Comparative analysis of the secondary structure of non-histone chromatin proteins HMGB1 and HMGB2

E V Chikhirzhina¹, D N Osinnikova² and A M Polyanichko¹,²

¹Institute of Cytology, RAS, Russia
²Faculty of Physics, St Petersburg University, Russia

E-mail: e.chikhirzhina@incras.ru

Abstract. The non-histone chromosomal proteins HMGB1 and HMGB2 were found in the cells of all studied eukaryotes. They are involved in cell decision and many biological processes such as replication, transcription, repair, etc. In this work, the secondary structure of the HMGB1 and HMGB2 proteins was studied by the circular dichroism method. It was shown that, despite the high homology between them, the secondary structure of these proteins is different. The revealed structural features, most likely, should influence their functions in the cell nucleus, in particular, the interaction with DNA and other proteins.

1. Introduction

Proteins HMGB1 and HMGB2 belong to the group of non-histone chromosomal proteins with high electrophoretic mobility (High Mobility Group). Proteins of this group are the most common non-histone chromatin proteins and are found in the cells of all studied eukaryotes; however, their role in chromatin is still not fully understood. Members of this family are actively involved in the structural organization of chromatin and play an important role in the regulation of various cellular processes [1]. All HMGB proteins contain at least one structurally conserved DNA-binding domain known as the HMGB domain, the presence of which is the main criterion for a protein to be attributed to this family. The HMGB domain is a motif of about 80 aa and is found in many transcription factors. Single-domain proteins (lymphoid enhancer-binding factor LEF-1, sex-determining factor SRY and proteins of the SOX subfamily, chromatin modeling factors BAF57 and PB1, and many others) bind to DNA in a site-specific manner [1]. Multidomain HMGB proteins (HMGB1-4, mitochondrial factors mtTF1 and ABF2, RNA polymerase I transcription factor UBF, Drosophila DSP1 protein, etc.) are characterized by the absence of sequence-specificity when interacting with DNA [1]. A characteristic feature of these proteins is the ability to recognize and bind to DNA regions with different structural features (for example, bends, cross structures, hairpins).

The primary structures of the HMGB1 and HMGB2 proteins are rather close: among 250 amino acid residues, only 37 are different (in 18 cases, this is a conservative substitution). Both proteins consist of two DNA-binding domains connected by a short linker, a short N-terminal fragment, and a structurally disordered C-terminal region [2]. One of the established differences in the primary structure of HMGB1/2 proteins is the size of their negatively charged C-terminal fragments, consisting of a continuous sequence of dicarboxylic amino acids. These proteins differ, first of all, in the length of this fragment: 30 amino acid residues in HMGB1 and 20 in HMGB2. It should be noted that it is this region of the protein that simulates its interaction with DNA [1]. Apparently, the differences in the primary structure of the C-terminal region of the HMGB1 and HMGB2 molecules lead to differences in the
functions of the proteins themselves. These regions do not have an ordered spatial organization and are believed to be able to interact with various histones, regulating their binding with DNA [3]. Earlier it was demonstrated that some structural properties of HMGB1/2 proteins resemble those of natively unfolded proteins [4].

It is known that both proteins are actively involved in the regulation of major nuclear processes, such as transcription, replication, recombination, and DNA repair. However, depending on the post-translational modifications and the redox state of the HMGB1 protein, it can function in nucleus, in cytoplasm, or leave the cell [1,5]. The functions performed by this protein are closely related to its localization. The extracellular activity of HMGB1 ranges from the first stages of damage / inflammation to tissue regeneration [6], affects the rate of tumor development, the migration of mesenchymal stem cells, promotes the synthesis of cytokines (chemokines that increase the rate of stem cell migration [7]). As for the HMGB2 protein, there are no data on its migration neither in the cytosol nor in intercellular space. Apparently, when the localization of HMGB1 changes, the HMGB2 protein can replace it in the nucleus. According to the earlier published data, the rate of development of malignant tumors and the spread of metastases in animals, as well as the proliferation and differentiation of embryonic stem cells, directly depends on the level of expression of the HMGB1 and HMGB2 proteins [1, 2, 8].

Despite the other high homology in the primary structures of the HMGB1 and HMGB2 proteins, there is currently no information on the differences in their secondary and tertiary structures. In addition, the presence in the cell of two proteins with close amino acid composition also suggests some difference in their functions.

2. Materials and methods
Nuclear proteins HMGB1 (M = 26 500 Da) and HMGB2 (M = 26 000 Da) were isolated from calf thymus by extraction with 5% perchloric acid with subsequent precipitation using 5,5 V precooled acidic acetone at -20 °C overnight as described earlier [9]. The isolated individual proteins were purified by FPLC. The purity of the protein was tested by sodium dodecyl sulfate polyacrylamide gel electrophoresis [10]. Spectra of circular dichroism (CD) of the proteins were registered using dichrograph Cary 60 (USA) in 1-cm quartz cells within the range of 200–260 nm. The degree of α-helicity was calculated based on the molar ellipticity at 222 nm \( \theta_{222} \) [11]

\[
\alpha\% = \frac{-[\theta_{222}]+3000}{39000}
\]

3. Results and discussion
Protein functions are closely related to their conformational state. To study the conformational features of the HMGB1 and HMGB2 proteins, we used circular dichroism, which allows us to detect differences in the secondary structure of homologous proteins. We analyzed CD spectra of the HMGB1 and HMGB2 proteins in far ultraviolet region, which allow us to determine their secondary structure under various conditions (Figure 1).

In a neutral solvent (water, pH 6.0), both proteins demonstrate similar spectra. The shape of the spectra is typical for presence α-helical segments and regions with a “random coil” structure. At the same time, the quantitative analysis of the spectra shows that under these conditions there are more α-helical segments in the HMGB1 protein (~25%) compared to HMGB2 (15%).

In order to assess the ability of the studied proteins to form additional α-helical structures, we analyzed the CD spectra of proteins in a solution of 80% ethanol, which stimulates the formation of helical structures in protein molecules. Under these conditions, we observe a shift of both negative CD bands to shorter wavelengths compared to the spectrum of a solution in a neutral solvent. Moreover, the intensity of the CD signal of HMGB1 in this region increases only slightly, while the CD of HMGB2 almost doubles. The α-helicity estimate gives 30% and 40% for HMGB1 and HMGB2 respectively.

An increase in the ionic strength of a solution to 1.5 M NaCl leads to the formation of more compact forms of the protein structures. Under these conditions, the studied proteins are characterized by different CD spectra and, therefore, differ in their conformation. The CD spectrum of HMGB2 is
characterized by deeper negative bands indicating higher $\alpha$-helical content in HMGB2 protein than in the HMGB1 protein (~70% vs ~40%).

Figure 1. CD spectra of HMGB1 (solid lines) and HMGB2 (dashed lines) in different solvents: black lines – water, pH 6.0; blue lines – 80% ethanol; red lines – 1.5 M NaCl

As it was shown earlier [12, 13], the length of the acid tail of HMGB1 is sufficient to bind the adjacent HMGB domain within the same protein molecule. The latter provides two consequences: (1) one of the HMGB domains of the HMGB1 molecule is unable to interact with DNA blocked by the tail, and (2) the tail itself is unable to establish intermolecular contacts. In the case of the HMGB2 protein, the shorter tail is not long enough for efficient intramolecular interaction, which makes it possible to tightly bind HMGB domains to DNA and stimulate intermolecular interactions of acidic tails. Most likely, binding of the C-terminal tail to the HMGB-domain restrict to some extent the conformational flexibility of the HMGB-domain in HMGB1 resulting. In HMGB2 on contrary, both HMGB-domains maintain their conformational flexibility, resulting in almost twice higher degree of helicity of the whole protein in the 1.5 M NaCl.

4. Conclusions

Analysis of the CD spectra of proteins showed that under physiological conditions, the HMGB1 protein is characterized by a more ordered secondary structure than HMGB2. At the same time, HMGB2 is able to form more $\alpha$-helical regions than HMGB1, which indicates a greater conformational flexibility of the HMGB2 protein. We believe that such flexibility contributes to the structural adaptation of the HMGB2 protein to a much greater extent than for the HMGB1 protein.

Both proteins are involved in various processes: replication, repair, etc. However, the presence in the cell of both proteins that are so close in primary and secondary structure cannot be accidental. There is a huge pool of experimental data on the extracellular functions of HMGB1 [1]. Under certain conditions (change in post-translational modifications or redox status), HMGB1 leaves the nucleus, moves into the
cytoplasm, and then goes out into the extracellular space [1,5]. All these processes are associated with various diseases, ranging from cardiac pathologies and ending with violations in the development of the fetus, and hence with damage to cells, with their death [6-8]. Consequently, the amount of HMGB1 protein in the cell decreases, and HMGB2 continues to function normally. Most likely, it is the ratio of the proteins in the cell nucleus that is the most important factor for their functioning. Thus, it is important to understand the subtle structural differences between the HMGB1 and HMGB2 proteins, which undoubtedly affects the mechanisms of their interaction with DNA.

References
[1] Chikhirzhina E V, Starkova T, Beljajev A, Polyanichko A M and Tomilin A N 2020 Inter. J. Mol. Sci. 21 7948
[2] Read C M, Cary P D, Crane-Robinson C, Driscoll P C and Norman D G 1993 Nucl. Acids Res. 21 3427-36
[3] Chikhirzhina E V, Starkova T Yu and Polyanichko A M 2021 Biophysics 66 373-8
[4] Polyanichko A M, Rodionova T J, Vorob'ev V I and Chikhirzhina E V 2011 Cell Tis. Biol. 5 114-9
[5] Chikhirzhina E V E V, Polyanichko A M and Starkova T Yu 2020 Tsitologia 62 716-25
[6] Kuznik B I, Khavinson V K, Linkova N S and Sall T S 2017 Usp. Fiziol. Nauk 48 40-55
[7] Raucci A, Di Maggio S, Scavello F, D’Ambrosio A, Bianchi M E and Capogrossi M C 2019 Cell. Mol. Life Sci. 76 211-29
[8] Bagherpoor A J et al. 2017 Stem Cells Dev. 26 328-40
[9] Chikhirzhina E V, Polyanichko A M, Skvortsov A N, Kostyleva E I, Houssier C, Vorobyev V I 2002 Mol. Biol. (Moscow) 36 412-8
[10] Laemmli U K 1970 Nature 227 680-5
[11] Kohlstaedt L A and Cole RD 1994 Biochemistry 3 570-5
[12] Cato L, Stott K, Watson M and Thomas J O 2008 J. Mol. Biol. 384 1262-72
[13] Polyanichko A M, Leonenko Z V, Cramb D, Wieser H, Vorob'ev V I and Chikhirzhina E V 2008 Biophysics 53 202-6