Biochemical and Biophysical Research Communications 443 (2014) 370–375

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Solution structure of CEH-37 homeodomain of the nematode *Caenorhabditis elegans*

Sunjin Moon [a], Yong Woo Lee [b], Woo Taek Kim [b], Weontae Lee [a,*]

[a] Structural Biochemistry and Molecular Biophysics Lab, Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, Seoul 120-749, Republic of Korea
[b] Department of Systems Biology, College of Life Science and Biotechnology, Yonsei University, Seoul 120-749, Republic of Korea

ARTICLE INFO

Article history:
Received 12 November 2013
Available online 19 December 2013

Keywords:
Solution structure
NMR spectroscopy
Nematode *Caenorhabditis elegans*
CEH-37
Telomere-binding protein
Homeodomain

ABSTRACT

The nematode *Caenorhabditis elegans* protein CEH-37 belongs to the paired OTD/OTX family of homeobox-containing homeodomain proteins. CEH-37 shares sequence similarity with homeodomain proteins, although it specifically binds to double-stranded *C. elegans* telomeric DNA, which is unusual to homeodomain proteins. Here, we report the solution structure of CEH-37 homeodomain and molecular interaction with double-stranded *C. elegans* telomeric DNA using nuclear magnetic resonance (NMR) spectroscopy. NMR structure shows that CEH-37 homeodomain is composed of a flexible N-terminal region and three α-helices with a helix-turn-helix (HTH) DNA binding motif. Data from size-exclusion chromatography and fluorescence spectroscopy reveal that CEH-37 homeodomain interacts strongly with double-stranded *C. elegans* telomeric DNA. NMR titration experiments identified residues responsible for specific binding to nematode double-stranded telomeric DNA. These results suggest that *C. elegans* homeodomain protein, CEH-37 could play an important role in telomere function via DNA binding.

1. Introduction

Members of the well-conserved OTD/OTX family of homeodomain (HD) proteins have been implicated in the development and patterning of the central nervous system and sensory structures in multiple species [1–3]. The nematode *Caenorhabditis elegans* CEH-37 protein belongs to the OTD/OTX family of homeobox proteins. CEH-37 is transiently expressed in AWB olfactory neurons and it is required for their development. Functional studies suggest that embryonic expression of *CEH-37* in AWB neurons is sufficient to trigger the expression of *lim-4*, which maintains its expression via auto-regulation [1]. Previous report suggests that CEH-37 functions as a telomere-binding protein in *C. elegans* and it is primarily co-localized at chromosome ends *in vivo* [4]. CEH-37 is also required for chromosome stability *in vivo* together with *mrt-2*, which is a checkpoint-protein gene. *CEH-37* encodes a protein that is 278 amino acids long and contains a homeodomain motif in the N-terminal region. Typically, the HD consists of 60 highly conserved residues and forms an N-terminal arm and a third helix, which recognizes specific DNA sequences that bear a TAAT core [5–8]. Consistent with this, CEH-37 specifically binds double-stranded *C. elegans* telomeric DNA with TTAGGC repeats, which is different from that of other eukaryotic species [4].

It has been proposed that the CEH-37 HD could play an important role as a Myb-like domain of telomere-binding protein although CEH-37 has low sequence identity with that of other OTX-related proteins. Therefore, it is of essence to know the structure and molecular interaction with telomeric DNA of CEH-37 in understating the molecular mechanism, which underlies its biological function as a transcriptional activator and telomere-binding protein. Here, we present the high-resolution solution structure of the CEH-37 HD by nuclear magnetic resonance (NMR) spectroscopy. In addition, we determined the detailed molecular interactions between CEH-37 HD and double-stranded *C. elegans* telomeric DNA by data from NMR titration, steady state XNOE, and fluorescence experiments. Our data provide strong evidence that CEH-37 HD binds specifically to the *C. elegans* double-stranded telomeric DNA although CEH-37 is a homeodomain protein [9–11].

2. Materials and methods

2.1. Cloning, protein purification, and sample preparation

The cDNA fragment encoding the *C. elegans* CEH-37 HD comprising residue from Pro38 to Pro104 was amplified by PCR. The amplified cDNA fragment was subcloned into the *Bam*HI and *Xho*I sites of the modified expression vector pET32a (Novagen)

Abbreviations: NMR, nuclear magnetic resonance; TBP, telomere-binding protein; HD, homeodomain; *C. elegans*, *Caenorhabditis elegans*; NOE, nuclear Overhauser effect; HSQC, heteronuclear single quantum coherence.

* Corresponding author. Address: Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, Republic of Korea. Fax: +82 2 363 2706.
E-mail address: wlee@spin.yonsei.ac.kr (W. Lee).

0006-291X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2013.11.133

fused with N-terminal hexahistidine affinity tags and TEV cleavage sites. The resultant plasmids were transformed into *Escherichia coli* strain BL21 (DE3). All plasmid DNAs were grown in both LB and M9 minimal media at 37 °C to an $OD_{600} \approx 0.6$. Cells were harvested by centrifugation and stored at −70 °C.

For NMR experiments, CEH-37 HD proteins were labeled with ^{15}N or $^{13}C/^{15}N$ by cultivating in M9 minimal medium containing $^{15}NH_4Cl$ (Cambridge Isotope Laboratories Inc.) and/or U-$^{13}C_6$-glucose (Cambridge Isotope Laboratories Inc.). The labeled proteins were extracted, isolated, and subjected to affinity chromatography. After affinity chromatography, we removed the TRX-His tag (pET32a) by incubation with tobacco etch virus protease for 12 h at 25 °C. The purified proteins were then subjected to size-exclusion chromatography on Superdex 75 (Amersham Biosciences) at the final buffer (10 mM HEPES, pH 7.5, 500 mM NaCl, 2 mM DTT, 0.01% NaN_3).

2.2. NMR experiments and structure determination of CEH-37 HD

All NMR experiments were recorded on Bruker 500 MHz and 800 MHz spectrometers equipped with a *z*-shielded gradient triple-resonance cryoprobe. Three-dimensional NMR data have been served for sequential assignments, which are HNCO, HNCACB, CBCA(CO)NH, HNCA, and HCCH-TOCSY experiments [12–17]. To determine the three-dimensional structure, ^{15}N-edited NOESY (τ_m = 150 ms) and ^{13}C-edited NOESY-HSQC (τ_m = 150 ms) experiments were performed [18–21]. Backbone dihedral angle restraints were derived from Hα, $^{13}C\alpha$, $^{13}C\beta$, and ^{13}CO chemical shifts with the program TALOS. All NMR data were processed using the XWINNMR program and NMRpipe/NMRDraw software. Structure calculations for CEH-37 HD were carried out using CYANA 2.2.5. A total of 100 structures were calculated, and 20 structures with the lowest target-function values were chosen for detailed analysis. A summary of NMR-derived restraints and structural statistics of CEH-37 HD is shown in Table 1. The final structures were retained and validated by the program PROCHECK [22]. Analysis with PROCHECK-NMR shows that more than 99% of residues have backbone dihedral angles in allowed regions of the Ramachandran plot, with 0.1% of residues in disallowed regions. The residues in the generously allowed and the disallowed regions are Asn41 and Arg42, which are located in the N-terminal tail region. Final structures were analyzed and visualized using the programs PyMOL and MOLMOL [23]. The electrostatic surface potential was calculated with the program APBS. NMR chemical shifts were deposited in the Biological Magnetic Resonance Bank (BMRB) under accession code 19600. The atomic coordinates of the final 20 structures and the energy-minimized average structure of CEH-37 HD were deposited in the Protein Data Bank (PDB) under accession code 2mgq. The following servers were used for sequence and structural analyses: BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), DALI (http://ekhidna.biocenter.helsink.fi/dali-server), and SCOP (http://scop.mrc-lmb.cam.ac.uk/scop).

2.3. NMR titration with telomeric DNA and XNOE experiments

NMR titration experiments were performed by recording a series of two-dimensional (2D) ^{15}N-^{1}H HSQC spectra on uniformly ^{15}N-labeled CEH-37 HD (0.5 mM) in the presence of different amounts of double stranded *C. elegans* telomeric DNA ($(TTAGGC)_2$). Samples were prepared in 10 mM HEPES at pH 7.5 containing 100 mM NaCl, 2 mM DTT, and 0.01% NaN_3. Chemical-shift perturbations were calculated using the equation $\Delta\delta_{AV} = [(\Delta\delta_{1H})^2 + (\Delta\delta_{15N}/5)^2]^{1/2}$, where $\Delta\delta_{AV}$, $\Delta\delta_{1H}$, and $\Delta\delta_{15N}$ are the average, proton, and ^{15}N chemical-shift changes, respectively. To investigate the dynamic properties of both free and DNA bound form of CEH-37 HD, the steady-state heteronuclear ^{1}H,^{15}N-NOE (XNOE) was measured [24,25]. The XNOE data were collected for $2048(t_2) \times 128(t_1)$ dimensions with a 3-s recycle delay. Both unsaturated and saturated XNOE spectra were acquired by interleaving pulse sequences, and they were separately processed for analysis. The heteronuclear XNOE values were determined by the ratio of the peak heights produced by the unsaturated versus saturated XNOE spectra, and error values were calculated using the equation $\sigma_{NOE}/NOE = [(\sigma I_{sat}/I_{sat})^2 + (\sigma I_{unsat}/I_{unsat})^2]^{1/2}$, where I_{sat} and I_{unsat} represent the cross-peak intensities, and σI_{sat} and σI_{unsat} represent the root-mean-square noise value of the background in proton-saturated and proton-unsaturated spectra, respectively.

Table 1
NMR structural statistics.

Distant restraints in the structure calculation	
All	1172
Short range {|i − j| = 1}	675
Medium range {$2 \leqslant \|i-j\| \leqslant 5$}	310
Long range {|i − j| > 5}	187
Hydrogen bond restraints[a]	22
Dihedral angle restraints	
All	66
Φ	33
Ψ	33
Residual violations	
CYANA target functions, Å	4.21 ± 0.11
NOE upper distance constrain violation	
Ave, Å (<0.1 A)	0.0191 ± 0.0005
Dihedral angle constrain violations	
Ave,° (<5°)	2.56 ± 0.05
Vander Waals violations	
Ave, Å (<0.2 A)	10.1 ± 0.4
RMS deviations from the average coordinate[b], Å	
Backbone atoms	0.31 ± 0.11
All heavy atoms	1.13 ± 0.12
Ramachandran statistics, % of all residues	
Residues in most favored regions	72.3%
Residues in additional allowed regions	26.5%
Residues in generously allowed regions	1.1%
Residues in disallowed allowed regions	0.1%

[a] Two restraints per one hydrogen bond.
[b] RMSD values for residues 51–102 a.a.

2.4. Fluorescence experiments

Double-stranded oligonucleotides (5′-TTAGGCTTAGGC-3′ and 5′-GCCTAAGCCTAA-3′) were synthesized and purified using polyacrylamide gel electrophoresis (PAGE) (Operon). Binding affinity between *C. elegans* telomeric DNA and CEH-37 HD were measured on a model RF-5301PC spectro-fluorophotometer (Shimadzu, Kyoto, Japan). CEH-37 HD in NMR buffer (10 mM HEPES,100 mM NaCl, 2 mM DTT, and 0.01% NaN_3 at pH 7.5) was titrated to a molar ratio of 1:1 (CEH-37 HD: *C. elegans* telomeric DNA) using a 2 ml thermostat cuvette. Samples were excited at 280 nm, and emission spectra were recorded for light scattering effects from 270 to 500 nm [26]. We calculated the K_d value using the equation $\log(F_0 - F/F) = \log(1/K_d) + n\log[\text{ligand}]$, where F_0 and F represent fluorescence intensity of the protein at 344 nm in the absence and presence of DNA, respectively. The number n represents the ligand binding site of the protein.

3. Results and discussion

3.1. Solution structure of CEH-37 homeodomain

All backbone and side-chain resonances were completely assigned using data from HNCA, CBCACONH, HNCACB, HCCH-TOCSY, and ^{15}N-edited NOESY experiments. The 2D [^{1}H-^{15}N] HSQC