Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray analysis of human Brn-5 transcription factor in complex with DNA

Jose Henrique Pereira,[a] Sung Chul Ha[a] and Sung-Hou Kim[a,b]*

[a]Department of Chemistry, University of California, Berkeley, California 94720, USA, and [b]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

Correspondence e-mail: shkim@cchem.berkeley.edu

Received 24 December 2007
Accepted 30 January 2008

The Brn-5 protein plays an important role in the control of cellular development and belongs to a class of transcription factors that usually contain two domains: the POU homeodomain (POU_{HD}) and the POU-specific domain (POU_S). Since high-quality crystals suitable for crystallographic studies of the proteins of this class are difficult to obtain, all the known structural information available is for POU_{HD} and/or POU_S. This paper describes several critical steps that allowed the production of high-quality crystals of the full-length Brn-5 protein complexed with its cognate DNA.

1. Introduction

The structural characterization of three mammalian transcriptional regulators, **P**it-1, **O**ct-1 and Oct-2, and a *Caenorhabditis elegans* developmental modulator, ***u**nc-86*, led to the identification of a DNA-binding motif referred to as the **POU** domain. The POU-family proteins play a vital role in controlling cell-fate determination and the timing of cellular events in a number of tissues (Wu *et al.*, 2001). Several members of the POU family have been identified as being transcription factors involved in various aspects of transcriptional regulation (Rosenfeld, 1991; Ruvkun & Finney, 1991). The proteins in this family interact with DNA *via* a two-part binding domain, consisting of an amino-terminal conserved 75–82-amino-acid POU-specific (POU_S) domain, separated by a flexible linker (variable both in length, 15–55 amino acids, and sequence) from a 60-amino-acid POU homeodomain (POU_{HD}) (Herr & Cleary, 1995; Sturm & Herr, 1988). In the brain-5 protein (Brn-5), POU_S is composed of 75 amino acids, the linker is 21 amino acids in length and POU_{HD} is 60 amino acids in length. Both DNA-binding domains (POU_S and POU_{HD}) contain a helix–turn–helix (HTH) motif, a structure that is common to a broad class of DNA-binding proteins (Pabo & Sauer, 1992). POU proteins bind to a DNA octamer with sequence 5′-ATGCAAAT-3′. The octamer element can be described as a bipartite site, with POU_S contacting the ATGC sequence and POU_{HD} contacting the AAAT sequence (Laughon, 1991; Verrijzer *et al.*, 1992).

Brn-5 is widely expressed, with the highest levels being found in the developing human brain and spinal cord in the early days of embryogenesis. In the adult, Brn-5 is most abundant in the brain; however, it is also found in multiple tissues outside the central nervous system, including kidney, lung, heart, adrenal skin, testis and anterior pituitary (Andersen *et al.*, 1993). Brn-5 appears to enhance prolactin (PRL) gene expression directly and indirectly *via* the activation of pituitary-specific positive transcription factor 1 (Pit-1) gene expression (Toda *et al.*, 2008). The Brn-5 transcription factor binds with high affinity to a corticotrophin-releasing hormone gene-promoter sequence (CRH element) 5′-GCATAAATAAT-3′ (Andersen *et al.*, 1993; Gruber *et al.*, 1997).

The crystallization of protein–DNA complexes is generally challenging, especially when the protein consists of multiple domains and the best DNA sequence for forming the complex crystal is not known (Tan *et al.*, 2000). Thus, all the structural information available to date for this family is for the POU_{HD} and/or POU_S DNA-binding regions (PDB codes 1hdp, 1pog, 1pou, 1cqt, 1ocp, 1oct, 1gt0, 1hf0, 1e3o, 2cpj,

Table 1
Data-collection statistics of a crystal of Brn-5 in complex with CHR-14AT.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 100.23, b = 112.11, c = 181.38
Resolution (Å)	50–2.8 (2.9–2.8)
Total no. of observations	188408 (12654)
No. of unique observations	46442 (4218)
Multiplicity	4.1 (3.0)
Data completeness (%)	90.6 (84.2)
$\langle I/\sigma(I)\rangle$	14.4 (1.8)
$R_{\rm merge}$† (%)	0.089 (0.52)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\sum_{hkl}$ denotes the sum over all reflections and $\sum_i$ is the sum over all equivalent and symmetry-related reflections.

1au7, 1ic8, 1o4x and 2h8r). Here, we present the first example of the crystallization of full-length Brn-5 complexed with its cognate DNA fragment.

2. Materials and methods

2.1. Cloning, expression and purification

The primers (IDTDNA, Coralville, Iowa, USA) for PCR amplification from genomic DNA contained an *Nde*I restriction site in the forward primer (5′-CATATGATGCCAGGGATCAGCCAGTC-3′) and a *Bam*HI site in the reverse primer (5′-GGATCCTAGAGGGATCTGAAAGACGTTCAGC-3′). PCR was performed using Deep Vent Polymerase (New England Biolabs, Inc., Beverly, Massachusetts, USA) and human genomic DNA. The PCR product was cloned into pCR-BluntII-TOPO vector (Invitrogen Corp., Carlsbad, California, USA) and the Brn-5 gene insert was confirmed by DNA sequencing. The amplified TOPO vector was restricted with *Nde*I and *Bam*HI and the gene insert was purified by agarose gel electrophoresis extraction. This insert was ligated into pSKB3 vector (a gift from Steve Burley, Rockefeller University, New York) digested with *Nde*I and *Bam*HI and transformed into DH5α. A plasmid containing the gene insert was confirmed and then transformed into BL21 (DE3)/pSJS1240 (Kim *et al.*, 1998).

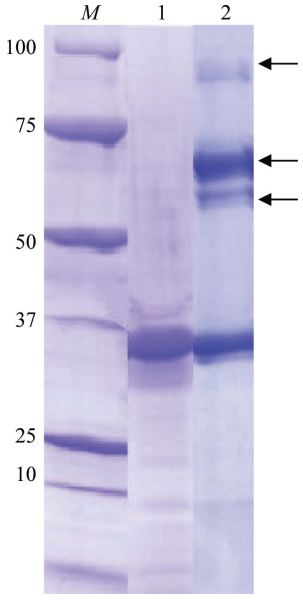


Figure 1
A 4–20% nonreducing SDS–PAGE of purified Brn5. Lane *M*, molecular-weight markers (kDa). Lane 1, sample of Brn-5 after the last purification step. Only a single band was observed prior to crystallization trials. Lane 2, sample of the Brn-5–DNA complex from crystallization drops after two weeks. The higher molecular-weight bands indicated by arrows suggest protein aggregates.

The Brn-5 transcription factor was expressed with a six-histidine tag at the N-terminus in *Escherichia coli* strain BL21 (DE3)/pSJS1240 cells upon auto-induction in ZYM media (Studier *et al.*, 1990). The bacteria were lysed using a Microfluidizer processor (Microfluidics, Newton, Massachusetts, USA) in buffer *A* [50 m*M* HEPES pH 7, 0.1 *M* NaCl, 1 m*M* PMSF, 10 µg ml^{-1} DNAse, one Roche Protease Inhibitor Cocktail Tablet (EDTA-free)] and cell debris was pelleted by centrifugation at 21 000*g* for 20 min at 277 K in a Sorvall centrifuge. The lysate was then spun in a Beckman ultracentrifuge with a Ti45 rotor at 90 000*g* for 30 min at 277 K to remove membrane proteins. The Brn-5 transcription factor was affinity-purified from the soluble fraction using a 10 ml HiTrap Chelating HP column (GE Healthcare, Piscataway, New Jersey, USA); elution was achieved with a linear gradient from 30 m*M* to 1 *M* imidazole in 20 column volumes. The eluted sample was dialyzed against 50 m*M* HEPES pH 7, 0.1 *M* NaCl to remove imidazole. The His tag was cleaved with mTEV (Parks *et al.*, 1994) for 2 h at room temperature. The target protein was finally purified by ion-exchange chromatography using a 5 ml Hi-Trap S-Sepharose column (GE Healthcare, Piscataway, New Jersey, USA) in buffer containing 50 m*M* HEPES pH 7.0, 0.1 *M* NaCl using a linear salt gradient from 0.1 to 1 *M* NaCl in 20 column volumes. The final purity of the sample was determined by SDS–PAGE, the monodispersity was measured by dynamic light scattering (DynaPro 99; Wyatt Technologies, Santa Barbara, California, USA) and the molecular weight was confirmed by MALDI–TOF mass spectrometry (Voyager DE; Applied Biosystems, Foster City, California, USA).

2.2. Crystallization of Brn-5 in complex with DNA

The protein was concentrated to 40 mg ml^{-1} in 50 m*M* HEPES pH 7, 0.1 *M* NaCl using a Centricon YM10 unit (Millipore, Bedford, Massachusetts, USA). Oligonucleotides were synthesized on the 1 µmol scale and purified to remove small-molecule impurities (IDTDNA, Coralville, Iowa, USA). The oligonucleotides were resuspended in 20 m*M* Tris–HCl pH 7.5, 10 m*M* $MgCl_2$. Oligonucleotide pairs were annealed together in equimolar ratios by heating to 363 K for 10 min and gradual cooling to room temperature. The protein–DNA complexes were formed by adding the oligonucleotides to the protein solution in a 1:1.5 (protein:DNA) molar ratio. The final concentration of Brn-5 used for crystallization trials was 20 mg ml^{-1}. The Brn-5–DNA complex was screened using the sparse-matrix method (Jancarik & Kim, 1991) with a Hydra-Plus-One Robot (Apogent Technologies, Sunnyvale, California, USA) and a Phoenix Robot (Art Robbins Instruments, Sunnyvale, California, USA) using the following crystallization screens: Crystal Screens I and II, PEG/Ion, Lite and Index Screens (Hampton Research, Aliso Viejo, California, USA) and Sigma Complex Screen (Sigma-Aldrich, St Louis, Missouri, USA).

2.3. X-ray data collection

Crystals were placed in a cryosolution containing 20%(*v*/*v*) glycerol in reservoir solution and then flash-frozen in liquid nitrogen. The data sets for the Brn-5–DNA complex were collected at the Macromolecular Crystallography Facility beamline 5.0.2 of the Advanced Light Source at Lawrence Berkeley National Laboratory (LBNL). The diffraction data were recorded using an ADSC-Q315 detector. The best native data set was collected using a 95° oscillation

range, $\Delta\varphi = 1^\circ$ and a wavelength of 0.979 Å. Processing and scaling were accomplished using *HKL*-2000 (Otwinowski & Minor, 1997). The statistics of the crystallographic data are summarized in Table 1.

3. Results and discussion

The Brn-5 transcription factor used in this work is 303 residues in length with a molecular weight of 32 839.6 Da. The recombinant Brn-5 was expressed with an mTEV-cleavable N-terminal sequence MGSSHHHHHHDYDIPTTENLYQGH for affinity purification. The sequence ENLYFQG is cleaved by mTEV protease between Q and G with high specificity, therefore the residues GH remained at the N-terminus of the protein. The protein was subjected to immobilized metal-affinity chromatography, mTEV cleavage and ion-exchange (S-Sepharose) purification. The purified Brn-5 sample was analyzed on an SDS–PAGE gel (Fig. 1, lane 1) and dynamic light-scattering (DLS) experiments showed a low polydispersity value of 0.14. Values below 0.3 are correlated with crystallization success (Ferre-D'Amare & Burley, 1997; Reményi *et al.*, 2001).

The initial crystallization trials of Brn-5–DNA were performed using several lengths of DNA (from 11 to 22 bp) containing the CRH element in the sequence (Fig. 2). These experiments did not yield any crystals. Crystallization drops were analyzed in nonreducing SDS–PAGE two weeks after setup (the sample was prepared in sample buffer without β-mercaptoethanol and was not boiled prior to running on the gel). The nonreducing SDS–PAGE gel showed several high-molecular-weight bands (Fig. 1, lane 2), in contrast to the results obtained with the fresh sample after purification (Fig. 1, lane 1). The higher molecular-weight bands suggest that non-native disulfide bonds are formed by the oxidation of cysteines. Similar results were observed with the POU protein Oct-1 (Reményi *et al.*, 2001). The three cysteines in Brn-5 (Cys119, Cys186 and Cys283) were replaced by serines using site-directed mutagenesis and the redox-insensitive triple mutant of Brn-5 showed the same capacity for binding DNA in an electrophoresis mobility-shift assay (EMSA) as the wild-type protein (data not shown). Expression, purification and preparation of the redox-insensitive Brn-5 are the same as outlined for the native protein.

The redox-insensitive Brn-5 was complexed with DNA sequences (11–22 bp). Small crystals were obtained of the complex between Brn-5 and the 14 bp CRH-14 DNA; however, these crystals showed poor diffraction quality. Optimization of the CRH-14 oligonucleotide was made by varying the ends and the sequence and two new DNA sequences, CRH-14OH (insertion of one overhang into the original sequence) and CRH-14AT (every nucleotide modified to A or T except for the CRH element), produced large crystals (Figs. 3*a* and 3*b*). The optimum crystallization conditions for the Brn-5–CRH-14OH and Brn-5–CRH-14AT complexes were found with reservoir solutions consisting of 0.1 *M* $MgCl_2$, 0.1 *M* sodium cacodylate pH 6.0, 15% polyethylene glycol 4000 and of 0.1 *M* potassium/sodium tartrate tetrahydrate, 0.05 *M* magnesium chloride, 20% polyethylene glycol 3350, respectively. Under these conditions, crystals were produced in droplets composed of 0.5 µl each of protein and reservoir solution that were equilibrated against 50 µl reservoir solution in the sitting-drop configuration in a period of two weeks. The sequences of the oligonucleotides used to crystallize Brn-5 are shown in Fig. 2.

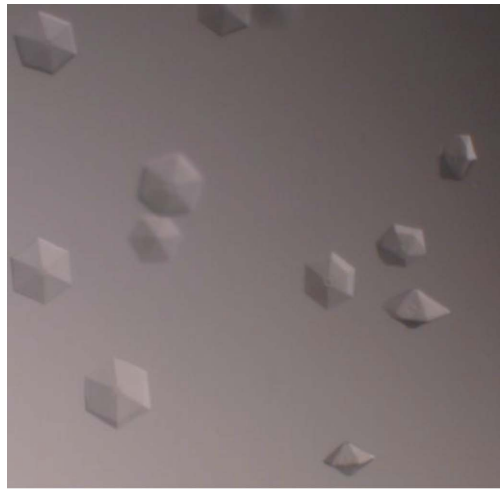
The best X-ray diffraction quality was observed for the Brn-5–CRH-14AT complex. The crystals diffracted to 2.8 Å resolution using synchrotron radiation at LBNL (Fig. 4). The Brn5–CRH-14AT crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit-

5′-CGCATAAATAATGA-3′
3′-GCGTATTTATTACT-5′
(*a*)

5′-ACGCATAAATAATG-3′
3′-GCGTATTTATTACT-5′
(*b*)

5′-AGCATAAATAATAA-3′
3′-TCGTATTTATTATT-5′
(*c*)

Figure 2
DNA sequences used to crystallize the transcription factor Brn-5 in complex with DNA. (*a*) CRH-14. Crystallization of this first DNA sequence with Brn-5 yielded very small crystals that diffracted poorly. (*b*) CRH-14OH and (*c*) CRH-14AT. Optimization of the DNA sequence produced large crystals that were suitable for X-ray analysis.



(*a*)

(*b*)

Figure 3
Crystals of Brn-5 in complex with CRH elements. (*a*) Brn-5–CRH-14OH complex crystals. Even though these crystals were of good size (150 × 150 × 100 µm) and were well shaped, the diffraction resolution was only 6 Å. (*b*) Brn-5–CRH-14AT complex crystals. After exhaustive crystallization trials with changes in the sequence and length of the oligonucleotide, the crystals of Brn-5–CRH-14AT (250 × 40 × 40 µm) diffracted to 2.8 Å resolution.

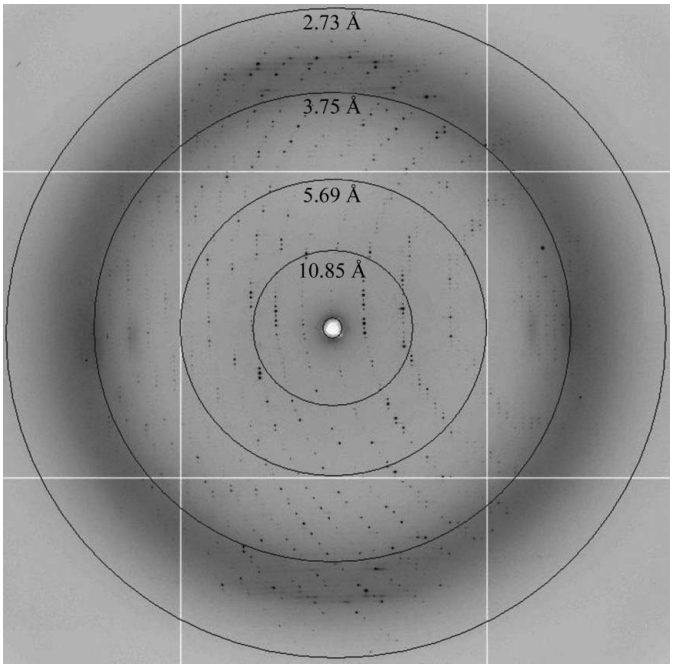


Figure 4
Diffraction pattern from a crystal of the Brn-5–CRH14-AT complex obtained using synchrotron radiation at the Advanced Light Source, Lawrence Berkeley National Laboratory (BL-5.0.2).

cell parameters $a = 100.23$, $b = 112.11$, $c = 181.38$ Å, $\alpha = \beta = \gamma = 90^\circ$. The Matthews coefficient ($V_M = 3.20$ Å^3 Da^{-1}) suggested the presence of four Brn-5–CRH-14AT complexes per asymmetric unit and a solvent content of 64%. The data-collection statistics are given in Table 1. The absorbance ratio ($A_{280\,nm}/A_{260\,nm}$) observed for Brn-5 (without DNA) and dissolved crystals of Brn-5–CRH-14AT were 1.68 and 0.67, respectively. The crystals of Brn-5–CRH-14AT were washed three times in the reservoir solution prior to measuring the spectra in order to completely remove any nonspecifically associated nucleic acids. This result clearly shows that the crystals are composed of the complex between Brn-5 and DNA.

Despite the formation of many protein–DNA complexes by POU-family members, only crystals of the DNA-binding domains have been obtained, possibly owing to the presence of unstructured regions that could affect the crystallization contacts. We were able to crystallize Brn-5 using the full-length protein. Site-specific mutations and a particular DNA sequence were necessary in order to avoid the aggregation of Brn-5 and to lead to successful crystallization. Crystals of the selenomethionine-derived protein have been prepared in order to collect data to solve the structure by anomalous dispersion. The structure of Brn-5 in complex with DNA will provide crucial information about protein–DNA interactions and will consequently provide a better understanding of the mechanisms involved in transcription regulation.

We are grateful to Drs Debanu Das and Rosalind Kim for their advice and help and to the staff at the Advanced Light Source, Lawrence Berkeley National Laboratory, which is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division of the US Department of Energy under Contract No. DE-AC03-76SF00098. This work was supported by grants from the National Research Council of Brazil (CNPq; 200969/2005-6), the US National Institutes of Health (GM62412) and the Korea Science and Engineering Foundation, Ministry of Science and Technology, Korea (M1064102000106N410200110).

References

Andersen, B., Schonemann, M. D., Pearse, R. V. II, Jenne, K., Sugarman, J. & Rosenfeld, M. (1993). *J. Biol. Chem.* **268**, 23390–23398.
Ferre-D'Amare, A. R. & Burley, S. K. (1997). *Methods Enzymol.* **276**, 157–166.
Gruber, C. A., Rhee, J. M., Gleiberman, A. & Turner, E. E. (1997). *Mol. Cell. Biol.* **17**, 2391–2400.
Herr, W. & Cleary, M. A. (1995). *Genes Dev.* **9**, 1679–1693.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S.-H. (1998). *Biotechnol. Lett.* **20**, 207–210.
Laughon, A. (1991). *Biochemistry*, **30**, 11357–11367.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pabo, C. O. & Sauer, R. T. (1992). *Annu. Rev. Biochem.* **61**, 1053–1095.
Parks, T. D., Leuther, K. K., Howard, E. D., Johnston, S. A. & Dougherty, W. G. (1994). *Anal. Biochem.* **216**, 413–417.
Reményi, A., Pohl, E., Schöler, H. & Wilmanns, M. (2001). *Acta Cryst.* D**57**, 1634–1638.
Rosenfeld, M. G. (1991). *Genes Dev.* **5**, 897–907.
Ruvkun, G. & Finney, M. (1991). *Cell*, **64**, 475–478.
Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.
Sturm, R. A. & Herr, W. (1988). *Nature (London)*, **336**, 601–604.
Tan, S., Hunziker, Y., Pellegrini, L. & Richmond, J. (2000). *J. Mol. Biol.* **297**, 947–959.
Toda, K., Yamamoto, D., Fumoto, M., Ikeshita, N., Herningtyas, E. H., Iida, K., Takahashi, Y., Kaji, H., Chihara, K. & Okimura, Y. (2008). *Mol. Cell. Endocrinol.* **280**, 20–29.
Verrijzer, C. P., Alkema, M. J., van Weperen, W. W., Van Leeuwen, H. C., Strating, M. J. & van der Vliet, P. C. (1992). *EMBO J.* **11**, 4993–5003.
Wu, R., Jurek, M., Sundarababu, S. & Weinstein, D. E. (2001). *Mol. Cell. Neurosci.* **17**, 683–695.