FEBS 21261 FEBS Letters 441 (1998) 161–164

Minireview

Linker histones: paradigm lost but questions remain

Sergey Belikov[1,a,b], Vadim Karpov[b]

[a]Department of Cell and Molecular Biology, Medical Nobel Institute, Karolinska Institute, S-171 77 Stockholm, Sweden
[b]W.A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 117984, Russia

Received 2 November 1998

Abstract Linker histones (LH) represent a diverse family of proteins that bind to nucleosomes and bring them together to form a 30-nm chromatin fiber. Although the structure of the globular domain of linker histones H1 and H5 has been solved, the details of its interaction with the nucleosome are not understood in full. Recent data on the location of LH in nucleosome are discussed here.

Key words: Chromatin; Linker histone; Nucleosome

1. Linker histones

The nucleosome is a fundamental unit of chromatin; it contains the nucleosome core and linker DNA of variable length to which linker histones (LHs) bind [1]. The nucleosome core (146 bp of DNA wrapped around a histone octamer) can be obtained by extensive digestion of chromatin with micrococcal nuclease. The association of a single molecule of LH with the nucleosome results in the protection of an additional 20–22 bp of DNA after digestion with micrococcal nuclease. In vitro removal of LH or of its globular domain (see below) leads to loss of this protection, whereas the reassociation restores it [2].

Unlike core histones, which show a high level of sequence conservation, LHs diverge significantly both in sequence and structure. Metazoan LH has a well defined three-domain structure: the central globular domain (GLH) and N- and C-terminal domains [3]. GLH consists of a three-helix bundle (helices I–III) with a β-hairpin at the C-terminus [4,5] (see Fig. 1). Based on a striking similarity to the helix-turn-helix proteins [6], it was proposed that helix III of GLH would bind in the major groove of DNA, whereas the β-hairpin would interact close to the adjacent minor groove thus comprising the LH-DNA primary binding site as shown in Fig. 1 [4,5]. The secondary binding site is more diffuse and was identified as a cluster of highly conserved basic residues located on the opposite side of the molecule [4]. Both binding sites are essential for correct LH binding and chromatosome formation ([7], but see also [8]).

N- and C-LH tails comprise roughly half of the molecule. The C-terminal tail is highly enriched in basic residues [9], it can adopt a segmental α-helical conformation [10] and is known to be involved in chromatin condensation. Indeed, histone H5 (but not GH5 alone) has been shown to bind to the DNA entering and exiting the nucleosome, forming a stem-like structure ~30 bp in length [11]. Almost nothing is known about the structure and role of the N-terminal LH tail, which contains regions rich in basic residues as well as regions rich in prolines and alanines.

Interestingly, in *Tetrahymena* histone H1 lacks the globular domain entirely and is very similar to the 'C-terminal domain' of metazoan LH; it is solely responsible for chromatin condensation [12]. Even so, the elimination of this 'incomplete' LH has a minor effect on the morphology and transcription of *Tetrahymena* knockout strains [13]. Another lower eukaryote, *Saccharomyces*, has no linker histones and, thus, a very short linker DNA [1]. However, a LH-like histone was recently found in yeast [14] that behaves as a bona fide LH in vitro [15]; its functions and involvement in the organization of yeast chromatin remain obscure.

2. Paradigm lost

It has been believed for a long time that LH binds in a quasi-symmetrical manner simultaneously to the nucleosomal pseudo-dyad and to the linker DNA at the site where DNA helices entering and exiting the nucleosome cross [1]. This symmetrical, on-axis, placement would also imply symmetrical protection of the additional 20 bp of DNA at each end in the chromatosome after micrococcal nuclease digestion (2×10 bp). Actually, this model cannot be considered purely 'symmetrical', since the LH molecule itself is highly asymmetrical (see above). DNase I footprinting studies of the chromatosome and dinucleosome containing LHs [16] and UV-induced pyrimidine dimer formation [17] were interpreted as strong support for the on-axis model (see [18] for review).

Neutron scattering studies of chromatosomes [19], however, suggested a more interior location of GLH in the nucleosome. The resolution of the structure of GLH by crystallography [4] and NMR spectroscopy [5] raised further questions about the adequacy of the 'symmetrical' model.

Furthermore, it was shown that at least one chromatosome, specifically the one reconstituted on the somatic 5S gene from *Xenopus borealis*, is very 'asymmetric'. The DNA extensions of the chromatosome were found to be 5 bp on one side and 15 bp on the other [20]. At the same time no protection against hydroxyl radicals or DNase I cleavage by LH was observed at the dyad or at any other site of the chromatosomal [20] or dinucleosomal DNA [21]. Finally, LH/DNA cross-linking [22,23] and site specific cleavage experiments [24] showed GH5 to be located asymmetrically inside the upper

[1]Corresponding authors.
E-mail: sergey.belikov@cmb.ki.se; karpov@genome.eimb.rssi.ru

Abbreviations: LH(s), linker histone(s); GLH, globular domain of linker histone; GH1, globular domain of histone H1; GH5, globular domain of histone H5

PII: S0014-5793(98)01512-9

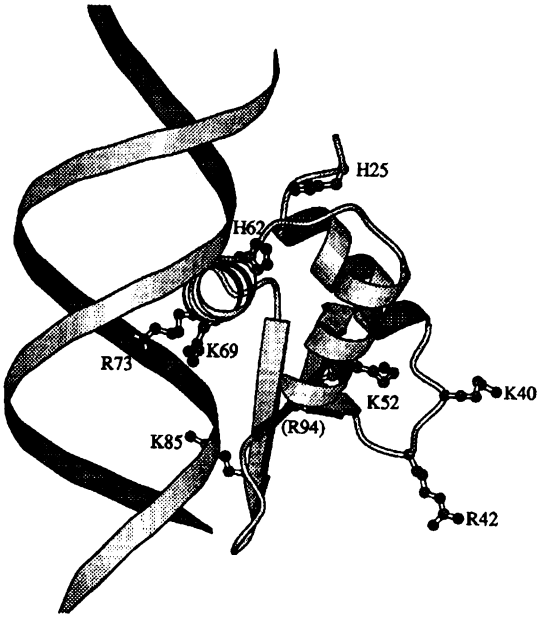

Fig. 1. A model for GH5 binding to DNA. Highly conserved lysines and arginines and two histidines crosslinkable in chromatin are shown (See [7] for details). Reproduced with permission from [7].

gyre of the DNA at position ~6.5 helical turns from the dyad of the nucleosome (as defined by Richmond and coworkers [25]) and in close contact with the histone octamer. Specific crosslinking/cleavage was shown to occur on the same side of the chromatosome where protection against micrococcal nuclease is 15 bp. Micrococcal nuclease, the only available tool for addressing the positions of chromatosome borders, is rather imprecise (see, for example, [26]), and therefore the protection pattern observed for the *Xenopus* 5S nucleosome may reflect the limitations imposed by a strong cleavage specificity of micrococcal nuclease. It has been shown, however, for several other reconstituted chromatosomes that LHs protect chromatosomal DNA asymmetrically, in all cases with ~20 bp protected on one side and no protection on the other [27–30]. DNA-LH crosslinking data both in vivo [31] and in vitro [30] argue for a highly asymmetrical association of LH with DNA. It was also shown that the winged helix transcription factor HNF3, whose structure is isomorphous with that of GLH [6], binds to nucleosomal DNA asymmetrically [32].

If so, there should be a DNA sequence signal(s) strong enough to: (i) select one out of two possible off-axis locations for LH on the DNA; and (ii) determine by targeted LH binding the translational positioning of histone octamer by restricting the nucleosome mobility [21,33]. Such putative sequences were found while analyzing the DNA sequence of a large set of chicken erythrocyte chromatosomes, namely an NGGR (R=purine) quartet always found at one, but not at both, termini of chromatosomal DNA, and an out-of-phase AAA/TTT triplet located ~70 bp away, i.e. close to the dyad [34]. Interestingly, the NGGR sequence is not present at the proposed place in the 5S RNA nucleosome from *X. borealis*. This argues that the NGGR sequence signal is not the only signal directly or indirectly responsible for LH positioning/binding. However, since the arrangement of LHs in chromatin appears to be regular (as supported by chemical crosslinking [35], but see also [36]), it might be sufficient for one nucleosome to have such a signal to determine the orientation of LHs for a nucleosomal array, or even for a whole chromatin domain.

Thus, the results obtained for the 5S nucleosome from *X. borealis* and for nucleosomes reconstructed on other sequences (see above) are in conflict with DNase I footprinting [16,37] and pyrimidine formation data [17] that show preservation of a specific pattern in nucleosome cores and in chromatosomes and thus support symmetrical extending of an extra ~20 bp of chromatosome DNA. If so, it is worth considering these data in more detail. Except for particular concerns relevant to the method used, both these experiments had one essential feature in common: they dealt with the bulk population of native nucleosomes and thus, the patterns are inevitably artificially symmetrized about the DNA midpoint. Such data can be interpreted equally well as indicative of either symmetrical LH binding, or equal occupancy of two

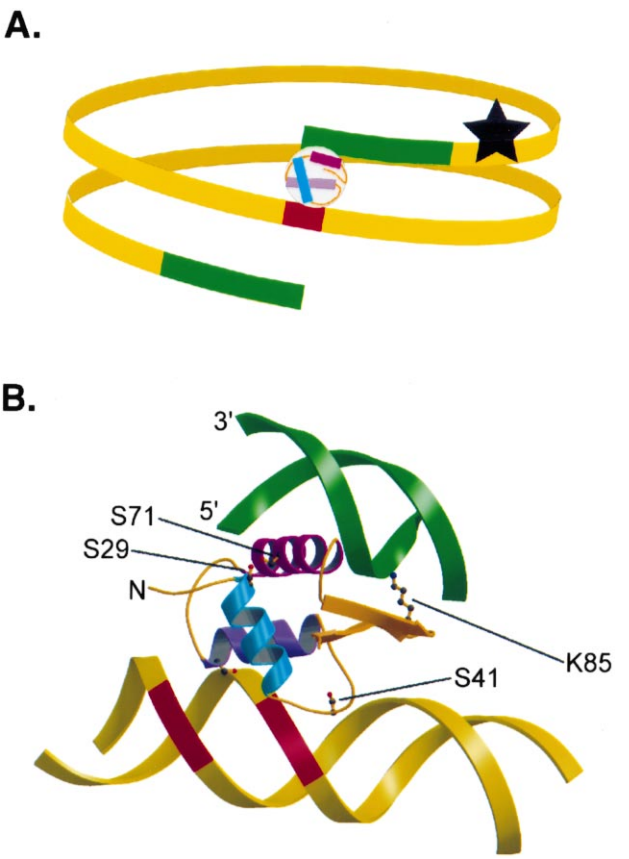

Fig. 2. Schematic representation of the GH5 location in the chromatosome (A) and a more detailed model (B). The GH5 (circle with helix I, cyan; helix II, purple; and helix III, magenta; turns and β-strands in orange) is located between the two DNA gyres. The black star in A shows the contact of GLH as proposed by Pruss and coworkers [23]. The 3 bp around the core particle dyad are in red, the 10 bp of chromatosome 'extensions' are in green, all other DNA is in yellow. Note that both figures are schematic and do not take into account the DNA distortions near the dyad and exit/entry region [17]; note also that only the globular domain is shown. In A the GH5 is located such that its C-terminus is pointing away from the plane of the figure. Reprinted with permission from Nature [40], © 1998 Macmillan Magazines Ltd.

slightly off-axis sites [38]. Another piece of data that again can be interpreted as supporting or disputing the symmetrical (as well as asymmetrical) model was extracted from the database containing DNA sequences isolated from the chromatosomes or nucleosome dimers containing histones H1 and H5. It shows an out-of-phase occurrence of the AAA/TTT triplet 95 bp from the limit of micrococcal nuclease digestion (see [34] for details). On the other hand, in nucleosomal core DNA the same translational marker was found in the immediate vicinity of the dyad, i.e. 70–75 bp away from the limit of micrococcal nuclease [39]. The authors stated that "there is no compelling reason why such a translational marker should be conserved in the two particles". However, there is no reason why the data obtained cannot be interpreted in another way, namely as indicating preservation of AAA/TTT triplet position (location) with respect to the histone octamer and thus for asymmetrical protection of DNA in the chromatosome (95−75=20). The question of whether protection of DNA on both sites is symmetrical or asymmetrical is therefore still open, while the imprecise nature and obvious sequence specificity of micrococcal nuclease further add to the uncertainty.

Recently the binding site and orientation of the globular domain of LH was mapped using site specific DNA-protein photocrosslinking on the bulk population of chicken chromatosomes that were first depleted of endogenous LH and then reconstituted with either GH5 or different GH5 mutants [40]. A model was proposed postulating that GH5 links together one terminus of DNA and the central part of the chromatosome (the dyad). Helix III within the primary site binds to DNA termini via the major groove around position 7.5–8 (in concordance with crystallography and NMR data [4,5]), while the secondary binding site interacts with chromatosomal DNA close to its midpoint (see Fig. 2). This location of GLH suggests an orientation of the LH C-terminus which allows simultaneous interaction with both entering and exiting nucleosomal linkers, bridging between them as was observed by electron microscopy [11]. The model assumes symmetric (10+10 bp) extensions of chromatosomal DNA, however, the authors themselves stated that the "data are compatible with either symmetric or asymmetric extensions". Since LH interacts only with one DNA terminus, protection against micrococcal nuclease action from the other side of chromatosomal DNA could result from the interaction with the histone octamer altered by LH binding. Indeed, it was recently shown that accommodation of LH in the nucleosome results in substantial redistribution of core histone-DNA contacts [29,41].

This model is in perfect accord with almost all biochemical data available so far (see above), but differs from the model proposed by Wolffe and coworkers for 5S *X. borealis* nucleosome which places GLH further into the nucleosome and in close proximity to the histone octamer (position ~6.5) [23]. The data obtained on the 5S chromatosome may of course reflect some specific features of the sequence of the *X. borealis* somatic 5S gene (see, for example, [8]). However, the interpretation of crosslinking [22,23] and site specific cleavage experiments [24] depends crucially on the assumption that the histone octamer adopts a single translational position within the sequence used. This critical point has recently been questioned [42–44], and since both sides have strong arguments in support of their conclusions, more experiments need to be done to clarify this question.

3. But questions remain

I only say, suppose the supposition.Lord Byron, *Don Juan*, part I, stanza LXXXVII

In conclusion, the concept of asymmetrical location of LH histones is consistent with all previous experimental data obtained so far and thus seems to be beyond reasonable doubt. Several questions of great importance to our understanding of chromatin structure, however, still remain unclear.

(1) *Is the location of LHs along the chromatin fiber unidirectional or random (or potentially more complex, for example, occurring as pairs)*? We suppose that LHs are always located at the same end of successive nucleosomes within a domain or at least in rather long arrays of nucleosomes. This supposition is indirectly supported by the LH-LH crosslinking experiments [35] but needs to be addressed by more direct means.

2. *What is the orientation of the LH placement in a nucleosome with respect to direction of transcription*? A possibility that LH is always located on the 3′ end of the transcribed nucleosome with respect to the direction of transcription (as has been shown for somatic-type *X. borealis* 5S RNA nucleosome [30]) seems very attractive; however, this is no more than a speculation and again must be verified directly.

The list of the questions could continue. The successful resolution of these questions may give us a clue to how the chromatin fiber is organized and how LHs specifically repress and activate transcription.

Acknowledgements: We are indebted to Drs. D. Rhodes and S. Muyldermans for making their data available to us before publication. We wish to thank Drs. T. Klenka, S. Muyldermans, D. Rhodes and A. Wolffe for critical reading of the manuscript and valuable comments. S.B. acknowledges support from the Swedish Cancer Foundation (Grants 3745-B97-02VAA and 2222-B97-13XCC) and the Royal Swedish Academy of Sciences (Grant 12682). The work was financially supported by the Russian Foundation for Fundamental Research (Grant 96-04-50484), the Russian state program 'Frontiers in Genetics' and the Department of Energy (USA).

References

[1] van Holde, K.E. (1989) Chromatin, Springer-Verlag, New York.
[2] Allan, J., Hartman, P.G., Crane-Robinson, C. and Aviles, F.X. (1980) Nature 288, 657–679.
[3] Hartman, P.G., Chapman, G.E., Moss, T. and Bradbury, E.M. (1977) Eur. J. Biochem. 77, 45–51.
[4] Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. and Sweet, R.M. (1993) Nature 362, 219–223.
[5] Cerf, C., Lippens, G., Ramakrishnan, V., Muyldermans, S., Segers, A., Wyns, L., Wodak, S.J. and Hallenga, K. (1994) Biochemistry 33, 11079–11080.
[6] Clark, K.L., Halay, E.D., Lai, E. and Burley, S.K. (1993) Nature 364, 412–417.
[7] Goytisolo, F.A., Gershman, S.-E., Yu, X., Rees, C., Graziano, V., Ramakrishnan, V. and Thomas, J.O. (1996) EMBO J. 15, 3421–3429.
[8] Hayes, J.J., Kaplan, R., Ura, K., Pruss, D. and Wolffe, A. (1996) J. Biol. Chem. 271, 25817–25822.
[9] Baxevanis, A.D. and Landsman, D. (1997) Nucleic Acids Res. 25, 272–273.
[10] Clark, D.J., Hill, C.S., Martin, S.R. and Thomas, J.O. (1988) EMBO J. 7, 69–75.
[11] Hamiche, A., Schultz, P., Ramakrishnan, V., Oudet, P. and Prunell, A. (1996) J. Mol. Biol 257, 30–42.
[12] Wu, M., Allis, C.D., Richman, R., Cook, R.G. and Gorovsky, M.A. (1986) Proc. Natl. Acad. Sci. USA 83, 8674–8678.
[13] Shen, X. and Gorovsky, M.A. (1996) Cell 86, 457–483.
[14] Landsman, D. (1996) Trends Biochem. Sci. 21, 287–288.

[15] Patterton, H.G., Landel, C.C., Landsman, D., Peterson, C.L. and Simpson, R.T. (1998) J. Biol. Chem. 273, 7268–7276.
[16] Staynov, D.Z. and Crane-Robinson, C. (1988) EMBO J. 7, 3685–3691.
[17] Pehrson, J.R. (1989) Proc. Natl. Acad. Sci. USA 86, 9149–9153.
[18] Crane-Robinson, C. (1997) Trends Biochem. Sci. 22, 75–77.
[19] Lambert, S., Muyldermans, S., Baldwin, J., Ibel, K. and Wyns, L. (1991) Biochem. Biophys. Res. Commun. 179, 810–816.
[20] Hayes, J.J. and Wolffe, A.P. (1993) Proc. Natl. Acad. Sci. USA 90, 6415–6419.
[21] Ura, K., Hayes, J.J. and Wolffe, A.P. (1995) EMBO J. 14, 3752–3765.
[22] Hayes, J.J., Pruss, D. and Wolffe, A.P. (1994) Proc. Natl. Acad. Sci. USA 91, 7817–7821.
[23] Pruss, D., Bartholomew, B., Persinger, J., Hayes, J.J., Arents, G., Moundrianakis, E.N. and Wolffe, A.P. (1996) Science 274, 614–617.
[24] Hayes, J.J. (1996) Biochemistry 35, 11931–11937.
[25] Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. and Klug, A. (1984) Nature 311, 532–537.
[26] Hörz, W. and Altenburger, W. (1981) Nucleic Acids Res. 9, 2643–2658.
[27] Wong, J., Li, Q., Levi, B.Z., Shi, Y.-B. and Wolffe, A.P. (1997) EMBO J. 23, 7130–7145.
[28] An, W., Leuba, S.H., van Holde, K. and Zlatanova, J. (1998) Proc. Natl. Acad. Sci. USA 95, 3396–3401.
[29] Guschin, D., Chandler, S. and Wolffe, A.P. (1998) Biochemistry 37, 8629–8636.
[30] Sera, T. and Wolffe, A.P. (1998) Mol. Cell. Biol. 18, 3668–3680.
[31] Belikov, S.V., Belgovsky, A.I., Partolina, M.P., Karpov, V.L. and Mirzabekov, A.D. (1993) Nucleic Acids Res. 21, 4796–4802.
[32] 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.
[33] Pennings, S., Meersseman, G. and Bradbury, E.M. (1994) Proc. Natl. Acad. Sci. USA 91, 10275–10279.
[34] Travers, A. and Muyldermans, S.V. (1996) J. Mol. Biol 257, 486–491.
[35] Lennard, A.C. and Thomas, J.O. (1985) EMBO J. 4, 3455–3462.
[36] Nikolaev, L.G., Glotov, B.O., Dashkevich, V.K., Barbashov, S.F. and Severin, E.S. (1983) FEBS Lett. 125, 20–24.
[37] Simpson, R.T. (1978) Biochemistry 17, 5524–5531.
[38] Staynov, D.Z. and Proykova, Y.G. (1998) J. Mol. Biol. 279, 59–71.
[39] Satchwell, S.C. and Travers, A.A. (1989) EMBO J. 8, 229–238.
[40] Zhou, Y.-B., Gershman, S.E., Ramakrishnan, V., Travers, A. and Muyldermans, S. (1998) Nature 395, 402–405.
[41] Lee, K.-M. and Hayes, J.J. (1998) Biochemistry 37, 8622–8628.
[42] Howe, L. and Ausio, J. (1998) Mol. Cell. Biol. 18, 1156–1162.
[43] Panetta, G., Butinelli, M., Flaus, A., Richmond, T.J. and Rhodes, D. (1998) J. Mol. Biol. 282, 683–697.
[44] An, W., van Holde, K. and Zlatanova, J. (1998) Nucleic Acids Res. 26, 4042–4047.