Crystallization and preliminary X-ray analysis of the cardiac transcription factor complex of NKX2.5 and TBX5 with DNA

Lagnajeet Pradhan, Sunil Gopal
and Hyun-Joo Nam*

Department of Bioengineering, University of
Texas at Dallas, Richardson, TX 75080, USA

Correspondence e-mail: hnam@utdallas.edu

Received 18 January 2014

Accepted 26 March 2014

Heart development depends on timely expression of genes regulated by combinatorial interactions of master cardiac transcription factors. To elucidate the molecular basis of their interactions, a ternary complex of cardiac transcription factors, NKX2.5 and TBX5, and their target DNA was studied using X-ray crystallography. Here, the purification, crystallization and preliminary X-ray crystallographic analyses of the NKX2.5 homeodomain and TBX5 DNA-binding domain complex with a DNA element from the -252 promoter region of the atrial natriuretic factor are reported. The crystal diffracted to $2.88\text{ \AA}$ resolution and belonged to space group $P2_1$, with unit-cell parameters $a = 69.30$, $b = 77.78$, $c = 77.60\text{ \AA}$, $\beta = 108.31^\circ$. Two sets of ternary complexes are present in an asymmetric unit with a solvent content of 54%.

1. Introduction

Heart morphogenesis in a mammalian system is controlled by combinatorial interactions of the core cardiac transcription factors (TFs) such as TBX5 and NKX2.5 (Olson, 2006). A handful of cardiac TFs can orchestrate the complex spacio-temporal regulation of heart development through combinatorial interactions (Durocher & Nemer, 1998). They interact cooperatively and synergistically to activate downstream targets (Bruneau *et al.*, 2001; Hiroi *et al.*, 2001). Because of their key roles in heart development, mutations in these genes are linked to congenital heart disease (CHD), the most common birth defect in humans (Hoffman, 1995). CHD is also a life-threatening progressive disorder which continues long after surgical correction of structural defects (Toko *et al.*, 2002).

NKX2.5 and TBX5 recognize and bind to the specific DNA motifs through their centrally located DNA-binding domains: the homeodomain (HD; amino-acid residues 138–197) in NKX2.5 and the T-box domain (TBD; amino-acid residues 58–239) in TBX5. The amino-terminal and carboxy-terminal domains contain regulatory elements involved in interactions with other transcription factors or transcription machinery. The binary complex structures of NKX2.5 HD in complex with DNA and TBX5 TBD with DNA have previously been reported (Pradhan *et al.*, 2012; Stirnimann *et al.*, 2010).

NKX2.5 and TBX5 co-occupy promoters/enhancers of downstream targets and physically interact with each other, but details of their interactions are not known (Hiroi *et al.*, 2001). One of their downstream targets, atrial natriuretic factor (ANF), contains closely spaced TBX5 and NKX2.5 binding motifs at the -252 site of its proximal promoter, providing a convenient platform for physical interactions between the two factors. The NKX2.5 HD and the TBX5 protein containing TBD and the N-terminal domain were shown to be necessary and sufficient for physical and functional interactions. With only DNA-binding domains of both proteins, partial activities were observed. Nonetheless, DNA-binding domains are important components for physical interactions and synergistic activation (Hiroi *et al.*, 2001). To understand the mechanisms of combinatorial interactions among cardiac transcription factors, we initiated structural studies of the TBX5 TBD and NKX2.5 HD in complex with ANF-252 DNA using X-ray crystallographic methods. Here, we describe the purification and crystallization of the ternary complex and preliminary X-ray crystallographic results.

Figure 1

Purification of the NKX2.5 HD–TBX5 TBD–ANF-252 DNA ternary complex. (a) An elution profile of TBX5 TBD and NKX2.5 HD with ANF-252 motif from a Superdex 75 column. (b) An elution profile of TBX5 TBD and NKX2.5 HD mixture without DNA. Peak 1 is for the TBX5 TBD–NKX2.5 HD–ANF-252 complex, peaks 2 and 4 are for TBX5 TBD, and peaks 3 and 5 for NKX2.5 HD. (c) A 15% SDS–PAGE analysis of eluted fractions in (a) and (b). Peak fractions as labeled in (a) and (b) are shown in lanes 1–5. In peak fraction 1, TBX5 TBD (23.8 kDa) and NKX2.5 HD (7.6 kDa) bands are marked with arrows. (d) An agarose gel analysis of peak fraction 1. The DNA marker (lane 1), ANF-252 DNA (lane 2) and peak fraction 1 (lane 3) are shown. The gel was stained with ethidium bromide.

Figure 2

Crystal images at various PEG 400 concentrations. Optical images of crystals are shown with the PEG 400 concentration as a percentage at the bottom. The approximate dimensions of the crystals from the 26% PEG 400 condition are $200 \times 80 \times 80 \mu\text{m}$.

2. Materials and methods

2.1. Protein expression and purification

Cloning of the NKX2.5 HD expression construct and purification of the protein have previously been reported (Genis *et al.*, 2008; Pradhan *et al.*, 2012). The DNA coding amino acids 58–238 of TBX5 protein was cloned between the *Nde*I and *Bam*HI sites of pET28 vector (Novagen). The expressed protein contains an N-terminal six-histidine tag and a thrombin cleavage site derived from the vector. TBX5 TBD was produced based on a previously described protocol with minor modifications (Stirnemann *et al.*, 2010). For protein production, *Escherichia coli* BL21 (RIPL) cells were transformed with the plasmid and grown at 37°C in Luria–Bertani broth. Protein expression was induced by the addition of 0.5 mM IPTG and the cells were grown overnight at 24°C. The cells were lysed by sonication for 5 min in 50 mM Tris pH 7.5, 0.5 M NaCl, 20% glycerol, 2 mM β -mercaptoethanol (β -ME), and cell debris was pelleted by centrifugation at 27 000g for 30 min. TBX5 TBD proteins in the supernatant were purified by standard affinity chromatography using a cobalt Sepharose column (New England Biolabs). The column was washed

three times in series with wash buffers A, B and C. Wash buffer A consisted of 20 mM Tris pH 7.5, 1 M NaCl, 20% glycerol, 2 mM β -ME, 15 mM imidazole; wash buffer B consisted of 20 mM Tris pH 7.5, 0.5 M NaCl, 10% glycerol, 2 mM β -ME, 15 mM imidazole; and wash buffer C consisted of 20 mM Tris pH 7.5, 0.25 M NaCl, 5% glycerol, 2 mM β -ME, 15 mM imidazole. The protein was eluted by adding 20 mM Tris pH 7.5, 0.25 M NaCl, 5% glycerol, 2 mM β -ME, 250 mM imidazole and was concentrated using Amicon Ultra-15 (Millipore) with a molecular-weight cutoff value of 10 kDa. It was further purified by size-exclusion chromatography using a Superdex 75 column (GE Healthcare) with a buffer consisting of 20 mM HEPES pH 7.5, 0.15 M NaCl, 5 mM MgCl_2 , 2 mM DTT.

Oligonucleotides for co-crystallization experiments were synthesized and purified using reverse-phase chromatography (Sigma Genosys). The DNA was annealed in 10 mM Tris pH 8.0, 100 mM NaCl, 10 mM MgCl_2 first by heating the solution to 95°C for 10 min and cooling slowly to room temperature in a water bath.

For protein–DNA complex purification, the double-stranded DNA and the purified proteins TBX5 TBD and NKX2.5 HD were mixed roughly in a molar ratio of 1:1.5:2. The protein–DNA complex was

purified by size-exclusion chromatography using buffer consisting of 20 mM Tris pH 8.0, 0.2 M NaCl, 5 mM MgCl₂, 2 mM DTT and a Superdex 75 column. The purified ternary complex was analyzed by SDS-PAGE, concentrated to a final protein concentration of ~10 mg ml⁻¹ and stored at -80°C.

2.2. Crystallization of the ternary complex

Crystallization conditions for the ternary complex of NKX2.5 HD, TBX5 TBD and DNA were screened by the sparse-matrix method using crystallization screens such as NeXtal Nucleix Suite (Qiagen), Crystal Screen, Crystal Screen 2 and Natrix (Hampton Research). The hanging-drop vapor-diffusion method was used with crystallization drops consisting of 1 µl of protein/DNA sample solution and 1 µl of crystallization solution at room temperature. To keep the solutions under reducing conditions, 10 mM DTT was added to the reservoir solutions.

2.3. X-ray data collection, processing and preliminary structure determination

For data collection, crystals were cryoprotected using 50 mM HEPES pH 7.0, 100 mM KCl, 10 mM CaCl₂, 30% polyethylene glycol (PEG) 400, 10 mM DTT and flash-cooled in liquid nitrogen. The X-ray diffraction data were collected on the A1 beamline of the Cornell High Energy Synchrotron Source (CHESS) using an ADSC Quantum 210 CCD detector. For data collection, a crystal-to-detector distance of 200 mm, an oscillation angle of 1° and an exposure time of 20–30 s per image were used. The data were indexed and processed using the HKL-2000 suite (Otwinowski & Minor, 1997), and additional analyses were performed using programs from the PHENIX suite (Adams *et al.*, 2010).

A molecular-replacement search was performed using the NKX2.5 HD (PDB entry 3rkq, Pradhan *et al.*, 2012) and TBX5 TBD (PDB entry 2x6v, Stirnimann *et al.*, 2010) as search models. The program

Table 1

Crystal data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	0.977
Space group	<i>P</i> 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 69.30, <i>b</i> = 77.78, <i>c</i> = 77.60, <i>β</i> = 108.31
Resolution range (Å)	47.06–2.88 (2.99–2.88)
<i>V</i> _M (Å ³ Da ⁻¹)	2.69
No. of unique reflections	17571 (966)
Completeness (%)	98.8 (89.0)
<i>R</i> _{merge} †	0.083 (0.494)
<i>R</i> _{meas} ‡	0.099 (0.585)
Multiplicity	3.5 (3.6)
<i>I</i> /σ(<i>I</i>)	11.4 (2.3)

† $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is a single value of the measured intensity of the hkl reflection, $\sum_i$ is the sum over all i measurements of the hkl reflection and $\langle I(hkl) \rangle$ is the mean of all measured values of the intensity of the hkl reflections. ‡ $R_{\text{meas}} = \sum_{hkl} [N(hkl) / [N(hkl) - 1]]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is a single value of the measured intensity of the hkl reflection, $\sum_i$ is the sum over all i measurements of the hkl reflection, $\langle I(hkl) \rangle$ is the mean of all measured values of the intensity of the hkl reflection and $N(hkl)$ is the multiplicity of the reflection hkl .

Phaser (McCoy *et al.*, 2007) was used for the search and the structure was refined using the PHENIX suite.

3. Results and discussion

The NKX2.5 HD contains an additional Gly residue at the N-terminus after cleaving off the MBP fusion protein. The TBX5 TBD was expressed with an additional 20 amino-acid residues (MGSSHHHHHHSSGLVPAGSH) at the N-terminus of the protein which comprise the affinity-tag and thrombin cleavage sequence. NKX2.5 HD and TBX5 TBD with respective molecular masses of 7.6 and 23.8 kDa were purified to over 90% homogeneity as judged by SDS-PAGE (data not shown). The purified proteins were then mixed with the double-stranded DNA at a ratio of two NKX2.5 HDs and 1.5 TBX5 TBDs per DNA-binding site to ensure full occupation of the DNA-binding sites with the proteins. The ternary complex of NKX2.5 HD, TBX5 TBD and DNA was purified by size-exclusion chromatography. The chromatography profile showed a distinct high-molecular-mass peak observed only in the presence of ANF-252 DNA (Figs. 1a and 1b). This peak represents the ternary complex as determined by SDS-PAGE and DNA gel electrophoresis (Figs. 1c and 1d). The DNA gel analysis showed electrophoretic mobility shift by the proteins (Fig. 1d).

Several sets of DNA oligonucleotides containing the ANF-252 sequence were used. To help crystallization, DNA double strands with various lengths (19–21) and overhangs were used. Crystals of the ternary complexes were obtained from the initial crystallization screening condition consisting of 50 mM HEPES pH 7.0, 100 mM KCl, 10 mM CaCl₂, 10% PEG 400, 10 mM DTT, but showed poor diffraction quality and high mosaicity. After optimization of crystallization conditions, the best quality crystals were obtained with blunt-ended 19-mer ANF-252 DNA duplex (TCTCACACCTTTGAA-GTGG/CCACTTCAAAGGTGTGAGA). Fig. 2 shows a series of crystal images with varying concentrations of the precipitant PEG 400. The reservoir solution of the optimal crystallization condition consisted of 50 mM HEPES pH 7.0, 100 mM KCl, 10 mM CaCl₂, 26% PEG 400, 10 mM DTT. Crystallization drops consisted of 1 µl sample solution and 1 µl reservoir solution. The drops were equilibrated against 500 µl reservoir solutions using the hanging-drop vapor-diffusion method at room temperature.

Figure 3
A diffraction image from a ternary complex crystal of NKX2.5 HD-TBX5 TBD-ANF-252 DNA. The concentric circles indicate resolution ranges of the diffraction and the inset shows a magnified view of the boxed area. Diffraction was observed beyond 3.0 Å resolution.

The data for the ternary complex were collected on the A1 beamline of CHESS from a thin plate-shaped crystal which diffracted to 2.88 Å resolution (Fig. 3). The data were indexed and scaled in space group $P2_1$, with unit-cell parameters $a = 69.30$, $b = 77.78$, $c = 77.60$ Å, $\beta = 108.31^\circ$. Systemic absences at $h = k = 0$, $l = 2n + 1$ indices were evident. The detailed data statistics are presented in Table 1. Considering the unit-cell dimensions and the molecular mass of the NKX2.5 HD–TBX5 TBD–DNA complex, two sets of ternary complexes were expected in the asymmetric unit, with a solvent content of 54% ($V_M = 2.69$ Å³ Da^{−1}; Matthews, 1968).

The phases for the data were determined by molecular replacement using NKX2.5 HD–DNA (Pradhan *et al.*, 2012) and TBX5 TBD–DNA (Stirnemann *et al.*, 2010) complex structures as search models and the program *Phaser*. Unambiguous solutions were obtained by searching the two TBX5 TBD solutions first ($Z = 14.2$ and 12.2, respectively) and then positioning NKX2.5 HD molecules ($Z = 34.8$ and 31.0, respectively). When the models were positioned in a unit cell, the DNA containing binding sites for both NKX2.5 and TBX5 formed a continuous helix, and subsequent iterative rounds of structural refinement using *PHENIX* showed ordered density for all base pairs. Crystallographic refinement for this structure is in progress.

A high-resolution crystal structure of the NKX2.5 HD–TBX5 TBD–DNA ternary complex will provide a detailed picture of cardiac transcription factors in a combinatorial interaction state. This will provide insight into one of the major regulatory mechanisms of heart development processes and malfunctions associated with pathogenic NKX2.5 and TBX5 mutations.

We are grateful to the staff at the A1 beamline of the Cornell High Energy Synchrotron Source for assistance during X-ray data collection. Use of the Cornell High Energy Synchrotron Source was supported by NSF award DMR-0936384 and NIH/NIGMS award GM103485. This work is supported by a grant to H-JN from the American Heart Association (13BGIA13960001).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
 Bruneau, B. G., Nemer, G., Schmitt, J. P., Charron, F., Robitaille, L., Caron, S., Conner, D. A., Gessler, M., Nemer, M., Seidman, C. E. & Seidman, J. G. (2001). *Cell*, **106**, 709–721.
 Durocher, D. & Nemer, M. (1998). *Dev. Genet.* **22**, 250–262.
 Genis, C., Scone, P., Kasahara, H. & Nam, H.-J. (2008). *Acta Cryst.* **F64**, 1079–1082.
 Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R. & Komuro, I. (2001). *Nature Genet.* **28**, 276–280.
 Hoffman, J. I. (1995). *Pediatr. Cardiol.* **16**, 155–165.
 Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
 Olson, E. N. (2006). *Science*, **313**, 1922–1927.
 Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
 Pradhan, L., Genis, C., Scone, P., Weinberg, E. O., Kasahara, H. & Nam, H.-J. (2012). *Biochemistry*, **51**, 6312–6319.
 Stirnemann, C. U., Ptchelkine, D., Grimm, C. & Müller, C. W. (2010). *J. Mol. Biol.* **400**, 71–81.
 Toko, H. *et al.* (2002). *J. Biol. Chem.* **277**, 24735–24743.