Journal of the Korean Magnetic Resonance Society **2018**, *22*, 119-124
DOI 10.6564/JKMRS.2018.22.4.119

Structural assessment of the tetramerization domain and DNA-binding domain of CP2c

Ku-Sung Jo[1,†], Ki-Sung Ryu[1,†], Hee-Wan Yu[1,¶], Seu-Na Lee[1,§], Ji-Hun Kim[2], Eun-Hee Kim[3], Chae-Yeon Wang[1], Chan-Gil Kim[1], Chul Geun Kim[4], and Hyung-Sik Won[1,*]

[1]Department of Biotechnology, College of Biomedical and Health Science, Konkuk University, Chungju, Chungbuk 27478, Republic of Korea
[2]College of Pharmacy, Chungbuk National University, Cheongju, Chungbuk 28160, Republic of Korea
[3]Protein Structure Group, Korea Basic Science Institute, Ochang, Chungbuk 28119, Republic of Korea
[4]Department of Life Science and Research Institute for Natural Sciences, College of Natural Sciences, Hanyang University, Seoul 04763, Republic of Korea

Received Dec 1, 2018; Revised Dec 13, 2018; Accepted Dec 14, 2018

Abstract Although the transcription factor CP2c has been recently validated as a promising target for development of novel anticancer therapy, its structure has not been solved yet. In the present study, the purified recombinant protein corresponding to the tetramerization domain of CP2c appeared to be well folded, whereas the Elf-1 domain showed a largely unfolded conformation. Particularly, the Elf-1 domain, which contains the putative DNA-binding region, showed a conformational equilibrium between relatively less-ordered and well-ordered conformers. Interestingly, addition of zinc shifted the equilibrium to the relatively more structured conformer, whereas zinc binding decreased the overall stability of the protein, leading to a promoted precipitation. Likewise, a dodecapeptide that has been suggested to bind to the Elf-1 domain also appeared to shift the conformational equilibrium and to destabilize the protein. These results constitute the first structural characterization of the CP2c domains and newly suggest that zinc ion might be involved in the conformational regulation of the protein.

Keywords CP2c, Elf-1 domain, zinc binding, conformational regulation, NMR

Introduction

Grh/CP2-family proteins, which are transcriptional regulators conserved in multicellular organisms, are divided into two individual divisions, Grh (Grainyhead; also known as Elf-1 or NTF-1) and CP2 (CCAAT-binding protein 2).[1,2] Grh was first discovered in *Drosophila* and three Grh-like (Grhl) subfamily proteins were found in human. In case of CP2, which was originally identified as a murine α-globin gene regulator, six human homologs were identified.[3,4] CP2c (also referred to as LSF or LBP-1c), the central transcription factor of CP2 isoforms, is ubiquitously expressed to participate in diverse cellular processes, including hematopoiesis, immune responses, the cell cycle and neural development. For those pleiotropic actions, CP2c specifically regulates the expression of various target

† These authors contributed equally to this work
¶ Current address: R&D center, TiumBio Co. Ltd., Seongnam, Gyeonggi 13493, Republic of Korea
§ Current address: Department of Life Sciences, Korea University, Seoul 02841, Republic of Korea
* Address correspondence to: **Hyung-Sik Won,** Department of Biotechnology, College of Biomedical and Health Science, Konkuk University, Chungju, Chungbuk 27478, Republic of Korea; Tel: 82-43-840-3589; E-mail: wonhs@kku.ac.kr

Figure 1. Schematic diagram showing the putative domain organization of CP2c. Constructed regions for structural studies in this study are illustrated by gray bars. TAD, transactivation domain; BD, basic domain; TD, tetramerization domain.

genes, including interleukins and thymidylate synthases. In particular, tetrameric organization of CP2c is known to be responsible for its DNA binding, whereas dimeric Grhl binds to target DNA sites.[1,2,5] However, both the two proteins are known to share a similar DNA binding domain that is predicted to resemble the p53 DNA-binding domain.[1,2] For CP2c, the residues 63-244 and 305-403 regions are denoted as the putative DNA binding domain (Elf-1 domain in Fig. 1) and the tetramerization domain (TD in Fig. 1).[4,5] The multifunctional aspects of CP2c are also mediated by molecular interactions with diverse binding partners such as PIAS1 in erythroid cells and p66α in chromatin remodeling complexes.[3,6,7] Consequently, the functional regulation of CP2c is also associated with various human diseases, such as Alzheimer's disease, AIDS, drepanocytosis, and cancer. In particular, oncogenic function of CP2c was recently validated in hepatocellular carcinoma,[5,8] which led to the discovery of a small-molecule anticancer agent, FQI1,[5,9] targeting CP2c. In addition, a peptide library screening by phage display has previously yielded several peptide sequences as promising binders to CP2c,[4] although their functional utilities have not yet been explored.

Despite those functional significances of CP2c, its detailed structure, which could definitely provide valuable information for developing novel anticancer therapy, is not available yet. Therefore, we initiated structural study of CP2c in solution, by attempting recombinant production of the CP2c subdomains. In this paper, a preliminary characterization of the tetramerization domain (TD) and the Elf-1 domain is reported. In addition, potential involvement of zinc in the conformational regulation of CP2c is suggested.

Experimental Methods

Protein preparation - DNA fragments encoding residues 63-244 (Elf1D) and 305-403 (TD) of CP2c were PCR-amplified from the previously constructed recombinant plasmids, pGEX-4T1-CP2c,[4] as templates, using appropriate forward and reverse primers that contained the *Nde*I and *Xho*I restriction sites, respectively. The PCR products digested with the restriction enzymes were ligated into the *Nde*I/*Xho*I-digested pCold-1 vector to express the protein with an N-terminally hexahistidine tag. The *E. coli* BL21(DE3)pLysS cells transformed with the constructed plasmids were cultured at 37 °C. When the A_{600} of cell growth reached about 0.6, protein expression was induced for 12 h at 18 °C, by adding IPTG at a final concentration of 0.5 mM. The cells were then harvested by centrifugation, followed by disruption by sonication. From the supernatants, the expressed target proteins were purified via sequential applications of nickel-affinity chromatography (HisTrap HP column, GE Healthcare), ion-exchange chromatography (HiTrap SP FF column, GE Healthcare), and gel-permeation chromatography (HiLoad 16/600 Superdex 75, GE Healthcare). As the purified proteins severely precipitated during the next step for cleavage of the tagged histidines, final samples for analysis were prepared with the hexahistidine tag-fused proteins. The protein concentration was spectrophotometrically deduced using its molar absorptivity at 280 nm, which was predicted from the amino acid sequence.

Analytical gel-filtration - Elution volumes of purified proteins and molecular weight standards in

gel-filtration were measured on a HiLoad 16/60 Superdex^{TM} 75 (GE Healthcare) gel-permeation column pre-equilibrated with the standard buffer (50 mM sodium phosphate, pH 7.4, 150 mM NaCl, 1 mM DTT), at a 1 ml/min flow rate. From the individual elution volumes in gel-filtration, hydrodynamics sizes (apparent molecular weights) of the sample proteins were calculated as described previously.[10,11]

Circular dichroism (CD) spectroscopy - The standard far-UV CD spectra of the purified protein solutions dissolved in 15 mM sodium phosphate buffer (pH 7.4) containing 50 mM NaCl were measured on a Jasco J-710 spectropolarimeter at room temperature, using a 0.1 cm path-length cell, with a 1 nm bandwidth and a 4 sec response time. Three individual scans taken from 260 to 190 nm were summed and averaged, followed by the subtraction of solvent CD signals.

Nuclear magnetic resonance (NMR) spectroscopy - The [^{15}N]-enriched protein sample for NMR was produced by cultivating the transformed *E. coli* cells in M9 minimal media supplemented with [^{15}N]NH_4Cl as a sole nitrogen source. NMR measurements were performed with a Bruker Biospin Avance II 800 spectrometer equipped with a cryoprobe, at 298 K. The conventional 2D-[^{1}H/^{15}N]-TROSY spectra of the [^{15}N]-enriched samples were measured. For titration to Elf1D, zinc sulfate and a dodecapeptide (sequence: N-acetyl-HKFHQHRLPHLA-amide)[4] dissolved in the standard buffer were used. Obtained NMR data were processed using NMRPipe software and the spectra were visualized and analyzed using NMRViewJ program.

Results

Characterization of the tetramerization domain - Our initial attempt to produce recombinant protein for structural study of full-length CP2c (residues 1-503) was not successful using *E. coli* expression system (data not shown). Alternatively, therefore, we conducted subcloning of various regions of CP2c for *E. coli* expression (Fig. 1). Although most of the constructed proteins formed inclusion bodies in *E. coli* or readily precipitated *in vitro*, the tetramerization domain (TD; residues 305-403) could be obtained as a soluble protein and was subjected to NMR measurement. Consistent with its far-UV CD spectrum, which indicated a predominantly α-helical conformation (note the strong negative bands near 208 and 222 nm; inset in Fig. 2), the [^{1}H/^{15}N]-TROSY spectrum that showed a good dispersion of resonances (Fig. 2) was also indicative of a well-ordered conformation of the domain. However, this recombinant protein showed a low solubility and appeared to readily form high-order oligomers beyond decamer; i.e., the purified protein was eluted at void volume of gel-permeation chromatography after overnight storage at 4 °C (data not shown). The fact that the number of peaks observed in the NMR spectrum (Fig. 2) was quite fewer than that expected from its number of residues was also attributable to such a highly oligomerizing or aggregating tendency of the protein.

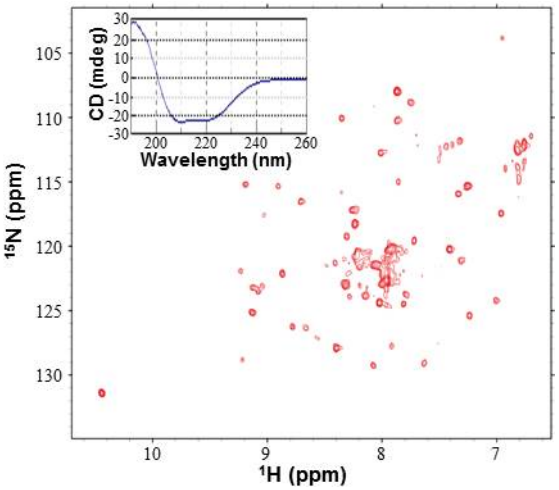

Figure 2. 2D-[^{1}H/^{15}N]-TROSY spectrum of CP2c-TD (0.1 mM). Inset shows its far-UV CD spectrum obtained with 20 μM sample.

Conformational equilibrium of the DNA-binding domain - As an another soluble construct, the purified Elf1D showed a distinctive behavior in the

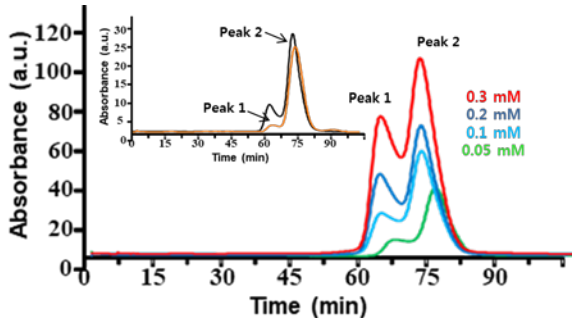

Figure 3. Gel-filtration profiles of the purified Elf1D at various concentrations. Inset shows the elution profiles of separated peak-1 and peak-2 fractions.

gel-filtration analysis characterized by two distinct elution peaks (Fig. 3). In addition, separated each fraction was split again into two peaks, where the portion of later elution was always higher than that of the first elution (inset in Fig. 3). The hydrodynamic size deduced from the later elution (22.6 kDa) was consistent with the theoretical monomer weight (22.9 kDa) of the protein, whereas the estimated hydrodynamic size for the first elution (39.2 kDa) was close to a dimeric size. Usually, however, a fast elution in gel-filtration can be also achieved by an elongated shape or unfolded conformation of a protein. In this regard, it was noteworthy that the far-UV CD spectrum indicated a large portion of disordered conformation reflected by a strong negative band near 200 nm (black in Fig. 4A), whereas the [^{1}H/^{15}N]-TROSY spectrum was indicative of a well-folded conformation yielding a good spectral dispersion (black in Fig. 4B). Collectively, therefore, the results could be interpreted as a conformational equilibrium of the isolated Elf1D between a predominantly disordered and a relatively well-structured conformer, where the former and the latter correspond to the fast and late elution, respectively, in the gel-filtration profile (Fig. 3). Interestingly, many resonances in the [^{1}H/^{15}N]-TROSY spectrum appeared to accompany small satellite peaks, which might also reflect a conformation equilibrium entailing a slow-rate of conformational exchange.

Zinc and a peptide binding to Elf1D - Although

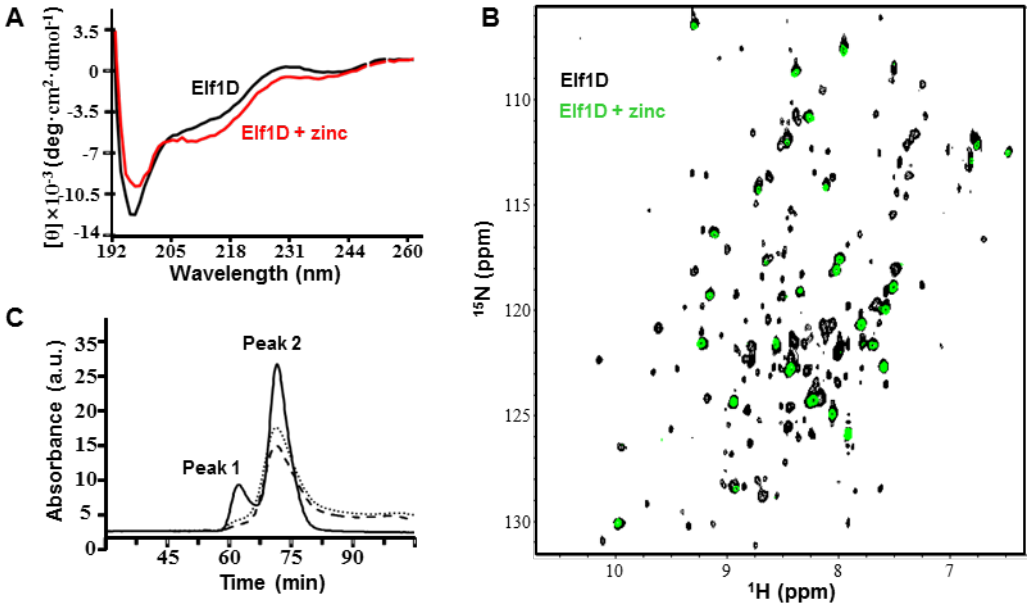

Figure 4. Monitoring of zinc binding to Elf1D. (A) Far-UV CD spectra of Elf1D (20 μM) in the absence (black) and presence of 2 equimolar zinc (red). (B) 2D-[^{1}H/^{15}N]-TROSY spectra of Elf1D (150 μM) in the absence (black) and presence of 2 equimoloar zinc (green). (C) Gel-filtration elution profiles of Elf1D (70 μM) in the absence (solid line) and presence of 2 equimolar zinc (dashed line for peak-1 fraction and dotted line for peak-2 fraction).

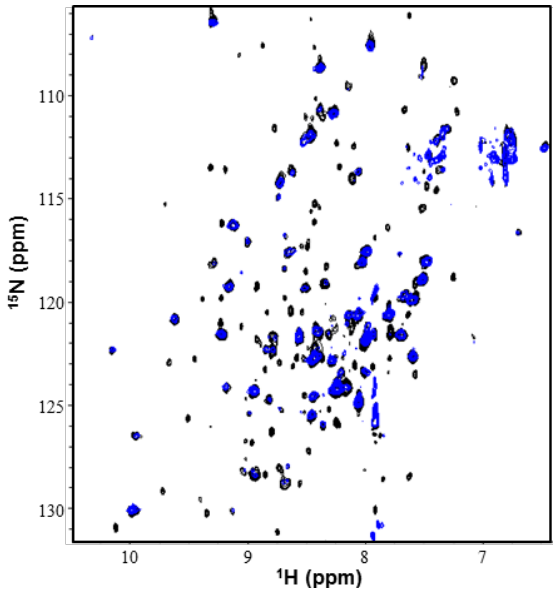

Figure 5. 2D-[^{1}H/^{15}N]-TROSY spectrum of Elf1D (150 μM) in the absence (black) and presence of a binding peptide (300 μM; blue).

CP2c-Elf1D could be obtained as a soluble recombinant protein, it also showed a low solubility less than 0.2 mM and a tendency to precipitate gradually during storage and NMR measurements. Although our screening of additives (such as various metals and amino acids) to increase the protein stability was not successful, we found conversely that the addition of zinc to the protein solution promoted the protein precipitation. Interestingly, far-UV CD spectrum of Elf1D showed a characteristic change in the presence of zinc (Fig. 4A), indicating the increased contents of β-strands (note the signal intensification centered at 216 nm) at the expense of disordered region (note the significant signal reduction around 200 nm). In addition, in the gel-filtration with zinc, the formerly split peaks without zinc merged into a single elution peak that corresponded to the more well-structured conformer. Unfortunately, in the presence of zinc, high-quality NMR spectrum of Elf1D could not be obtained due to the lowered solubility and promoted precipitation during measurement. However, in the presence of zinc, it was observed that the aforementioned small satellite peaks in the [^{1}H/^{15}N]-TROSY spectrum disappeared upon the addition of zinc, leading to the homogeneous shapes of the formerly heterogeneous resonances. This type of spectral change was also observed for a known peptide binding to Elf1D (Fig. 5), which also promoted the protein precipitation. Collectively, therefore, the present results suggest that zinc ion can bind to the Elf1D of CP2c, thereby shifting its conformational equilibrium to the more structured conformer.

Discussion

The tetramerization domain (TD) and DNA-binding domain of CP2c (Elf1D) constitutes the structural requirements for DNA binding of CP2c. In the present study, the TD was revealed to adopt a predominantly α-helical conformation that exerts a high propensity for oligomerization. In contrast, the Elf1D appeared to undergo a conformational equilibrium between largely disordered one and a predominantly β-stranded conformer. Upon the binding of zinc to Elf1D, the equilibrium could be shifted to the conformer more structured but having lower stability. In this regard, it is noteworthy that the DNA-binding domains of Grh/CP2-family proteins share a structural homology to the p53 DNA-binding domain.[1,2] The p53 tumor suppressor requires zinc ion for its site-specific DNA binding and proper transcriptional activation.[12] Furthermore, it has been suggested that zinc modulates folding and misfolding of p53 at its DNA-binding domain. However, zinc is not involved in regulation of the CP2c-homologous Grh proteins.[2] Likewise, p53 functions as tetramers, whereas Grhl proteins forms dimers for their DNA binding. In particular, the tetramerization of p53 is mediated by its TD that adopts an α-helical conformation. Collectively, the present results showing the α-helical conformation of TD and zinc binding to Elf1D of CP2c suggest that the mode of action and conformational regulation of CP2c would closely resemble those of p53 rather than Grhl proteins. Additionally, the present results verified the binding of a known dodecapeptide to Elf1D. Therefore, structural and functional regulation of CP2c by zinc and/or the binding peptide would be

worthy of further investigation in detail.

Acknowledgements

This study was supported by the National Research Foundation (NRF) (grant no. 2016R1A2B4009700). The use of NMR and CD machines was supported by the Korea Basic Science Institute (Ochang, Korea) under the R&D program (project No. D38700) supervised by the Ministry of Science and ICT.

References

1. U. Hansen, L. Owens, and U. H. Saxena, *Cell Cycle* **8**, 2146 (2009)
2. Q. Ming, Y. Roske, A. Schuetz, K. Walentin, I. Ibraimi, K. M. Schmidt-Ott, and U. Heinemann, *Nucleic Acids Res.* **46**, 2082 (2018)
3. K. -S. Jo, H. -R. Jo, C. G. Kim, C. -G. Kim, and H. -S. Won, *J. Kor. Magn. Reson. Soc.* **18**, 30 (2014)
4. H. C. Kang, B. M. Chung, J. H. Chae, S. -I. Yang, C. G. Kim, and C. G. Kim, *FEBS J.* **271**, 1265 (2005)
5. P. K. Santhekadur, D. Rajasekaran, A. Siddiq, R. Gredler, D. Chen, S. E. Schaus, U. Hansen, P. B. Fisher, and D. Sarkar, *Am. J. Cancer Res.* **2**, 269 (2012)
6. H. C. Kang, J. H. Chae, K. S. Choi, J. H. Shin, C. G. Kim, and C. G. Kim, *Nucleic Acids Res.* **38**, 5456 (2010)
7. M. Y. Kim, J. S. Kim, S. H. Son, C. S. Lim, H. Y. Eum, D. H. Ha, M. A. Park, E. J. Baek, B. -Y. Ryu, H. C. Kang, V. N. Uversky, and C. G. Kim, *Nucleic Acids Res.* **46**, 4933 (2018)
8. B. K. Yoo, L. Emdad, R. Gredler, C. Fuller, C. I. Dumur, K. H. Jones, C. Jackson-Cook, Z. Su, D. Chen, U. H. Saxena, U. Hansen, P. B. Fisher, and D. Sarkar, *Proc. Natl. Acad. Sci. USA* **107**, 8357 (2010)
9. T. J. Grant, J. A. Bishop, L. M. Christadore, G. Barot, H. G. Chin, S. Woodson, J. Kavouris, A. Siddiq, R. Gredler, X. -N. Shen, J. Sherman, T. Meehan, K. Fitzgerald, S. Pradhan, L. A. Briggs, W. H. Andrews, D. Sarkar, S. E. Schaus, and U. Hansen, *Proc. Natl. Acad. Sci. USA* **109**, 4503 (2012)
10. K. -S. Jo, D. -W. Sim, E. -H. Kim, D. -H. Kang, Y. -B. Ma, J. -H. Kim, and H. -S. Won, *J. Kor. Magn. Reson. Soc.* **22**, 64 (2018)
11. Y. -S. Lee, J. Lee, K. -S. Ryu, Y. Lee, T. -G. Jung, J. -H. Jang, D. -W. Sim, E. -H. Kim, M. -D. Seo, K.W. Lee, and H. -S. Won, *J. Mol. Biol.* **427**, 3850 (2015)
12. S. N. Loh, *Metallomics* **2**, 442 (2010)