Reeves R, Gronenborn AM, Clore GM. The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol*. 1997;4(8):657-65.
316. Allan J, Hartman PG, Crane-Robinson C, Aviles FX. The structure of histone H1 and its location in chromatin. *Nature*. 1980;288(5792):675-9.
317. Roque A, Orrego M, Ponte I, Suau P. The preferential binding of histone H1 to DNA scaffold-associated regions is determined by its C-terminal domain. *Nucleic Acids Res*. 2004;32(20):6111-9.

318. Chen YH, Yang JT, Chau KH. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. *Biochemistry*. 1974;13(16):3350-9.
319. Arrondo JL, Goni FM. Structure and dynamics of membrane proteins as studied by infrared spectroscopy. *Prog Biophys Mol Biol*. 1999;72(4):367-405.
320. Arrondo JL, Echabe I, Iloro I, Hernando MA, de la Cruz F, Goni FM. A bacterial TrwC relaxase domain contains a thermally stable alpha-helical core. *J Bacteriol*. 2003;185(14):4226-32.
321. Arrondo JL, Castresana J, Valpuesta JM, Goni FM. Structure and thermal denaturation of crystalline and noncrystalline cytochrome oxidase as studied by infrared spectroscopy. *Biochemistry*. 1994;33(38):11650-5.
322. Dyson HJ, Rance M, Houghten RA, Wright PE, Lerner RA. Folding of immunogenic peptide fragments of proteins in water solution. II. The nascent helix. *J Mol Biol*. 1988;201(1):201-17.
323. Kennedy DF, Crisma M, Toniolo C, Chapman D. Studies of peptides forming 3(10)- and alpha-helices and beta-bend ribbon structures in organic solution and in model biomembranes by Fourier transform infrared spectroscopy. *Biochemistry*. 1991;30(26):6541-8.
324. Miick SM, Martinez GV, Fiori WR, Todd AP, Millhauser GL. Short alanine-based peptides may form 3(10)-helices and not alpha-helices in aqueous solution. *Nature*. 1992;359(6396):653-5.
325. Reisdorf WC Jr, Krimm S. Infrared amide I' band of the coiled coil. *Biochemistry*. 1996;35(5):1383-6.
326. Churchill ME, Travers AA. Protein motifs that recognize structural features of DNA. *Trends Biochem Sci*. 1991;16(3):92-7.
327. Manning GS. Theory of H1-mediated control of higher orders of structure in chromatin. *Biopolymers*. 1979;18(12):2929-42.
328. Wilson RW, Bloomfield VA. Counterion-induced condensation of deoxyribonucleic acid. a light-scattering study. *Biochemistry*. 1979;18(11):2192-6.
329. Bracken C, Iakoucheva LM, Romero PR, Dunker AK. Combining prediction, computation and experiment for the characterization of protein disorder. *Curr Opin Struct Biol*. 2004;14(5):570-6.
330. Fink AL. Natively unfolded proteins. *Curr Opin Struct Biol*. 2005;15(1):35-41.
331. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*. 2005;6(3):197-208.
332. Koop R, Di Croce L, Beato M. Histone H1 enhances synergistic activation of the MMTV promoter in chromatin. *EMBO J*. 2003;22(3):588-99.
333. Roque A, Iloro I, Ponte I, Arrondo JL, Suau P. DNA-induced secondary structure of the carboxyl-terminal domain of histone H1. *J Biol Chem*. 2005;280(37):32141-7.
334. Uversky VN. What does it mean to be natively unfolded? *Eur J Biochem*. 2002;269(1):2-12.
335. Uversky VN. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci*. 2002;11(4):739-56.
336. Uversky VN. Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding crossroads: which way to go? *Cell Mol Life Sci*. 2003;60(9):1852-71.
337. Munishkina LA, Cooper EM, Uversky VN, Fink AL. The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J Mol Recognit*. 2004;17(5):456-64.
338. Hansen JC, Lu X, Ross ED, Woody RW. Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J Biol Chem*. 2006;281(4):1853-6.
339. Minton AP. Implications of macromolecular crowding for protein assembly. *Curr Opin Struct Biol*. 2000;10(1):34-9.
340. van den Berg B, Wain R, Dobson CM, Ellis RJ. Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell. *EMBO J*. 2000;19(15):3870-5.
341. Qu Y, Bolen DW. Efficacy of macromolecular crowding in forcing proteins to fold. *Biophys Chem*. 2002;101-102:155-65.
342. Sasahara K, McPhie P, Minton AP. Effect of dextran on protein stability and conformation attributed to macromolecular crowding. *J Mol Biol*. 2003;326(4):1227-37.
343. Flaugh SL, Lumb KJ. Effects of macromolecular crowding on the intrinsically disordered proteins c-Fos and p27(Kip1). *Biomacromolecules*. 2001;2(2):538-40.
344. Dedmon MM, Patel CN, Young GB, Pielak GJ. FlgM gains structure in living cells. *Proc Natl Acad Sci U S A*. 2002;99(20):12681-4.
345. Uversky VN, M Cooper E, Bower KS, Li J, Fink AL. Accelerated alpha-synuclein fibrillation in crowded milieu. *FEBS Lett*. 2002;515(1-3):99-103.

346. Shtilerman MD, Ding TT, Lansbury PT Jr. Molecular crowding accelerates fibrillization of alpha-synuclein: could an increase in the cytoplasmic protein concentration induce Parkinson's disease? *Biochemistry*. 2002;41(12):3855-60.
347. Guinier A and Fournet G. Small angle scattering of X-rays. 1955. Wiley, New York.
348. Debye P. Molecular weight determination by light scattering. *J. Phys. Colloid Chem.* 1947;51:18-32.
349. Calmettes P, Durand D, Desmadril M, Minard P, Receveur V, Smith JC. How random is a highly denatured protein? *Biophys Chem.* 1994;53(1-2):105-13.
350. Glatter O and Kratky O. Small angle X-ray scattering. 1982. Academic Press. London.
351. Semisotnov GV, Kihara H, Kotova NV, Kimura K, Amemiya Y, Wakabayashi K, Serdyuk IN, Timchenko AA, Chiba K, Nikaïdo K, Ikura T, Kuwajima K. Protein globularization during folding. A study by synchrotron small-angle X-ray scattering. *J Mol Biol.* 1996;262(4):559-74.
352. Perez J, Vachette P, Russo D, Desmadril M, Durand D. Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering. *J Mol Biol.* 2001;308(4):721-43.
353. Millett IS, Doniach S, Plaxco KW. Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. *Adv Protein Chem.* 2002;62:241-62.
354. Goto Y, Fink AL. Conformational states of beta-lactamase: molten-globule states at acidic and alkaline pH with high salt. *Biochemistry.* 1989;28(3):945-52.
355. Semisotnov GV, Rodionova NA, Razgulyaev OI, Uversky VN, Gripas' AF, Gilmanshin RI. Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers.* 1991;31(1):119-28.
356. Privalov PL, Gill SJ. Stability of protein structure and hydrophobic interaction. *Adv Protein Chem.* 1988;39:191-234.
357. Griko YV, Privalov PL. Thermodynamic puzzle of apomyoglobin unfolding. *J Mol Biol.* 1994;235(4):1318-25.
358. Peng ZY, Kim PS. A protein dissection study of a molten globule. *Biochemistry.* 1994;33(8):2136-41.
359. Redfield C, Smith RA, Dobson CM. Structural characterization of a highly-ordered 'molten globule' at low pH. *Nat Struct Biol.* 1994;1(1):23-9.
360. Ptitsyn OB. Molten globule and protein folding. *Adv Protein Chem.* 1995;47:83-229.
361. Privalov PL. Intermediate states in protein folding. *J Mol Biol.* 1996;258(5):707-25.
362. Muramatsu N, Minton AP. Tracer diffusion of globular proteins in concentrated protein solutions. *Proc Natl Acad Sci U S A.* 1988;85(9):2984-8.
363. Zhou HX. Protein folding and binding in confined spaces and in crowded solutions. *J Mol Recognit.* 2004;17(5):368-75.
364. Roque A, Ponte I, Suau P. Macromolecular crowding induces a molten globule state in the C-terminal domain of histone H1. *Biophys J.* 2007;93(6):2170-7.
365. Gurley LR, D'Anna JA, Barham SS, Deaven LL, Tobey RA. Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. *Eur J Biochem.* 1978;84(1):1-15.
366. Voet D, Voet JG and Pratt CW. *Fundamentals of Biochemistry.* Upgrade edition 2002. John Wiley & Sons, Inc.
367. Marks F. *Protein Phosphorylation.* 1996. VCH Weinheim, New York, Basel, Cambridge, Tokyo.
368. Johnson LN, Lewis RJ. Structural basis for control by phosphorylation. *Chem Rev.* 2001;101(8):2209-42.
369. Vetter SW, Leclerc E. Phosphorylation of serine residues affects the conformation of the calmodulin binding domain of human protein 4.1. *Eur J Biochem.* 2001;268(15):4292-9.
370. Ojala PM, Yamamoto K, Castanos-Velez E, Biberfeld P, Korsmeyer SJ, Makela TP. The apoptotic v-cyclin-CDK6 complex phosphorylates and inactivates Bcl-2. *Nat Cell Biol.* 2000;2(11):819-25.
371. Chehin R, Iloro I, Marcos MJ, Villar E, Shnyrov VL, Arrondo JL. Thermal and pH-induced conformational changes of a beta-sheet protein monitored by infrared spectroscopy. *Biochemistry.* 1999;38(5):1525-30.
372. Cordeiro Y, Machado F, Juliano L, Juliano MA, Brentani RR, Foguel D, Silva JL. DNA converts cellular prion protein into the beta-sheet conformation and inhibits prion peptide aggregation. *J Biol Chem.* 2001;276(52):49400-9.
373. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. *Biochemistry.* 2002;41(21):6573-82.
374. Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, Brown CJ, Lawson JD, Dunker AK. Protein flexibility and intrinsic disorder. *Protein Sci.* 2004;13(1):71-80.

375. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Higgs KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. Intrinsically disordered protein. *J Mol Graph Model*. 2001;19(1):26-59.