Journal of Cellular Biochemistry 49:266–271 (1992)

Extremely Divergent Histone H_4 Sequence From *Trypanosoma cruzi*: Evolutionary Implications

G. Cecilia Toro, Christer Wernstedt, Carlos Medina, Nicolás Jaramillo, Ulf Hellman, and Norbel Galanti

Department of Cell Biology and Genetics, School of Medicine, University of Chile, Casilla 70061, Santiago 7, Chile (G.C.T., N.J., N.G.); Department of Biology, Faculty of Sciences, University of Chile (C.M.); and Ludwig Institute for Cancer Research, Uppsala Branch, S-751 24, Uppsala, Sweden (C.W., U.H.)

Abstract *Trypanosoma cruzi* presents six histones electrophoretically resolved in three gel systems. Indirect evidence shows that one of these histones, named e, corresponds to H_4 in other species. We present evidence that histone e is H_4 by sequencing its amino terminal end. The amino terminal of *T. cruzi* histone H_4, unlike that of other H_4s examined thus far is not blocked. Moreover, this protein presents two variants. This partial amino acid sequence of *T. cruzi* histone H_4 differs greatly from homologous sequences of human, yeast, or *Tetrahymena*.

Since the conservatism of the core histones (H_2A, H_2B, H_3, and H_4) is clearly illustrated by comparative sequence analyses, the data shown here demonstrates that *T. cruzi* histone H_4 is the most divergent reported. Quantitative analysis of the data suggests that the rate of substitutions in the histone H_4 amino terminal sequence varies among different lineages. We postulate a slow-down in the evolutionary rate of histone H_4 amino terminal domain in the metazoa branch related perhaps to the appearance of a novel function for this domain. © 1992 Wiley-Liss, Inc.

Key words: histone H_4, chromatin, amino terminal domain, molecular clock, evolution

Eukaryotic chromatin contains, in addition to DNA, a core of histone proteins (H_2A, H_2B, H_3, and H_4) that are highly conserved among different phyla (Behe, 1990). Nevertheless, significant differences have been found in the number, as well as in the electrophoretic mobilities, of protist histones when they are compared with the histones of higher eukaryotes (Rizzo, 1985; Gorovsky, 1986; Toro and Galanti, 1988).

While all of the core histones are important for nucleosome formation, histones H_3 and H_4 play a central role. An octamer containing H_3 and H_4 forms a structure with topological and hydrodynamic properties similar to that of a nucleosome (Simon et al., 1978).

Trypanosoma cruzi chromatin is organized in nucleosomes and the characterization of the histones has shown that these proteins differ in various aspects when compared to those of other eukaryotes (Astolfi et al., 1980; Toro and Galanti, 1990). Moreover *T. cruzi* chromatin, in contrast to vertebrate, is physically and enzymatically fragile (Rubio et al., 1980; Hecker and Gander, 1985).

During many years, the fact that histone H_4 sequence of pea differs from that of mammals by only two conservative substitutions in 102 residues led to the belief that few, if any, substitutions could be tolerated in its sequence (De Lange et al., 1969; McGhee and Felsenfeld, 1980; Kayne et al., 1988). Nevertheless, it was shown that the amino terminal domain of histone H_4 in *Tetrahymena* presents a high number of substitutions (Claiborne et al., 1979). Moreover, Kayne et al. (1988) found that different deletions in the amino terminal sequence of histone H_4 in yeast did not alter its basic functions. They also found that yeast with nucleosomes containing histones H_4 with altered amino terminal sequences were viable but they were not able to reproduce sexually. These results proved unambiguously that yeast histone H_4 could accept substitutions in the amino terminal domain. Moreover, these works defined "in vivo" different functions for the amino terminal domain of histone H_4 and described a unique function for histone H_4 amino terminus not shared by either H_2A, H_2B, and H_3 amino terminal domains.

However, more recent works of Megee et al. (1990) and Durrin et al. (1991) have shown that

Received February 12, 1992; accepted February 19, 1992.

Address reprint requests to Norbel Galanti, Department of Cell Biology and Genetics, School of Medicine, University of Chile, Casilla 70061, Santiago 7, Chile.

© 1992 Wiley-Liss, Inc.

certain positions at the amino terminal sequence of yeast histone H_4 which are involved in post-translational modifications do not accept substitutions without the loss of particular functions.

In agreement with the observations of dramatic changes in the electrophoretic mobilities of histones from protists (Rizzo, 1985) and from *T. cruzi* (Rubio et al., 1980; Toro and Galanti, 1990), we report a highly divergent amino terminal sequence in *T. cruzi* histone H_4.

Although histone H_4 was pointed out as a classical example of a protein which could not be used for the estimation of phylogenetic relationships due to its slow rate of substitutions, *T. cruzi* histone H_4, with a very divergent amino terminal sequence in a domain described structurally (Van Holde, 1989) and functionally (Megee et al., 1990; Durrin et al., 1991) as a unit, permitted us to test the molecular clock hypothesis in relation to this domain.

METHODS

Cell Culture

T. cruzi epimastigotes strain Tulahuen were grown in Diamond medium (Diamond, 1968) supplemented with 2.5% fetal calf serum at 28°C. Cells were routinely harvested by centrifugation at day 7 of culture, washed three times in phosphate-buffered saline (PBS), and used immediately.

Preparation of Chromatin

Basically, the procedure of Stein et al. (1975) was used, with some modifications (Toro and Galanti, 1990).

Extraction of Histones

The procedure of Panyin and Chalkey (1969) as described previously (Toro and Galanti, 1990) was applied.

Electrophoretic Blotting Procedures

Histones (30 μg) were loaded onto gels containing 15% acrylamide, 0.9 N acetic acid, 6 M urea, and 0.38% Triton DF-16, as indicated by Alfageme et al. (1974). After electrophoresis, the gels were soaked in transfer buffer:25 mM Tris-HCl, pH 8.4 (Walsh et al., 1988) for 5 min. During this time a polyvinylidenedifluoride (PVDF) membrane was rinsed with 100% methanol for 3 s, immersed in water for 2 min to elute the methanol, and then equilibrated in transfer buffer. The gel, sandwiched between a sheet of PVDF membrane and several sheets of blotting paper, was assembled into a blotting apparatus (Sartorius, Semi Dry Transfer Cell) and electroeluted for 25 min at 14 V in transfer buffer. The PVDF membrane was stained directly with 0.1% Coomassie brilliant blue in 50% methanol for 2 min and destained in 1% acetic acid-50% methanol. Bands corresponding to histone H_4 (band "e") and its two variants were cut out and sequenced by automated Edman degradation in an Applied Biosystems Protein sequencer (Applied Biosystems, Foster City, CA)

Histone separation was also performed by narrow-bore reversed phase high performance liquid chromatography (HPLC) on a 2.1 × 100 mm, 3 μm support (μRPC C2/C18 from Pharmacia, Uppsala Sweden). The column was operated at a flow rate of 100 μl/min at room temperature and the eluate was monitored at 215 nm.

Sequence Comparisons

The amino terminal sequence of band e protein was screened against the Protein Identification Resource (PIR) of the National Biomedical Research Foundation with the FASTP program (Lipman and Pearson, 1985). The statistical significance of the scores obtained with FASTP was evaluated with the program RDF (Lipman and Pearson, 1985). According to this method, a z value higher than 6 s.d. above mean is probably significant and z is undoubtedly significant when higher than 10 s.d. *Trypanosoma cruzi* sequence was included in histone H_4 alignment (Wells and McBride, 1989) using the FASTP program. Based on this sequence alignment, a distance matrix was constructed. Unrooted phylogenetic trees were obtained with the protein parsimony program PROTPARS of J. Felsenstein's PHYLIP package, version 3.1.

RESULTS

Figure 1 shows the electrophoretic pattern of histones obtained from *T. cruzi* (lane 1) and sea urchin (lane 2) in a Triton-acid-urea system. Band e is resolved in three bands, named e_1, e_2, and e_3, considering their mobilities in the gel.

The amino terminal sequence of these proteins transferred to PVDF membranes gave basically the same sequences in 30 cycles of Edman degradation. Nevertheless, the amino acid in position 4 of bands e_1 and e_2 was unidentified, probably due to a post-translational modification. This amino acid was identified as a lysine

in the variant e_3. In the sequence shown in Figure 2, amino acids number 31–41 were sequenced after isolation of histone H_4 by reversed phase HPLC.

The screening of the PIR database with the amino-terminal sequence gave the highest scores with histone H_4 (Table I). None of the other histones nor any other protein gave scores close to those obtained with histone H_4. An alignment including histones H_4 from distantly related lineages is displayed in Figure 2. Two regions, according to the degree of sequence conservation, can be seen in the alignment. The limit between these regions, following this criterion, can be set in two alternative positions. One of these positions (amino acid 33 in human sequence numeration) coincides roughly with the accepted limit between the amino terminal domain and the globular domain (Van Holde, 1989). Thus, we chose this position as the carboxy end of the sequences compared between different taxa. The number of amino acid differences, excluding positions with gaps, and the percent differences are presented in Table II.

Based on alignment from Figure 2, three most parsimonious unrooted trees, each one with 29 inferred nucleotide substitutions, were obtained

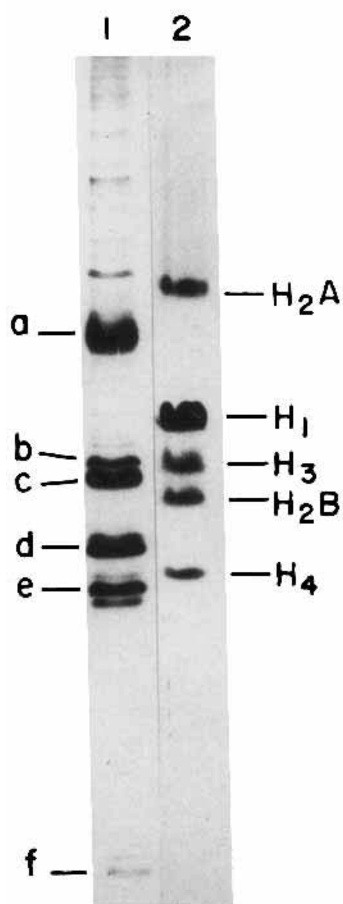

Fig. 1. Triton-acid-urea gel of histones extracted from *T. cruzi* epimastigotes (lane 1) and histones from sea urchin sperm cells (lane 2). Order of migration for sea urchin sperm cells: H_2A, H_1, H_3, H_2B, and H_4. a, b, c, d, e, and f, *T. cruzi* histones.

TABLE I. Statistical Significance (z) According to RDF Program of Scores Obtained From Sequence Comparisons*

Proteins compared with *T. cruzi* band e amino-terminal sequence	z (s.d. above mean)
Histone H_4, Baker's yeast	**14.44**
Histone H_4, human	**13.04**
Histone H_4, *Tetrahymena* thermophila	**8.86**
Histone H_3, pea	−0.77
Histone H_2A, human	−1.91
Histone H_2B, fruit fly	0.34
Histone H_1, trout	1.05
HMG-17, chicken	−0.85

*The scores (not shown) were calculated with FASTP. Significant z values are boldfaced.

TABLE II. Differences Between Histone H_4 Amino-Terminal Sequences*

	(1)	(2)	(3)	(4)
(1) Human	—	1	8	15
(2) Yeast	3	—	8	14
(3) *Tetrahymena*	27	27	—	19
(4) *Trypanosoma*	50	47	63	—

*The values are based on alignment from Figure 2, including 30 positions without gaps from amino acid 1 to 33 (human sequence numeration). Upper right: number of differences. Lower left: percent differences.

```
                                      ¬       ¬
      ** * *  *   **   *         ***   **********
(1)  SGRGKGGK-GLGKGGAKRH-RKVLRDNIQGITKPAIRRLARRGGV
(2)  SGRGKGGK-GLGKGGAKRH-RKILRDNIQGITKPAIRRLARRGGV
(3)  AG-GKGGK-GMGKVGAKRHSRKSNKASIEGITKPAIRRLARRGGV
(4)      AKGKKSGEAKGTQKRQ-KKILRENVRGITRGSIRRLARRGGV
```

Fig. 2. Alignment of histone H_4 amino-terminal sequences (1), human; (2), yeast; (3), *Tetrahymena thermophila;* (4), *Trypanosoma cruzi.* *, positions conserved in all sequences. ¬, possible COOH-ends of amino terminal domain.

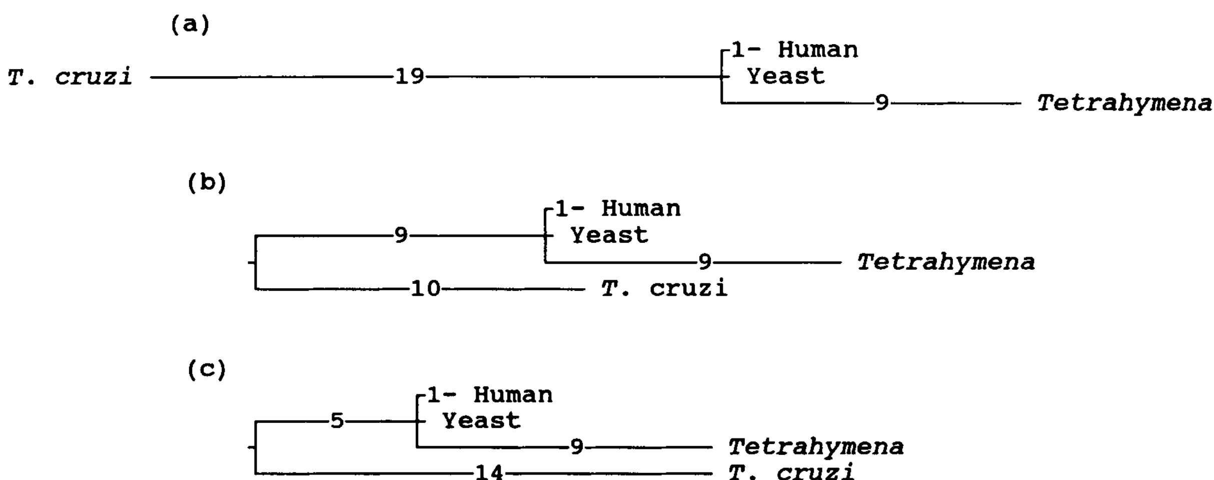

Fig. 3. A most parsimonious tree constructed from Figure 2 alignment including those positions mentioned in Table II legend. Number of inferred nucleotide substitutions are indicated in the branches. (a) Unrooted tree. (b) Rooted tree postulating an increase in the evolutionary rate in the ciliates branch. (c) Rooted tree postulating a slow-down in the evolutionary rate in fungi and metazoa branches.

with PROTPARS. One of these trees is displayed in Figure 3a. The interpretation we made of it is essentially valid, however, for all three trees. Supposing a constant evolutionary rate of substitutions, it is not possible to root Figure 3a tree in such a way as to have similar rates of change in all lineages. By accepting a nonconstant evolutionary rate but minimizing the number of rate changes in the tree, two different root positions are possible (Fig. 3b, 3c) in agreement with the distant phylogenetic position estimated by other authors for the flagellates (McLaughlin and Dayhoff, 1973; Woese, 1987).

DISCUSSION

Sequence comparisons clearly identify *T. cruzi* bands e (e_1, e_2, and e_3) as a histone H_4 (Table I). *T. cruzi* sequence is, however, the most divergent of all known histones H_4 (Table II). This is in agreement with the unusual pattern of *T. cruzi* observed in polyacrylamide gels (Toro and Galanti, 1990) and with the early divergence of the flagellates within the eukaryotic lineage (Woese et al., 1990). As can be seen in Table II, differences are higher between *Trypanosoma* and *Tetrahymena* than between *Trypanosoma* and human or yeast. If we accept that *Trypanosoma* diverged earlier than the other three lineages, whatever the order of divergence among these three lineages might be, a slow-down in the evolutionary rate of human and yeast branches or an acceleration in the *Tetrahymena* branch might have occurred. According to these alternatives, two different root positions are possible if we adjust the evolutionary rate of the *Trypanosoma* branch to one of the other branch rates. These solutions are displayed in Figure 3b,c. We must emphasize that both are extreme alternatives that do not exclude others less schematic. According to the Figure 3b solution, flagellate divergence occurred earlier than ciliates-metazoa divergence by a factor of 10 (ten nucleotide substitutions/one nucleotide substitution). According to the Figure 3c solution, flagellate divergence occurred earlier than ciliates-metazoa divergence by a factor of 1.56 (14 nucleotide substitutions/9 nucleotide substitutions). Accepting that ciliates-metazoa divergence cannot be more recent than 600 million years (metazoa radiation, Cloud and Glaessner, 1982) and flagellate divergence cannot be older than 1,400 million years (origin of eukaryotes, Schopf and Oehler, 1976), a value of 1.56 for the date quotient is much more compatible with that range than a value of 10. Thus, we conclude that there has been a slow-down in the evolutionary rate of histone H_4 amino-terminal domain in the lineage leading to metazoa and that this slow-down excluded the ciliates, possibly because they diverged earlier than metazoa. The slow-down could be related to the arousal of a novel function for the amino terminal domain.

It has been postulated that the central role of histone H_4 in the nucleosome makes it extremely intolerant to any mutational change (Kayne et al., 1988). However, both in nature

and in the laboratory, substitutions and deletions in histone H_4 do not affect its basic structural function in organizing the nucleosome. The finding that histone H_4 N-terminal deletions cause some unfolding of chromatin "in vivo" (Kayne et al., 1988), which is similar to that shown by chromatin trypsinized "in vitro" (Whitlock and Stein, 1978), has suggested that the N-terminal domain of histone H_4 participates in the compactation of chromatin. Trypanosomes present some characteristics in their *chromatin* which resemble this behavior. Thus, experimental observations show that *T. cruzi* chromatin differs from that of vertebrates in its stability, being both physically and enzymatically fragile (Rubio et al., 1980; Hecker and Gander, 1985). Consequently, *T. cruzi* chromatin is less stable at nucleosomal and higher order of organization than higher eukaryotes chromatin.

On the other hand electron microscopy analysis show that the ultrastructure and the condensation behavior of chromatin from *T. cruzi* are different from rat chromatin. Condensation into compact fibers (solenoid) was complete for rat chromatin at 100 mM salt concentration, while chromatin of *T. cruzi* showed less condensation (tangle formation) under identical experimental conditions (Hecker and Gander, 1985).

Considering the data showed in this work, it might be postulated that the extremely divergent N-terminal domain of histone H_4 from *T. cruzi* might contribute to explain this behavior.

We conclude that a) in *T. cruzi* a histone H_4 is present; b) this histone H_4 is resolved in acid-polyacrylamide gels in three variants; c) the amino terminal of *T. cruzi* histone H_4 is not blocked; d) these three histones H_4 present an extremely divergent amino acid sequence in the amino terminal domain; e) the following ten amino acids corresponding to the globular domain are conservative to all histones H_4 known; and finally f) most probably there are two evolutionary behaviors in the history of this domain, one with a high rate of amino acid substitutions up to ciliates, and later a slow-down in the evolutionary rate in the metazoa branch, related perhaps to the appearance of a novel function for this domain.

ACKNOWLEDGMENTS

This research was supported in part by Re-Entry Grant ID 890378 of the World Health Organization; Project DTI, University of Chile, number B3212-9013, Project FONDECYT 096-89 Chile, and SAREC (Swedish Agency for Research Cooperation). N. Jaramillo is a recipient of Research Training Grant TDR900179 from UNDP/World Bank/WHO. We thank Dr. C. Connelly for corrections.

NOTE ADDED IN PROOF

During the preparation of this article we take notice by personal communication of Dr. Hermann Hecker from the Swiss Tropical Institute, Basel, Switzerland, about similar data sequence obtained from *Trypanosoma brucei brucei* histone H_4.

REFERENCES

Alfageme CR, Zweidler A, Mahowald A, Cohen LH (1974): Histones of *Drosophila* embryos. J Biol Chem 249:3729–3736.

Astolfi Filho S, Martins De Sa C, Gander ES (1980): On the chromatin structure of *Trypanosoma cruzi.* Molec Biochem Parasitol 1:45–53.

Behe MJ (1990): Histone deletion mutants challenge the molecular clock hypothesis. TIBS 15:374–376.

Claiborne UC, Glover, Gorovsky MA (1979): Amino acid sequence of Tetrahymena histone H_4 differs from that of higher eukaryotes. Proc Natl Acad Sci USA 76:585–589.

Cloud P, Glaessner MF (1982): The Ediacarian period and system: Metazoa inherit the Earth. Science 217:783–792.

De Lange RJ, Fambrough DM, Smith EL, Bonner J (1969): Calf and Pea histone IV. III. Complete amino acid sequence of pea seedling histone IV: comparison with the homologous calf thymus histone. J Biol Chem 244:5669–5679.

Diamond LS (1968): Improved method for the monoaxeni cultivation of *Entamoeba histolytica Schaudinn* and *E. histolytica*-like amoeba with trypanosomatids. J Parasitol 54:715–719.

Durrin LK, Nan RK, Kayne PS, Grunstein M (1991): Yeast histone H_4 N-terminal sequence is required for promoter activation "in vivo." Cell 65:1023–1031.

Gorovsky MA (1986): Ciliate chromatin and histones. In Gall JG (ed): "The molecular Biology of Ciliated Protozoa." New York: Academic Press, pp 227–261.

Hecker H, Gander ES (1985): The Compaction Pattern of the Chromatin of *Trypanosomes.* Biol Cell 53:199–208.

Kayne PS, Kim UJ, Han M, Mullen JR, Yoshizaki F, Grunstein M (1988): Extremely Conserved Histone H_4 N Terminus Is Dispensable for Growth but Essential for Repressing the Silent Mating Loci in Yeast. Cell 55:27–39.

Lipman DJ, Pearson WR (1985): Rapid and sensitive protein similarity searches. Science 227:1435–1441.

Mc Laughlin PJ, Dayhoff MO (1973): Eukaryote Evolution: A view Based on Cytochrome *c* Sequence Data. J Mol Evol 2:99–116.

McGhee JD, Felsenfeld G (1980): Nucleosome structure. Annu Rev Biochem 49:1115–1156.

Megee PC, Morgan BA, Mittman BA, Smith MM (1990): Genetic analysis of histone H_4: Essential role of lysines subject to reversible acetylation. Science 247:841–845.

Panyim S, Chalkley R (1969): High resolution acrylamide

gel electrophoresis of histones. Arch Biochem Biophy 130: 337–346.

Rizzo PJ (1985): Histones in Protistan Evolution. BioSystem 18:249–262.

Rubio J, Rosado Y, Castañeda M (1980): Subunit structure of *Trypanosoma cruzi* chromatin. Can J Biochem 58:1247–1251.

Schopf JW, Oehler DZ (1976): How old are the Eukaryotes? Science 193:47–49.

Simon RH, Camerini-Otero RD, Felsenfeld G (1978): An octamer of histones H_3 and H_4 forms a compact complex with DNA of nucleosome size. Nuch Acids Res 5:4805–4818.

Stein GS, Mans RJ, Gabbay EJ, Stein JL, Davis J, Adawadkar PD (1975): Evidence for fidelity of chromatin reconstitution. Biochemistry 14:1859–1966.

Toro GC, Galanti N (1988): H_1 Histone and Histone variants in *Trypanosoma cruzi*. Exp Cell Res 174:16–24.

Toro GC, Galanti N (1990): *Trypanosoma cruzi* Histones: Further Characterization and Comparison with Higher Eukaryotes. Biochem Int 21:481–490.

Van Holde KE (1989): Secondary and tertiary structure of histones. In Rich A (ed): "Chromatin." New York: Springer-Verlag, pp 148–162.

Walsh MJ, Mc Dougall J, Wittman-Liebold B (1988): Extended N-terminal sequencing of proteins of archaebacterial ribosomes blotted from two-dimensional gels onto glass fiber and poly (vinylidene difluoride) membrane. Biochemistry 27:6867–6876.

Wells D, McBride C (1989): A comprehensive compilation and alignment of histones and histone genes. Nucl Acids Res 17:r311–r346.

Whitlock JP, Stein A (1978): Folding of DNA by histones which lack their NH_2-terminal regions. J Biol Chem 253: 3857–3861.

Woese CR (1987): Bacterial evolution. Microbiol Rev 51:221–271.

Woese CR, Kandler O, Wheelis ML (1990): Towards a natural system of organism: Proposal for the domains Archaea, Bacteria and Eucarya. Proc Natl Acad Sci USA 87:4576–4579.