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

Crystallization and preliminary X-ray analysis of the NKX2.5 homeodomain in complex with DNA

Caroli Genis,[a] Peyton Scone,[a] Hideko Kasahara[b] and Hyun-Joo Nam[a]*

[a]Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32610, USA, and [b]Department of Physiology and Functional Genomics, University of Florida, Gainesville, FL 32610, USA

Correspondence e-mail: hnam@mbi.ufl.edu

Received 4 August 2008
Accepted 14 October 2008

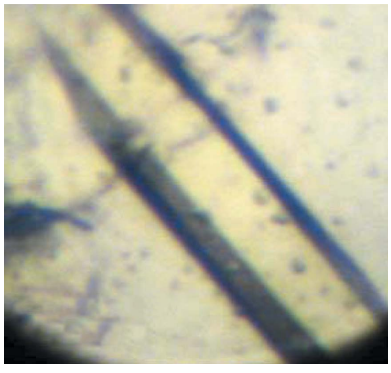
As part of an effort to elucidate the molecular basis for the pathogenesis of NKX2.5 mutations in congenital heart disease using X-ray crystallography, the NKX2.5 homeodomain has been crystallized in complex with a specific DNA element, the −242 promoter region of atrial natriuretic factor. Crystals of the homeodomain–DNA complex diffracted X-rays to 1.7 Å resolution and belonged to space group $P6_5$, with unit-cell parameters $a = b = 71.5$, $c = 94.3$ Å. The asymmetric unit contained two molecules of the NKX2.5 homeodomain and one double-stranded oligonucleotide.

1. Introduction

Congenital heart disease is the most common birth defect in humans (Hoffman, 1995). The causes of congenital heart disease are multifactorial and both genetic and environmental factors have been implicated (Schott *et al.*, 1998; Garg *et al.*, 2003). Genetic analyses have revealed that mutations in cardiac transcription factors, including the homeoprotein NKX2.5, cause congenital heart disease (Benson *et al.*, 1999) and it has been estimated that approximately one to a few percent of CHD patients carry NKX2.5 mutations (Harvey *et al.*, 2002).

The NKX2.5 protein belongs to the NK family of homeoproteins, which play roles in the cell-type specification of many different tissues. NKX2.5 is preferentially expressed in cardiomyocytes and contributes to cardiac development and maturation (Akazawa & Komuro, 2005). A substantial number of NKX2.5 mutations identified in congenital heart-disease patients are located in its homeodomain (HD; amino acids 138–197), indicating its functional importance (Benson *et al.*, 1999; Goldmuntz *et al.*, 2001; Gutierrez-Roelens *et al.*, 2002; Kasahara & Benson, 2004; Schott *et al.*, 1998). The HD of NKX2.5 is highly homologous to the HDs of other NK-class protein members and binds to a consensus DNA motif, C/GAAG. One of the known downstream targets of NKX2.5 is atrial natriuretic factor (ANF) and NKX2.5 binds to the −242 region of its promoter (ANF −242; Durocher *et al.*, 1997). The ANF −242 region contains two palindromic NKX2.5-binding motifs and occupation of both binding sites by NKX2.5 has been reported (Kasahara *et al.*, 2001).

Among the various NK-class proteins, the NMR structures of *Drosophila* vnd/NK2 HD and rat NKX2.1 HD are the only three-dimensional structures that have been determined to date (Gruschus *et al.*, 1999; Esposito *et al.*, 1996); no crystal structure of an NK-family protein is currently available. Here, we present the purification, crystallization and preliminary X-ray crystallographic analysis of the NKX 2.5 HD in complex with ANF −242 DNA containing palindromic binding sites.



© 2008 International Union of Crystallography
All rights reserved

2. Materials and methods

2.1. Protein expression and purification

The DNA encoding amino acids 138–197 of NKX2.5 was cloned into the *Eco*RI and *Hin*dIII sites of the pMALc2 vector (New England Biolabs) to express the protein as a maltose-binding protein (MBP) fusion protein. A tobacco etch virus (TEV) protease-cleavage site was introduced between MBP and NKX2.5 HD by inclusion of the protease-cleavage sequence into the forward primer. An HD construct with the oxidizable Cys193 substituted by Ser (amino acids 138–197; C193S) was generated by site-directed mutagenesis (Stratagene). For protein production, *Escherichia coli* BL21 (DE3) cells were transformed with plasmids and grown at 310 K in Luria broth. Protein expression was induced by the addition of 0.3 m*M* IPTG for 3 h. The bacteria were lysed by sonication for 4 min in 20 m*M* Tris pH 7.5, 0.2 *M* NaCl, 1 m*M* EDTA, 1 m*M* DTT (column binding buffer) and cell debris was pelleted by centrifugation at 27 000*g* for 30 min. The soluble MBP-NKX2.5 HD in the supernatant was purified by standard affinity chromatography using an amylose–agarose column (New England Biolabs). After the MBP-fusion protein had bound to the amylose–agarose column, the MBP tag was removed by incubating the column in 3 ml (one column volume) of column binding buffer containing 100 µg TEV protease for ∼14 h at 277 K. NKX2.5 HD was eluted from the column using 50 ml column binding buffer. Eluted fractions were further purified by cation-exchange chromatography using a Hi-Trap SP column (GE Healthcare) with a linear salt gradient from 50 m*M* to 1 *M* NaCl at pH 8.0. Since the eluted fractions contained high salt, the samples were diluted fourfold with 20 m*M* Tris pH 8.0.

The oligonucleotides for DNA cocrystallization were synthesized and purified using reverse-phase chromatography (Sigma Genosys). The double-stranded DNA was annealed in 10 m*M* Tris pH 8.0, 100 m*M* NaCl, 10 m*M* $MgCl_2$ by heat denaturation at 368 K for 10 min followed by slow cooling to room temperature.

For protein–DNA complex purification, double-stranded DNA and NKX2.5 HD protein purified by affinity and cation-exchange chromatography were mixed in the ratio of roughly 1.5 NKX2.5 HD per DNA-binding site and concentrated using Amicon Ultra MW 10 000 (Millipore). The protein–DNA complex was purified by size-exclusion chromatography on a Superdex75 gel-filtration column (GE Healthcare) using 20 m*M* Tris pH 8.0, 0.2 *M* NaCl, 10 m*M* DTT as elution buffer. The purified complex was analyzed by SDS–PAGE using a PhastSystem (GE Healthcare), concentrated to a final protein concentration of ∼10 mg ml^{-1} and stored at 193 K.

2.2. Crystallization of the NKX2.5 HD–DNA complex

Crystallization conditions for the NKX2.5 HD–DNA complex were screened by the sparse-matrix method using Nextal Nucleix Suite (Qiagen), Crystal Screens I and II and Natrix Screen (Hampton Research) crystallization screens. The hanging-drop vapor-diffusion method was employed using crystallization drops containing 1 µl protein–DNA sample and 1 µl reservoir solution equilibrated against 500 µl reservoir solution at room temperature. 10 m*M* DTT was added to the reservoir solutions.

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

For data collection, crystals were cryoprotected using a reservoir solution containing 30% 2-methyl-2,4-pentanediol and flash-frozen in liquid nitrogen. X-ray diffraction data were collected on the 22-ID beamline of the South East Regional Collaborative Access Team (SER-CAT) facility at the Advanced Photon Source (APS), Argonne National Laboratory using a MAR 300 CCD detector. For data collection, a crystal-to-detector distance of 200–300 mm, an oscillation angle of 1° and an exposure time of 1–5 s per image were used. The data were indexed and processed using *HKL*-2000 (Otwinowski & Minor, 1997).

Twinning analysis was carried out using the programs *TRUNCATE* and *DETWIN* (Collaborative Computational Project, Number 4, 1994) and the Twinning Server (http://nihserver.mbi.ucla.edu/Twinning; Yeates, 1997). The molecular-replacement search was performed using *Phaser* (McCoy *et al.*, 2007) with the crystal structure of the MSX-1 transcription factor HD–DNA complex (PDB code 1ig7; Hovde *et al.*, 2001) as a search model.

3. Results and discussion

NKX2.5 HD was overproduced as an MBP-fusion protein. After protease cleavage, the cleaved protein contained an additional Gly residue at the N-terminus, resulting in a molecular weight of 7.6 kDa

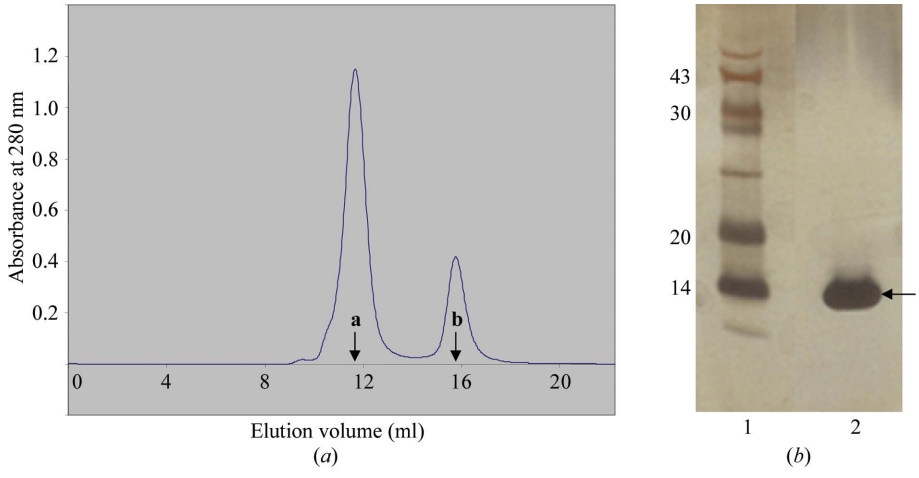


Figure 1
Purification of the NKX2.5 HD–ANF −242 DNA complex. (*a*) Elution profile of the NKX2.5 HD–ANF −242 DNA complex from a Superdex75 column. The estimated molecular weights of a globular protein at time points **a** and **b** were ∼30 and ∼14 kDa, respectively. (*a*) 20% SDS–PAGE of purified NKX2.5 HD. Lane 1 contains low-molecular-weight markers (kDa) and lane 2 contains a sample of NKX2.5 HD after the last purification step.

Table 1
Crystal data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.00
Space group	$P6_5$
Unit-cell parameters (Å)	$a = b = 71.5$, $c = 94.3$
Resolution range (Å)	40–1.7 (1.76–1.70)
V_M (Å^3 Da^{-1})	2.61
Completeness (%)	98.7 (99.5)
No. of unique reflections	29800 (2976)
R_{merge}† (%)	0.068 (0.356)
Redundancy	10.7 (9.9)
$I/\sigma(I)$	58.7 (7.9)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

for the HD. The protein was then subjected to cation-exchange purification, taking advantage of the high pI (calculated pI of 11.4; Gasteiger *et al.*, 2005); after analysis by SDS–PAGE, it was judged to be over 90% pure (data not shown).

The purified protein was then mixed with double-stranded DNA at a ratio of 1.5 NKX2.5 HD per DNA-binding site in order to ensure full occupation of the DNA-binding sites with NKX2.5 HD protein. The protein–DNA complex was purified by size-exclusion chromatography and the chromatography profile showed that the majority of the protein was present as higher molecular-weight complexes (Fig. 1*a*) with an estimated molecular weight of ∼30 kDa, corresponding to two HD molecules and one DNA duplex. Uncomplexed protein eluted later as a ∼14 kDa protein. This elution time point is earlier than that expected for NKX2.5 HD monomers (7.6 kDa), suggesting the presence of a dimeric form or an elongated shape of the uncomplexed protein. The purity of the protein after the last purification step was estimated to be over 95% based on SDS–PAGE analysis (Fig. 1*b*).

More than ten oligonucleotides of various lengths (8–24) including the ANF −242 motif were used for co-purification and crystallization. Crystals were obtained within 1–3 d, but all showed poor diffraction quality and high mosaicity. To improve the quality of the crystals several additives were added to the reservoir solution, but many reacted with the added reducing agent, thereby hindering the crystallization process. To circumvent this problem, mutant NKX.2.5 HD in which the oxidizable Cys193 was replaced by Ser [NKX2.5 HD (C193S)] was constructed. An NMR experiment showed this protein to have similar binding characteristics as the wild-type NKX2.5 HD (Fodor *et al.*, 2005). The best-quality crystals were obtained with NKX2.5 HD (C193S) and a 19-mer: a blunt-ended ANF −242 DNA duplex (TGAAGTGGGGGCCTCTTGA/TCAAGAGGCCCCCA-CTTCA; Fig. 2). The reservoir solution of the optimum crystallization condition for the NKX2.5 HD (C193S)–DNA complex consisted of 50 m*M* Tris pH 7.0, 5 m*M* $MgCl_2$, 15% polyethylene glycol monomethyl ether 550. Crystallization drops contained 1 µl sample and 1 µl reservoir solution. The drops were equilibrated against 500 µl reservoir solution using the hanging-drop vapor-diffusion method at room temperature.

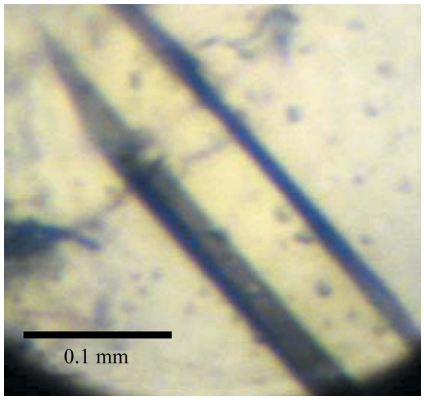


Figure 2
A hexagonal crystal of the NKX2.5 HD–ANF −242 complex. The crystal dimensions are 400 × 80 × 80 µm.

The diffraction data for the NKX2.5 HD (C193S)–DNA complex were collected on the APS 22-ID beamline at Argonne National Laboratory from a long rod-shaped crystal that diffracted to 1.7 Å resolution (Fig. 3). The data were initially indexed and scaled in two possible hexagonal space groups: $P6_522$ and $P6_122$. Considering the unit-cell parameters and the molecular weight of the NKX2.5 HD (C193S)–DNA complex, one HD domain and one half of the double-stranded DNA were expected to be present in the asymmetric unit, with a solvent content of 58% ($V_M = 2.61$ Å^3 Da^{-1}; Matthews, 1968).

After many unsuccessful attempts using various HD structures and search programs including *AMoRe* (Navaza, 1994) and *MOLREP* (Vagin & Teplyakov, 1997), the phases for the NKX2.5 HD (C193S)–ANF −242 data were determined by molecular replacement using the MSX-1 HD–DNA complex structure and the program *Phaser* (McCoy *et al.*, 2007). An unambiguous solution was obtained in space group $P6_522$ ($Z = 6.6$) but not in $P6_122$ ($Z = 4.6$). When the model was positioned in a unit cell, a pair of DNA molecules related by twofold symmetry along the *b* axis formed a continuous helix. Subsequent iterative rounds of structural refinement using the rigid-body, simulated-annealing, conjugate-gradient energy-minimization and *B*-factor refinement options of *CNS* (Brünger *et al.*, 1998) failed to reduce the crystallographic *R* factor below 35% in this space group. The data were then reprocessed in the lower symmetry space group $P6_5$ and examined for merohedral twinning. The *P*6 point group was chosen over the lower symmetry groups *P*3, *P*312 and *P*321 because clear systematic absences indicated the presence of a $6_1/6_5$ screw axis.

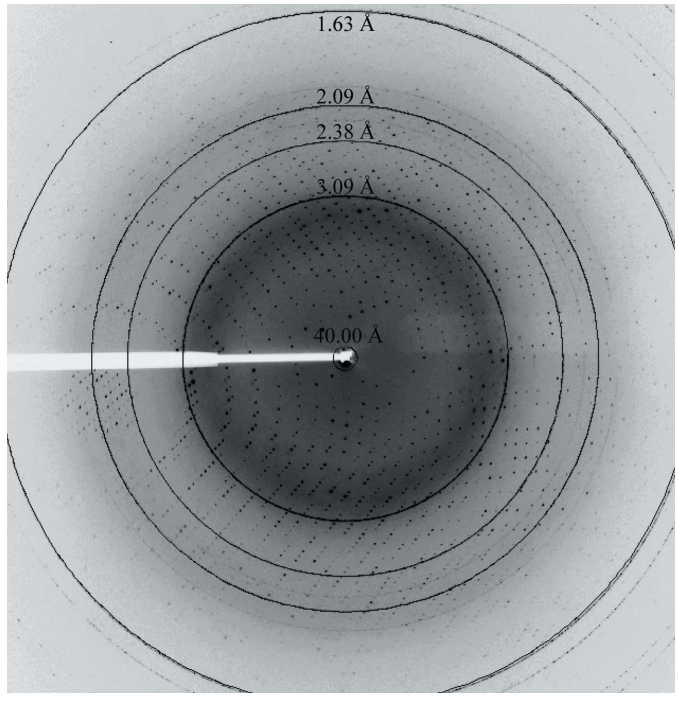


Figure 3
Diffraction image of a crystal of the NKX2.5 HD–ANF −242 DNA complex obtained using synchrotron radiation at the Advanced Photon Source, Argonne National Laboratory (beamline 22-ID).

Data-processing statistics showed similar R_{merge} values for both the $P6_5$ (6.8%) and $P6_522$ (7.0%) space groups. The detailed data statistics for space group $P6_5$ are presented in Table 1. Twinning analysis of the data (Yeates, 1997) in space group $P6_5$ showed the crystal to be merohedrally twinned with a large twinning fraction (0.47). Subsequent refinement of the NKX2.5 HD (C193S)–DNA structure using the appropriate options in the *CNS* program suite (Brünger *et al.*, 1998) yielded better refinement statistics with *R*-factor values (both R_{cryst} and R_{free}) below 28%. In particular, the maps generated using *CNS-TWIN* options in $P6_5$ showed ordered density for all base pairs. Crystallographic refinement of this structure is in progress.

A high-resolution crystal structure of NKX2.5 HD will provide a detailed picture of NKX2.5–DNA interactions. This will enable the prediction of aberrations in the DNA binding of various mutant NKX2.5 proteins that have been identified in human patients with congenital heart disease.

We are grateful to Drs Kenneth K. O, Mavis Agbandje-McKenna, Ellen Weinberg and L. Govindasamy for their advice and help and to the staff at the SER-CAT 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory for assistance during X-ray data collection. Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science under Contract No. W-31-109-Eng-38. This work was supported by grants to H-JN from the American Heart Association (0535161B) and to HK from the National Institutes of Health (HL081577).

References

Akazawa, H. & Komuro, I. (2005). *Pharmacol. Ther.* **107**, 252–268.
Benson, D., Silberbach, G., Kavanaugh-McHugh, A., Cottrill, C., Zhang, Y., Riggs, S., Smalls, O., Johnson, M., Watson, M., Seidman, J., Seidman, C., Plowden, J. & Kugler, J. (1999). *J. Clin. Invest.* **104**, 1567–1573.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Durocher, D., Charron, F., Warren, R., Schwartz, R. & Nemer, M. (1997). *EMBO J.* **16**, 5687–5696.
Esposito, G., Fogolari, F., Damante, G., Formisano, S., Tell, G., Leonardi, A., Di Lauro, R. & Viglino, P. (1996). *Eur. J. Biochem.* **241**, 101–113.
Fodor, E., Mack, J. W., Maeng, J. S., Ju, J. H., Lee, H. S., Gruschus, J. M., Ferretti, J. A. & Ginsburg, A. (2005). *Biochemistry*, **44**, 12480–12490.
Garg, V., Kathiriya, I., Barnes, R., Schluterman, M., King, I., Butler, C., Rothrock, C., Eapen, R., Hirayama-Yamada, K., Joo, K., Matsuoka, R., Cohen, J. & Srivastava, D. (2003). *Nature (London)*, **424**, 443–447.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.
Goldmuntz, E., Geiger, E. & Benson, D. (2001). *Circulation*, **104**, 2565–2568.
Gruschus, J., Tsao, D., Wang, L., Nirenberg, M. & Ferretti, J. (1999). *J. Mol. Biol.* **289**, 529–545.
Gutierrez-Roelens, I., Sluysmans, T., Gewillig, M., Devriendt, K. & Vikkula, M. (2002). *Hum. Mutat.* **20**, 75–76.
Harvey, R., Lai, D., Elliott, D., Biben, C., Solloway, M., Prall, O., Stennard, F., Schindeler, A., Groves, N., Lavulo, L., Hyun, C., Yeoh, T., Costa, M., Furtado, M. & Kirk, E. (2002). *Cold Spring Harbor Symp. Quant. Biol.* **67**, 107–114.
Hoffman, J. (1995). *Pediatr. Cardiol.* **16**, 155–165.
Hovde, S., Abate-Shen, C. & Geiger, J. H. (2001). *Biochemistry*, **40**, 12013–12021.
Kasahara, H. & Benson, D. (2004). *Cardiovasc. Res.* **64**, 40–51.
Kasahara, H., Usheva, A., Ueyama, T., Aoki, H., Horikoshi, N. & Izumo, S. (2001). *J. Biol. Chem.* **276**, 4570–4580.
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.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Schott, J., Benson, D., Basson, C., Pease, W., Silberbach, G., Moak, J., Maron, B., Seidman, C. & Seidman, J. (1998). *Science*, **281**, 108–111.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.