Biochimica et Biophysica Acta xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Linker histones: History and current perspectives☆

C. Crane-Robinson

Biophysics Laboratories, School of Biology, University of Portsmouth, PO1 2DT, UK

ARTICLE INFO

Article history:
Received 27 July 2015
Received in revised form 7 October 2015
Accepted 8 October 2015
Available online xxxx

Keywords:
Chromatin
Nucleosome
Linker histone

ABSTRACT

Although the overall structure of the fifth histone (linker histone, H1) is understood, its location on the nucleosome is only partially defined. Whilst it is clear that H1 helps condense the chromatin fibre, precisely how this is achieved remains to be determined. H1 is not a general gene repressor in that although it must be displaced from transcription start sites for activity to occur, there is only partial loss along the body of genes. How the deposition and removal of H1 occurs in particular need of further study. Linker histones are highly abundant nuclear proteins about which we know too little. This article is part of a Special Issue entitled: Histone H1, edited by Dr. Albert Jordan.

The functional role of the fifth histone (H1, for short in what follows) has proven the most difficult to understand. Whilst the structure of the octameric 147 bp core particle [1] rapidly led to an enhanced appreciation of core histone function, in particular regarding the roles of their different structural domains, no equivalent structure has been obtained for an H1-containing nucleosome. The main reason is that reconstitution of the very basic H1 molecule with an extended core particle of high net negative charge is subject to artifactual non-specific interactions and, furthermore, there is no obvious functional assay for correct binding. Additionally, chromatin-bound H1 is known to rapidly exchange [2,3] (in vivo, at least) so the conditions for generating a stable in vitro H1-containing chromatosome are not yet defined. Even more uncertain are methods for generating H1-containing higher order chromatin structures, such as the so-called 30 nm supercoil (30 nm fibre). So the primary issues in need of resolution are to understand precisely how H1 is bound to the nucleosome and how this helps the formation and stabilisation of the 30 nm fibre. Only then will we be able to fully comprehend how chromatin structures are controlled to facilitate the various transactions to which the DNA is subject.

1. Structure of linker histones

It was early appreciated that H1 plays a role in condensing chromatin and studies of its free solution structure gave hints as to how it achieves this. Limited proteolysis showed that most H1 species consist of 3 domains [4]. The central globular domain (GH1, ~80 residues, diameter ~ 3 nm) is the only folded element in the ~200 amino acid protein in free solution. It adopts the winged helix fold [5]. The short N-terminal domain (NTD, ~35 residues) is very basic but only in its second half: in contrast, the N-terminal part is somewhat apolar, even acidic. The NTD is disordered in free solution and does not play an obvious role in chromatin condensation [6] but is subject to post-translational modifications (PTMs), so may play a regulatory/signalling role in H1 function. Linker histones specific for condensed chromatin typically have a shorter NTD but in oocyte specific H1s it is often longer. The C-terminal domain (CTD) is also disordered in free solution but is much longer: ~100 residues in the canonical mammalian species but >200 residues in some oocyte-specific linker histones. Since the CTD is very lysine rich it is assumed to play the main role in condensing the nucleosomal fibre, a function modulated by phosphorylation at multiple sites, i.e. the addition of negative charges [6–9]. An important but difficult question to which a definitive answer is needed is the extent to which the NTD and CTD adopt defined secondary structures when bound into the chromatin fibre [10,11].

2. H1s and their variants

The canonical mammalian replication-dependent (RD) H1 subtype genes (H1.1 to H1.5 and the testis-specific H1t in humans) are encoded within multiple histone clusters together with those of the core histones, synthesised only during S-phase [12]. So despite the substantial structural and functional distinctions between the core and linker histones, the H1s are an intrinsic component of the chromatin of higher eukaryotes — quite different from the situation in yeast (*Saccharomyces cerevisiae*) where there are only two copies of each of the core histone genes [13,14] and the presumed linker histones look very different from mammalian H1s.

As with the core histones H2A and H3, in mammalian cells there are also a number of replacement (variant) H1s, synthesised throughout the cell cycle from genes that sometimes contain introns and generate polyA+ mRNAs. Variant H1s play very specific roles: for example, H5

☆ This article is part of a Special Issue entitled: Histone H1, edited by Dr. Albert Jordan.

http://dx.doi.org/10.1016/j.bbagrm.2015.10.008
1874-9399/

from the nucleated erythrocytes of birds, and its mammalian homologue H1.0, are both a feature of highly compacted chromatin, probably playing a condensing role by including multiple arginines (largely absent from the canonical mammalian H1 subtypes). Oocyte-specific variants, for example *Xenopus* B4 and mammalian H1oo, have CTDs with a substantial admixture of acidic residues (E and D) — amino acids totally absent from the CTDs of canonical mammalian H1s.

3. H1 PTMs

Extensive phosphorylation of serine and threonine in the NTD and especially the CTD of canonical H1s was documented early and linked to chromatin condensation at mitosis [15]. ADP-ribosylation of the NTD was also early recognised [16] and might be thought of a component of the histone code but H1s were initially not thought to be acetylated or methylated at lysines. Mass spectrometry has shown this not to be the case [17]: for example, K26 of human H1.4 in the NTD can be methylated and the adjacent residue S27 phosphorylated. The same H1 subtype can be acetylated at K34, a residue on the border of the NTD and globular domains. Other modifications have been noted within GH1, notable citrullination of R54 in ESCs [18], a conserved residue within DNA binding Site II (see below).

4. Histone H1 in the 30 nm fibre

The first and much vexed question was the location of GH1 on the nucleosome. On the assumption that each nucleosome carries a single H1 molecule [19] and the observation that GH1 alone is able to protect an extra 10 bp at each end of the core particle, giving the 167 bp chromatosome, it was proposed that GH1 binds symmetrically to the DNA on the dyad axis, making contact with both exiting duplexes and the central gyre: a 3-contact model [4]. Although a symmetrical location was supported by DNaseI footprinting experiments on native chromatin [20], it provoked the question as to what determines the orientation of the linker histone on the nucleosome, so the model was widely challenged both experimentally and by molecular dynamics calculations of preferred orientations [21–23]. General agreement emerged that GH1/5 carries just two DNA binding sites [24,25]: Site I on Helix 3 and Site II on the Loop between Helix 1 and Helix 2. This led to proposed interactions being with the central DNA gyre and just one of the exiting duplexes, i.e. a slightly off-axis binding site — as found in a DNA-GH5 crosslinking experiment [26]. Such locations did not however offer any explanation for symmetrical MNase protection of 2 × 10 bp beyond the core particle.

To extend knowledge of how H1 (in particular GH1/5) is bound to the nucleosome, i.e. define where it sits in the compacted state, neutron scattering contrast measurements were made of the 30 nm fibre containing native and deuterated H1 [27]. Plots of the square of the radius of gyration against the inverse contrast made it clear that the increased contrast in the deuterated sample must come from internally located H1. Whatever the precise model of H1 binding, this indicated that the DNA exit points of the nucleosomes must be on the inside of the fibre. Although this result did not lead to a description of nucleosome connectivity within the fibre (continuous solenoid vis-à-vis cross-linker models), it excluded all models with outward facing nucleosomes.

Appreciation of a possible third DNA binding site on GH1/5 [28,29, 22,30] led to continuing interest in a symmetrical location and this was supported by hydroxyl radical footprinting studies (giving higher resolution than DNaseI [20]) that showed a clear shadow from GH1 protection right on the dyad [31]. As the latter studies were conducted on reconstituted short oligo-nucleosomes, it was also possible to investigate the linker DNA. Strikingly, this showed protection by H1 of ~50 bp (170 Å) of linker with an alternating period of 10 bp, implying extended shielding from one side — presumed to result from the CTD as deletion of the 35-residue NTD did not affect the protection. The success of these footprinting experiments [31] may well have been due to the use of a chaperone (NAP1) for loading the H1s onto the chromatin template, an approach full of perspective.

Further evidence for the location of GH1 came from a cryo-EM study [32], at ~11 Å resolution, of crosslinked 12-mer arrays of H1.4-containing nucleosomes, which appear as three tetranucleosomal units stacked to form a zig-zag 2-start helix, each unit having fully extended DNA linkers. Fitting the crystal structure of the 167 bp repeat tetranucleosome (that lacks H1) [33] to the averaged cryo-images allowed an assessment of the location of H1. Well-defined density was observed between the exiting DNA duplexes: fitting this to the crystal structure of free GH5 [5] led to a 3-contact model for the globular domain in which Site III, (comprising the two β-strands of Loop 3), is in contact with the central DNA gyre of the nucleosome and Sites I and Site II make contact with the two exiting DNA duplexes. This location put GH1 somewhat asymmetrically placed, the Site III contact being about one quarter of a turn (~8 Å) from the dyad axis.

Even more recently, however, a crystal structure of a reconstituted GH5-containing 167 bp chromatosome at 3.5 Å resolution was published [34] showing the H5 globular domain located essentially symmetrically on the dyad axis, with 3 contact regions: Loop 3 (i.e. Site III, e.g. S90 and V87) is in contact with the central DNA gyre (major groove); Loop 1 (e.g. R42) is in contact with the minor groove of one exiting DNA duplex and Helix 3 (e.g. Q67) is edge on into the minor groove of the other exiting duplex, in addition to several K/R-phosphate contacts. This is essentially the symmetrical model originally proposed in our 1980 article [4] but this does not imply that all H1 globular domains are necessarily placed so symmetrically. The authors of the GH5 chromatosome structure themselves previously used several other physical techniques, i.e. not crystallography, to study the nucleosomal binding of the globular domain of *Drosophila* H1 [35] and concluded that binding is asymmetric: contact is made with the central DNA gyre and with one exiting duplex. This conclusion was reinforced in their latest publication [34] in which they compared the sedimentation of a 12-nucleosome array carrying GH5 or *Drosophila* GH1: the GH5 array was seen to be substantially more compact, implying that the binding mode of linker histones is variable and can influence the higher order folding of chromatin. This is unquestionably an attractive hypothesis: the existence of multiple subtypes and variants of the linker histones are clearly related to disparate functions that must surely imply variations of structure in the resulting chromatin.

A primary aim of structural studies of H1 in the context of the fibre has always been to understand how it compacts the nucleosomal beads-on-a-string structure. Since H1-containing nucleosomal arrays adopt a rigid zig-zag structure that compacts to form the 30 nm fibre [36–39], the first assumption was that GH1 directs the exit angles of the DNA from the nucleosome and the basic CTD covers the linker DNA, the resulting charge reduction permitting folding and compaction. Such a general principle of H1 involvement in folding could hold for all models of nucleosome connectivity within the fibre. However, such a mechanism suggests that without H1 the 30 nm fibre might not form at all — but this is not the case: the presence of divalent cations is sufficient to compact an H1-depleted regular nucleosomal array [40]. A possible explanation for this is seen in AFM observations that the DNA duplexes exiting from nucleosomes lacking H1 cross at right angles, provided 4–10 mM Mg++ is present [41]. Thus Mg++ could help order the DNA architecture within the fibre but how such divalent cations might be sufficient to counteract repulsions between linker DNAs on the inside of the fibre remains an open question.

The precise geometry of the nucleosomes and the trajectory of the linker DNA in the folded 30 nm chromatin fibre has been the subject of much study and dispute. The 167 bp repeat tetranucleosome, lacking H1, [33], as well as the recent cryo-EM structure [32] showed a stacked arrangement with fully extended DNA linkers. This has lent support to models in which the linker DNA between adjacent nucleosomes crisscrosses the fibre and several arrangements are physically possible [42]. In vitro studies of longer fibres typically use arrays of a core

Please cite this article as: C. Crane-Robinson, Linker histones: History and current perspectives, Biochim. Biophys. Acta (2015), http://dx.doi.org/10.1016/j.bbagrm.2015.10.008