Multidomain integration in the structure of the HNF-4α nuclear receptor complex

Vikas Chandra1, Pengxiang Huang1, Nalini Potluri1, Dalei Wu1, Youngchang Kim2 & Fraydoon Rastinejad1

The hepatocyte nuclear factor 4α (HNF-4α; also known as NR2A1) is a member of the nuclear receptor (NR) family of transcription factors, which have conserved DNA-binding domains and ligand-binding domains1,2. HNF-4α is the most abundant DNA-binding protein in the liver, where some 40% of the actively transcribed genes have a HNF-4α response element1-4. These regulated genes are largely involved in the hepatic gluconeogenic program and lipid metabolism5-6. In the pancreas HNF-4α is also a master regulator, controlling an estimated 11% of islet genes5. HNF-4α protein mutations are linked to maturity-onset diabetes of the young, type 1 (MODY1) and hyperinsulinaemic hypoglycaemia8-11. Previous structural analyses of NRs, although productive in elucidating the structure of individual domains, have lagged behind in revealing the connectivity patterns of NR domains. Here we describe the 2.9 Å crystal structure of the multidomain human HNF-4α homodimer bound to its DNA response element and coactivator-derived peptides. A convergence zone connects multiple receptor domains in an asymmetric fashion, joining distinct elements from each monomer. An arginine target of PRMT1 methylproline protrudes directly into this convergence zone and sustains its integrity. A serine target of protein kinase C is also responsible for maintaining domain–domain interactions. These post-translational modifications lead to changes in DNA binding by communicating through the tightly connected surfaces of the quaternary fold. We find that some MODY1 mutations, positioned on the ligand-binding domain and hinge regions of the receptor, compromise DNA binding at a distance by communicating through the interjunctional surfaces of the complex. The overall domain representation of the HNF-4α homodimer is different from that of the PPAR-γ–RXR-α heterodimer, even when both NR complexes are assembled on the same DNA element. Our findings suggest that unique quaternary folds and interdomain connections in NRs could be exploited by small-molecule allosteric modulators that affect distal functions in these polypeptides.

We previously reported the only high-resolution structural example of a multidomain NR complex, that of the PPAR-γ–RXR-α heterodimer on its DNA response element12. To understand the extent of domain integration in other NRs, here we analyse the crystal structure of the complex of HNF-4α, an obligate homodimer, bound to its DNA element and coactivator-derived peptides. HNF-4α uses the linear domain arrangement shown in Fig. 1a. Our efforts to crystallize full-length HNF-4α were unsuccessful. However, by proteolytically probing its DNA-assembled complex, we identified an extended segment comprising the DNA-binding domain (DBD)–hinge–ligand-binding domain (LBD) portions and corresponding to residues 46–368 (Supplementary Fig. 1). Cloning, expression and purification of the stable DBD–hinge–LBD multidomain segment made it possible to obtain well-diffracting crystals of a complex with its consensus response element and coactivator (NCOA2) peptide. Electron density maps for all the critical junctions of the complex are shown in Supplementary Figs 2–5. The response element consists of a direct repeat of AGGTCA half-sites with one base-pair spacing (DR1). The DR1 is the major consensus binding site for both HNF-4α and PPAR-γ–RXR-α13-15.

X-ray diffraction data was collected to 2.9 Å resolution, and the structure refined (see Supplementary Table 1). The crystal asymmetric unit contains two independent representations of the HNF-4α–homodimer–DNA–peptide complex. The electron density map from one complex and the comparison of the two complexes is shown in Fig. 1b, c. The two representations are nearly identical, with root mean squared deviations of less than 2.0 Å over all their atoms. The LBD and DBD portions match their previously determined isolated structures (Supplementary Figs 6 and 7). Both DBDs are in register with their half-sites, interacting with the major grooves (Supplementary Figs 8 and 9). Helix 12 of the LBDs is in the active conformation, and a coactivator LXXLL peptide is bound to each LBD.

The HNF-4α homodimer shows a striking and complex pattern of interfacial junctions. A central zone incorporates surfaces from both LBDs, the DBD of the upstream subunit, and the hinge region of the downstream subunit. This domain convergence zone suggests a path of communication between the conserved domains through their coupled surfaces (Fig. 2a). The LBDs, symmetrical in their mutual interactions when viewed in isolation, cooperate in a highly asymmetric fashion to straddle the surface of only the upstream DBD (Fig. 1d). As a result, the overall complex appears partitioned towards the upstream half of the DR1, and adopts a highly asymmetrical organization for a homodimeric transcription factor. A previous study suggested that HNF-4α homodimers could bind asymmetrically to their DNA response elements6. The resulting quaternary arrangement creates precisely the correct DBD-to-DBD distances needed to match the geometric constraints of the two AGGTCA half-sites and their intervening spacer. At the same time, the quaternary organization renders both LBD pockets and their coactivator-interacting surfaces unencumbered, allowing free access to both ligands and LXXLL elements, respectively.

The interface that forms between the DBD of the upstream subunit and the hinge region of the downstream subunit is one important domain–domain interface of the complex, and is reminiscent of an interaction we described previously in the PPAR-γ–RXR-α complex12. The resulting arrangement places the two DBDs in a solid head-to-tail arrangement that extends their combined footprint to match their DR1 contact surface perfectly. The manner in which the two LBDs cooperate to interact with the upstream DBD suggests that the physical integration of all three domains may be required for high-affinity DNA binding (Fig. 1d). To test this idea, we measured the DNA affinity of the HNF-4α protein when various segments of the protein were removed (Fig. 1d.) Measuring first the DNA-binding affinity of the HNF-4α that only contains its DBD and hinge portions, we observed very weak binding to DR1 with an equilibrium dissociation constant (Kd) of approximately 6,000 nM. When the LBD portion of the receptor is contained within the polypeptide, the complex showed a 75-fold enhanced affinity for DR1, with a Kd of approximately 80 nM (see Fig. 1e and Supplementary Table 2). These results are consistent with

1Metabolic Signaling and Disease Program, Sanford-Burnham Medical Research Institute, Orlando, Florida 32827, USA. 2Biosciences Division, Structural Biology Center, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA.

©2013 Macmillan Publishers Limited. All rights reserved
or both (HNF-4 DBD-hinge (DBD) region to a (FL) receptor. However, removal of the LBD reduces the affinity of the resulting fractional DNA bound. Removal of the AB domain (a DNA binding.

LBDs in a complex on top of the upstream DBD enables high-affinity DR1 prime refer to the upstream-positioned subunit. c, Ser 78, positioned on the back side of the DNA recognition helix, also enables the convergence of the DBD and the LBDs.

Figure 1 | Overall organization of the HNF-4α homodimer on DNA. a, Linear depiction of the HNF-4α protein domains. b, Electron density map (2F, − F,) for one of the two HNF-4α homodimer–DNA complexes of the crystallographic asymmetric unit. c, Superposition of the two independent HNF-4α homodimer–DNA complexes in the asymmetric unit. The two homodimeric complexes are coloured orange/yellow in one case, and blue/purple in the other. The ligands in the LBD are shown in blue, Zn is shown in grey, and the coactivator peptides are shown in red/brown. Numbers with prime refer to the upstream-positioned subunit. d, The positioning of both LBDs in a complex on top of the upstream DBD enables high-affinity DR1 DNA binding. e, Contribution of various receptor domains to the DNA-binding affinity of HNF-4α. DNA binding is measured using fluorescent polarization studies with a 5’ fluorescein isothiocyanate (FITC)-labelled DR1. The x-axis shows concentration of the HNF-4α protein, and the y-axis shows the fractional DNA bound. Removal of the AB domain (ΔAB), F domain (ΔF) or both (ΔAABF) does not alter the DNA affinity compared to the full-length (FL) receptor. However, removal of the LBD reduces the affinity of the resulting DBD-hinge (DBD) region to a Kd of approximately 6,000 nM, whereas the presence of the LBD together with the DBD hinge (AABAF) allows DNA binding with a Kd of 82 nM (see also Supplementary Table 2).

Figure 2 | Domain–domain contacts of HNF-4α. a, Circle indicates the convergence centre, where four domains (both LBDs, the upstream DBD and the downstream hinge) come together. b, Arg 91, located on the surface of the DBD, inserts deeply into a pocket at the base of the LBD–LBD surface. Numbers with prime refer to the upstream-positioned subunit. c, Ser 78, positioned on the back side of the DNA recognition helix, also enables the convergence of the DBD and the LBDs.

Both HNF-4α LBDs have electron density profiles consistent with a trapped fatty acid, with the size indicating a myristic acid derived from Escherichia coli, in which the protein was expressed (Supplementary Fig. 2)25,26. The fatty acid is believed to lend structural integrity to the HNF-4α/γ subfamily. Linoleic acid has been shown to be an exchangeable and potentially endogenous ligand of HNF-4α, although this molecule does not confer transcriptional activity17. A stabilizing fatty acid, or a silent molecule that cannot switch receptor activity on and off, raises the question of how HNF-4α activity is otherwise regulated.

The activities of NRs can be regulated by a variety of post-translational modifications (PTMs)19. In the case of HNF-4α, two PTMs are well described for their ability to regulate receptor properties19,20. These modifications control the receptor’s ability to bind DNA, and by extension its ability to regulate gene expression. We identified the quaternary sites of these PTMs within HNF-4α. The first site, Arg 91, is a target of PRMT1, an enzyme that adds up to two methyl groups to the arginine side chain. Arg 91 methylation produces a marked enhancement in the DNA-binding activity of HNF-4α19. The second site, Ser 78, is phosphorylated by protein kinase C (PKC), which disrupts the ability of HNF-4α to bind DNA22. Therefore, taken together, these two PTMs act as on and off switches for regulating receptor activity.

Arg 91 methylation substantially enhances DNA affinity, but is not positioned to influence DNA binding directly from its location on the DBD farthest from the DNA. Figure 2b shows how its side chain deeply protrudes into the LBD–LBD cooperating surface that we described earlier as the receptor’s multidomain convergence zone. There is a cavity directly above the side chain of Arg 91 to accommodate the two extra methyl groups, and the extension of the side chain through
methylation would more firmly ‘glue’ the DBD junctional interface with both LBDs. Therefore, this PTM acts to bias allosterically the receptor to bind DNA, by stabilizing the interdomain junctions associated with the final productive DNA complex.

We next examined the location of Ser 78, the site of PKC phosphorylation in a number of NRs. Along with HNF-4α, other NRs including FXR (also known as NR1H4), RAR-α (also known as NR1B1), VDR (also known as NR1I1), PPAR-α (also known as NR1C1), PXR (also known as NR1I2) and TR2 (also known as NR2C1) are similarly targeted by PKC, which in each case phosphorylates a similarly positioned serine on the DBD. Curiously, this serine always resides on the ‘wrong side’ of the DNA recognition helix, as is the case in HNF-4α, where it seemingly cannot participate directly in DNA binding (Fig. 2c). Yet Ser 78 phosphorylation nevertheless weakens receptor–DNA binding substantially. Our structure indicates that Ser 78 is positioned to engage the receptor’s interfacial connections so as to reduce DNA binding allosterically. Figure 2c suggests how Ser 78 phosphorylation weakens receptor–DNA binding substantially.

Figure 3 | Disease-linked mutations in HNF-4α. a, Summary of MODY1 and hyperinsulinaemic hypoglycaemia point mutations identified clinically in human populations. b, c, The MODY1 mutations, in many cases (residues in red) map to the convergence centre of the receptor domains (blue circle in b). d, DNA-affinity measurements of the wild-type (WT) and mutant receptors, as described in Fig. 1e. See also Supplementary Fig. 10 for studies of other disease-linked mutations.

We next asked whether some point mutations linked to MODY1 and hyperinsulinaemic hypoglycaemia are similarly positioned in sensitive interjunctional surfaces (Fig. 3a). For R76W and R80W mutations (associated with hyperinsulinaemic hypoglycaemia), there is a simple explanation for receptor dysfunction, as this pair of arginine residues directly contacts the AGGTCA half-sites (Supplementary Figs 9 and 10). V255M alters a residue that points into the LBD pocket, the only residue doing so among all the mutations associated with MODY1 and hyperinsulinaemic hypoglycaemia mutations. We found a number of mutations that lie at the sensitive domain–domain junctions of the complex. Sites such as R127W, D126Y, D126H and R125W locate to the downstream hinge region where they form domain–domain arrangements with the upstream DBD (Supplementary Figs 4 and 10). Mutational changes in this hinge site would misalign the interaction between domain–domain surfaces required to bridge the two DBDs into register with their successive AGGTCA half-sites. Indeed, we find that these mutant proteins substantially compromised DNA affinity (Supplementary Fig. 10). This loss of DNA binding also translates to a reduction in transcriptional activity.

We next examined MODY1 mutations I314F and R324H, and their adjacent residues (R322A, Q318A, D316A and N315A), which were found to be on the LBD and at the multidomain convergence centre of the complex (Fig. 3b–d). These mutations reduced the DNA affinity and transcriptional activity of the receptor (Fig. 3d and Supplementary Fig. 11).

Our investigation of PTMs and MODY1 mutations shows that changes introduced in the LBD, the hinge region, or in the DBD away from the DNA interface, still affect the DNA-binding properties of the receptor at a distance, by communicating through the interdomain junctions of the quaternary fold. It is interesting to note the subtlety of a single PTM or a single amino-acid mutational change, and the large distance with which these signals travel across the polypeptide to modulate DNA binding. Therefore, the domain convergence centre should be seen as both a sensitive centre for receiving signals, and an allosteric transmission system for propagating signals. At the same time, it is important to note that the two subunits of the homodimer are in altogether different environments owing to the asymmetric nature of the two subunits. PTM sites such as Ser 78 or Arg 91 will have a large influence on the complex only if they occur in the upstream DBD. In the same way, some MODY1 mutations would appear to be damaging if located in one, but not the other, subunit of the homodimer. Owing to the considerable numbers of genes being actively controlled by HNF-4α in the liver and pancreas, the loss of even a fractional population of functional homodimers caused by heterozygous mutations can cause disease.

As both the HNF-4α homodimer and PPAR-γ–RXR-α complexes target DR1, we asked if their quaternary architectures were related. The common DR1 is expected to establish a similar DBD–DBD spacing in these complexes. Figure 4a, b shows the PPAR-γ–RXR-α heterodimer and the HNF-4α homodimer in an identical way, based on the layout of their common DR1 sequences. Figure 4c shows the superposition of these complexes when their DR1 sequences are aligned to match. Indeed, the DBDs occupy nearly identical positions in both DR1 complexes. Nevertheless, the higher-order quaternary arrangements are distinct for these two complexes (Fig. 4). In HNF-4α, the LBDs are biased towards the upstream DBD, whereas in the PPAR-γ–RXR-α
complex the LBDs are biased towards the downstream RXR-α DBD. Moreover, the PPAR-γ–RXR-α complex has its own type of domain convergence centre, which is not identical to that in the HNF-4α complex.

The structural comparison indicates that the DNA response element type is not the only driver of quaternary structure in NRs. Receptor organization seems to be highly dependent on the constellation of non-conserved amino acids on these LBD surfaces, and on the length and sequence of the hinge segments, which are unique to NR members. We also note that DNA recognition is not identical in these two complexes. PPAR-γ uses its hinge region to recognize an additional six base-pair segments located upstream to the DR1 core element, establishing the polarity of subunits in that heterodimer. HNF-4α subunits do not use their hinge regions for DNA recognition, nor do they contact sequences outside the core DR1.

Our crystallographic findings with both NR complexes do not support the notion of a ‘common architecture’ for full-length NRs.23 Our findings also dispel the view that NR polypeptides are arrays of ‘domains on a string’, each of which confers its own independent function without physical and functional integration. The repertoire of quaternary structures in the NR family is likely to be diverse, even though both the DBDs and LBDs are conserved. This expectation stems from the fact that neither hinge regions nor LBD surface residues are conserved in the NR family, yet these features are the key drivers of quaternary folding. The multiple response element configurations used in the NR family are another driver of quaternary organization.

Mounting evidence points to the importance of interdomain communication in the NR family. For oestrogen receptors, the activities of ligands are influenced by the response elements, and DNA can also influence coactivator binding.24 In the glucocorticoid receptor, small conformational changes in the DBD propagate across the receptor to influence the LBD, and in the androgen receptor there is also evidence of DBD-to-LBD communication.21,26 Our findings reveal that PTMs can modulate the interdomain connections in the quaternary fold. It has been reported that certain PPAR-γ ligands can selectively block the phosphorylation of PPAR-γ2, indicating communications between the LBD pocket and the site of phosphorylation.27,28 Ser 273 in PPAR-γ is positioned within a domain–domain junction of the PPAR-γ–RXR-α complex (Supplementary Fig. 12). From its position, the phosphorylation state of Ser 273 can communicate with the DNA-binding pocket of the PPAR-γ–RXR-α heterodimer.

For HNF-4α, small molecules directed at the sensitive interjunctional sites may prove to be beneficial for treating MODY1 patients in which the DNA-binding properties have been mutationally compromised. To find these molecules, high-throughput screening efforts must target the complete architecture of this receptor and not just the isolated LBD. We have identified two locations in the quaternary structure of the HNF-4α complex that appear to be accessible for the binding of small-molecule allosteric modulators (Supplementary Fig. 13). An expanded understanding of the physical connectivity between LBDs, DBDs and other domains in the NR family should expand and better guide the discovery of receptor modulators with therapeutic value.

**METHODS SUMMARY**

The HNF-4α protein segment lacking the AB and F regions, corresponding to residues 46–368 (National Center for Biotechnology Information (NCBI) accession NM_178849), was expressed in E. coli. The purified protein was then combined with synthetic duplex DNA (DR1) and coactivator peptide, further purified as a complex, and subjected to crystallization trials. X-ray diffraction data were collected at the Argonne National Laboratory SBC-CAT 19ID beamline and the structure was solved by molecular replacement, using the previously determined structures of the HNF-4α LBD and DBD. Biochemical studies measuring DNA affinity and nuclease protection experiments were performed to develop a small-molecule allosteric modulator (Supplementary Fig. 14).

Received 9 October 2012; accepted 28 January 2013.

**Published online 13 March 2013.**

1. Sladek, F. M., Zhong, W. M., Lai, E. & Darnell, J. E. Jr. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev.* 4, 2353–2365 (1990).

2. Mangelsdorf, D. J. & Evans, R. M. The RXR heterodimers and orphan receptors. *Cell* 83, 841–850 (1995).

3. Bolotin, E. et al. Integrated approach for the identification of human hepatocyte nuclear factor 4α target genes using protein binding microarrays. *Hepatology* 51, 642–653 (2010).

4. Wallerman, O. et al. Molecular interactions between HNF4α, FOXA2 and GABP identified at regulatory DNA elements through ChIP-seq sequencing. *Nucleic Acids Res.* 37, 7498–7508 (2009).

5. Yoon, J. C. et al. Control of hepatic glucose homeostasis through the transcriptional coactivator PGC-1. *Nature* 413, 131–138 (2001).

6. Fang, B.,bane-Padros, D., Bolotin, E., Jiang, T. & Sladek, F. M. Identification of a binding motif specific to HNF4 by comparative analysis of multiple nuclear receptors. *Nucleic Acids Res.* 40, 5343–5356 (2012).

7. Odorn, D. T. et al. Control of pancreas and liver gene expression by PPAR transcription factors. *Science* 303, 1376–1381 (2004).

8. Ryeff, G. U. Mutations in the human genes encoding the transcription factors of the nuclear receptor. *Hepatology* 41, 379–441 (2005).

9. Easterling, C. E. et al. Hyperinsulinemic hypoglycemia caused by HNF4A gene mutations. *Eur. J. Endocrinol.* 162, 987–992 (2010).

10. Chandra, V. et al. Structure of the intact PPAR-γ–RXR-α nuclear receptor complex on DNA. *Nature* 456, 2953–2967 (2008).

11. Nielsen, R. et al. Genome-wide profiling of PPAR-γ–RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev.* 22, 2965–2976 (2008).

12. Jiang, G., Lee, U. & Sladek, F. M. Proposed mechanism for the stabilization of nuclear receptor DNA binding via protein dimerization. *Mol. Cell. Biol.* 17, 6546–6554 (1997).

13. Wiely, G. B. et al. Hepatocyte nuclear factor 4 is a transcription factor that constitutively binds fatty acids. *Structure* 10, 1225–1234 (2002).

14. Dhe-Paganon, S., Duda, K., Iwamoto, M., Chi, Y. I. & Shoelson, S. E. Crystal structure of the HNF4α ligand binding domain in complex with endogenous fatty acid ligand. *J. Biol. Chem.* 277, 37973–37976 (2002).

15. Yuan, X. et al. Identification of an endogenous ligand bound to a native orphan nuclear receptor. *PloS ONE* 4, e5609 (2009).

16. Weigel, N. L. & Moore, N. L. Steroid receptor phosphorylation: a key modulator of multiple receptor functions. *Mol. Endocrinol.* 21, 2311–2319 (2007).

17. Barrero, M. J. & Malik, S. Two functional modes of a nuclear receptor-recruited arginine methyltransferase in transcriptional activation. *Mol. Cell.* 24, 233–243 (2006).

18. Sun, K. et al. Phosphorylation of a conserved serine in the deoxyribonucleic acid binding domain of nuclear receptors alters intracellular localization. *Mol. Endocrinol.* 21, 1297–1311 (2007).

---

**Figure 4** Comparison of the HNF-4α homodimer and the PPAR-γ–RXR-α heterodimer complexes on DR1 DNA. **a.** The PPAR-γ–RXR-α heterodimer on DR1. **b.** The HNF-4α homodimer on DR1. **c.** Their overlap when the DR1 sequences are superimposed, showing the distinct domain–domain arrangements in these two complexes. The two complexes are shown in an identical fashion with respect to the DNA sequence facing the viewer.

©2013 Macmillan Publishers Limited. All rights reserved
21. Gineste, R. et al. Phosphorylation of farnesoid X receptor by protein kinase C promotes its transcriptional activity. *Mol. Endocrinol.* **22**, 2433–2447 (2008).

22. Lu, P. et al. Structural basis of natural promoter recognition by a unique nuclear receptor, HNF4α. *J. Biol. Chem.* **283**, 33685–33697 (2008).

23. Rochel, N. et al. Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. *Nature Struct. Mol. Biol.* **18**, 564–570 (2011).

24. Hall, J. M., McDonnell, D. P. & Korach, K. S. Allosteric regulation of estrogen receptor structure, function, and coactivator recruitment by different estrogen response elements. *Mol. Endocrinol.* **16**, 469–486 (2002).

25. Meijling, S. H. et al. DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **324**, 407–410 (2009).

26. Helsen, C. et al. Evidence for DNA-binding domain–ligand-binding domain communications in the androgen receptor. *Mol. Cell. Biol.* **32**, 3033–3043 (2012).

27. Choi, J. H. et al. Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPARγ by Cdk5. *Nature* **466**, 451–456 (2010).

28. Choi, J. H. et al. Antidiabetic actions of a non-agonist PPARγ ligand blocking Cdk5-mediated phosphorylation. *Nature* **477**, 477–481 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by National Institutes of Health grants R01 DK094147 and R01 DK097475.

**Author Contributions** V.C. expressed, purified and crystallized the complex. P.H., along with V.C., solved and refined the structure and carried out the mutational binding studies. Y.K. collected, processed and reduced the X-ray diffraction data, and assisted with the molecular replacement search in structure determination. N.P., along with V.C., made the expression constructs for crystallization and for DNA-binding studies. D.W. carried out the transcription assays. F.R. supervised the work and wrote the manuscript.

**Author Information** Data have been deposited in the Protein Data Bank under accession 4IQR. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.R. (frastinejad@sanfordburnham.org).
METHODS

Expression, purification and crystallization. The HNF-4α proteins used in this study reference the NCBI sequence for HNF4, accession NM_178849. All HNF-4α (human) constructs, including FL (1–464), ΔABF (46–368), ΔF (1–368), ΔAB (46–464) and DBD (46–126) and others were expressed from pET46 Ek/LIC vector in BL21(DE3) E. coli cells (Novagen). The HNF-4α ΔABF (human) construct was used in crystallization experiments. Cells were induced with 0.5 mM IPTG at 17°C for 16 h, and lysed in 20 mM Tris (pH 7.5), 500 mM NaCl, 20 mM imidazole and 10% glycerol. The purification used involved His-Bind resin (Novagen), SP-Sepharose column (GE Healthcare) and gel filtration (Superdex 200). Purified HNF-4α was then combined with oligonucleotide (DR1) in a 1:1.5 molar ratio. The oligonucleotide strands 5'-GGAACTAGGTCAAAGGTCAG-3' and 5'-CCTGACCTTTGACCTAGTTC-3' were purified and annealed. Coactivator peptide (EKHKILHRLLQDSY) was also added in a 3:3 molar ratio. A final gel-filtration step was carried out to remove any excess DNA. Crystals were grown with 5–10% PEG 3350, 25 mM MgCl₂, 25 mM MES pH 6.5 at 4°C.

Data collection and structure determination. Diffraction data was collected at the Argonne National Laboratory SBC-CAT 19ID beamline and the structure was solved by molecular replacement. X-ray data were collected at a wavelength of 0.9793 Å, at 100 K. The backbone dihedral (Ramachandran) angles for the amino acids in the final coordinates conform to preferred, allowed and disallowed statistics of 90.5%, 8.2% and 1.3%, respectively. The details of X-ray diffraction data collection, structure solution and refinement are in Supplementary Table 1.

Fluorescence polarization assay. All the mutants used in DNA-binding assays were prepared using QuikChange Site-Directed Mutagenesis (Stratagene). The DNA strands 5'-GGAACTAGGTCAAAGGTCAG-3' and 5'-CCTGACCTTTGACCTAGTTC-3' were annealed to make the DR1 for binding studies. For binding assays, 2 nM fluorescein-conjugated DNA (5'-end conjugation on the top strand) was incubated with purified HNF-4α protein for 2 h at room temperature (23°C). Protein concentration was varied by serial dilution in binding buffer (20 mM Tris-Cl, pH 7.8, 20 mM NaCl, 8% glycerol, 10 mM DTT). The fluorescence polarization signals were recorded using 96-well black polystyrene plates on FlexStation 3. The data were later converted to fluorescence anisotropy values and normalized. The K_d of each construct and mutants were calculated by fitting the curve in KaleidaGraph 4.1.

Interactions of the A/B and F domains within the HNF-4α complex. For the study shown in Supplementary Fig. 1D, the AB domain (residues 1–45) and F domain (369–465) of human HNF-4α were expressed from pET46 Ek/LIC vector in BL21(DE3) E. coli cells (Novagen). Each protein was purified using a HisTrap column, followed by size-exclusion chromatography on a HiLoad 16/60 Superdex 200 column. The final samples were prepared in 10 mM phosphate (pH 7.0) and 100 mM NaCl buffer for fluorescein labelling. Ten micrograms of each peptide, HNF-4α-AB, HNF-4α-F and PGC-1α LXXLL-motif peptide (AEEPSLLKLLAY, synthetically made and purchased from AnaSpec) were incubated with fluorescein-5-EX succinimidyl ester at a molar ratio of 1:2 in 100 mM potassium phosphate (pH 7.0) coupling buffer at 37°C for 60 min, quenched by adding 100 mM Tris (pH 8.0) and incubated for 30 min at room temperature. The labelled peptides were purified using Sephadex G-15 column for fluorescence polarization measurements. The HNF-4α ΔABF–DNA complex was prepared for the peptide-binding assays as described earlier. Five nanomolar of each fluorescein-labelled peptide was incubated with HNF-4α–DNA complex of serially diluted concentrations for 2 h at room temperature. Recording of fluorescence polarization signals and data processing were conducted as described earlier.

Transcription reporter assays. For the transcriptional reporter studies shown in Supplementary Fig. 11, we used both HEK293T and COS-7 cells that were seeded in 24-well plates and 1 day later transfected with 400 ng of the pCMV-Tag1-HNF-4α wild-type or mutant plasmid, 100 ng of apoCIII-pTKLuc reporter (a gift from D. Kelly) and 10 ng of pRL (control Renilla luciferase) using jetPEI reagent (Polyplus) according to the manufacturer’s protocol. Luciferase activity was measured 48 h after transfection using the Dual-Glo Luciferase Assay System (Promega) and data were normalized by the relative ratio of firefly and Renilla luciferase activity.