

Chromatin structure: Linking structure to function with histone H1

Jonathan Widom

A recent study has determined the position and orientation in the nucleosome of the H1 variant 'linker histone' H5; the results focus attention on the unknown function of this highly abundant nuclear protein, and highlight the question of whether H1 is primarily an architectural or a gene-regulatory protein.

Addresses: Department of Biochemistry, Molecular Biology and Cell Biology, and Department of Chemistry, Northwestern University, Evanston, Illinois 60208-3500, USA.
E-mail: j-widom@nwu.edu

Current Biology 1998, 8:R788–R791
<http://biomednet.com/elecref/09609822008R0788>

© Current Biology Ltd ISSN 0960-9822

Histone H1 is one of the most abundant proteins in the nucleus. Remarkably, despite years of study, the most basic questions concerning H1, such as its location in chromatin and its function, remain controversial or unanswered. Histone H1 has long been known as the 'linker histone', on the basis of classic studies that showed H1 to be somehow located on, or to contact, the linker DNA that goes between the nucleosomes that are the basic structural units of chromatin. A new study has revealed the position and orientation of the conserved globular domain of an H1 variant with respect to the nucleosome. The results contradict other recent studies, and draw attention to the unknown function of the protein.

H1 has long been considered a stoichiometric structural component of chromatin, present in one molecule per nucleosome [1–3]. Digestion of chromatin by a non-specific nuclease yields oligonucleosomes and mononucleosomes, which, upon further digestion of their linker DNA, are cleaved down to resistant subnucleosomal particles called 'chromatosomes'. These contain approximately 168 base pairs of DNA, the core histone octamer and one molecule of H1, and are readily isolated as intact particles. With more extensive digestion, the chromatosomal DNA is reduced to approximately 147 base pairs in length, with concomitant loss of histone H1, yielding the nucleosome core particle. These results implied that chromatin should be considered as a chain of chromatosomes joined together by the linker DNA, with the chromatosome being the fundamental unit of chromatin structure.

Although the structure of the nucleosome core particle has recently been determined at high resolution by X-ray crystallography [4], structures of intact chromatosomes or higher architectural levels of chromatin are not presently available. H1 can be released from chromatin by increasing

the salt concentration, leaving nucleosomes otherwise intact. Electron micrographs reveal that, when H1 is present, linker DNA enters and exits each nucleosome from a single region, whereas when H1 is removed, the sites of DNA entry and exit become plainly distinct. H1 can bind cooperatively to pairs of DNA segments, and it has a preference for binding also to other DNA structures containing closely juxtaposed DNA segments. These and other observations led to the view that H1 is located over the nucleosomal dyad-symmetry axis, binding simultaneously to a pair of DNA segments entering and leaving the nucleosome [1–3].

Histone H1 shows a greater degree of evolutionary variability than the notoriously conservative core histones. Most organisms express several different H1 variants in different cell types and developmental states. H5 is the predominant variant in mature chicken erythrocytes, a common source of chromatin for studies *in vitro*. H1 proteins have an unusual domain organization: a short, extended amino-terminal domain and a long, extended carboxy-terminal 'tail' domain, both of which are highly basic, flank a central, approximately 80 residue folded 'globular' domain. The structures of the globular domains of H1 and H5 — 'GH1' and 'GH5' — have been determined [2]. They are members of the 'winged helix' class of DNA-binding domains, although in contrast to other members of that family GH1 and GH5 contain a distinct, additional cluster of positively-charged amino acids that form a second DNA-binding surface, on a side of the protein opposite the primary DNA-binding site [5]. The globular domain on its own can provide increased protection of approximately 168 base pairs of DNA in a chromatosome-like particle, suggesting that it is this domain of H1 that is located over the nucleosomal dyad, at the DNA entry/exit region on the surface of the chromatosome.

It came as a great surprise when a pair of papers appeared that suggested a very different location for globular domains of H1 and H5 in the nucleosome [6,7]. These papers suggested that the globular domain is located asymmetrically in the nucleosome, packing between the final superhelical turn of the DNA and the histone octamer, near where the DNA leaves the core particle. This model leaves the second DNA-binding surface of GH1 exposed and thus free to interact with other DNA — perhaps on an adjacent nucleosome in the folded chromatin fiber. In accord with this view, these investigators found that the second DNA-binding site on GH1 was not required for the binding of H1 to its site in a nucleosome [8] — in apparent contradiction to a key conclusion of the study that showed the existence of this second site [5].

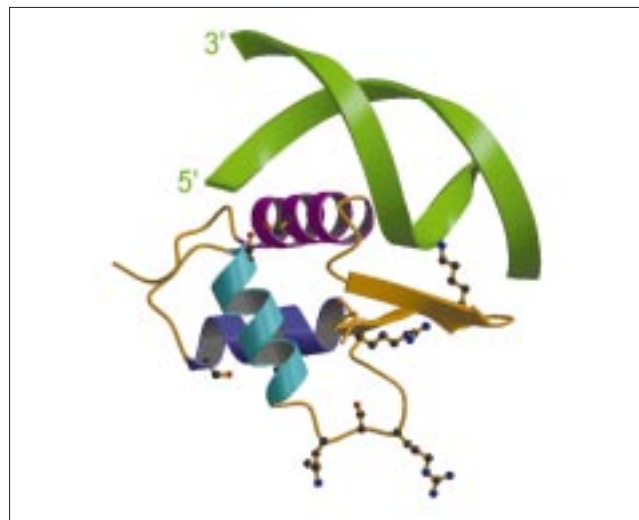
Now there has been another surprising twist to this tale. Zhou *et al.* [9] have recently reported a quite different location for GH5, and defined its orientation. The authors probed the location of GH5 reconstituted into natural chromatosomes depleted of H1 and H5. By site-directed mutagenesis, single cysteine residues were introduced at locations distributed over the surface of GH5. A photoactivatable crosslinking reagent was covalently attached to these cysteine residues, and the derivatized GH5 reconstituted into the chromatosomes. Only a subset of the derivatized GH5 domains retained the ability to bind to the chromatosomes. Exposure to UV light activated the crosslinking moiety, yielding for some of the mutants a modest but adequate fraction of reconstituted chromatosomes that contained DNA–protein crosslinks. The DNA backbone adjacent to the crosslink was hydrolyzed, and the site of the cleavage determined by running the end-labeled DNA strand on a sequencing gel.

From the results of this crosslinking experiment, Zhou *et al.* [9] infer that GH5 bridges between the chromatosomal dyad and one DNA terminus. Helix III of GH5 — which in classical winged-helix domains is the DNA-recognition helix — binds in the major groove at one end of the chromatosomal DNA, with the protein's 'wing' residues contacting the DNA backbone over the adjacent minor groove (Figure 1). The second DNA-binding surface, unique to the GH1 subfamily of winged-helix domains, contacts the central superhelical turn of the chromatosomal DNA, adjacent to the dyad axis (Figure 2). In this model, the carboxy-terminal domain of intact H1/H5 would extend out in a direction that allows for interactions with both the entering and exiting linker DNA segments.

This new model [9] shares one important feature with the earlier studies [6–8] with which it largely conflicts: both suggest that GH5 is located asymmetrically — although the nature of this asymmetry is different. The asymmetric location, together with the inherent asymmetry of the protein itself, emphasizes the possibility of a directionality, or polarity, to the folding of the chromatin fiber. This would fit with reported experimental evidence that is consistent with an overall polarity for stretches of chromatin. Both models also leave open the question of how nucleosomes bind only one molecule of H1, when there appear to be two, non-overlapping H1-binding sites. One possibility is that the GH5 actually sits slightly further over the dyad axis, occluding the second site.

What is the explanation for the discrepancy between these studies? Zhou *et al.* [9] suggest that the problem lies in nucleosome positioning, which is important in these experiments because it defines the reference point from which the location of the H1 is measured. The two earlier studies [6,7] both used reconstituted nucleosomes, rather than natural chromatosomes (all three studies subsequently

Figure 1

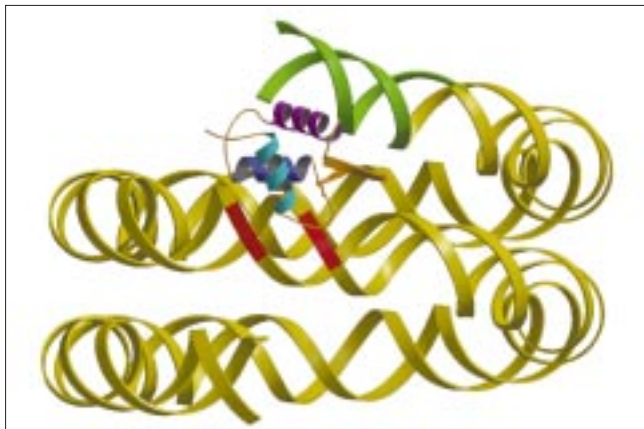


Interaction of GH5 DNA-binding site 1 with DNA, modeled by homology to the co-crystal structure of the related transcription factor HNF3. Binding site 1 comprises residues from helix III and an additional lysine sidechain (illustrated) from the adjacent 'wing' domain. Positively-charged sidechains contributing to DNA-binding site 2, on the lower surface of the protein, are shown. Serine sidechains that were mutated to cysteines and used for attachment of the crosslinker are also shown. See [5,9] for additional details. (Figure courtesy S. Muyldermans and T. Hamelrijck.)

reconstituted the H1). These were prepared with a 'nucleosome-positioning' DNA sequence, with the intention of creating a population of homogeneously positioned nucleosomes. It is now recognized that DNA sequence-directed nucleosome positioning is statistical, not absolute, with the probability of occupancy of a single preferred site linked to the free energy of histone–DNA interactions. The nucleosome-positioning sequence used in the earlier studies has a quite modest free energy compared to random sequence DNA, and thus is expected to allow a substantial population of alternative positions [10]. Zhou *et al.* [9] thus argue that the set of crosslinks or cleavages mapped in the earlier studies cannot be unambiguously assigned to a particular site within a nucleosome. In contrast, their new study used natural chromatosomes, the positioning of which is defined by the micrococcal nuclease that was used to liberate the chromatosomes in the first place.

The question of H1's function is even more controversial and uncertain. A conventional view is that H1 is primarily a structural protein. In addition to its role in organizing the chromatosomal DNA, H1 also contributes importantly to folding of the nucleosome filament into the next higher level of structure, the 30 nanometre chromatin fiber [1–3]. In the absence of H1, chromatin folds into irregular clumps, rather than into the more highly-ordered 30 nanometre fibers characteristic of native chromatin, and

Figure 2



Location of GH5 on a model of the nucleosome core-particle DNA superhelix, as deduced from the new study by Zhou *et al.* [9]. The DNA backbone at the nucleosomal dyad-symmetry axis is highlighted. The orientation of GH5 is the same as illustrated in Figure 1. (Figure courtesy S. Muyldermans and T. Hamelrijck.)

the folding that can occur requires a higher than usual salt concentration. These observations imply that H1 contributes to the free energy of chromatin folding, and that it helps select a single folded state from among the set of compact states that may be reached in its absence. Other structural and biophysical studies suggest that the 30 nanometre fiber is stabilized in part by contacts between H1 molecules on neighboring nucleosomes. Specific phosphorylation of H1 correlates with the further folding of 30 nanometre chromatin fibers during mitosis.

The past several years have produced a series of surprising new results that force one to rethink the function of H1. First, it is no longer certain that H1 is an essential protein (see [3])! The complete genomic DNA sequence of the yeast *Saccharomyces cerevisiae* has been determined and found to include only one coding region for a protein with significant homology to the conserved globular domain of the H1 proteins of multicellular eukaryotes. This protein has now been eliminated by knocking out the gene, and the resulting mutant cells remain viable, although they show detectable alterations in gene regulation. This is likely not to be a peculiarity of yeast. Physical studies of H1-depleted chromatin *in vitro* show that it is capable of at least some degree of higher-order folding (although the extent to which such folding is native-like remains in doubt).

No candidate H1-like proteins have been identified in the embryogenic stages of *Drosophila* development, and the one identified candidate H1-like protein present during *Xenopus* early development can be eliminated, with little evident consequence for nuclear assembly or, indeed, for the development of the organism. Finally, many cases have been discovered in which the bulk nucleosome

repeat length of an organism or cell type — a measure of the amount of DNA per nucleosome, averaged over the entire genome — is less than the chromosomal DNA content, implying that the chromosome is not, after all, a fundamental unit of chromatin structure.

The available structural data also suggest an alternative interpretation of the role of H1. While the structural similarity of GH1 to the winged-helix transcription factors could of course simply be a distinct use of a successful protein structural motif, it could also indicate that H1 has a direct role in gene regulation. In accord with this view, there is competition between H1 and the transcription factor HNF3 for binding to a nucleosomal transcription factor target site in the albumin gene enhancer [11].

H1 could act indirectly as a regulatory protein through its effects on chromatin folding. Chromatin folding hinders transcription factor access and elongation by RNA and DNA polymerases so, by contributing to chromatin folding stability, H1 could act as a general repressor of transcription. H1 variants differ in their affinity for DNA and chromatin, so they could vary in their ability to stabilize chromatin folding and repress gene activity. Different H1 variants are, in some cases, segregated in blocks in the genome, providing a molecular basis for differential folding stability and repression of entire chromosome domains. An alternative mechanism is suggested by the preference of H1 for binding to methylated DNA, which points to a model in which H1 binding and histone deacetylation — which is linked to methylation — combine to yield effective gene repression through stabilization of folded chromatin domains.

H1 could also act as a regulatory protein through its effects on nucleosome positioning or dynamics. DNA in a nucleosome may need to uncoil from an end to allow polymerase elongation or transcription factor access, and the equilibrium constants for such ‘site exposure’ processes decrease strongly as the amount of DNA that must be uncoiled is increased [3]. By biasing nucleosome positioning, perhaps through its known modest DNA sequence preferences, H1 could modulate the time-averaged accessibility of nucleosomal DNA sites. Alternatively, by binding to the linker DNA segments entering and leaving the nucleosome, H1 could suppress such spontaneous ‘site exposure’ uncoiling events altogether, or alter the fraction of time that the DNA is uncoiled and accessible.

Importantly, however, H1 is not a ‘glue’ that seals two turns of DNA in the nucleosome, as is often imagined. At least *in vitro*, H1 molecules are found to be in free exchange. The effects of H1 on gene regulation at the level of individual nucleosomes should therefore be considered as a problem of coupled chemical equilibria. A related question is whether H1 — when it is present in a

cell type or organism at all — remains, in the time average, bound to the chromatin of active genes. Several studies suggest that H1 does remain present in active chromatin, although in moderately decreased amounts; it may also be bound in an altered fashion.

In summary, this new study represents an important step forward in understanding H1 and its variants: the location and orientation in chromatin of this highly abundant nuclear protein are now at least roughly established. But many key questions remain. Is H1 primarily an architectural protein, a regulatory protein, or both? And, how does it function? What are the roles of the other domains, and the variants and posttranslational modifications? The pace of research in this area seems certain to pick up dramatically!

Acknowledgements

I thank Venki Ramakrishnan for helpful discussions and Serge Muyldermans and Thomas Hamelrijck for generously producing the figures.

References

1. Widom J: **Toward a unified model of chromatin folding.** *Annu Rev Biophys Chem* 1989, **18**:365-395.
2. Ramakrishnan V: **Histone structure and the organization of the nucleosome.** *Annu Rev Biophys Biomol Struct* 1997, **26**:83-112.
3. Widom J: **Structure, dynamics, and function of chromatin *in vitro*.** *Annu Rev Biophys Biomol Struct* 1998, **27**:285-327.
4. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ: **Structure of the nucleosome core particle at 2.8Å resolution.** *Nature* 1997, **389**:251-260.
5. Goytisolo FA, Gerchman S, 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**:3421-3429.
6. Pruss D, Bartholomew B, Persinger J, Hayes J, Arents G, Moudrianakis EN, Wolffe AP: **An asymmetric model for the nucleosome: a binding site for linker histones inside the DNA gyres.** *Science* 1996, **274**:614-617.
7. Hayes JJ: **Site-directed cleavage of DNA by a linker histone-Fe(II) EDTA conjugate: localization of a globular domain binding site within a nucleosome.** *Biochemistry* 1996, **35**:11930-11937.
8. Hayes JJ, Kaplan R, Ura K, Pruss D, Wolffe A: **A putative DNA binding surface in the globular domain of a linker histone is not essential for specific binding to the nucleosome.** *J Biol Chem* 1996, **271**:25817-25822.
9. Zhou Y-B, 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**:402-405.
10. Lowary PT, Widom J: **New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning.** *J Mol Biol* 1998, **276**:19-42.
11. 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**:244-254.