Multiple Binding Modes between HNF4α and the LXXLL Motifs of PGC-1α Lead to Full Activation*‡

Received for publication, August 5, 2009, and in revised form, October 2, 2009. Published, JBC Papers in Press, October 21, 2009, DOI 10.1074/jbc.M109.052506

Geun Bae Rha†, Guangteng Wu†, Steven E. Shoelson‡, and Young-In Chi†‡

From the †Department of Molecular and Cellular Biology, Center for Structural Biology, University of Kentucky, Lexington, Kentucky 40536, and the ‡Joslin Diabetes Center and Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215

Hepatocyte nuclear factor 4α (HNF4α) is a novel nuclear receptor that participates in a hierarchical network of transcription factors regulating the development and physiology of such vital organs as the liver, pancreas, and kidney. Among the various transcriptional coregulators with which HNF4α interacts, peroxisome proliferation-activated receptor γ (PPARγ) coactivator 1α (PGC-1α) represents a novel coactivator whose activation is unusually robust and whose binding mode appears to be distinct from that of canonical coactivators such as NCoA/SRC/p160 family members. To elucidate the potentially unique molecular mechanism of PGC-1α recruitment, we have determined the crystal structure of HNF4α in complex with a fragment of PGC-1α containing all three of its LXXLL motifs. Despite the presence of all three LXXLL motifs available for interactions, only one is bound at the canonical binding site, with no additional contacts observed between the two proteins. However, a close inspection of the electron density map indicates that the bound LXXLL motif is not a selected one but an averaged structure of more than one LXXLL motif. Further biochemical and functional studies show that the individual LXXLL motifs can bind but drive only minimal transactivation. Only when more than one LXXLL motif is involved can significant transcriptional activity be measured, and full activation requires all three LXXLL motifs. These findings led us to propose a model wherein each LXXLL motif has an additive effect, and the multiple binding modes by HNF4α toward the LXXLL motifs of PGC-1α could account for the apparent robust activation by providing a flexible mechanism for combinatorial recruitment of additional coactivators and mediators.

Hepatocyte nuclear factor 4α (HNF4α) is a unique member of the nuclear receptor (NR) family that plays a critical role in early vertebrate development and metabolic regulation (1). HNF4α is highly expressed in liver, kidney, intestine, and pancreas, and its crucial role in liver and endocrine pancreas has been demonstrated by genome-wide expression profiling (2) and conditional deletion in mice (3–6). In addition, dysfunction of HNF4α is associated with various metabolic disorders, and mutations in HNF4α cause a dominantly inherited form of diabetes referred to as MODY1 (maturity-onset diabetes of the young) (7). This form of diabetes is mainly characterized by early onset and a defect in insulin production or release, further underscoring the pivotal role of HNF4α in human pancreatic β-cell function and glucose homeostasis (5, 8).

Like other members of the NR superfamily, HNF4α has a modular structure comprising an all-cysteine zinc finger DNA binding domain, a lipophilic ligand binding domain (LBD), and additional domains with activation function (Fig. 1A). HNF4α functions exclusively as a homodimer despite its sequence similarity to retinoid X receptor α (9) and exerts its function through various macromolecular interactions utilizing its modular domains. Its full activity is achieved through the interactions of the HNF4α homodimer with DNA and various coactivators, including NCoA/SRC-1/p160 family members and others that, in turn, recruit the histone acetyltransferase activity containing general coactivators such as CBP/p300 and pCAF and mediators bringing the remainder of the basal transcriptional machinery (10, 11).

The PPARγ coactivators (PGC-1α, PGC-1β, and PGC-1-related coactivator) are also recruited by NRs for enhanced transactivation (12, 13). This unique set of coactivators possess, in addition to the LXXLL motif-containing activation domains, the arginine/serine-rich (RS) region and the RNA-recognition motif, which are involved in RNA processing and mRNA export and further distinguish them from other coactivators such as NCoA/SRC-1/p160 family members (14). PGC-1α, the best studied member of the family, was originally identified through its interaction with PPARγ and its role in thermogenesis within the muscle and brown fat (15, 16). PGC-1α is now known to be a key player in the maintenance of glucose, lipid, and energy homeostasis and has been implicated in pathogenic conditions such as obesity, diabetes, neurodegeneration, and cardiomyopathy (13). The PGC-1α/HNF4α partnership has been demonstrated as being particularly important in various metabolic functions, including glucose homeostasis, lipid metabolism, and energy production.

© 2009 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.
Characterization of HNF4α/PGC-1α Interactions

(A) DNA Binding Domain
HNF4α
AF-1  Zinc finger  Hinge
Ligand Binding Domain
Fatty Acid Binding  AF-2  E-domain
Transactivation

(B) Schematic drawings and sequence alignments of the HNF4α and PGC-1α functional regions.

FIGURE 1. Schematic drawings and sequence alignments of the HNF4α and PGC-1α functional regions. A, functional domain structure of HNF4α and the sequence alignment of the LBD of NRs whose crystal structures in complex with the PGC-1α LXXL motifs are known. The residues making contact with the LXXL motif are highlighted in boxes (green and pink), including two charge clamp residues (pink boxes). B, functional domain structure of PGC-1α and its three NR-boxes containing the LXXL motif or its derivatives. Relative positions of the residues comprising the LXXL motif and the flanking regions by the conventional numbering system are shown at the bottom. The characteristic serine and leucine residues of the class III NR-box are also indicated by circles. Regulatory phosphorylation sites are also indicated. These figures are not on the same scale and thus not proportional to the polypeptide chain length of each protein. The abbreviations used are as follows: AF, activation function; AD, activation domain; NRs, nuclear receptor boxes, LXXL motifs; ID, inhibitory domain; RS, serine/arginine-rich region; and RRM, RNA recognition motif.

pathways, including mitochondrial biogenesis, gluconeogenesis and insulin secretion, and onset of diabetes (17–20).

The main interaction between the NRs and coactivators, including PGC-1α, occurs via a short α-helical segment containing the LXXL sequence motif of the coactivators (NR-box) (21) and helix 12 (AF-2) of the NR-LBDs, which has been heavily targeted for therapeutic interventions (22). Many coactivators have multiple LXXL motifs, and inspections of the PGC-1α sequence reveal three putative LXXL motifs within its N-terminal region (Fig. 1B). These multiple LXXL motifs provide selectivity against various NRs and allow diversity in the spectrum of NR/coactivator interactions (23, 24). PGC-1α activates multiple NRs, including PPARγ (25), retinoid X receptor-α (26), estrogen receptor-α (ERα) (27), and estrogen-related receptor (ERR-α and -γ) (28–30); however, its interaction with HNF4α appears to be novel (31). The PGC-1α activation of HNF4α is unusually robust, with both a tight physical interaction and several orders of magnitude greater increase in transcriptional activity than those mediated by other coactivators

LXXL motif (40, 41) and the PPARγ/PGC-1α LXXL motif (42) provide further evidence for diverse modes of PGC-1α recognition consistent with NR-specific recruitment of additional elements of the transcriptional apparatus.

EXPERIMENTAL PROCEDURES

Plasmid Constructs—Because it has been suggested that the hinge region of HNF4α might also be needed for interaction with PGC-1α (43), we used residues 120–368 of human recombinant HNF4α, containing the hinge region and the LBD, and residues 74–219 of human recombinant PGC-1α, containing all three LXXL motifs (Fig. 1). DNA encoding the fragment of HNF4α with a His6 tag and a tobacco etch virus cleavage site was cloned into the pCDF Duet (Novagen) vector, whereas DNA encoding the fragment of PGC-1α with a GST tag and a tobacco etch virus cleavage site was cloned to pET41a (Novagen) vector. These vectors were cotransformed into Escherichia coli to generate the complex in vivo.
Sample Preparation and Crystallization—Plasmids harboring HNF4α and PGC-1α were cotransformed into BL21 (DE3) under the selection of streptomycin and kanamycin. Both proteins were expressed, and the pure complex was obtained following sequential affinity chromatography on Talon beads (Clontech), tobacco etch virus protease digestion, ion exchange, and gel filtration. Dynamic light scattering data indicated a homogeneously populated complex favorable for crystallization. The pure complexes were dissolved in the optimal freezing conditions (20% ethylene glycol) diffracted to 2.2 Å. The presence of both proteins in the crystals was confirmed by SDS-PAGE followed by silver staining (Fig. 2B). The crystals belong to the space group P4_2_2 with the unit cell dimensions a = b = 113.049 Å, c = 57.322 Å. The best crystals after optimal freezing conditions (20% ethylene glycol) diffracted to 2.2 Å resolution with the synchrotron radiation (collected at beamline SER-CAT 22ID, Advanced Photon Source, Argonne, IL). There is one HNF4α-LBD-ligand-PGC-1α-NR-box fragment ternary complex (half of the functional dimeric complex) in the crystal asymmetric unit resulting in 40.5% solvent content. Oscillation images (every 0.5°) were collected at 100 K, and the data were processed using the HKL2000 software package (Table 1) (45).

Structure Determination and Refinement—The structure was solved by the molecular replacement using MOLREP (46). Our previous structure of human HNF4α-LBD (Protein Data Bank accession code 1PZL) (47), after deleting a bound ligand and the SRC-1 peptide, was used as a search model. The best solution was obtained with a correlation coefficient of 50.3%, 17.0% above the second best solution. The R_cry value after the molecular replacement was 0.51. After one round of rigid-body refinement by the CNS program (48), R_cry and R_free dropped to 0.43 and 0.44, respectively. Difference maps were calculated, and the electron densities corresponding to the PGC-1α fragment and the bound fatty acid were clearly visible (Fig. 2C). Model building was done with O (49), and the refinement continued by simulated annealing using CNS, with restraints placed on bond lengths, bond angles, nonbonded contacts, and temperature factors of neighboring atoms. Inclusion of individual atomic temperature factors and addition of solvent atoms during the later stages of refinement were accompanied by a substantial decrease in R_free values. TLS (translation/libration/screw rotation) refinement was applied to the final models using Refmac5 (50) with rigid body assignments recommended by the TLSMD server (51). Structure validation tools such as MolProbity (52) and WHAT_CHECK (53) were used to assess the quality of the structure; the final refinement statistics are provided in Table 1.

NR-box Mutant Generation—The QuickChange multisite-directed mutagenesis kit (Stratagene) was used to generate the constructs with mutations in the LXXLL motifs (to LXXAA) according to the manufacturer’s instructions. The plasmid templates used in the mutagenesis protocol were pET41a GST-PGC-1α-(74–219) and pCMV Sport6 PGC-1α-(full-length). All constructs with the mutated sequences were verified by DNA sequencing.

GST Pulldown Assays—HNF4α (120–368) and human PPARγ (187–477) GST fusion proteins were produced in E. coli, although full-length wild-type and mutated 35S-labeled PGC-1α proteins were produced with a TntT reticulocyte lysate in vitro transcription and translation kit (Promega). About 1 μg of fusion protein was mixed with 30 μl of the in vitro translate in a binding buffer containing 50 mM Tris buffer (pH 8.0), 100 mM NaCl, 0.5 mM EDTA, and 0.1% Nonidet P-40. Binding was performed for 5 h at 4 °C, and the mixtures were loaded onto the beads containing glutathione. Immobilized GST protein was detected by SDS-PAGE and Western blotting using Coomassie Blue staining and probing with a GST-specific antibody (GE Healthcare). For PPARγ and PGC-1α binding in the supplemental material, 1 μM trilgiltazone was added as an agonist for PPARγ. The beads were then washed multiple times with the binding buffer and resuspended in SDS-PAGE sample buffer. After electrophoresis, the 35S-labeled proteins were detected by autoradiography.

Cell Culture—HeLa cell line was cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum (Invitrogen), 50 units/ml penicillin G, 50 μg/ml streptomycin (Sigma), and 0.1 mM nonessential amino acids (Invitrogen).

Transient Transfection and Luciferase Reporter Assays—The luciferase constructs were prepared as described previously (54). Briefly, the HNF4α-binding element (−64 to −52) from the human Hnf1α promoter (−298 to the first AGT) was subcloned into the firefly luciferase reporter vector pGL3-Basis (Promega) (pGL3-HNF1α). The full-length cDNA of human PGC-1α (wild-type or mutants) was subcloned into the pCMV Sport6 vector (Invitrogen). For transcription assays, HeLa cells were transfected using Opti-MEM and Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer’s recommendations. A total of 20 ng of pcDNA3 HNF4α and PCMV Sport6 PGC-1α wild-type or mutants, 50 ng of pGL3-HNF1α, and 5 ng of pRL-TK (control Renilla luciferase vector) were used for transfection of 1 × 10^5 cells seeded on 24-well plate 1 day before transfection. Forty eight h after transfection, cells were washed with phosphate-buffered saline and lysed with luciferase lysis buffer supplied with the luciferase assay kit (Promega). Luciferase activity was measured using Dual-Luciferase assay system (Promega) and LMax luminometer (Molecular Devices). All values were normalized by the relative ratio of firefly luciferase activity and Renilla luciferase activity. At least three independent transfections were performed in quadruplicate. For PPARγ transfection in the supplemental material, a total of 25 ng of pcDNA3.1 PPARγ and 250 ng of pCMV Sport6 PGC-1α wild-type or mutants, 25 ng of pGL3 PP1RE (AB)_2, and 10 ng of pRL-TK (control Renilla luciferase vector) were used for transfection of 1 × 10^5 cells seeded on a 24-well plate 1 day before transfection. Five h after transfection, medium was removed, and 1 μM trilgiltazone was added. For statistical analysis, data were expressed as means ± S.E. of at least three independent groups. Statistical significance was determined by one-
way analysis of variance followed by Student-Newman-Keuls method using SigmaStat 3.1 software (Systat Software, San Jose, CA). $p < 0.01$ denoted the presence of a statistically significant difference.

Isothermal Titration Calorimetry (ITC)—Protein/protein interactions were quantitatively measured by ITC using a high precision VP-ITC titration calorimeter (GE Healthcare). Both purified proteins were prepared in crystallization optimal buffer containing 20 mM BisTris (pH 6.0), 100 mM NaCl, and 5% glycerol. These samples were degassed for 5 min and pre-cooled to 5 °C below the actual experimental temperature of 20 °C prior to sample loading. A defined amount (typically 1 ml) of HNF4α recombinant protein was titrated against 1:20 molar (or higher for weaker interactions) concentration of PGC-1α wild type or the mutant proteins in the thermostated cell by means of 25 individual injections via a syringe. No exogenous ligands were added for the experiments. The stoichiometry, thermodynamic parameters, and binding affinities were calculated by fitting the data using the software Origin 7.0 (MicroCal). The binding isotherm was fitted by a nonlinear least square regression using the One Set of Sites model. The binding Gibbs free energy ($\Delta G$) was calculated from enthalpy changes ($\Delta H$) and association constant ($K_a$) through the equation, $\Delta G = -RT \ln K_a$, where $R$ was the gas constant and $T$ was the absolute temperature in Kelvin. Meanwhile, the stoichiometry ($n$) of the interaction was determined from the One Set of Sites model.

RESULTS

Overall Structure of the HNF4α-LBD-PGC-1α Complex—The crystallization methods and structure determination are described under “Experimental Procedures”; typical crystals and the final refinement statistics are shown in Fig. 2A and Table 1, respectively. Even though our gel filtration analysis suggests that all three LXXLL motifs are needed for stable complex formation (supplemental Fig. S1), the crystal structure of HNF4α-LBD-PGC-1α complex structure shows only one LXXLL motif bound to the canonical binding pocket (Fig. 2, C and D). The overall structure is thus very similar to that of the HNF4α-LBD-SRC-1 peptide complex we reported previously (supplemental Fig. S2) (47).

Despite the presence of three NR-boxes in PGC-1α protein used for crystallization (Fig. 2B), only one NR-box is bound to HNF4α. Electron density is visible only for the residues of the core LXXLL motif within the canonical coactivator binding
Characterization of HNF4α/PGC-1α Interactions

Lys-194 (H3) (Fig. 3, A and B). This charge clamp made by two invariant residues of NRs (glutamate and lysine) is a universally conserved feature that provides the capping interactions at each end and at the same time complement the dipole moment of the LXXLL motif-containing α-helix. Recently, it has been suggested that PGC-1α physically and functionally interacts with the hinge region between the central DNA binding domain and the LBD of ER (27, 31), thyroid hormone receptor (35), and HNF4α (43) (amino acids 160–175 of HNF4α in fact is the first helix of the LBD, see Fig. 1A). However, these proposed additional interactions are not observed in our crystal structure despite the presence of well ordered HNF4α in this region. The additional binding surface made by the H8-H9 loop in the recently solved ERRα/PGC-1α NR3 peptide structure (41) is not observed in our structure either.

Another unique feature of each NR-LBD structures is the makeup of its lipophilic ligand binding pocket. Our previous crystal structures of HNF4α-LBD alone (57) and in complex with the coactivator SRC-1 peptide (47) revealed free fatty acids (FFAs) as tightly associated ligands for HNF4α, making it constitutively active. Because it was not possible to remove the FFAs without at least partially denaturing the proteins, we believe FFAs in HNF4α serve as structural ligands resembling prosthetic groups. Our new structure in complex with the PGC-1α LXXLL motif shows similarly sequestered bacterial FFAs in the ligand binding pocket (supplemental Fig. S4). The FFAs are anchored in the ligand binding pocket via an interaction between the fatty acid headgroup and the side chain guanidinium group of Arg-226. The carboxyl groups of the FFA headgroup form additional hydrogen bonds with the headgroup form additional hydrogen bonds with the hydroxyl nidinium group of Arg-226. The carboxyl groups of the FFA headgroup form hydrogen bonds with the carboxyl groups of the FFA carbon chains and the protein residues lining the pocket. Both monomers adopt the “closed” active conformation, very similar to the previous crystal structure of HNF4α-LBD in complex with an LXXLL motif peptide from SRC-1 (47). These structures suggest that HNF4α might not require an external ligand for activation, consistent with the high basal activation properties of HNF4α (58, 59).

With regard to binding stoichiometry of NR-coactivator complexes, a 2:1 complex between the PPARγ homodimer and SRC-1 bearing two NR-boxes has been reported (60). This crystal structure and the related studies with ER homodimers (61) showed that a single p160 protein (e.g. SRC-1) containing multiple LXXLL motifs can interact, in a highly cooperative manner, with both subunits of NRs through two different LXXLL motif interactions from the same SRC-1 molecule with the canonical binding pocket of each NR monomer, rendering a 2:1 stoichiometry (this also proves that variant L

### TABLE 1

Data and refinement statistics

| Parameter | Total | Outer shell |
|-----------|-------|-------------|
| Resolution | 40.2 to 2.2 Å | 2.28 to 2.20 Å |
| Rmerge<sup>a</sup> | 9.0% | 34.9% |
| Completeness | 98.4% | 87.0% |
| I/σ(I) | 29.1 | 4.5 |
| Redundancy | 10.7 | 4.6 |

<sup>a</sup> Rmerge = \( \sum_i |\langle h_i(h) \rangle - \langle \langle h \rangle \rangle | / \sum_i |\langle h_i(h) \rangle | \), where \( h_i(h) \) is the intensity of reflection \( h_i \).

*b 5% of the reflection data excluded from refinement.
tide chain surrounding dimeric HNF4α molecules to pass through, and the LXXLL motif bound to each HNF4α monomer should belong to individual PGC-1α molecule, suggesting a 1:1 molar ratio of the complex within the asymmetric unit (supplemental Fig. S3).

**Structural Evidence for Binding Different PGC-1α LXXLL Motifs—**
There was a surprise in our structure. Unlike the structure of PPARγ/PGC-1α fragment (42), which shows the selection of a single bound LXXLL motif, a close inspection of our electron density map reveals that the bound LXXLL motif is not a single sequence but an averaged structure of more than one LXXLL motif (Fig. 3, A and B). Although the electron densities for the side chains of a major helix of HNF4α (H7 as a representative portion) are well defined at 2.2 Å resolution (Fig. 3C), those for the side chains of PGC-1α are poorly defined except for the leucine residues that constitute the only common residues within the three NR-boxes (Fig. 3, A and B). The initial $F_0 - F_c$ difference electron density map calculated with an all-Ala model of the PGC-1α fragment showed positive peaks at three leucine positions (Fig. 3A), and the final $2F_0 - F_c$ map showed no detectable electron densities for the side chains of PGC-1α, except for the leucine residues (Fig. 3B). No electron densities were visible for the remaining side chains even at the 0.5σ contour level, and the remaining side chains were modeled as alanines in the final structure (Fig. 3B). Furthermore, the temperature factors of the PGC-1α atoms in the final structure are much higher (almost twice as high) than those of the HNF4α atoms, including H3, H4, and H12 regions that make direct contact with the PGC-1α fragment, suggesting a mixture of more than one PGC-1α fragment with local structural variations.

Our final structure contains only two α-helical turns for PGC-1α that encompass the LXXLL motif and a few flanking residues at either end (peptide numbering −2 to +7; Fig. 1B), only extending to the residues that are tethered by the charge clamp (Fig. 3B). By convention, residues in the NR-boxes are numbered so that the first leucine of the LXXLL motif represents the +1 position, and the resi-
Characterization of HNF4α/PGC-1α Interactions

The overall transactivation potential of the wild-type PGC-1α and each of the NR-box mutants. For NR-box mutants, we mutated one, two, or all three of the individual LXXLL motifs of PGC-1α to LXXAA and tested their functional and biochemical properties. The wild-type and mutated proteins were expressed in HeLa cells, and their transcripational activities were measured using a luciferase reporter system. As shown in Fig. 4, the LL to AA substitutions in any one of the NR-boxes (mt-NR1, mt-NR2, and mt-NR3) reduced transactivation potential to different extents. The most pronounced effect was observed by the mutant that substitutes NR2, followed by NR3 (both with more than 70% reduction), and with the least effect due to substituting NR1. These findings suggest that NR2 and NR3 play dominant roles, but all three NR-boxes are required for full activation. These findings are consistent with the earlier observation that point mutations of the core residues of PGC-1α NR2 diminish but do not abolish the interaction with HNF4α (18). A similar pattern of NR-box preference (NR2 over NR3) was noticed for the interactions of ERRα-mediated transcription and glucocorticoid receptor-mediated transcription by PGC-1α (39, 62), although a dissimilar pattern (NR3 over NR2) was observed for the interactions between ERRα and PGC-1α even though both NR2- and NR3-boxes of PGC-1α are required for full activation (34, 39). We also repeated a similar set of experiments for the activation of PPARγ-mediated transcription by PGC-1α, and the results are shown in the supplemental Fig. S5. The data also suggest the requirement of all three NR-boxes for full activation. However, the NR-box preference by PPARγ appears to be different from that of HNF4α (i.e. NR2 > NR1 > NR3 versus NR2 > NR3 > NR1), and, unlike HNF4α, mutating all three NR-boxes of PGC-1α does not completely abolish binding and transactivation toward PPARγ indicating a potential involvement of other regions of PGC-1α for interaction with PPARγ.

Double NR-box mutations (mt-NR1+2, mt-NR2+3, and mt-NR1+3) were also quite informative as to the contribution of each NR-box and their inter-dependence. They all showed very limited transactivation, suggesting that one LXXLL motif is not sufficient for optimal transactivation. Only when more than one LXXLL motif is involved (e.g. PGC-1α wild type or...
Characterization of HNF4α/PGC-1α Interactions

mt-NR1 in Fig. 4) can significant transcriptional activity be measured. This suggests that the individual LXXLL motifs might act in a synergistic way, especially between NR2 and NR3. More evidence comes from the following observations. Even though the additional substitution in NR1 had little effect on NR2- or NR3-mediated transactivation (mt-NR2 versus mt-NR1+2 or mt-NR3 versus mt-NR1+3 in Fig. 4), the additional substitution in NR3 showed further reduction in the NR2-mediated transactivation (mt-NR2 versus mt-NR2+3), further underscoring a potential synergistic activity between NR2 and NR3. The coinvolvement of more than one NR-box in the construct failed to detect any additional interactions. Our data suggest that only one LXXLL motif of coactivators and the AF-2 helix of the current model of transcriptional regulation by NR can be described by combinatorial regulation involving multiprotein complexes (65, 66). These orderly events of complex formations are dynamic, yet sufficiently precise in their molecular recognition utilizing suitably formed molecular surfaces (platform) as well as chemical moieties (e.g. phosphorylation and methylation) and binding motifs (e.g. LXXLL motif and DNA/RNA-binding motifs). Among these recognitions, LXXLL motif-mediated molecular interactions between NRs and the NR-boxes of coactivators are best characterized and have provided excellent opportunities for therapeutic intervention (22, 55).

Finally, to further examine these interactions in more detail, the binding energetic of each NR-box was measured by ITC experiments with the recombinant proteins of HNF4α-LBD (residues 120–368 containing the hinge region) and PGC-1α wild-type or the mutants (residues 74–219) bearing the mutations within the LXXLL motifs. Purification of each protein to homogeneity and accurate measurement of its concentration enabled quantitative determination of binding affinity and associated thermodynamic parameters upon binding through the ITC experiments. The raw ITC data best fit the model for a single class of binding sites and were analyzed accordingly.

In agreement with the results of the transcription assays and the pulldown experiments, the quantitative analysis of in vitro bindings also showed a similar pattern of NR-box preference for binding (Fig. 5B and Table 2). The binding of HNF4α-LBD to the NR2 of PGC-1α ($K_a = 3.33 \times 10^5$ M$^{-1}$) was 6-fold stronger than the binding to the NR3 of PGC-1α ($K_a = 0.55 \times 10^5$ M$^{-1}$). The interaction with NR1 by itself, in the context of the NR2+3-substituted mutant (mt-NR2+3), was too weak to be measured by this method. Likewise, substitution of all three NR-boxes completely ablated the binding, and no heat change was detected. More importantly, the additive effects of each NR-box on overall binding affinities were also observed in these measurements. The native binding affinity by the wild type was considerably higher than any NR-box alone or a combination of two (around 2-fold increase). The exact mechanism of synergistic enhancement is not known. Again, the presence of additional weak binding sites cannot be ruled out. However, if the second NR-box-binding site does exist, then its binding appears to be too weak to be detected in our crystal structure (crystallography is not suited to observe weak intermolecular interactions unless the interaction sites are fully occupied within the crystal lattice). In all cases, ITC measured the global thermodynamic parameters of a system that is consistent with a combined contribution by the hydrophobic interactions (positive $\Delta TS$ or negative $-\Delta TS$; entropically favorable) and the formation of hydrogen bonds through the charge clamp (negative $\Delta H$; enthalpically favorable). Finally, the best fit of the binding stoichiometry ($N$) was one within experimental error in all cases, confirming a 2:2 stoichiometry of the complex because the functional unit of HNF4α is a homodimer.

**DISCUSSION**

The current model of transcriptional regulation by NR can be described by combinatorial regulation involving multiprotein complexes (65, 66). These orderly events of complex formations are dynamic, yet sufficiently precise in their molecular recognition utilizing suitably formed molecular surfaces (platform) as well as chemical moieties (e.g. phosphorylation and methylation) and binding motifs (e.g. LXXLL motif and DNA/RNA-binding motifs). Among these recognitions, LXXLL motif-mediated molecular interactions between NRs and the NR-boxes of coactivators are best characterized and have provided excellent opportunities for therapeutic intervention (22, 55).

Various crystal structures of the NR-LBDs bound to agonists and peptides containing the LXXLL motifs have revealed a highly conserved binding mode of coactivator recruitment utilizing a single LXXLL motif of coactivators and the AF-2 helix of NR-LBDs (55). However, why there are multiple LXXLL motifs in the coactivators (within a relatively close proximity) and how they are used during the NR-mediated transcription activation have not been concisely explained, and our findings provide some answers to these questions.

Our data suggest that only one LXXLL motif of PGC-1α is bound to HNF4α despite having all three available for interac-
tions and that the bound LXXLL motif is an averaged structure of more than one NR-box. This is the first time multiple binding modes (thus a lack of strict selectivity) toward the LXXLL motifs have been structurally observed. This discovery was further corroborated by the ensuing functional studies that indicate that each LXXLL motif (notably NR2 and NR3) is capable of binding in a ligand-independent manner, but driving only minimal transactivation. Even though NR2 of PGC-1α appears to function as a primary site of interaction for HNF4α, the neighboring NR-boxes have an additive role in binding and overall transactivation, and the optimal transactivation requires more than one LXXLL motif (especially NR2 and NR3). Furthermore, the full scale integrity of the interactions and full transactivation require all three LXXLL motifs.

It is well known that each NR coactivator has specific receptor preferences, and each LXXLL motif interacts with NR in a different manner, which provides diversity in the spectrum of NR/coactivator interactions for NR-specific responses upon specific ligand binding in a signal- and tissue-specific manner (23, 24). PGC-1α is recruited by many different NRs, and each NR exhibits a differential preference for the NR-boxes of PGC-1α. For example, the NR3 binds more efficiently with ERRα and ERRγ (34, 39–41), whereas the NR2 appears to be the major contributor to the PGC-1α interaction with the LBD of many NRs, including ERα (27), TRα (35), LXR (38), and PPARγ (67). Our findings prove that PGC-1α NR2 serves as a primary NR-box for HNF4α interaction, although others predominantly use a selected NR-box proven by complete annihilation of its functional activity by the mutations on a single NR-box (26, 38).
Previously, a series of systematic studies have shown that the flanking sequences right outside the LXXLL motif are the key determinants of the binding affinity and specificity of coactivators for their NR partners (24, 69), and these LXXLL motifs have been divided into four different classes based on immediate flanking residue sequences (especially by the residues at positions $i_1$ and $i_2$ from the LXXLL motif; Fig. 1B) (23, 70). According to this classification, the amino acid sequence of PGC-1α NR2 is a prototypical example of the class III of NR-boxes that contains Ser and Leu/Ile/Val at position $i_2$ and $i_1$, respectively (Fig. 1B). On the other hand, NR3 could be considered as an atypical LXXYL motif-containing the NR-box instead of LXXLL with the inverse orientation. In fact, the crystal structures of PGC-1α NR3 displayed the right orientation of the LXXYL motif when it binds to ERRα (40, 41). A similar finding was made with an LXXYL motif from another NR coactivator when the crystal structure of SRC2 NR3 in complex with ERRα displayed the LXXYL motif in right orientation (71). Fur-

### TABLE 2

|          | $K_d$ ($\times 10^3$ M$^{-1}$) | $\Delta H$ (kcal/mol) | $-T \Delta S$ (kcal/mol) | $N^b$ |
|----------|-------------------------------|------------------------|--------------------------|-------|
| WT       | 8.14 ± 1.14                   | -2.01 ± 0.15           | -5.92                    | 1.14 ± 0.06 |
| mt-NR1   | 5.13 ± 0.80                   | -1.95 ± 0.20           | -5.60                    | 1.01 ± 0.06 |
| mt-NR2   | 0.68 ± 0.20                   | -2.45 ± 3.70           | -3.17                    | 0.95 ± 0.12 |
| mt-NR3   | 4.31 ± 0.62                   | -1.49 ± 0.18           | -5.33                    | 1.13 ± 0.07 |
| mt-NR1+2 | 0.55 ± 0.09                   | -1.37 ± 0.28           | -4.98                    | 0.94 ± 0.10 |
| mt-NR1+3 | 3.33 ± 0.22                   | -1.52 ± 0.04           | -6.04                    | 1.04 ± 0.02 |
| mt-NR2+3 | N/A                           | N/A                    | N/A                      | N/A   |
| mt-NR1+2+3 | N/A                        | N/A                    | N/A                      | N/A   |

$^a$ Data were determined at 20 °C and at pH 6.0 as described under “Experimental Procedures.” The reported values contain some degree of errors indicated by the standard deviation. N/A means not applicable; WT means wild type.

$^b$ The apparent stoichiometry from the curve fitting data is shown.

Previously, a series of systematic studies have shown that the flanking sequences right outside the LXXLL motif are the key determinants of the binding affinity and specificity of coactivators for their NR partners (24, 69), and these LXXLL motifs have been divided into four different classes based on immediate flanking residue sequences (especially by the residues at positions $i_1$ and $i_2$ from the LXXLL motif; Fig. 1B) (23, 70). According to this classification, the amino acid sequence of PGC-1α NR2 is a prototypical example of the class III of NR-boxes that contains Ser and Leu/Ile/Val at position $i_2$ and $i_1$, respectively (Fig. 1B). On the other hand, NR3 could be considered as an atypical LXXYL motif-containing the NR-box instead of LXXLL with the inverse orientation. In fact, the crystal structures of PGC-1α NR3 displayed the right orientation of the LXXYL motif when it binds to ERRα (40, 41). A similar finding was made with an LXXYL motif from another NR coactivator when the crystal structure of SRC2 NR3 in complex with ERRα displayed the LXXYL motif in right orientation (71). Fur-

**FIGURE 6.** Model scheme of “expanded combinatorial recruitment” for enhanced transactivation. When the coactivator exerts multiple binding modes, it provides a flexible mechanism for combinatorial recruitment and renders an elevated number of opportunities to recruit various additional coactivators and mediators (depicted by various small objects on top of the NR-PGC-1α complex) that can bring the main transcription machinery to the transcription initiation site. Dimeric NRs are shown in red, although PGC-1α in multiple orientations are shown in green. The magenta crescents represent the mediator complex, and the blue objects represent the main transcriptional machinery, including RNA polymerase II (Pol II) and the associated general transcription factors. The remaining small symbols represent various coactivators that can be assembled and some of which have chromatin remodeling activities. This model is focused on the LXXLL motif-mediated interactions between PGC-1α and NRs, and the potential involvements of different regions of two proteins for additional interactions are not included.
thermore, additional structures containing the Leu/Phe or Leu/Met swapping within the NR-box were observed for their complexes with androgen receptor, ERα, and retinoid X receptor α (56, 72–75). However, unlike NR2 and NR3, PGC-1α NR1 does not fall into any distinct class of the NR-boxes. PGC-1α NR1 contains an LXXVI motif and does not match any of the consensus sequences of the NR-box classification. Even though NR1 has an inverse LXXLL sequence and the inverted binding mode cannot be ruled out, none of such binding mode has been previously observed.

From these findings, it appears that some NRs have an exclusive selectivity for the LXXLL motif, although others have a loose selectivity. This flexibility in NR/coactivator recognitions can be attributed to rather simple structural features driving the LXXLL motif-mediated complementary interactions between an amphipathic α-helix and a hydrophobic groove. A small variation in each NR binding pocket can generate a unique local environment that allows NR-specific recruitment of coactivators. These LXXLL motifs could have functional redundancy in activation of NR-mediated transcription; however, our findings suggest that the LXXLL motifs are not functionally equivalent. Instead, they have unique properties and exert additive effects for optimal transactivation when they are recruited to NRs. The exact molecular mechanism of this additive effect is not known; however, multiple binding modes by HNF4α toward PGC-1α LXXLL motifs can provide some answers.

Recognition of NR-boxes in more than one alternative way should render various molecular surfaces of the complex that can serve as multiple platforms for more effective and versatile recruitment of additional coactivators, mediators, and the remainder of the main transcriptional machinery (Fig. 6). In this model, differential arrangements of the transactivation complex can be made and provide different platforms for adaptable and more efficient recruitment of the primary coactivators and the secondary complexes with the coactivator-associated proteins such as CBP/p300, CARM1, and PRMT1, which can further modify the target proteins (68). In addition, these multiple binding modes by HNF4α should have an advantage when it has to compete with other NRs for binding to the LXXLL motifs of common coactivators. It is also possible that the HNF4α dimer could simultaneously recognize the NR-boxes from two different coactivators, which can further augment the overall transactivation potential exerted by individual coactivators. All together, these combinatorial effects should lead to enhanced transactivation and could account for the apparent robust activation of HNF4α-mediated transcription by PGC-1α (Fig. 6).

Acknowledgments—We thank the staff at SER-CAT beamlines 22-ID and 22-BM for data collection. We are grateful to Verna Frasca for assistance in analyzing ITC data and Dan Noonan for assistance in the transcriptional reporter assays.

REFERENCES
1. Sladek, F. M., and Seidel, S. D. (2001) in Nuclear Receptors in Genetic Disease (Burris, T. P., and McCabe, E. R. B., eds) Academic Press, New York
2. 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., and Young, R. A. (2004) Science 303, 1378–1381
3. Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M., and Gonzalez, F. J. (2001) Mol. Cell. Biol. 21, 1393–1403
4. Parviz, F., Matullo, C., Garrison, W. D., Savatski, L., Adamson, J. W., Ning, G., Kaestner, K. H., Rossi, J. M., Zaret, K. S., and Duncan, S. A. (2003) Nat. Genet. 34, 292–296
5. Gupta, R. K., Vatamaniuk, M. Z., Lee, C. S., Flaschen, R. C., Fulmer, J. T., Matschinsky, F. M., Duncan, S. A., and Kaestner, K. H. (2005) J. Clin. Invest. 115, 1006–1015
6. Miura, A., Yamagata, K., Kakei, M., Hatakeyama, H., Takahashi, N., Fukui, K., Nammo, T., Yoneda, K., Inoue, Y., Sladek, F. M., Magnunson, M. A., Kasai, H., Miyagawa, J., Gonzalez, F. J., and Shimomura, I. (2006) J. Biol. Chem. 281, 5246–5257
7. Yamagata, K., Furuta, H., Oda, N., Kaisaki, P. J., Menzel, S., Cox, N. J., Fajans, S. S., Signorini, S., Stoffel, M., and Bell, G. I. (1996) Nature 384, 458–460
8. Wang, H., Macchel, P., Antinozzi, P. A., Hagenfeldt, K. A., and Wollheim, C. B. (2000) J. Biol. Chem. 275, 35953–35959
9. Bogan, A. A., Dallas-Yang, Q., Ruse, M. D., Jr., Maeda, Y., Jiang, G., Nepomuceno, I., Scanlan, T. S., Cohen, F. E., and Sladek, F. M. (2000) J. Mol. Biol. 302, 831–851
10. McKenna, N. J., and O’Malley, B. W. (2002) Cell 108, 465–474
11. Barrero, M. J., and Malik, S. (2006) Mol. Cell 24, 233–243
12. Handschin, C., and Spiegelman, B. M. (2006) Endocr. Rev. 27, 728–735
13. Lin, J., Handschin, C., and Spiegelman, B. M. (2005) Cell Metab. 1, 361–370
14. Knutti, D., and Kralli, A. (2001) Trends Endocrinol. Metab. 12, 360–365
15. Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R. C., and Spiegelman, B. M. (1999) Cell 98, 115–124
16. Puigserver, P., Wu, Z., Park, C. W., Graves, R., Wright, M., and Spiegelman, B. M. (1998) Cell 92, 829–839
17. Rhee, J., Inoue, Y., Yoon, J. C., Puigserver, P., Fan, M., Gonzalez, F. J., and Spiegelman, B. M. (2003) Proc. Natl. Acad. Sci. U.S.A. 100, 4012–4017
18. Yoon, J. C., Puigserver, P., Chen, G., Donovan, J., Wu, Z., Rhee, J., Adelmant, G., Stafford, J., Kahn, C. R., Granner, D. K., Newgard, C. B., and Spiegelman, B. M. (2001) Nature 413, 131–138
19. Soyal, S., Krempler, F., Oberkofer, H., and Patsch, W. (2006) Diabetologia 49, 1477–1488
20. Yoon, J. C., Xu, G., Deeney, J. T., Yang, S. N., Rhee, J., Puigserver, P., Levens, A. R., Yang, R., Zhang, C. Y., Lowell, B. B., Berggren, P. O., Newgard, C. B., Bonner-Weir, S., Weir, G., and Spiegelman, B. M. (2003) Dev. Cell 5, 73–83
21. Heery, D. M., Kalkhoven, E., Hoare, S., and Parker, M. G. (1997) Nature 387, 733–736
22. Gronemeyer, H., Gustafsson, J. A., and Laubet, V. (2004) Nat. Rev. Drug Discov. 3, 950–964
23. Chang, C., Norris, J. D., Gres, H., Paige, L. A., Hamilton, P. T., Kenan, D. I., Fowlkes, D., and McDonnell, D. P. (1999) Mol. Cell. Biol. 19, 8262–8239
24. McInerney, E. M., Rose, D. W., Flynn, S. E., Westin, S., Mullen, T. M., Krones, A., Inostroza, J., Torcha, J., Noile, R. T., Assa-Munt, N., Milburn, M. V., Glass, C. K., and Rosenfeld, M. G. (1998) Genes Dev. 12, 3357–3368
25. Vega, R. B., Huss, J. M., and Kelly, D. P. (2000) Mol. Cell. Biol. 20, 1868–1876
26. Delerive, P., Wu, Y., Burris, T. P., Chin, W. W., and Suen, C. S. (2002) J. Biol. Chem. 277, 3913–3917
27. Tcherepanova, I., Puigserver, P., Norris, J. D., Spiegelman, B. M., and McDonnell, D. P. (2000) J. Biol. Chem. 275, 16302–16308
28. Mootha, V. K., Handschin, C., Arlow, D., Xie, X., St Pierre, J., Sihag, S., Yang, W., Altshuler, D., Puigserver, P., Patterson, N., Willy, P. J., Schulman, I. G., Heyman, R. A., Lander, E. S., and Spiegelman, B. M. (2004) Proc. Natl. Acad. Sci. U.S.A. 101, 6570–6575
29. Schreiber, S. N., Knutti, D., Brogli, K., Uhmann, T., and Kralli, A. (2003) J. Biol. Chem. 278, 9013–9018
30. Hentschke, M., Süssens, U., and Borgmeyer, U. (2002) Biochem. Biophys. Res. Commun. 299, 872–879
31. Bourdoncle, A., Labesse, G., Marguèron, R., Castet, A., Cavaillé, V., and
