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

Crystallization of hepatocyte nuclear factor 4α (HNF4α) in complex with the HNF1α promoter element

Peng Lu, Jianguo Liu, Manana Melikishvili, Michael G. Fried and Young-In Chi*

Department of Molecular and Cellular Biochemistry and Center for Structural Biology, University of Kentucky, Lexington, KY 40536, USA

Correspondence e-mail: ychi@uky.edu

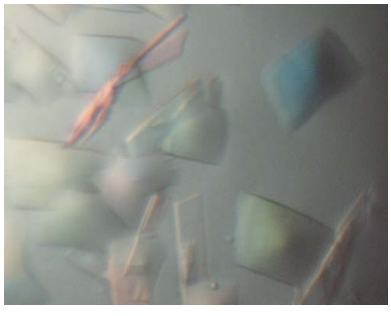
Received 29 January 2008
Accepted 14 March 2008

Hepatocyte nuclear factor 4α (HNF4α) is a member of the nuclear receptor superfamily that plays a central role in organ development and metabolic functions. Mutations on *HNF4α* cause maturity-onset diabetes of the young (MODY), a dominant monogenic cause of diabetes. In order to understand the molecular mechanism of promoter recognition and the molecular basis of disease-causing mutations, the recombinant HNF4α DNA-binding domain was prepared and used in a study of its binding properties and in crystallization with a 21-mer DNA fragment that contains the promoter element of another MODY gene, *HNF1α*. The HNF4α protein displays a cooperative and specific DNA-binding activity towards its target gene-recognition elements. Crystals of the complex diffract to 2.0 Å using a synchrotron-radiation source under cryogenic (100 K) conditions and belong to space group *C*2, with unit-cell parameters $a = 121.63$, $b = 35.43$, $c = 70.99$ Å, $\beta = 119.36^\circ$. A molecular-replacement solution has been obtained and structure refinement is in progress. This structure and the binding studies will provide the groundwork for detailed functional and biochemical studies of the MODY mutants.

1. Introduction

Hepatocyte nuclear factor 4α (HNF4α) is a tissue-specific transcription factor that plays an essential role in early vertebrate development and embryonic survival. It regulates the expression of a wide variety of essential genes, including those involved in liver and pancreatic cell differentiation (Li *et al.*, 2000; Odom *et al.*, 2004), embryogenesis and early development (Duncan *et al.*, 1994; Lausen *et al.*, 2000), glucose metabolism (Stoffel & Duncan, 1997), lipid homeostasis (Hayhurst *et al.*, 2001) and amino-acid metabolism (Schrem *et al.*, 2002). Mutations in *HNF4α* cause a dominantly inherited form of diabetes known as maturity-onset diabetes of the young (MODY; Yamagata *et al.*, 1996). These mutations cause the loss of function of the gene product (Lausen *et al.*, 2000), which leads to impaired insulin secretion and defects in metabolic pathways (Miura *et al.*, 2006).

HNF4α is a prototypical member of a unique nuclear receptor superfamily (NR2A1; Nuclear Receptors Nomenclature Committee, 1999) and exclusively functions as a homodimer (Jiang *et al.*, 1995), despite its sequence homology to retinoic X receptor (RXR), which can readily heterodimerize with a related nuclear receptor (Szanto *et al.*, 2004). HNF4α consists of distinctive functional domains including a DNA-binding domain (DBD), a ligand-binding domain (LBD) and additional domains with transcription-activation functions (AF; Schrem *et al.*, 2002). However, the identity of its *bona fide* ligand is still under dispute (Hertz *et al.*, 1998; Petrescu *et al.*, 2005), even though its apparent ligand has been identified from structural studies (Dhe-Paganon *et al.*, 2002; Wisely *et al.*, 2002). HNF4α-DBD contains two zinc-finger motifs that specifically recognize and bind as a homodimer to a direct repeat of two hexameric half-sites separated by one (DR1; in the majority) or two nucleotides (DR2) (Jiang *et al.*, 1995; Rajas *et al.*, 2002). Five MODY1 missense mutations (on four different residues) are found within the region of our HNF4α-DBD



construct (Fig. 1*a*) and an additional MODY mutation is found in the HNF4α-binding site within the promoter of another MODY (MODY3) culprit gene *HNF1*α (Fig. 1*b*; Gragnoli *et al.*, 1997). Analysis of the structural consequences of each amino-acid substitution should be instructive as to the functional role of each residue. In order to elucidate the molecular basis of HNF4α function and the monogenic causes of diabetes, we have prepared and crystallized the human HNF4α DNA-binding domain in complex with a high-affinity HNF1α promoter element containing the HNF4α recognition sequence.

2. Materials and methods

2.1. Construction, expression and purification of HNF4α DNA-binding domain

The cDNA harboring the full-length human HNF-4αB splice variant (Kritis *et al.*, 1996) was a kind gift from Dr Steve Shoelson of Joslin Diabetes Center. A fragment of human HNF4α cDNA (amino acids 46–126) was subcloned by standard PCR into a pET41a vector (GE Healthcare). HNF4α was overexpressed in *Escherichia coli* BL21-Gold (Novagen) with induction by 0.5 m*M* IPTG at an OD_{600} of 0.8–1.0 at 310 K and harvested after culturing for an additional 3–4 h. No zinc solution was added during the purification since there should be a sufficient amount of Zn atoms in the medium to be incorporated into the protein. The cells were lysed by sonication and the expressed GST-fusion proteins were isolated with the use of glutathione-agarose beads (Invitrogen) in bulk plus washing in the presence of 0.6 *M* NaCl to prevent nonspecific binding to bacterial DNA. HNF4α was released by thrombin digestion from the resin after overnight incubation at 277 K and was further purified by ion-exchange chromatography (Mono-S FPLC). Thrombin digestion produced a two-residue remnant (Gly-Ser) at the N-terminal end (Fig. 1*a*). The purified protein was estimated to be at least 98% pure as judged by staining with Coomassie on 8–25% gradient SDS–PAGE gel (Fig. 2). Fractions were pooled and stored at 193 K as a 10%(*v*/*v*) glycerol stock.

2.2. Gel filtration of HNF4α DNA-binding domain

Gel filtration was performed on a Superdex 75 HR 10/30 column (GE Healthcare) equilibrated with running buffer containing 20 m*M* Tris pH 7.5, 200 m*M* NaCl, 1 m*M* EDTA and 10 m*M* 2-mercaptoethanol. Elution was performed at a flow rate of 0.5 ml min^{-1}. The apparent molecular weight of HNF4α-DBD was determined using the same column calibrated previously with a range of reference proteins (Bio-Rad): thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), equine myoglobin (17 kDa) and vitamin B_{12} (1.4 kDa). Blue dextran was used to determine the void volume of the column.

2.3. Preparation of DNA oligomers

Tritylated oligonucleotides were purchased from the Midland Certified Reagent Company (Midland, Texas, USA) and further purified by reverse-phase HPLC on a C8 XTerra prep column (Waters) using a linear 5–50%(*v*/*v*) acetonitrile gradient in 50 m*M* triethylamine acetate buffer pH 7.0. Excess mobile phase containing acetonitrile was removed using HiTrapQ (GE Healthcare) and the trityl groups were removed with 80%(*v*/*v*) acetic acid. The deprotected oligonucleotides were precipitated with 75%(*v*/*v*) ethanol, dissolved in water for concentration measurement by A_{260} and lyophilized before storage at 193 K. Double-stranded DNAs were generated for crystallization by heating equimolar amounts of complementary oligonucleotides to 358 K for 10 min and slowly cooling to 277 K. The annealing buffer condition was 20 m*M* Tris pH 8.0, 200 m*M* NaCl and 1 m*M* EDTA.

2.4. Electrophoretic mobility-shift assay (EMSA)

Single-stranded oligonucleotides 1 and 2 (Fig. 1*b*) were dissolved in 10 m*M* Tris (pH 7.5 at 293 K) and 1 m*M* EDTA. Oligonucleotide 1 was 5′-end labeled with ^{32}P as described by Maxam & Gilbert (1977). Labeled oligonucleotide 1 was mixed with a 1.1-fold molar excess of oligonucleotide 2 and the samples were heated to 363 K and cooled slowly to 293 K. DNA was transferred by dialysis into binding buffer [10 m*M* Tris pH 8.0, 1 m*M* EDTA, 100 m*M* NaCl, 1 m*M* $MgCl_2$, 1 m*M* DTT, 4%(*v*/*v*) glycerol]. DNA concentrations were measured by absorbance using a molar extinction coefficient $\varepsilon_{260}{}^{1\,\mathrm{cm}}$ of 2.57×10^5. Samples were stored at 253 K until use. EMSAs were carried out as described by Hellman & Fried (2007) using 10%(*w*/*v*) polyacrylamide gels cast and run in 45 m*M* Tris–borate, 2.5 m*M* EDTA (pH 8.3 at 293 K). Autoradiographic images were captured on storage phosphor screens (type GP, GE Healthcare), detected with a Typhoon phosphorimager and quantitated with *Image-Quanta* software (GE

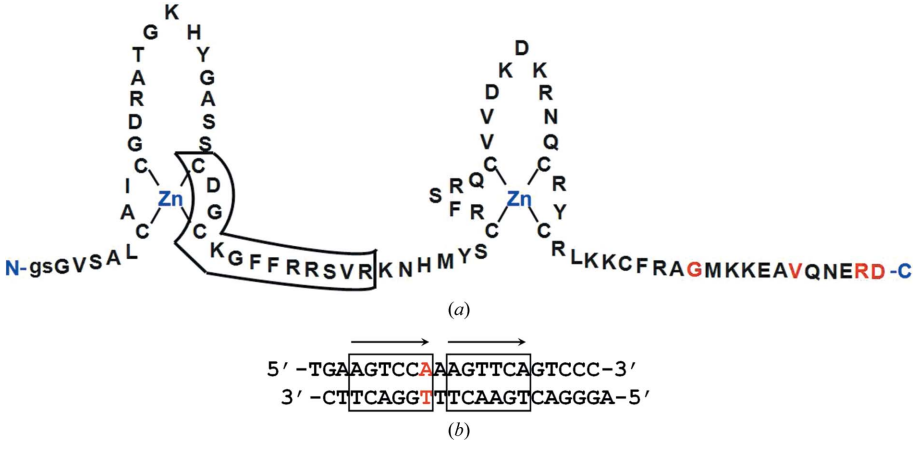


Figure 1
The protein and DNA constructs used in characterization and crystallization. (*a*) Human HNF4α-DBD protein sequence. The predicted DNA-recognition helix is boxed and the MODY mutations are shown in red. Cloning artifacts from the expression vector are indicated by lower case letters at the N-terminal end. (*b*) The natural HNF1α promoter sequence (−46 to −66; Gragnoli *et al.*, 1997) containing the HNF4α recognition site used in crystallization and binding studies. Two direct-repeat half-sites are boxed and the MODY mutation is shown in red.

Healthcare). Data from serial dilution experiments were analyzed using the equations

$$\ln\frac{[P_nD]}{[D]} = n\ln[P] + \ln K_{\rm obs} \quad (1)$$

and

$$[P] = [P]_{\rm tot} - n[P_nD]. \quad (2)$$

Here, n is the binding stoichiometry, $[P_nD]$ and $[D]$ the concentrations of complex and free DNA, respectively, and $K_{\rm obs} = [P_nD]/[P]^n[D]$. When both n and $K_{\rm obs}$ are unknown (as in this case), iterative calculation of n and $[P]$ using (1) and (2) results in convergence on self-consistent values of n and $K_{\rm obs}$ (Adams & Fried, 2007; Fried & Crothers, 1984).

2.5. Dynamic light-scattering measurement

The effective molecular radius and the homogeneity/monodispersity of the complex within various particular buffer conditions were measured using the Solubility Screening Kit (Jena Biosciences) in conjunction with a Dynapro-99 dynamic light-scattering instrument (Proterion Corporation) and a DynaPro-MSTC200 microsampler (Protein Solutions). The results were analyzed using *DYNAMICS* v.5.26.60 (Protein Solutions). 20 µl of sample was inserted into the cuvette with the temperature control set to 293 K. The light-scattering signal was collected at a wavelength of 830.7 nm. Protein concentrations were about 2 mg ml^{-1} in each buffer and an average of 15 readings were recorded for each measurement.

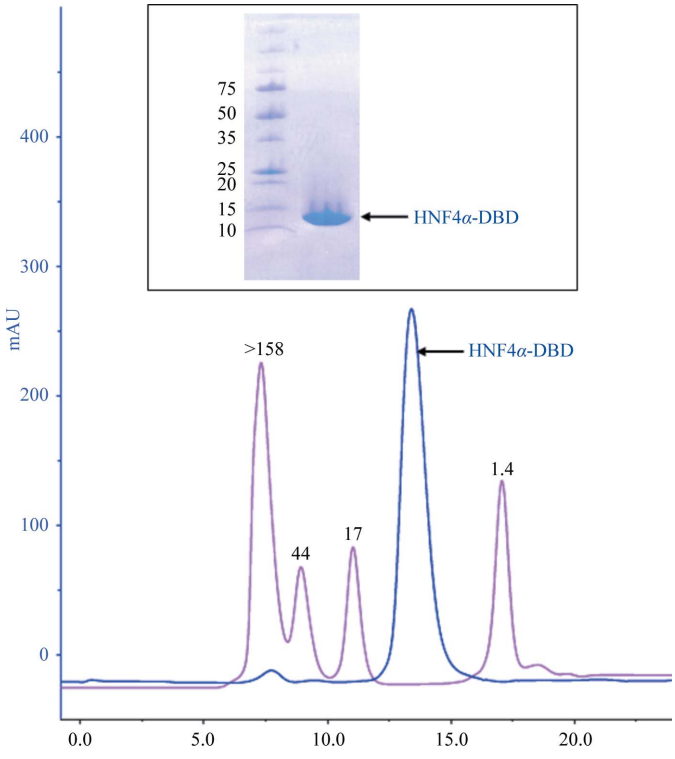


Figure 2
Gel filtration of HNF4α-DBD with the molecular-weight standard samples (labelled in kDa). HNF4α-DBD is homogeneous and appears to be a monomer in solution. The SDS–PAGE of purified HNF4α-DBD along with the molecular-weight standard is shown in the inset.

2.6. Crystallization and optimization

Protein–DNA complexes were dialyzed in 20 m*M* Tris pH 7.5, 75 m*M* NaCl and 1 m*M* DTT at 277 K for 2.5 h and concentrated to at least 10 mg ml^{-1}. The initial crystallization trials were carried out at 295 K in 24-well plates using the hanging-drop vapor-diffusion method with a sparse-matrix screen (Jancarik & Kim, 1991) and similar commercially available versions such as Crystal Screens I and II (Hampton Research), Natrix and PEG/Ion Screens (Hampton Research), Cryo I and II (Molecular Dimensions) and Wizard I and II (deCODE Genetics).

Drops consisting of 0.5 µl protein–DNA solution were mixed with an equal volume of reservoir solution and equilibrated against 500 µl reservoir solution. Although many different DNA constructs varying

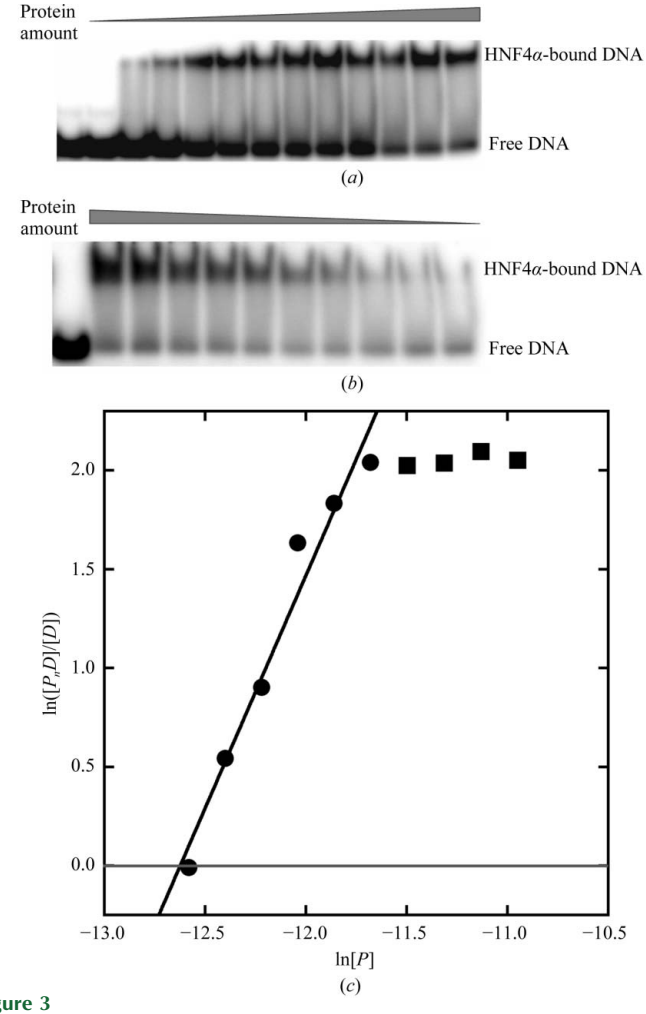


Figure 3
Binding of HNF4α-DBD to duplex 21-mer target DNA detected by EMSA. Reactions were carried out in 10 m*M* Tris (pH 8.0 at 293 K), 1 m*M* EDTA, 100 m*M* NaCl, 1 m*M* $MgCl_2$, 1 m*M* DTT, 4% glycerol, 100 µg ml^{-1} BSA. (*a*) Forward titration. All samples contained 0.21 µ*M* duplex 21-mer. Samples contained HNF4α-DBD protein, from the second lane, at 0.41, 0.82, 1.23, 2.05, 2.87, 3.69, 4.51, 5.33, 6.56, 8.2, 12.0 and 20.5 µ*M*, respectively. (*b*) Serial dilution. The first lane contained reference DNA; the initial sample (second lane) contained 0.2 µ*M* duplex 21-mer and 17.9 µ*M* HNF4α-DBD protein. Each succeeding lane contained an aliquot of the previous sample diluted 1.2-fold. (*c*) Determination of stoichiometry and association constant. Graph of $\ln[P_nD]/[D]$ as a function of $\ln[P]$. The slope about the mid-point of the reaction (where $\ln[P_nD]/[D] = 0$) gives the stoichiometry ($n = 2.3 \pm 0.2$). The data used for this determination are indicated by filled circles. The data near binding saturation (filled squares) deviate from the linear relationship and were excluded from the fit. A fit of (1) to the linear data returned an intercept value of $\ln[P] = -12.58 \pm 0.03$, equivalent to $K_{\rm obs} = 8.48 \pm 0.67 \times 10^{10}\ M^{-2}$.

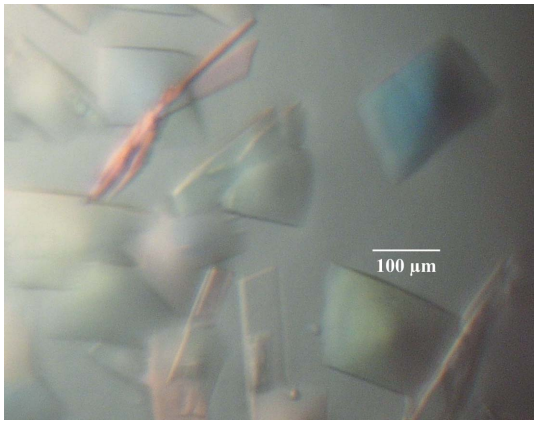


Figure 4
Typical crystals of the HNF4α–DNA complex.

in length and the nature of the ending (blunt end *versus* overhang) were screened, diffraction-quality crystals were only reproducibly obtained using the overhang 21-mer shown in Fig. 1(*b*) (the two HNF4α direct-repeat recognition sites are indicated by the boxes). Conditions yielding small crystals were further optimized by variation of the crystallization parameters and additives. The final condition, which produced somewhat flat bipyramidal crystals at 295 K, contained 26%(*v*/*v*) PEG 4000, 80 m*M* magnesium acetate and 50 m*M* sodium citrate pH 4.8.

2.7. Data collection and processing

The crystals were transferred into mother liquor containing an additional 15%(*v*/*v*) glycerol as a cryoprotectant before being directly plunged into liquid nitrogen and stored for data collection. The native data were collected at 100 K at APS (SER-CAT 22BM) using a MAR 225 CCD detector and an oscillation angle of 1° with 2 s exposure and were processed using *HKL*-2000 (Otwinowski & Minor, 1997). The wavelength used was 0.92017 Å.

3. Results and discussion

Recombinant HNF4α-DBD (amino acids 46–126; Fig. 1*a*) was purified to homogeneity and mixed with pure DNA for subsequent studies. Gel-filtration experiments showed that the HNF4α-DBD protein existed as a monomer in solution (Fig. 2). Purified HNF4α-DBD protein forms a single complex with DNA containing its target sequence (Figs. 3*a* and 3*b*). Serial dilution analysis (Figs. 3*b* and 3*c*) revealed that the stoichiometry of the complex was 2:1 HNF4α:dsDNA, with an association constant K_{obs} of $8.48 \pm 0.67 \times 10^{10}\ M^{-2}$. The corresponding monomer equivalent dissociation constant was $3.43 \pm 0.13 \times 10^{-6}\ M$. The formation of a 2:1 complex without the accumulation of detectable levels of the 1:1 intermediate indicates that binding is cooperative. These features will serve as a reference when we study the effects of MODY mutations on DNA binding in the near future.

Dynamic light scattering (DLS) is a useful tool to monitor protein solubility behavior and to predict favorable crystallization conditions (Wilson, 2003). We used the Solubility Screening Kit (Jena Biosiences) in conjunction with DLS (Jancarik *et al.*, 2004) in order to identify the optimal buffer conditions for complex formation and crystallization. The best polydispersity value of 0.06 was obtained with a buffer containing 20 m*M* Tris pH 7.5 and 75 m*M* NaCl and this optimal buffer was used for subsequent crystallization.

For crystallization, purified HNF4α 46–126 and various DNAs were simply mixed in a 2:1.2 molar ratio, dialyzed against the optimal binding buffer (20 m*M* Tris pH 7.5, 75 m*M* NaCl) and concentrated using 10 kDa molecular-weight cutoff concentrators (Millipore). The protein–DNA concentration was 10 mg ml^{-1} for initial screenings and 20 mg ml^{-1} for final optimization. Crystals with the overhang 21-mer DNA (Fig. 1*b*) were grown at 295 K using the hanging-drop

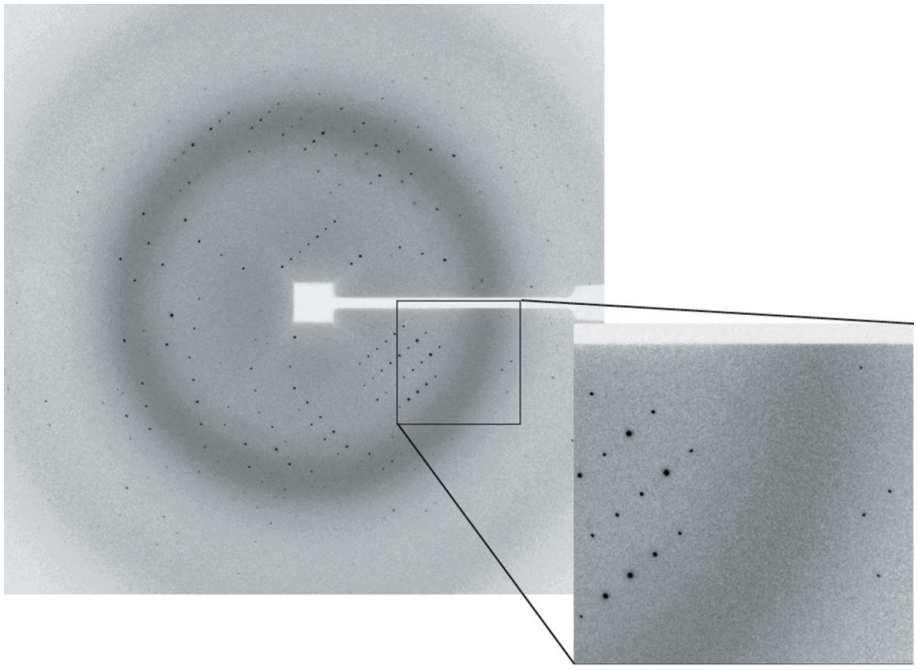


Figure 5
A typical X-ray diffraction pattern from a crystal of the HNF4α–DNA complex. A small section near the water ring is enlarged and shown in the inset. The overall mosaicity of the crystal was 0.3°.

Table 1
Data-collection statistics for the HNF4α–DNA crystal.

Values in parentheses are for the highest resolution bin.

Space group	*C*2
Unit-cell parameters (Å, °)	a = 121.63, b = 35.43, c = 70.99, β = 119.36
Resolution (Å)	30.29–2.00 (2.07–2.00)
Observed reflections	94972
Unique reflections	17094
Redundancy	5.6 (3.6)
Completeness (%)	93.9 (72.4)
$I/\sigma(I)$	12.1 (4.5)
$R_{\rm merge}$† (%)	6.4 (30.4)
Matthews coefficient (Å^3 Da^{-1})	2.12
Solvent content (%)	41.61
Molecules per ASU	One complex (2 HNF4α, 1 dsDNA)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of reflection hkl, $\sum_{hkl}$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection h.

vapor-diffusion method and the presence of the HNF4α–DNA complex in the crystals was confirmed by running SDS–PAGE and 0.5%(*w*/*v*) agarose gels (data not shown). Crystals initially appeared within 2 d and continued to grow until they reached average dimensions of 0.05 × 0.1 × 0.2 mm (Fig. 4). A range of solution conditions varying the pH, temperature and concentrations of additives such as organic solvents, divalent cations and polyamines were used to attempt to improve the crystal quality. The final optimized condition contains 26%(*v*/*v*) PEG 4000, 80 m*M* magnesium acetate and 50 m*M* sodium citrate pH 4.8. The best crystal diffracted to 2.0 Å at the synchrotron source and belongs to space group *C*2, with unit-cell parameters a = 121.63, b = 35.43, c = 70.99 Å, β = 119.36° (Fig. 5). The value of the Matthews coefficient (Matthews, 1968) is 2.12 Å^3 Da^{-1} for one complex (two HNF4α and one dsDNA) in the asymmetric unit and the estimated solvent content is 41.6% based on a protein specific density of 1.34. Final native data-collection statistics are summarized in Table 1.

The structure was determined by molecular replacement using the RXR–RAR–DNA complex structure (PDB code 1dsz) as a search model and the program *MOLREP* (Vagin & Teplyakov, 1997) from the *CCP*4 suite (Winn, 2003). An unambiguous solution was found that gave an initial *R* value of 51.4% and a correlation coefficient of 0.38 using data in the resolution range 15–3.0 Å. The subsequent $\sigma_{\rm A}$-weighted $2F_{\rm o} - F_{\rm c}$ map after rigid-body refinement clearly revealed density corresponding to the structural differences between the search model and the HNF4α–DNA complex. Model improvement and refinement of the structure are in progress.

We wish to thank the staff at Advanced Photon Source beamline 22-BM (SER-CAT) for their help with data collection and David Rodgers for the use of the tissue-culture and protein-production core facility. Use of the Advanced Photon Source is supported by the US Department of Energy. This work was funded by a Juvenile Diabetes Research Foundation grant (2004-503) and NIH Grant P20RR20171 from the COBRE program of the National Center for Research Resources to YIC and NIH grant GM070662 to MGF.

References

Adams, C. & Fried, M. G. (2007). *Protein Interactions: Biophysical Approaches for the Study of Multicomponent Systems*, edited by P. Schuck, pp. 417–446. New York: Academic Press.
Dhe-Paganon, S., Duda, K., Iwamoto, M., Chi, Y. I. & Shoelson, S. E. (2002). *J. Biol. Chem.* **277**, 37973–37976.
Duncan, S. A., Manova, K., Chen, W. S., Hoodless, P., Weinstein, D. C., Bachvarova, R. F. & Darnell, J. E. Jr (1994). *Proc. Natl Acad. Sci. USA*, **91**, 7598–7602.
Fried, M. G. & Crothers, D. M. (1984). *J. Mol. Biol.* **172**, 241–262.
Gragnoli, C., Lindner, T., Cockburn, B. N., Kaisaki, P. J., Gragnoli, F., Marozzi, G. & Bell, G. I. (1997). *Diabetes*, **46**, 1648–1651.
Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M. & Gonzalez, F. J. (2001). *Mol. Cell. Biol.* **21**, 1393–1403.
Hellman, L. M. & Fried, M. G. (2007). *Nature Protocols*, **2**, 1849–1861.
Hertz, R., Magenheim, J., Berman, I. & Bar-Tana, J. (1998). *Nature (London)*, **392**, 512–516.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Acta Cryst.* D**60**, 1670–1673.
Jiang, G., Nepomuceno, L., Hopkins, K. & Sladek, F. M. (1995). *Mol. Cell. Biol.* **15**, 5131–5143.
Kritis, A. A., Argyrokastritis, A., Moschonas, N. K., Power, S., Katrakili, N., Zannis, V. I., Cereghini, S. & Talianidis, I. (1996). *Gene*, **173**, 275–280.
Lausen, J., Thomas, H., Lemm, I., Bulman, M., Borgschulze, M., Lingott, A., Hattersley, A. T. & Ryffel, G. U. (2000). *Nucleic Acids Res.* **28**, 430–437.
Li, J., Ning, G. & Duncan, S. A. (2000). *Genes Dev.* **14**, 464–474.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Maxam, A. M. & Gilbert, W. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 560–564.
Miura, A., Yamagata, K., Kakei, M., Hatakeyama, H., Takahashi, N., Fukui, K., Nammo, T., Yoneda, K., Inoue, Y., Sladek, F. M., Magnuson, M. A., Kasai, H., Miyagawa, J., Gonzalez, F. J. & Shimomura, I. (2006). *J. Biol. Chem.* **281**, 5246–5257.
Nuclear Receptors Nomenclature Committee (1999). *Cell*, **97**, 161–163.
Odom, D. T., Zizlsperger, N., Gordon, D. B., Bell, G. W., Rinaldi, N. J., Murray, H. L., Volkert, T. L., Schreiber, J., Rolfe, P. A., Gifford, D. K., Fraenkel, E., Bell, G. I. & Young, R. A. (2004). *Science*, **303**, 1378–1381.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Petrescu, A. D., Hertz, R., Bar-Tana, J., Schroeder, F. & Kier, A. B. (2005). *J. Biol. Chem.* **280**, 16714–16727.
Rajas, F., Gautier, A., Bady, I., Montano, S. & Mithieux, G. (2002). *J. Biol. Chem.* **277**, 15736–15744.
Schrem, H., Klempnauer, J. & Borlak, J. (2002). *Pharmacol. Rev.* **54**, 129–158.
Stoffel, M. & Duncan, S. A. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 13209–13214.
Szanto, A., Narkar, V., Shen, Q., Uray, I. P., Davies, P. J. & Nagy, L. (2004). *Cell Death Differ.* **11**, *Suppl.* 2, S126–S143.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Wilson, W. W. (2003). *J. Struct. Biol.* **142**, 56–65.
Winn, M. D. (2003). *J. Synchrotron Rad.* **10**, 23–25.
Wisely, G. B., Miller, A. B., Davis, R. G., Thornquest, A. D. Jr, Johnson, R., Spitzer, T., Sefler, A., Shearer, B., Moore, J. T., Willson, T. M. & Williams, S. P. (2002). *Structure*, **10**, 1225–1234.
Yamagata, K., Furuta, H., Oda, N., Kaisaki, P. J., Menzel, S., Cox, N. J., Fajans, S. S., Signorini, S., Stoffel, M. & Bell, G. I. (1996). *Nature (London)*, **384**, 458–460.