Nucleosome Interaction Surface of Linker Histone H1c Is Distinct from That of H10

Eric M. George†, Tina Izard‡, Stephen D. Anderson§, and David T. Brown†‡

The fully organized structure of the eukaryotic nucleosome remains unsolved, in part due to limited information regarding the binding site of the H1 or linker histone. The central globular domain of H1 is believed to interact with the nucleosome core at or near the dyad and to bind at least two strands of DNA. We utilized site-directed mutagenesis and in vivo photobleaching to identify residues that contribute to the binding of the globular domain of the somatic H1 subtype H1c to the nucleosome. As was previously observed for the H10 subtype, the binding residues for H1c are clustered on the surface of one face of the domain. Despite considerable structural conservation between the globular domains of these two subtypes, the locations of the binding sites identified for H1c are distinct from those of H10. We suggest that the globular domains of these two linker histone subtypes will bind to the nucleosome with distinct orientations that may contribute to higher order chromatin structure heterogeneity or to differences in dynamic interactions with other DNA or chromatin-binding proteins.

The nucleosome is the fundamental repeating unit of eukaryotic chromatin (1, 2). Two molecules each of the four core histones, H2A, H2B, H3, and H4, form an octamer around which is wrapped 147 bp of DNA to form the nucleosome core particle (3, 4). In metazoans, one molecule of the H1 or linker class of histones is bound to nucleosomal core DNA and associates with the linker DNA (5–9). In higher organisms, the linker histones have a conserved tripartite structure consisting of a short, flexible N-terminal tail, a highly conserved central globular domain of ~80 residues, and a long (~100-residue) unstructured, highly basic C-terminal tail (10). The central domain is organized into a three-helix bundle containing a winged helix motif similar to that found in some transcription factors (11–13). This domain binds at or near the point at which DNA enters and exits the core particle and confers nucleosome resistance to an additional ~20 bp beyond the core particle (14–17). Via a mechanism that requires the C-terminal tail, linker histone binding promotes and/or stabilizes folding of nucleosomal arrays into higher order structures (5, 7, 18). It is now clear that chromatin in general, and specifically the H1 histones, are much more dynamic than previously thought and that regulation of the dynamic properties of the linker histones is a key aspect of many chromatin-mediated processes (19–21).

In mice, linker histones exist as a class of at least 11 nonallelic primary sequence variants or subtypes (22, 23). Seven of these are expressed in somatic tissues whereas the remaining subtypes are germ line-specific. Five of the somatic subtypes (H1a–H1e) are generally classified as replication-dependent and are expressed to varying degrees in most cell types. They display significant sequence heterogeneity in the N- and C-terminal tails but are very highly conserved (87–99% identity) within the globular core domain. The replication-independent or replacement subtype H10 is expressed to the highest level in terminally differentiated cells. The globular domain of this subtype displays considerable sequence divergence from those of the H1a–H1e subtypes (44–45% identity). The least characterized somatic subtype, H1x, is the most divergent in sequence and is expressed at low levels in most somatic tissues (24). There is a large amount of experimental evidence that individual subtypes differ from one another in their interactions with chromatin, supporting the concept of a functional significance to the evolutionarily conserved heterogeneity (24–26). However, a mechanistic explanation for how this is mediated is lacking.

X-ray crystallographic data of H1-containing nucleosomes has not been reported to date, and there is little direct structural information regarding the exact location of the linker histone on the nucleosome (16, 17, 27–29). This information is crucial to understanding how these proteins promote or stabilize chromatin condensation and how they might function to regulate gene expression at the nucleosomal level. Recently, we combined mutagenesis and photobleaching techniques to identify specific residues in the globular domain of the H11 subtype involved in binding of this domain to the nucleosome in vivo (30). Utilizing existing structural information on the nucleosome core particle (3) and the globular domain of the H11 ortholog H5 (11), we generated a map of the nucleosomal interaction surface of the H10 globular domain and were able to propose a model for binding of this subtype to the nucleosome (30).

The H10 subtype constitutes a minor fraction of the total linker histone composition of most cell types. Here, we sought to apply the same experimental approach to a representative of the main or replication-dependent class of H1 proteins that make up the bulk of the linker histone in somatic cells. The H1c subtype was chosen for this study partly because a high resolution NMR structure of the globular domain of the chicken...
ortholog is available (31). There are only seven differences, all conservative, between these globular domains of the mouse and chicken proteins. In addition, we had previously shown that, upon overexpression, the H1ο and H1c subtypes had very different effects on cell cycle progression and gene expression in cultured fibroblasts (32). Through domain-swapping experiments, we demonstrated that these differences were attributable to sequence or structural differences between the globular domain of these two proteins (33). Surprisingly, the results presented here indicate that, despite the strong conservation of overall structure of the globular domains of these two subtypes, there are more differences than similarities in the spatial orientation of the binding residues, indicating that they are likely to bind to the nucleosome with distinct geometries.

**EXPERIMENTAL PROCEDURES**

Constructs and Cell Lines—Plasmids MTH1οGFPneo and MTH1cGFPneo have been described previously (19). In these plasmids, the coding sequence for enhanced GFP2 is fused to the C terminus of the coding region for the histone, and expression is under control of the mouse metallothionein promoter.

Mutagenesis—Point mutations were introduced with the QuikChange mutagenesis kit (Stratagene) or by introduction of annealed oligonucleotides between restriction sites. Constructs were introduced into mouse BALB/c 3T3 cells, and multiple stable transfectants were isolated and analyzed as described (30). For fluorescent recovery after photobleaching (FRAP) assays, cultures were grown in the absence of inducer. Under these conditions, H1-GFP comprises less than 5% of the total H1 population (19).

FRAP Assays—For FRAP, cells were plated and observed in LabTekll chambered coverslips (Nalgene). FRAP was performed on a Zeiss 510 Meta LSM confocal microscope using the 488-nm line of an argon laser (nominal output 30 W). All experiments were done at 37 °C, and imaging was done with an Apochromat 63/1.4 DIC objective lens. Scanning was bidirectional at the highest possible rate using a ×4 zoom with a pinhole of 1 airy unit. Bleaching was achieved using three consecutive bleach scans of 49-ms duration at maximal laser power. The bleached region was a 2-μm spot that included euchromatin and heterochromatin but excluded nucleoli. Recovery was detected at 500–575 nm and recorded until the recovery signal reached a plateau by 90–100 s whereas that of H10-GFP occurred at 170–180 s. Each of the curves was fit to a single exponential function and heterochromatin but excluded nucleoli. Recovery was detected at 500–575 nm and recorded until the recovery signal reached a plateau by 90–100 s whereas that of H10-GFP occurred at 170–180 s. Each of the curves was fit to a single exponential function and yielded half-times for recovery, $t_{50}$, of 18.6 s for H1c and 35.1 s for H10, respectively. The results indicate that H1c binds to chromatin with significantly less affinity than does H1ο in 3T3 fibroblasts. We previously reported that H1c also binds significantly less tightly than any of the other variants examined (37).

Identification of H1c-binding Residues by Mutagenesis and FRAP.—The binding of linker histones to chromatin is primarily electrostatic in nature (38). We therefore created a comprehensive set of H1c-GFP mutant constructs in which each of the basic residues located on the surface of the globular domain of H1c was individually mutated to alanine (Fig. 2a). Each construct was stably transfected into mouse fibroblasts and analyzed by FRAP (Fig. 2, b–e and Table 1). The mutants fell into two distinct classes. Five mutants, K63A, K74A, K80A, K84A, and Table 1). The mutants fell into two distinct classes. Five mutants, K63A, K74A, K80A, K84A, and Table 1). The mutants fell into two distinct classes. Five mutants, K63A, K74A, K80A, K84A,

---

**TABLE 1**

Quantitative FRAP analysis of H1c-GFP constructs

| Construct          | $t_{50}$ (s) | $p$ value$^b$ |
|--------------------|--------------|---------------|
| Wild type H1c      | 18.6 ± 3.2   | $<10^{-4}$    |
| H1cK45A           | 19.7 ± 3.7   | 0.57          |
| H1cK51A           | 18.8 ± 3.8   | 0.92          |
| H1cK53A           | 19.1 ± 1.9   | 0.74          |
| H1cK63A           | 7.3 ± 1.7    | $<10^{-4}$    |
| H1cK74A           | 8.6 ± 1.8    | $<10^{-4}$    |
| H1cK80A           | 6.3 ± 1.5    | $<10^{-4}$    |
| H1cK84A           | 8.7 ± 1.4    | $<10^{-4}$    |
| H1cK89A           | 7.6 ± 0.6    | $<10^{-4}$    |
| H1cK96A           | 20.0 ± 3.4   | 0.63          |
| H1cK105A          | 18.9 ± 3.4   | 0.86          |
| H1cK108A          | 17.3 ± 2.5   | 0.35          |
| H1cK109A          | 17.2 ± 3.8   | 0.44          |
| H1cK63R           | 19.9 ± 2.5   | 0.24          |
| H1cK80R           | 18.9 ± 2.6   | 0.85          |
| H1cK84R           | 19.9 ± 3.8   | 0.49          |

$^a$ Values for $t_{50}$ were determined as described previously (30). Values are the mean ± S.D. from at least six independent measurements.

$^b$ Versus wild type H1c.
and K89A, reached a stable plateau by \( t_{50} \) values of 6–9 s (Fig. 2, b and c). The binding properties of these five mutants are significantly different from that of the wild type construct (Fig. 2g and Table 1). The remaining seven mutants, K45A, K51A, R53A, K96A, K105A, K108A, and K109A, displayed FRAP recovery curves that are not significantly different from the wild type construct (Fig. 2, d, e, and g, and Table 1).

We previously showed that the H1c-GFP protein behaves identically to the endogenous protein with respect to release by micrococcal nuclease digestion or by salt extraction (19). We also reported that these cell lines expressed H1c-GFP to \( \pm 5\% \) of the total H1 population and that expression of the fusion protein had no effect on cell-cycle progression or cell morphology. H1c-GFP co-localizes with DNA staining by Hoechst 33342 or propidium iodide, indicating that it is incorporated into both euchromatin and heterochromatin. We also demonstrated that H1-GFP constructs are fully functional in somatic cell nuclear transfer experiments (39). Proper positioning of linker histones on the nucleosome is most stringently assayed by detection of a 170-bp kinetic intermediate, termed the chromatosome, generated during micrococcal nuclease digestion of native chromatin (14, 15). To confirm that H1c-GFP is properly positioned, we separated mononucleosome particles containing H1c-GFP from those containing endogenous H1 by electrophoresis on nucleoprotein gels following micrococcal nuclease digestion (supplemental Fig. S1). DNA purified from H1c-GFP-containing particles and those containing endogenous H1 were each \( \sim 170–180 \) bp in length. We also demonstrated that three mutants that displayed fast FRAP recovery, K63A, K80A, and K84A, also confer protection of a discrete chromatosome-sized fragment of DNA (supplemental Fig. S1).

We also created and analyzed additional constructs in which the lysine residues at positions 63, 80, and 84 were mutated to arginine. FRAP analyses of these mutants each displayed a recovery curve that was not significantly different from that of wild type (Fig. 2f) as would be expected if these residues were interacting electrostatically with DNA. We conclude that lysine residues at positions 63, 74, 80, 84, and 89 are the major contributors to binding interactions between the globular domain of H1c and the nucleosome (Table 1 and Fig. 2g). Seven residues, lysines at positions 45, 51, 96, 105, 108, and 109, and arginine at position 53 are located on the surface of the domain but do not make major contributions to binding.

Mapping of Interaction Surface of H1c Globular Domain—To visualize the interaction surface of the H1c globular domain, we mapped the position of the major binding residues onto the NMR structure of the chicken ortholog (31) (Fig. 3a). The binding residues are not scattered around the surface of the domain but are clustered on one face of the domain. One major cluster consists of Lys-80, Lys-84, and Lys-89 in helix 3. The two remaining binding residues, Lys-63 and Lys-74, are oriented away from the major cluster and may constitute a single cluster or two independent binding sites. All but one of the residues identified as nonbinding are located on the opposite face of the domain, well separated from the binding residues. Note that the interaction surface of the H1c globular domain is quite different...
DISCUSSION

The linker histone family of proteins in mice and humans is a markedly heterogeneous collection of at least 11 nonallelic primary sequence subtypes or variants (23, 26). Results from gene knock-out studies indicate that there is some level of redundancy among the variants in their role as chromatin architectural proteins in that knock-out or knockdown of a single subtype could be compensated by up-regulation of others (40, 41). A considerable amount of experimental data indicates that individual variants differ in their expression patterns in different tissues and in their quantitative ability to condense chromatin (42 and references therein). These, and additional observations, have led to the view that a functional significance underlies the evolutionarily conserved linker histone heterogeneity of higher organisms (for recent reviews, see 24, 26). However, despite extensive and ongoing investigation, what is lacking is a mechanism for how this is mediated.

Here, we have used photobleaching and mutagenesis to map the nucleosome interaction surface of the globular domain of the replication-dependent H1c variant and to compare it with that determined previously for the replication-independent H10. The primary amino sequences for the globular domains of H1c and H10 and the results of this and our previous study (30) are aligned in Fig. 2a. Surprisingly, at only two positions, Lys-80 and Lys-84 in helix 3 of H1c, corresponding to Lys-69 and Lys-73 of H10, are the sequence and binding properties conserved. Two of the remaining three H1c-binding residues, Lys-74 and Lys-89, are not conserved in H10, having been replaced by neutral amino acids. Lys-63 of H1c was determined to be a major contributor to binding, but the conserved Lys-52 of H10 was previously shown not to be significantly involved in binding. Three additional binding residues in H10, His-25, Lys-47, and Arg-74, are not conserved in the H1c sequence. Conspicuously, mutation of residues Lys-96, Arg-53, Lys-105, or Lys-108 in H1c did not significantly affect binding. These residues are conserved in all of the major somatic linker histone subtypes of mouse and human (23, 26). Mutation of each of the corresponding residues in H10 significantly compromised binding, and three of these, Arg-52, Arg-94, and Lys-97, comprise a major DNA binding site as determined by molecular modeling (30). Thus, at the primary amino acid sequence level there are far more differences than similarities in the location of the binding residues for these two domains. These differences led to distinct nucleosomal surface interaction maps for the globular domains of these two linker histone subtypes (Fig. 3).

The stoichiometry and nucleosome binding geometry of the linker histones are expected to have a major influence on the structure of compacted chromatin (7–9). The structure of the secondary level of folding, the so-called 30-nm fiber, remains controversial (43, 44). In the one-start or interdigitated solenoid model (45), the fiber is coiled such that successive nucleosomes are adjacent to one another joined by bent linker DNA. This model is supported by electronmicrographic analysis of long nucleosomal arrays containing stoichiometric amounts of linker histone (46). In the two-start or zigzag model (47), consecutive nucleosomes are separated by relatively straight linker DNA such that each nucleosome is closest to its n + 1 neighbor. This model is supported by the crystal structure of a tetranucleosome compacted in the absence of H1 (48). It has been suggested that both structures may exist in vivo, with the more compact one-start structure representing H1-containing chromatin and the zigzag structure predominating in H1-depleted regions (8). This may be an oversimplification as cross-linking studies indicate that H1-containing nucleosomal arrays can adopt a zigzag conformation (49). Furthermore, a recent study provides evidence that both structures may be simultaneously present in a structurally heteromorphic chromatin fiber of 30 nm (50). Nevertheless, the binding geometry of the globular domain of H1 in the chromatosome is expected to contribute to the stability of the folded structure by constraining the entry/exit angle of the linker DNA. Monte Carlo simulations of chro-
Nucleosome Interaction Surface of Linker Histone H1c

molin fiber folding, utilizing the H1o-containing chromatome structure we previously proposed (30), were interpreted as evidence for an interdigitated one-start helix (51). However, the data presented here suggest that the globular domain of the H1c subtype will not be easily accommodated in this structure. We do not mean to suggest or imply that the linker histone subtype dictates the conformation of the 30-nm fiber. However, if the globular domains of individual H1 subtypes exhibit different binding geometries they may stabilize or be more easily accommodated in alternate structures.

Differences among H1 variants in the binding sites of the globular domain within the nucleosome might also influence how they interact within the dynamic network of chromatin-binding proteins (for recent reviews, see 52, 53). A number of proteins, including members of the various high mobility group families (54), the methylated DNA-binding protein MeCP2 (55), and certain transcription factors (56) interact directly with the nucleosome at or near the site of H1 binding.

Most studies addressing linker histone heterogeneity have focused on quantitative assessment of the binding affinity or chromatin condensing abilities of the various subtypes and conclude that the highly divergent C terminus is the primary domain responsible for these differences (24, 26, 37, 42). Here, we show, utilizing in vivo photobleaching of unperturbed chromatin, that there are qualitative differences in the binding surfaces of the conserved globular domains of two somatic H1 variants. We suggest that these two linker histone subtypes will bind to the nucleosome with distinct orientations that may contribute to higher order chromatin structure heterogeneity or to differences in dynamic interactions with other chromatin-binding proteins.

Acknowledgments—FRAP studies were performed at the Mississippi Functional Genomics Network Imaging Facility, University of Southern Mississippi, and we thank G. Shearer and B. Kang for assistance. We thank T. Smith and P. Vyas for technical assistance and D. Sittman and M. Hebert for critical reading of the manuscript and valuable discussions.

REFERENCES
1. Wolfe, A. (1998) Chromatin: Structure and Function, 3rd Ed., Academic Press, San Diego
2. van Holde, K. (1989) Chromatin, Springer-Verlag, New York
3. Luger, K., Maeder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Nature 389, 251–260
4. Woodcock, C. L. (2006) Curr. Opin. Struct. Biol. 16, 213–220
5. Thomas, F., Koller, T., and Klug, A. (1979) J. Cell Biol. 83, 403–427
6. Ramkrishnan, V. (1997) Crit. Rev. Eukaryot. Gene. Expr. 7, 215–230
7. Bednar, J., Horowitz, R. A., Grigoryev, S. A., Correll, S., Woodcock, C. L., and Schlick, T. (2001) Biochem. Cell Biol. 79, 289–304
8. Yu, A., Kamieniarz, K., and Schneider, R. (2008) Biochim. Biophys. Acta 178, 333–343
9. Brown, D. T. (2001) Genome Biol. 2, REVIEWS0006
10. Happel, N., and Doenecke, D. (2002) Gene 341, 1–12
11. Staynov, D. Z., and Crane-Robinson, C. (1998) EMBO J. 7, 3685–3691
12. Duggan, M. M., and Thomas, J. O. (2000) J. Mol. Biol. 304, 21–33
13. Zhou, Y. B., Gerchman, S. E., Ramkrishnan, V., Travers, A., and Muylermans, S. (1998) Nature 395, 402–405
14. Brown, D. T., Izard, T., and Misteli, T. (2006) Nature Struct. Mol. Biol. 13, 250–255
15. Cerf, C., Lippens, G., Ramkrishnan, V., Muylermans, S., Segers, A., Wyns, L., Wodak, S. J., and Hallenga, K. (1994) Biochemistry 33, 11079–11086
16. Brown, D. T., Alexander, B. T., and Sittman, D. B. (1996) Nucleic Acids Res. 24, 486–493
17. Brown, D. T., Gunjan, A., Alexander, B. T., and Sittman, D. B. (1997) Nucleic Acids Res. 25, 5003–5009
18. Thevlenaz, P., Ruttimann, U. E., and Unser, M. (1998) IEEE Trans. Image Process. 7, 27–41
19. Phair, R. D., Gorski, S. A., and Misteli, T. (2004) Methods Enzymol. 375, 393–414
20. Yellajoshula, D., and Brown, D. T. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 18568–18573
21. Thing, J. P., Sung, R., Ye, M., and Hendzel, M. J. (2005) J. Biol. Chem. 280, 27809–27814
22. Clark, D. L., and Kimura, T. (1990) J. Mol. Biol. 213, 260–269
23. Becker, M., Becker, A., Miyara, F., Han, Z., Kihara, M., Brown, D. T., Hager, G. L., Latham, K., Adashi, E. Y., and Misteli, T. (2005) Mol. Biol. Cell 16, 3887–3895
24. Fan, Y., Nikitina, L., Zhao, J., Fleury, T. J., Bhattacharyya, R., Bouhassira, E. E., Stein, A., Woodcock, C. L., and Skoulitchi, I. A. (2005) Cell 123, 1199–1212
25. Sancho, M., Diani, E., Beato, M., and Jordan, A. (2008) PLoS Genet. 4, e1000227
26. Clausell, J., Happel, N., Hale, T. K., Doenecke, D., and Beato, M. (2009) PLoS One 4, e8007243
27. Woodcock, C. L. (2005) Nat. Struct. Mol. Biol. 12, 639–646
28. Triemthel, D. J. (2007) Cell 128, 651–654
29. Finch, J. T., and Klug, A. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 1897–1901
30. Routh, A., Sandin, S., and Rhodes, D. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 8872–8877
31. Horowitz, R. A., Agard, D. A., Sedat, J. W., and Woodcock, C. L. (1994) J. Cell Biol. 125, 1–10
32. Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. (2005) Nature 436, 138–141
33. Dorago, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004) Science 306, 1571–1573
34. Grigoryev, S. A., Arya, G., Correll, S., Woodcock, C. L., and Schlick, T. (2009) Proc. Natl. Acad. Sci. U.S.A. 106, 13317–13322
35. Kepper, N., Foethke, D., Stehr, R., Wedemann, G., and Rippe, K. (2008)
Biophys. J. 95, 3692–3705
52. Zlatanova, J., Seebart, C., and Tomschik, M. (2008) Trends Biochem. Sci. 33, 247–253
53. McBryant, S. J., Lu, X., and Hansen, J. C. (March 23, 2010) Cell Res. 10.1038/cr2010.35
54. Catez, F., Yang, H., Tracey, K. J., Reeves, R., Misteli, T., and Bustin, M. (2004) Mol. Cell. Biol. 24, 4321–4328
55. Nikitina, T., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, I. C., Grigoryev, S. A., and Woodcock, C. L. (2007) J. Biol. Chem. 282, 28237–28245
56. Cirillo, L. A., McPherson, C. E., Bossard, P., Stevens, K., Cherian, S., Shim, E. Y., Clark, K. L., Burley, S. K., and Zaret, K. S. (1998) EMBO J. 17, 244–254