BMC Genomics

BioMed Central

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

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 pBluescript vector (Stratagene) and DNA sequence analysis was performed in both strands by the dideoxy method using T7 DNA polymerase (Amersham Biosciences) and 5'-[α-^{35}S]-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.

```
GGCACGAGCGAGCTCCTGCGCTCTGACACGCTCAGTCACCTCGCTCTCTCGTCCGGGCTT       60
AGCGGAGGCGCGGACGTTGGACCCACTGGCGGCAACC**ATGGGCCCGCGCCGCAAGCCGAG**      120
                                        **M  G  P  R  R  K  P  R**

**GACCCCGAGAAGGCGCCCCTCCAGCCCGGTTCCCGGACCCTCGCGACGCAGCTCCCGTCC**      180
**  T  P  R  R  R  P  S  S  P  V  P  G  P  S  R  R  S  S  R  P**

**AGGTAAGAGGCGGAAATTTCTGTGGCTTAAGGAGATAAAGAAACTGCAGAGGAGCACCGA**      240
**  G  K  R  R  K  F  L  W  L  K  E  I  K  K  L  Q  R  S  T  D**

**CCTCCTGCTAAGGAAGCTTCCTTTCAGCCGAGTTGTAAGAGAAATATGTGGGAAATTCAC**      300
**  L  L  L  R  K  L  P  F  S  R  V  V  R  E  I  C  G  K  F  T**

**TCGAGGTGTGGACCTATGTTGGCAAGCCCAAGCCTTGTTGGCCCTTCAAGAGGCAGCAGA**      360
**  R  G  V  D  L  C  W  Q  A  Q  A  L  L  A  L  Q  E  A  A  E**

**AGCATTTCTAGTTCACCTCTTTGAGGATGCCTATCTCCTCACCTTACATGCCGGCAGAGT**      420
**  A  F  L  V  H  L  F  E  D  A  Y  L  L  T  L  H  A  G  R  V**

**CACTATCTTCCCCAAGGACATACAGCTCACCAGGAGGATCCGAGGCATTGAGGGAGGACT**      480
**  T  I  F  P  K  D  I  Q  L  T  R  R  I  R  G  I  E  G  G  L**

**TGGCTGA**GCTGGCAGTGTGCTGTGTGTCCAGCCTACACGTGGGGCCTTCTGGAGCCACAG      540
**  G  ***

TGAGCCCAGTGGTGTCTGAGGAGCTGCCTGGACTCAGCTGGTTTTGAATAACATGTACTT      600
GTTGCCCTTTAAACTTTCTTTCTGAAGAAAAGACTGTAGCTTTTCTCTGTGACAGAGAGG      660
AAAGAAAGGACCATCTGCATATGCCCATGGACATGAGTGTTTGCAGACCACAGGTCAACT      720
TTGGGGTGTGAACAGCTCAACCATATAACTGTTTAAGGGTTTTGGACACACGAGATCTGC      780
CGCGTTAGGAGGGGATGTTGTTATGTTAGTCTTTCTGGTCTATGTACATCTTTTACCATA      840
TGTTTATTTGTACTTTCATTTAAAATATAAGGGAGGAAAATCTGGGTGAAATGTTTGAAT      900
CTGTTGTATGGAGAGCCCTGGCAGTTCCTTGTTAGTGCCTGTCAACAAACATGCAGCTTT      960
TCAAAACCAAACTAGAATTGATTACAAATCTTCCTTCCCTTCCATGCTAGGGCTTGAATC     1020
CAAAGCTTTGTACAGGCTAAGTGTGGACTATACTGTTGAGCCATATCCATAGCCTGAAAT     1080
GAACAGTCTTCTAATTCCTACCAGTAAATGCCTATCAATATCATCCAGAAAGCCTGTCAT     1140
AGAACTGGACATTTTATTCTAAGCTGTAAATCATATAAAATGTATCCTATCAGTCTGGAG     1200
TGTTAAAATGTGTTTTCAGTATTTTTACAGGAGTTTCTTTTTAGCTTATTTGGGGAGATG     1260
T**AATAAA**AAATTTCTTAATGAAAAATTAAAGTTACTGTCCTAGGTGTTACAAGCCTTTTT     1320
AGTAAATATTGAGCTCGTTTTTTACACCATTTGAAACAGCCAAAATGTAAACAAGAGATG     1380
TGAGTGGTTCGGATTTGCCACCGATGAGGTGAACACACACAAAATTTGCTCTTTTGTTGG     1440
CAGTAAAAGAAATGTAAGAGGTTCCAGGAAAATAGGCCACAATGCTCTTGGACTTGAGGT     1500
GAAAACAGCTGTGAGCACTAAACAGTCTGTGCTATTGATGGAGACATGCCAGCTGATTGA     1560
ACTGGAGTTCCAGGTTTTGCTATTT**AATAAA**CTTTACCCTTTGCAAAAAAAAAAAAAAAA     1620
AAA                                                              1623
```

Figure 1
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.

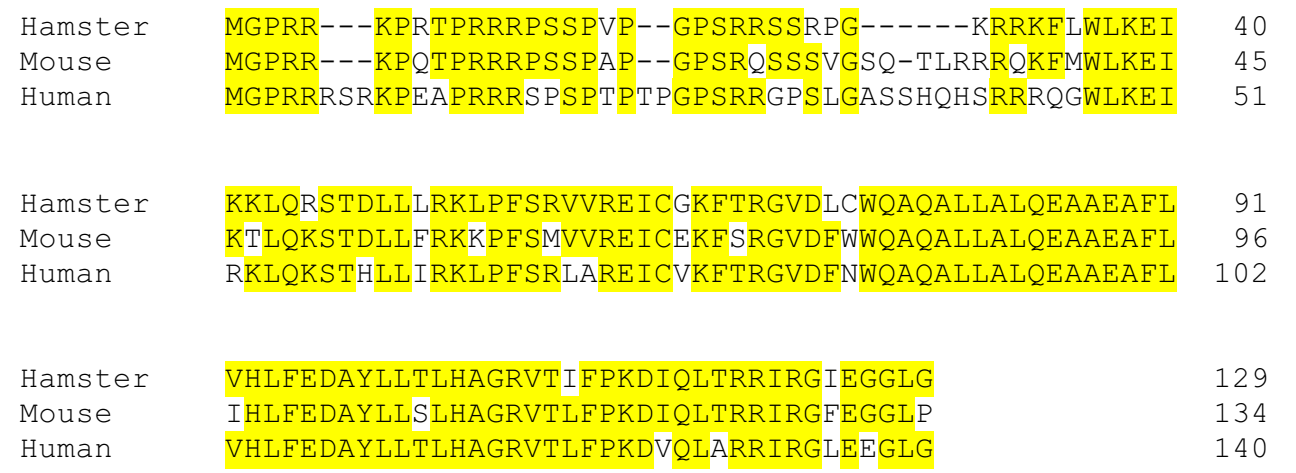

Figure 2
Alignment of hamster, mouse and human CENP-A amino acid sequences. The alignment has been shaded with identical residues. The C-terminal histone H3 fold domain (aa 37–129) is highly conserved. In contrast, several residues of the N-terminal tail are variable among the three species compared, 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

1. McIntosh JR: **Structural and mechanical control of mitosis progression** *Cold Spring Harbor Symp. Quant, Biol* 1991, **56:**613-619
2. Nicklas RB, Ward SC, 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
6. Henikoff S, Ahmad K, Malik HS: **The centromere paradox: stable inheritance with rapidly evolving DNA** *Science* 2001, **293:**1098-1102
7. 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
10. 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
11. 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
13. 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
14. 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
15. 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
16. 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
17. Malik HS, Vermaak D, Henikoff S: **Recurrent evolution of DNA-binding motifs in the Drosophila centromeric histone** *Proc. Natl. Acad. Sci. USA* 2002, **99:**1449-1454
18. Gilbert N, Allan J: **Distinctive higher-order chromatin structure at mammalian centromeres** *Proc. Natl. Acad. Sci. USA* 2001, **98:**11949-11954