Research article

Molecular cloning and sequence analysis of hamster CENP-A cDNA Javier Figueroa, Carlos Pendón and Manuel M Valdivia*

Address: Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad de Cádiz, 11510 Puerto Real, Cádiz, Spain E-mail: Javier Figueroa - fjavier.figueroa@uca.es; Carlos Pendón - carlos.pendon@uca.es; Manuel M Valdivia* - manuel.valdivia@uca.es *Corresponding author

Published: 3 May 2002

BMC Genomics 2002, 3:11

Received: 27 February 2002 Accepted: 3 May 2002

This article is available from: http://www.biomedcentral.com/1471-2164/3/11

© 2002 Figueroa et al; licensee BioMed Central Ltd. Verbatim copying and redistribution of this article are permitted in any medium for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: The centromere is a specialized locus that mediates chromosome movement during mitosis and meiosis. This chromosomal domain comprises a uniquely packaged form of heterochromatin that acts as a nucleus for the assembly of the kinetochore a trilaminar proteinaceous structure on the surface of each chromatid at the primary constriction. Kinetochores mediate interactions with the spindle fibers of the mitotic apparatus. Centromere protein A (CENP-A) is a histone H3-like protein specifically located to the inner plate of kinetochore at active centromeres. CENP-A works as a component of specialized nucleosomes at centromeres bound to arrays of repeat satellite DNA.

Results: We have cloned the hamster homologue of human and mouse CENP-A. The cDNA isolated was found to contain an open reading frame encoding a polypeptide consisting of 129 amino acid residues with a C-terminal histone fold domain highly homologous to those of CENP-A and H3 sequences previously released. However, significant sequence divergence was found at the N-terminal region of hamster CENP-A that is five and eleven residues shorter than those of mouse and human respectively. Further, a human serine 7 residue, a target site for Aurora B kinase phosphorylation involved in the mechanism of cytokinesis, was not found in the hamster protein. A human autoepitope at the N-terminal region of CENP-A described in autoinmune diseases is not conserved in the hamster protein.

Conclusions: We have cloned the hamster cDNA for the centromeric protein CENP-A. Significant differences on protein sequence were found at the N-terminal tail of hamster CENP-A in comparison with that of human and mouse. Our results show a high degree of evolutionary divergence of kinetochore CENP-A proteins in mammals. This is related to the high diverse nucleotide repeat sequences found at the centromere DNA among species and support a current centromere model for kinetochore function and structural plasticity.

Background

Centromeres provide the essential functions for chromosome segregation. They act as regulators of mitosis and meiosis by controlling attachment of chromosomes to the spindle and regulating progression into anaphase [1–4]. In mammalian cells, the centromere is associated with large blocks of heterochromatin comprised of highly repetitive satellite DNA of unconserved sequence repeats among species. It have been hypothesized that centromere function is therefore not specified directly by DNA sequence, but rather by higher order DNA or chromatin structure [5]. Centromeric DNA and protein components have been suggested to evolve rapidly in evolution and be responsible for the organization of functional centromere of emerging species [6].

CENP-A is a 17 kD histone H3 variant with homologous polypeptides being described from human to yeast [7–9]. Different approaches have indicated that CENP-A is required for the assembly of kinetochore and activation of the centromere [10–12]. The human protein shows more than 60% sequence identity to H3 at the C-terminal histone fold domain. The highly divergent N-termini of CENP-A is localized outside the nucleosome and should serve to interact with other kinetochore proteins including histones [4,13]. This unique N-terminal charged region forms a major epitope in the induction of specific autoantibodies against centromere in autoimmune patients [14,15]. The overall homology between homologous CENP-A proteins and the evolutionarily unconserved satellite repeat DNAs at the kinetochore described for many species, constitute a major unresolved enigma in modern biology for understanding the structural organization and function of the centromere.

Results and Discussion

In order to look for CENP-A conserved motifs in closed species, we have cloned CENP-A cDNA from hamster and compare the deduced amino acid sequence with that of mouse and human. Sequencing was done on both strands of DNA and the sequence data was deposited in GenBank with the accession number AJ428867. As shown in Fig. 1, we isolated a 1623nt cDNA containing several polyadenylation signals and a poly-A tail. The coding region was found between nucleotides 98 and 484. The cDNA of hamster CENP-A contains an open reading frame encoding a polypeptide consisting of 129 amino acid residues. Sequence analysis of deduced hamster CENP-A polypeptide showed 79% and 73% homology with those described previously of mouse and human respectively. CENP-A is defined as a histone H3-like protein and as expected the C-terminal histone H3 fold domain of the hamster cDNA (residues 37-129) showed a 84 % homology with those of human and mouse CENP-A cDNAs. However at the N-terminal domain (residues 1–36) the hamster antigen contains 5 deleted amino acid residues in comparison to that of mouse (Fig. 2). Further, in rodents there are also 5 deleted amino acid residues in the N-terminal related to that of the human CENP-A amino acid sequence. Significantly, this include a serine residue at position 7 which have been described in human CENP-A as a Aurora B kinase phosphorylation site involved in completion of cytokinesis [16]. It is plausible that other serine residues found at the N-terminal region of hamster CENP-A may serve for similar function. Overall changes found in the hamster CENP-A make the protein five and eleven amino acid shorter than those of mouse and human polypeptides respectively (Fig. 2). These observations indicate a degree of evolutionary divergence of protein sequences between hamster, mouse and human CENP-A and suggest a significant structural plasticity in an essential kinetochore component.

As suggested in a recent report on centromere DNA sequences divergences found in eukaryotes, those changes observed in the N-terminal region of several CENP-A proteins, could be explained for their specific recruitment at nucleosomes containing different DNA satellite repeat [6,17]. The centromere drive model proposed by Malik et al. [17] imply changes during evolution on centromeric repeat sequences associated with those found on centromeric constitutive antigens such as CENP-A. This specific event could play a role at the distinctive higher-order of chromatin structure at mammalian centromeres [18]. Thus, understanding the relationships between CENP-A primary structure and DNA repeat organization at the centromeres may provide a clue into the structural requirements needed for assembly of a kinetochore.

Conclusions

Hamster CENP-A and its human and murine homologous proteins are highly conserved at the C-terminal domain. However they have significant amino acid differences at the N-terminal tail, including a serine residue at position 7 which is missing in hamster and mouse CENP-A, but it is phosphorylated by Aurora B kinase in humans for a role in cytokinesis. Further, the N-terminal domain of hamster CENP-A is five and eleven amino acid shorter than those homologous proteins of mouse and human respectively. Thus, the divergences found in the N-terminal tail of CENP-A in different species during evolution, according to a current centromere model could be correlated with those variations widely described on centromeric DNA repeat sequences [6,17].

Materials and Methods Isolation of a hamster CENP-A cDNA

A CHO cDNA expression library constructed in lambda ZAP (Stratagene) was screened with a human CENP-A cDNA probe resulting in the isolation of several positive clones. One of these resulted to be a hamster full length CENP-A clone described here. cDNA was subcloned in pB-luescript vector (Stratagene) and DNA sequence analysis was performed in both strands by the dideoxy method using T7 DNA polymerase (Amersham Biosciences) and 5'- $[\alpha$ -³⁵S]-labeled dATP (NEN Life Science Prod. Zaventem, Belgium). DNA sequences were analysed in EMBL databases using BLAST and Clustaw softwares.

Authors' contributions

JF carried out the cloning studies, specific characterization of the clones and participate in the sequencing analysis.

M G P R R R K P R GACCCCCGAGAAGGCGCCCCTCCAGCCGGTTCCCGGACCCTGCGAGGCAGCTCCCGTCC T P R R R P S S P V P G P S R R S S R P AGGTAAGAGGCGGGAAATTTCTGTGGGCTTAAGGAGATAAAGAAACTGCAGAGGAGCACCGA G K R R K F L W L K E I K K L Q R S T D CCTCCTGCTAAGGAAGCTTCCTTTCAGCCGAGCTGTAAGGAGAAACTGCGAGAGAAATCAC J L L R K L P F S R V V R E I C G K F T TCGAGGTGTGGCACCTATGTGGCAAGCCCAAGCCTTGTTGGCCCTTCAAGAGGCAGCAGA R G V D L C W Q A Q A L L A L Q E A A E AGCATTTCTAGTCACCTCTTTGAGGAGTGCCAAGCCTAACCATGCCGGCGAGGCATCAAGGCCAGAGGAAATCTAGCCGGCAGGGCATTGAGGGAGG	GGCZ	ACGA	GCG	GAGC	TCC	TGC	GCT	CTG	ACA	CGC	TCA GCA	GTC ACC	ACC ATC	TCG GGC	CTC	TCT CGC		'CCG	GGG	CTT	60 120
GACCCCGAGAAAGGCGCCCCCCCCAGCCGGTTCCCGGACCCCGCGACGCAGCTCCCGTCC 180 T P R R P S S P V P G P S R R S S R P S S R P S S R P S S R P S S R P S S R P S S R P S S R P S T D C CCCCCCGCGGAGAGCCCCTCTCTCAGGGAGAAACCCGAAGGCCCTAAGGGCAGCAGAAACCGGGAGGAGCACCGAGGGGAGCACCGAGGGGGG	11000	50/10			1100	IIO	0110	0011	010	000	0011	1100	M	G	P	R	R	K	P	R	120
T P R R P S S P V P G P S R R S S R P AGGTAAGAGGCGGAAATTTCTGTGGCTTAAGGAGATAAAGAAACTGCAGGGGGCACCGA G K R R K F L W L K E I K K L Q R S T D CCTCCTGCTAAGGAAGCTTCCTTTCAGCCGAGCGCAGGGGATGTGTAAGAGAAATATGTGGGAAAGCCTTCCTT	GAC	ccce	GAGA	AGG	CGC	CCC	TCC	AGC	CCG	GTT	ccc	GGA	CCC	TCG	CGA	CGC	AGC	TCC	CG	rcc	180
AGGTAAGAGGCGGAAATTTCTGTGGCTTAAGGAGATAAAGAAACTGCAGAGGAGCACCGA 240 G K R R K F L W L K E I K K L Q R S T D CCTCCTGCTAAGGAAGCTTCCTTTCAGCCGAGTTGTAAGAGAAATATGTGGGGAAATTCAC 300 L L L R K L P F S R V V R E I C G K F T TCGAGGTGTGGACCTATGTTGGCAAGCCCAAGCCTTGTTGGCCCCTCAAGAGGCAGCAGA 360 R G V D L C W Q A Q A L L A L Q E A A E AGCATTCTAGTTCACCTCTTTGAGGAAGCCCAACCCAGGGGATCCCACGCGGCATGAGGGGAGGACT 420 A F L V H L F E D A Y L L T L H A G R V CACTATCTTCCCCAAGGACATACAGCTCACCAGGAGGACCCGAGGGCATTGAGGGGAGGACT 480 T I F P K D I Q L T R R I R G I E G C L 400 TGGCCGAGCTGGCTGGCGGGGGCCTGCCTGGACTCAGCTGGGGGCCTTCTGGAGCCACAG 640 G * TGGGCTGACACCCCAGGAGGCCCCAGGCCTGCGGGCCTTCTGGAGCCACAG 600 G TTGGGGTGCTGGAGGCGCCCCAGGCCTGGGGCCTTCTGGAGCCACAGG 600 600 G * TGGGCTGACCACGCGAGTGTGTGTGTGCCCAGCCAGGCCACGGGGGCCTTCTGGAGCCACAGG 600 CTGGGGTGAGACCACCGGAGTGTGTTATGTCCCAGGCCATGTCGCAGACCACGAGGACACTC 720 TGGGGTGAGCGCGCGCGCGCGCCGCCGGGCCTTCTGGGAGACACGCGAGAGG 600 CGCGTTAGGAGGCGCCTGCCAGGCGGCTTCTGTGGACACACGAGGCCTTCCATGGAGCCACACGAGGCCTTCTGCAACAACTGGCCAACTGCCAACGGCCATATCCATGGAGACTTTCATTGCCCATGGAGCCCACACGGCGAGACTCCCACAGGCCTTGAGGCCCACACGCGAGCTTGACCCCACATGCCGGCGTTGCTGTAGGCCTTGCCAACGCGAAATGCGCCGGCGCTTGAGGCCTGCCAACGCGAAATTGCGCCCATATCCATGGACCACACAAAATGGAGCCTTCTATAGGCGCCAAATGCCAAAACGCGCGCG	т	Ρ	R	R	R	Ρ	S	S	Ρ	v	Ρ	G	Ρ	S	R	R	S	S	R	Ρ	
G K R K F L W L K E I K K L Q R S T D CCTCCTGCTAAGGAAGCTTCCTTTCAGCCGAGTGTGTAAGAGAAATATGTGGGAAATTCAC 300	AGG!	raa (GAGG	CGG	AAA	TTT	CTG	TGG	CTT.	AAG	GAG.	АТА	AAG	ААА	CTG	CAG	AGG	AGC	AC	CGA	240
CCTCCTGCTAAGGAAGCTTCCTTTCAAGCCGAGTTGTAAGAGAAATATGTGGGGAAATTCAC 300 L L L R K L P F S R V V R E I C G K F T TGGAGGTGTGGGACCTATGTTGGCAAGCCCAAGCCTTGTTGGCCCTTCAAGAGGGCAGCAGA R G V D L C W Q A Q A L L A L Q E A A E C CG K F T T L A L L A L Q E A A E C C C C W Q L L T L L T L <td< th=""><th>G</th><th>K</th><th>R</th><th>R</th><th>K</th><th>F</th><th>L</th><th>W</th><th>L</th><th>K</th><th>Ε</th><th>I</th><th>к</th><th>к</th><th>L</th><th>Q</th><th>R</th><th>S</th><th>Т</th><th>D</th><th></th></td<>	G	K	R	R	K	F	L	W	L	K	Ε	I	к	к	L	Q	R	S	Т	D	
L L R K L P F S R V V R E I C G K F T TCGAGGGTGTGGACCTATGTTGGCAAGCCCAAGCCCTAGGCCTTGTGGGCCCTTCAAGAGGCCAGCAGA R G V D L C W Q A Q A L A L Q E A A E 20 A C C C C W W L F E D A Y L L T L H A C R C <td>ССТО</td> <td>ССТС</td> <td>CTA</td> <td>AGG</td> <td>AAG</td> <td>CTT</td> <td>CCT</td> <td>TTC</td> <td>AGC</td> <td>CGA</td> <td>GTT</td> <td>GTA</td> <td>AGA</td> <td>GAA</td> <td>ATA</td> <td>TGT</td> <td>'GGG</td> <td>AAA</td> <td>TTC</td> <td>CAC</td> <td>300</td>	ССТО	ССТС	CTA	AGG	AAG	CTT	CCT	TTC	AGC	CGA	GTT	GTA	AGA	GAA	ATA	TGT	'GGG	AAA	TTC	CAC	300
TCGAGGTGTGGACCTATGTTGGCAAGCCCAAGCCTTGTTGGCCCTTCAAGAGGCAGCAGAA R G V D L C W Q A Q A L L A L Q E A A E360AGCATTTCTAGTCACCTCTTTGAGGATGCCTATCTCCTCACCTTACATGCCGGCAGAGT A F L V H L F E D A Y L L T L H A G R V420CACTATCTTCCCCAAGGACATACAGCTCACCAGGAGGATCCGAGGCATTGAGGGAGG	L	L	L	R	K	L	Ρ	F	S	R	v	v	R	Е	I	С	G	K	F	т	
R G V D L C W Q A Q A L L A L Q E A A E AGCATTTCTAGTTCACCTCTTTGAGGAGGCCTATCTCCCCACCTTACATGCCGGCGCGCGC	TCG	AGGI	GTG	GAC	СТА	TGT	TGG	CAA	GCC	CAA	GCC	TTG	TTG	GCC	CTT	CAA	GAG	GCA	GCI	AGA	360
AGCATTTCTAGTTCACCTCTTTGAGGATGCCTATCTCCTCACCTTACATGCCGGCAGAGAT420A F L V H L F E D A Y L L T L H A G R VCACTATCTTCCCCAAGGACATACAGCTCACCAGGAGGATCCGAGGGCATTGAGGGAGG	R	G	V	D	L	С	W	Q	A	Q	A	L	L	A	L	Q	Ε	A	A	Ε	
A F L V H L F E D A Y L L H A G R V CACTATCTTCCCCAAGGACATACAGCTCACCAGGAGGATCCGAGGACTTGAGGGAGG	AGC	ATT]	CTA	GTT	CAC	CTC	TTT	GAG	GAT	GCC	TAT	стс	CTC	ACC	TTA	CAT	'GCC	GGC	AGI	AGT	420
CACTATCTTCCCCCAAGGACATACAGCTCACCAGGAGGATCCGAGGCATTGAGGGAGG	A	F	L	v	H	L	F	Е	D	A	Y	L	L	т	L	Н	A	G	R	v	
TIFPKDIQLTRRIRGIEGGLTGGCTGAGCTGGCAGTGTGTGTGTGTGTGTGTGTGCAGCCTACACGTGGGGCCTTCTGGAGCACACAG G**600GTAGCCCAGTGGTGTTGCAAGGAGCTGCCTGGACTCAGCTGGGTTTTGAATAACATGTACTT600GTAGCCCTTAAACTTTCTTCTGGAGAAGAAAAGACTGTAGGTTTTGCAGACACAGGTCAACT720TTGGGGTGTGAACAGCTCAACCATATAGCCCATGGACATGAGGTTTTGGACACACAGGACACACG780CGCGTTAGGAGGGGATGTTGTATGTTAGTCATGTTAGGTCTATGGACACACGAGACTCGC780CGCGTTAGGAGGGGATGTTGTAATGTAGCTTTTCGGTCTATGGACACCACGAGACTCGC780CGCGTTAGGAGGGGAGTGTTGTAATGTAGCTTTTGGGCCAGACACACAC	CAC	TATC	TTC	.ccc	AAG	GAC	ATA	CAG	CTC.	ACC	AGG	AGG	ATC	CGA	GGC	ATT	'GAG	GGA	GGI	ACT	480
TGGCTGAGCTGGCAGTGTGCTGTGTGTGCCAGCCTACACGTGGGGCCTTCTGGAGCCACAG540G*TGAGCCCAGTGGTGTCTGAGGAGCTGCCTGGACTCAGCTGGTTTTGAATAACATGTACTT600GTTGCCCTTTAAACTTTCTTTCTGGAAGAAAAGACTGTAGCTTTTCTCTGTGACAGAGAGG660AAAGAAAGGACCATCTGCATATGCCCATGGACATGAGTGTTTGCAGACCACAGGTCAACT720TTGGGGTGGACAGGCCAACCATATAACTGTTTAAGGGTTTTGGACACACGAGATCTGC780CGCGTTAGGAGGGGATGTTGTTATGTTAGTACTTTCTGGTCTATGTACATCTTTTACCATA840TGTTTATTTGTACTTTCATTTAAAATATAAGGGAGGAAAACTGGGGGGAAATGTTTGAAT900CTGTTGTATGGAGAGGCCCTGGCAGTCCTTGTTAGTGCCTGTCAACAAACA	Т	I	F	Ρ	K	D	I	Q	L	Т	R	R	I	R	G	I	Ε	G	G	L	
G *TGAGCCCAGTGGTGTCTGAGGAGCTGCCTGGACTCAGCTGGTTTTGAATAACATGTACTT600GTTGCCCTTTAAACTTTCTTTCTGAAGAAAAGACTGTAGCTTTTCCTGTGACAGAGAGG600AAAGAAAGGACCATCTGCATATGCCCATGGACATGAGTGTTTGCAGACCACAGGTCAACT720TTGGGGTGTGAACAGCTCAACCATATAACTGTTTAAGGGTTTTGGAGACCACAGGACCACG780CGCGTTAGGAGGGGATGTTGTTATGTTAGTCTTCTGGTCTATGTACACCATGAGACTCGC780CGCGTTAGGAGGGGAGGTGTGTTGTTATGTTAGTCTTTCTGGTCTATGTACACCACAGAGATCTGC700CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTTAGGTGCAAAACATGCAGCTT900CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACA	TGG	CTGF	GCI	GGC	AGT	GTG	CTG	TGT	GTC	CAG	CCT.	ACA	CGT	GGG	GCC	TTC	TGG	agc	CAC	CAG	540
TGAGCCCAGTGGTGTCTGAGGAGCTGCCTGGACTCAGCTGGTTTTGAATAACATGTACTT600GTTGCCCTTTAAACTTTCTTTCTGGAGAGAAAGACTGTAGCTTTTCTCTGTGACAGAGAGG660AAAGAAAGACCATCTGCATATGCCCATGGACATGAGTGTTTGCAGACCACAGGTCAACT720TTGGGGTGTGAACAGCTCAACCATATAACTGTTTAAGGGTGTTTGGACACACGGAGATCTGC780CGCGTTAGGAGGGGATGTTGTTATGTTAGTCTTTCTGGTCTATGTACATCTTTTACCATA840TGTTTATTTGTACTTTCATTTAAAATATAAGGAGGAGAAATCTGGGTGAAATGTTTGAAT900CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGCAACAAACA	G	*																			
GTTGCCCTTTAAACTTTCTTTCTGAAGAAAAGACTGTAGCTTTTCTCTGTGACAGAGAGG660AAAGAAAGGACCATCTGCATATGCCCATGGACATGAGTGTTTGCAGACCACAGGTCAACT720TTGGGGTGTGAACAGCTCAACCATATAACTGTTTAAGGGTTTTGCAGACACAGGAGATCTGC780CGCGTTAGGAGGGGATGTTGTTATGTTAGTTAGTTAGTCTTTCTGGTCTATGTACATCTTTACCATA840TGTTTATTTGTACTTTCATTTAAAATATAAAGGGAGGAAAATCTGGGTGAAATGTTTGAAT900CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACA	TGA	GCCC	CAGI	GGT	GTC	TGA	GGA	GCT	GCC	TGG.	АСТ	CAG	CTG	GTT	TTG	AAT	'AAC	ATG	TAC	CTT	600
AAAGAAAGGACCATCTGCATATGCCCATGGACATGAGTGTTTGCAGACCACAGGTCAACT720TTGGGGTGTGAACAGCTCAACCATATAACTGTTTAAGGGTTTTGGACACACGAGATCTGC780CGCGTTAGGAGGGGATGTTGTTATGTTATGTTAGTCTTTCTGGTCTATGTACATCTTTTACCATA840TGTTTATTTGTACTTTCATTTAAAATATAAGGGAGAAAATCTGGGTGAAATGTTTGAAT900CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACA	GTT	GCCC	CTTT	'AAA	CTT	TCT	TTC	TGA	AGA	AAA	GAC	TGT	AGC	TTT	TCT	СТС	TGA	CAG	GAGA	AGG	660
TTGGGGTGTGAACAGCTCAACCATATAACTGTTTAAGGGTTTTGGACACACGAGATCTGC780CGCGTTAGGAGGGGATGTTGTTATGTTAGTTAGTCTTTCTGGTCTATGTACATCTTTTACCATA840TGTTTATTTGTACTTTCATTTAAAATATAAGGGAGGAGAAAATCTGGGTGAAATGTTTGAAT900CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACA	AAA	GAAA	AGGA	ACCA	TCT	GCA	TAT	GCC	CAT	GGA	CAT	GAG	TGT	TTG	CAG	ACC	ACA	GGI	CAR	ACT	720
CGCGTTAGGAGGGGATGTTGTTATGTTATGTTAGTCTTTCTGGTCTATGTACATCTTTTACCATA840TGTTTATTTGTACTTTCATTTAAAATATAAGGGAGGAAAATCTGGGTGAAATGTTTGAAT900CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACA	TTG	GGGI	GTG	GAAC	AGC	TCA	ACC	ATA	TAA	CTG	TTT.	AAG	GGT	TTT	GGA	CAC	ACG	AGA	TC	ſGC	780
TGTTTATTTGTACTTTCATTTAAAATATAAGGGAGGAAAATCTGGGTGAAATGTTTGAAT900CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACA	CGC	GTTA	AGGA	GGG	GAT	GTT	GTT	ATG	ΓΤΑ	GTC	TTT	CTG	GTC	TAT	GTA	CAT	CTI	TTA	CCA	ATA	840
CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACA	TGT	TAT	TTG	STAC	TTT	CAT	TTA.	AAA	TAT.	AAG	GGA	GGA	AAA	ТСТ	GGG	TGA	AAT	GTI	TGA	AAT	900
TCAAAACCAAACTAGAATTGATTACAAATCTTCCTTCCCTTCCATGCTAGGGCTTGAATC1020CAAAGCTTTGTACAGGCTAAGTGTGGACTATACTGTTGAGCCATATCCATAGCCTGAAAT1080GAACAGTCTTCTAATTCCTACCAGTAAATGCCTATCAATATCATCCAGAAAGCCTGTCAT1140AGAACTGGACATTTTATTCTAAGCTGTAAATCATATAAAATGTATCCTATCAGTCTGGAG1200TGTTAAAATGTGTTTTCAGTATTTTTACAGGAGTTTCTTTTAGCTTATTTGGGGAGATG1260TAATAAAAAATTTCTTAATGAAAAATTAAAGTTACTGTCCTAGGTGTTACAAGCCTTTT1320AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCAAAATGTAAACAAGAGATG1380TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	CTG	ГТGI	ATG	GAG	AGC	CCT	GGC	AGT	TCC	TTG	TTA	GTG	ССТ	GTC	AAC	AAA	CAI	GCA	GC	ГТТ	960
CAAAGCTTTGTACAGGCTAAGTGTGGACTATACTGTTGAGCCATATCCATAGCCTGAAAT1080GAACAGTCTTCTAATTCCTACCAGTAAATGCCTATCAATATCATCCAGAAAGCCTGTCAT1140AGAACTGGACATTTTATTCTAAGCTGTAAATCCATATAAAATGTATCCTATCAGTCTGGAG1200TGTTAAAATGTGTTTTCAGTATTTTACAGGAGTTTCTTTTAGCTTATTTGGGGAGATG1260TAATAAAAAATTTCTTAATGAAAAATTAAAGTTACTGTCCTAGGTGTTACAAGCCTTTT1320AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCCAAAATGTAAACAAGAGATG1380TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	TCA	AAA	CAA	ACT	AGA	ATT	GAT	TAC	AAA	ТСТ	TCC	TTC	ССТ	TCC	ATG	СТА	GGG	CTI	'GA <i>I</i>	ATC	1020
GAACAGTCTTCTAATTCCTACCAGTAAATGCCTATCAATATCATCCAGAAAGCCTGTCAT1140AGAACTGGACATTTTATTCTAAGCTGTAAATCATATAAAATGTATCCTATCAGTCTGGAG1200TGTTAAAATGTGTTTTCAGTATTTTTACAGGAGTTTCTTTTTAGCTTATTTGGGGAGATG1260TAATAAAAAATTTCTTAATGAAAAATTAAAGTTACTGTCCTAGGTGTTACAAGCCTTTTT1320AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCCAAAATGTAAACAAGAGATG1380TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	CAA	AGCI	TTG	STAC	AGG	СТА	AGT	GTG	GAC	TAT	ACT	GTT	GAG	CCA	TAT	ССА	TAG	CCI	'GA <i>I</i>	AAT	1080
AGAACTGGACATTTTATTCTAAGCTGTAAATCATATAAAATGTATCCTATCAGTCTGGAG1200TGTTAAAATGTGTTTTCAGTATTTTACAGGAGTTTCTTTTAGCTTATTTGGGGAGATG1260TAATAAAAAATTTCTTAATGAAAAATTAAAGTTACTGTCCTAGGTGTTACAAGCCTTTT1320AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCAAAATGTAAACAAGAGATG1380TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	GAA	CAGI	CTT	CTA	ATT	CCT	ACC	AGT.	AAA	TGC	СТА	ТСА	ATA	TCA	TCC	AGA	AAG	CCI	GTO	CAT	1140
TGTTAAAATGTGTTTTCAGTATTTTACAGGAGTTTCTTTTTAGCTTATTTGGGGAGATG1260TAATAAAAATTTCTTAATGAAAAATTAAAGTTACTGTCCTAGGTGTTACAAGCCTTTT1320AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCAAAATGTAAACAAGAGATG1380TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	AGA	ACTO	GAC	CATT	TTA	TTC	TAA	GCT	gta.	AAT	CAT.	АТА	AAA	TGT	ATC	СТА	TCA	GTC	TGC	GAG	1200
TAATAAA1320AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCCAAAATGTAAACAAGAGATG1380TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	TGT	raa <i>f</i>	ATG	STGT	TTT	CAG	TAT	TTT	TAC.	AGG.	AGT	TTC	TTT	TTA	GCT	TAT	TTG	GGG	GAGA	ATG	1260
AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCCAAAATGTAAACAAGAGATG1380TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	TAA!	[AA7	AAA	TTT	CTT	AAT	GAA.	AAA	TTA	AAG	TTA	CTG	TCC	TAG	GTG	TTA	CAA	GCC	TTT	ГТТ	1320
TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACAC	AGT	AAT	'ATT	'GAG	CTC	GTT	TTT	TAC	ACC	ATT	TGA	AAC	AGC	САА	ААТ	GTA	AAC	AAG	AGA	ΑTG	1380
CAGTAAAAGAAATGTAAGAGGTTCCAGGAAAATAGGCCACAATGCTCTTGGACTTGAGGT1500GAAAACAGCTGTGAGCACTAAACAGTCTGTGCTATTGATGGAGACATGCCAGCTGATTGA1560ACTGGAGTTCCAGGTTTTGCTATTTAATAAACTTTACCCTTTGCAAAAAAAA	TGA	GTGG	GTTC	CGGA	TTT	GCC	ACC	GAT	GAG	GTG	AAC	ACA	CAC	AAA	ATT	TGC	TCT	TTT	GT	ſGG	1440
GAAAACAGCTGTGAGCACTAAACAGTCTGTGCTATTGATGGAGACATGCCAGCTGATTGA1560ACTGGAGTTCCAGGTTTTGCTATTTAATAAACTTTACCCTTTGCAAAAAAAA	CAG	raa <i>f</i>	AGA	AAT	GTA	AGA	GGT	TCC	AGG	AAA	ATA	GGC	CAC	AAT	GCT	CTT	'GGA	LCTT	'GA(GGT	1500
ACTGGAGTTCCAGGTTTTGCTATTT AATAAA CTTTACCCTTTGCAAAAAAAAAAAAAAAAAAAA	GAA	AACA	AGCT	GTG	AGC	ACT	AAA	CAG	ГСТ	GTG	CTA	TTG	ATG	GAG	ACA	TGC	CAG	CTG	ATT	ГGА	1560
AAA 1623	ACT	GGAC	GTTC	CAG	GTT	TTG	CTA	TTT	AAT	AAA	CTT	TAC	CCT	TTG	CAA	AAA	AAA	AAA	AAA	AAA	1620
	AAA	-	-	_				-				-		-							1623

Figure I

Nucleotide sequence of hamster CENP-A cDNA. Deduced amino acids of 129 encoded codons are shown in black. Numbers in the right indicate nucleotide position. The asterisk designates the stop codon. Two consensus polyadenylation signal sequences are shown in black underlined.

Hamster	MGPRRKPRTPRRRPSSPVPGPSRRSSRPGKRRKFLWLKEI	40
Mouse	MGPRR <mark>KP</mark> QTPRRRPSSP <mark>APGPSR</mark> Q <mark>SSS</mark> V <mark>G</mark> SQ-TLRR <mark>RQKF</mark> MWLKEI	45
Human	MGPRRRSR <mark>KP</mark> EAPRRRSPSPTPTPGPSRRGP <mark>S</mark> LGASSHQHS <mark>RR</mark> RQG <mark>WLKEI</mark>	51
Hamster	KKLQR <mark>STDLL</mark> LRKLPFSRVVREICGKFTRGVDLCWQAQALLALQEAAEAFL	91
Mouse	KTLQKSTDLLFRKKPFSMVVREICEKFSRGVDFWWQAQALLALQEAAEAFL	96
Human	R <mark>KLQKST</mark> HLLIRKLPFSRLAREICVKFTRGVDFNWQAQALLALQEAAEAFL	102
Hamster	VHLFEDAYLLTLHAGRVT <mark>IFPKDIQLTRRIRG</mark> I <mark>EGGLG</mark>	129
Mouse	I <mark>HLFEDAYLL</mark> S <mark>LHAGRVTLFPKDIQLTRRIRG</mark> F <mark>EGGL</mark> P	134
Human	VHLFEDAYLLTLHAGRVTLFPKD <mark>VQL</mark> ARRIRG <mark>LE</mark> EGLG	140

Figure 2

Alignment of hamster, mouse and human CENP-A amino acid sequences. The alignment has been shaded with identical residues. The C-terminal histone H3 fold domain (aa 37–129) is highly conserved. In contrast, several residues of the N-terminal tail are variable among the three species compared, including several deletions as shown. The accession numbers are as follows: hamster AJ428867 (this work); mouse AF012708-10 and human U14518-19.

CP carried out the sequence alignment and databases comparison and participates in the writing and drafted the manuscript.

MMV conceived the study and its design coordination and writing.

All authors read and approved the final manuscript.

Acknowledgments

We thanks Dr. Kevin Sullivan of Scripps Research Institute, La Jolla CA, for supply us with a human CENP-A cDNA probe. This work was supported in part by grants from Comisión Interministerial de Ciencia y Tecnología (PB97-1354) and by Plan Andaluz de Investigación (CVI 167) of Junta de Andalucía. We are grateful to Unidad de Radioisótopos of University of Cádiz for their assistance in using their facilities.

References

- McIntosh JR: Structural and mechanical control of mitosis progression Cold Spring Harbor Symp. Quant, Biol 1991, 56:613-619
 Nicklas RB, Ward SC, Gorbsky GJ: Kinetochore chemistry is sen-
- Nicklas RB, Ward SČ, Gorbsky GJ: Kinetochore chemistry is sensitive to tension and may link mitotic forces to a cell cycle checkpoint J. Cell Biol 1995, 130:929-939
- 3. Rieder CL, Cole RW, Khodiakov A, Sluder G: The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores J. Cell Biol 1995, 130:941-948
- 4. Choo KHA: Domain organization at the centromere and neocentromere Developmental Cell 2001, 1:165-177
- 5. Karpen GH, Allshire RC: **The case for epigenetic effects on centromere identity and function** *Trends Genet* 1997, 13:489-496
- Henikoff S, Ahmad K, Malik HS: The centromere paradox: stable inheritance with rapidly evolving DNA Science 2001, 293:1098-1102
- Palmer DK, O'Day K, Wener MH, Andrews BS, Margolis RL: A 17kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones J. Cell Biol 1987, 104:805-815

- 8. Sullivan KF, Hechenberger M, Masri K: Human CENP-A contains a histone H3 fold domain that is required for targeting to the centromere J. Cell Biol. 1994, 127:581-592
- 9. Stoler S, Keith KC, Curnick KE, Fitzgerald-Hayes M: A mutation in CSE4 an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis *Genes & Dev* 1995, 9:573-586
- Figueroa J, Saffrich R, Ansorge W, Valdivia MM: Microinjection of antibodies to centromere protein CENP-A arrests cells in interphase but does nor prevent mitosis Chromosoma 1998, 107:397-405
- Howman EV, Fowler KJ, Newson AJ, Redward S, MacDonald AC, Kalitsis P, Choo KHA: Early disruption of centromeric chromatin organization in centromere protein A (CENP-A) null mice Proc. Natl. Acad. Sci. USA 2000, 97:1148-1153
- 12. Blower MD, Karpen GH: The role of Drosophila CENP-A/CID in kinetochore formation, cell-cycle progression and interactions with heterochromatin Nature Cell Biol. 2001, 3:730-739
- Shelby RD, Vafa O, Sullivan KF: Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites J. Cell Biol 1997, 136:501-513
- Moroi Y, Peebles C, Fritzler MJ, Steigerwald J, Tan EM: Autoantibody to centromere (kinetochore) in scleroderma sera Proc. Natl. Acad. Sci. USA, 1980, 77:1627-1631
- Valdivia MM, Figueroa J, Iglesias C, Ortiz M: A novel centromere monospecific serum to a human autoepitope on the histone H3-like protein CENP-A FEBS Letters 1998, 422:5-9
- Zeitlin SG, Shelby RD, Sullivan KF: CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis J. Cell Biol 2001, 155:1147-1157
- Malik HS, Vermaak D, Henikoff S: Recurrent evolution of DNAbinding motifs in the Drosophila centromeric histone Proc. Natl. Acad. Sci. USA 2002, 99:1449-1454
- Gilbert N, Allan J: Distinctive higher-order chromatin structure at mammalian centromeres Proc. Natl. Acad. Sci. USA 2001, 98:11949-11954