Transcriptional coactivator PGC-1α contains a novel CBP80-binding motif that orchestrates efficient target gene expression

Hana Cho,[1,2,8] Xavier Rambout,[1,2,8] Michael L. Gleghorn,[1,2,3,8] Phuong Quoc Thuc Nguyen,[4] Christopher R. Phipps,[1,2] Keita Miyoshi,[1,2,7] Jason R. Myers,[5] Naoyuki Kataoka,[6] Rudi Fasan,[4] and Lynne E. Maquat[1,2]

[1]Department of Biochemistry and Biophysics, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, USA; [2]Center for RNA Biology, University of Rochester, Rochester, New York 14642, USA; [3]School of Chemistry and Materials Science, College of Science, Rochester Institute of Technology, Rochester, New York 14623, USA; [4]Department of Chemistry, University of Rochester, Rochester, New York 14627, USA; [5]Genomics Research Center, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, USA; [6]Laboratory of Cell Regulation, Department of Applied Animal Sciences, Graduate School of Agriculture and Life Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan

Although peroxisome proliferator-activated receptor-γ (PPARγ) coactivator 1α (PGC-1α) is a well-established transcriptional coactivator for the metabolic adaptation of mammalian cells to diverse physiological stresses, the molecular mechanism by which it functions is incompletely understood. Here we used in vitro binding assays, X-ray crystallography, and immunoprecipitations of mouse myoblast cell lysates to define a previously unknown cap-binding protein 80 (CBP80)-binding motif (CBM) in the C terminus of PGC-1α. We show that the CBM, which consists of a nine-amino-acid α helix, is critical for the association of PGC-1α with CBP80 at the 5′ cap of target transcripts. Results from RNA sequencing demonstrate that the PGC-1α CBM promotes RNA synthesis from promyogenic genes. Our findings reveal a new conduit between DNA-associated and RNA-associated proteins that functions in a cap-binding protein surveillance mechanism, without which efficient differentiation of myoblasts to myotubes fails to occur.

Peroxisome proliferator-activated receptor-γ (PPARγ) coactivator 1α (PGC-1α), PGC-1β, and the PGC-related coactivator (PRC) comprise the PGC-1 family of transcriptional coactivators that broadly reprograms gene expression to mediate complex cellular adaptations (e.g., see Correia et al. 2015; Martínez-Redondo et al. 2015). These coactivators rapidly promote mitochondrial energy-generating functions in response to physiological stresses that include exercise, nutrient deprivation, and reactive oxygen species; i.e., cellular environments that become bioenergetically suboptimal with aging and disease (Villena 2015; Kadlec et al. 2016). In particular, elevated PGC-1α levels in the skeletal muscles of transgenic mice subjected to exercise or muscle injury promote oxi-

dative fiber formation at the expense of glycolytic fiber formation (Lin et al. 2002), improve exercise performance (Calvo et al. 2008), increase mass muscle and strength as well as resistance to muscle wasting (Ruas et al. 2012), and augment early steps in the activation and proliferation of adult muscle stem cells (Dinulovic et al. 2016).

Domain organization within PGC-1α and PGC-1β is largely conserved (Michael et al. 2001; Martínez-Redondo et al. 2015; Villena 2015; Long et al. 2016). First, an N-terminal activation domain (AD) binds several transcription factors, including nuclear receptors (NRs), and recruits histone acetyltransferases (HATs), such as p300, to loosen chromatin and facilitate transcription preinitiation complex formation. Second, a middle region contains

regulatory domains that bind additional transcription factors, including the myocyte-specific enhancer factor 2C (MEF2C) and the transcriptional suppressor p160. Third, a C-terminal region harbors one or two arginine–serine-rich regions (the RS domain) that recruit Mediator via its MED1 subunit to activate transcription and also bind RNA. This RS domain is followed by an RNA recognition motif (RRM) of unknown function.

Cap-binding protein 80 (CBP80) and CBP20, which constitute the prototypical heterodimeric cap-binding complex (CBC), bind the 7-methyl guanine "cap" structure that is added to the 5′ end of elongating RNA polymerase II (RNAPII)-synthesized transcripts, including pre-mRNAs and long noncoding RNAs such as enhancer RNAs (eRNAs) (Guttman et al. 2009). The CBC has known roles in virtually every step of mammalian gene expression (Gonatopoulos-Pournatzis and Cowling 2014): initiation of gene transcription, pre-mRNA splicing, pre-mRNA 3′ end processing, nuclear mRNA export to the cytoplasm, and the pioneer round of translation as well as nonsense-mediated mRNA decay.

More than 20 years ago, a yeast two-hybrid screen for human proteins that bind human CBP80 identified nuclear cap-binding protein-interacting protein 2 (NIP2) (Kataoka et al. 1995), which we report here is PGC-1β. We mapped the CBP80-binding domain of PGC-1β to an α helix in the extreme PGC-1β C terminus, for which we present the 2.68 Å-resolution X-ray crystal structure in complex with CBC–m^{7}GpppA. Given our findings that PGC-1β and PGC-1α bind CBP80 using an identical α helix, we focused on the functional significance of the PGC-1α CBP80-binding motif (CBM), since PGC-1α function is better characterized than PGC-1β function. We demonstrate that the PGC-1α CBM is required for PGC-1α to associate in cells with the cap of transcripts synthesized from PGC-1α-inducible gene targets in a way that is critical for the PGC-1α-mediated activation of gene transcription. Our findings define PGC-1α as a new conduit through which DNA-associated and RNA-associated proteins control the quality of PGC-1α CBM-responsive gene expression and myogenesis.

Results

PGC-1β interacts directly with CBP80 via a C-terminal nine-amino-acid α helix

Yeast two-hybrid results (Kataoka et al. 1995) had revealed that full-length human CBP80 interacts with the ∼270 C-terminal residues of human PGC-1β (Fig. 1A). To better define this interaction, we showed using GST pull-downs that GST-PGC-1β deletion variants lacking the 30 C-terminal amino acids 994–1023 fail to interact with CBP80 (Fig. 1A,B). Additional pull-downs demonstrated that PGC-1β amino acids 999–1023 are sufficient to interact with CBP80 (Supplemental Fig. S1A). The PSIPRED Web server (Buchan et al. 2013) predicted that an α helix spans PGC-1β amino acids 1010–1020 (Supplemental Fig. S1B). In support of this prediction, circular dichroism spectroscopy demonstrated that amino acids 1009–1021 exhibit

an α-helical structure (26% α-helicity), as evidenced by the presence of a pronounced negative band at 215–222 nm and a positive band at 190–195 nm (Supplemental Fig. S1C) in the presence of the secondary structure-inducing solvent trifluoroethanol (Nelson and Kallenbach 1986). Thus, we synthesized a set of peptides encoding PGC-1β amino acids 1009–1021 in which each nonalanine residue was individually mutated to alanine. Results of fluorescence polarization competition assays revealed that amino acids F1011, D1012, L1014, and L1015 are critical for PGC-1β binding to a preformed complex of CBP80–CBP20; i.e., the CBC (Supplemental Fig. S1D, see PGC-1β-3A, PGC-1β-4A, PGC-1β-6A, and PGC-1β-7A). Reduced affinity by mutating A1018 to a potential α-helical-disrupting glycine (Supplemental Fig. S1D, see PGC-1β-10G) provided additional evidence that PGC-1β binding to the CBC could rely on the integrity of this predicted C-terminal α helix, which we define as the CBM (see below).

To better define the protein–protein interface of PGC-1β bound to cap-bound CBC, we solved the structure of a synthetic PGC-1β peptide spanning amino acids 994–1023 cocrystallized with a complex of CBC–m^{7}GpppA at 2.68 Å resolution and with refinement statistics of 0.18/0.24 R_{model}/R_{free} (Supplemental Table S1). m^{7}GpppA is analogous to the cap in 5′–5′ linkage to the first transcribed adenine nucleotide of an RNAPII-synthesized RNA. The crystal structure of CBP80 in a complex of CBC–m^{7}GpppG (Protein Data Bank [PDB] ID 1H2V) (Mazza et al. 2002b) provided a molecular replacement search model for our structure, to which we added CBP20 from a different CBC–m^{7}GpppG crystal structure (PDB ID 1N52) (Calero et al. 2002). Structural refinement revealed electron density corresponding to an expected α helix at PGC-1β positions 1011–1019 (Fig. 1C; Supplemental Fig. S1E). Strikingly, residues F1011–D1012, L1014–L1015, and A1018 of the PGC-1β CBM identified in fluorescence polarization competition assays (Supplemental Fig. S1D) directly face CBP80 and are positioned on a lipped shelf generated by the first two α helices of CBP80 and the loop that joins them (Fig. 1C; Supplemental Fig. S1F). Binding is largely hydrophobic: CBM residue F1011 interdigitates with CBP80 residues L34, K37, and V38; CBM residue D1012 interacts with CBP80 loop residues, including S42 and C44, that join CBP80 α helices 1 and 2; CBM residue L1014 interacts with CBP80 residues H30, L34, and K37; CBM residue L1015 interacts with CBP80 residues L34, G52, L53, and V56; and CBM residue A1018 interacts with CBP80 residues H30 and V56 (Fig. 1C). Of note, PGC-1β residues 993–1010 and 1020–1023 are likely disordered, since they lacked discernable electron density, making them unsuitable for modeling. Additionally, the K41–C44 loop of CBP80 that interacts with CBM residue D1012 was difficult to model, consistent with its high B factors that reflect potential flexibility or structural variability.

Superimposing the previously published crystal structure of CBC–m^{7}GpppG (PDB ID 1N52) (Calero et al. 2002) on our crystal structure of PGC-1β CBM–CBC–m^{7}GpppA (Fig. 1C) suggests that PGC-1β CBM binding to CBP80 does not appreciably affect the conformation of the CBC (Supplemental Fig. S1G) and thus CBC binding
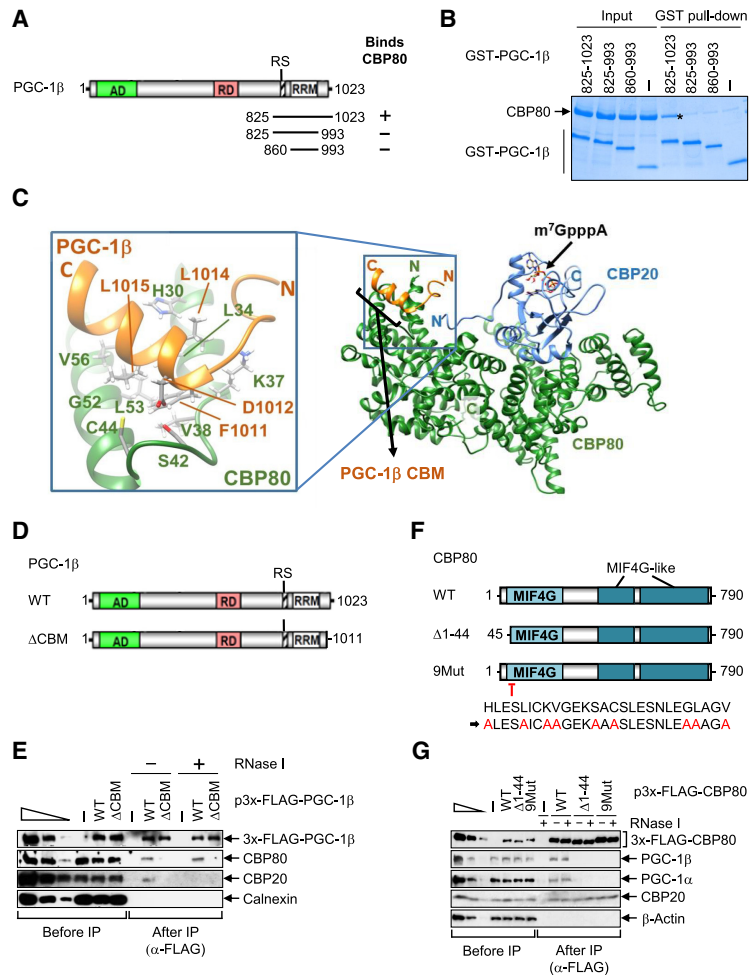
Figure 1. Defining the α-helical CBM of PGC-1β. (*A*) Diagram of human PGC-1β denoting some of the known or putative functional regions. (RD) Negative regulatory domain. Lines *below* the diagram show *Escherichia coli*-produced GST-PGC-1β regions tested for the ability to pull down baculovirus-produced CBP80. Numbers specify amino acids. (*B*) SDS–polyacrylamide separation and Coomassie blue staining of GST pull-downs of *E. coli*-produced GST alone (−) as a negative control or the denoted GST-tagged regions of human PGC-1β in the presence of baculovirus-produced human CBP80. The asterisk denotes the fragment of PGC-1β that binds CBP80. (*C*) Crystal structure of PGC-1β–CBC–m⁷GpppA, where the box encompasses the PGC-1β–CBP80 interface that is enlarged at the *left*. (Green) CBP80; (blue) CBP20; (black) m⁷GpppA; (golden brown) PGC-1β peptide, which includes the nine (FDSLLKEAQ) amino acids at residues 1011–1019 that comprise the CBM. The residues of CBP80 and PGC-1β that interact are detailed and labeled. CBM residue A1018 was omitted because it is hidden under the α-helix ribbon. (*D*) Diagrams (as in *A*) of human PGC-1β wild-type [PGC-1β(WT)] and PGC-1β(ΔCBM). (*E*) Western blots of lysates of MCF-7 cells transiently transfected with plasmid (p) encoding the specified Flag-tagged PGC-1β variant or Flag alone (−) before or after anti-Flag (α-Flag) immunoprecipitation, the latter in the presence (+) or absence (−) of RNase I. Here and in the following Western blots, calnexin served to control for variations in loading and immunoprecipitation specificity, and the *top left* triangle denotes threefold serial dilutions of samples to use for quantitative comparisons. (*F*) Diagrams of human CBP80 wild type, CBP80(Δ1–44), and CBP80(9Mut) denoting some of the known or putative functional regions (Shatsky et al. 2014). (MIF4G) Middle domain of eukaryotic initiation factor 4G (eIF4G). (*G*) Western blots of lysates of MCF-7 cells transiently transfected with plasmid encoding the specified Flag-tagged CBP80 variant or Flag alone (−) before or after anti-Flag immunoprecipitation, the latter in the presence (+) or absence (−) of RNase I. Here and in the following Western blots, β-Actin served to control for variations in loading and immunoprecipitation specificity.

to the cap. Moreover, unlike the CBP20 RRM, the PGC-1β RRM failed to bind the cap analog 7-methyl-GTP (Supplemental Fig. S1H), indicating that it does not directly bind RNA caps. Altogether, our results indicate that the PGC-1β CBM (1) binds CBP80 alone (Fig. 1A,B; Supplemental Fig. S1A) and CBP80 in complex with cap-free CBP20 (Supplemental Fig. S1D) or cap-bound CBP20 (Fig. 1C), (2) does not affect CBC binding to the cap (Supplemental Fig. S1G), and (3) does not directly bind the cap (Supplemental Fig. S1H).

PGC-1β interacts with CBP80 in cells via its CBM

Turning to a more physiological context, anti-Flag immunoprecipitations of lysates of human breast adenocarcinoma MCF-7 cells transiently expressing either 3x-Flag-tagged full-length human PGC-1β—i.e., wild-type Flag-PGC-1β [Flag-PGC-1β(WT)]—or a 3x-Flag-tagged human PGC-1β variant lacking the CBM and the remaining four C-terminal residues—i.e., Flag-PGC-1β (ΔCBM)—at the physiological level revealed that the CBM is required for PGC-1β to associate with CBP80 in

cells (Fig. 1D,E; Supplemental Fig. S1I; Supplemental Table S2). Importantly, the coimmunoprecipitation (co-IP) of CBP80 with Flag-PGC-1β(WT) was resistant to RNase I treatment, demonstrating that the interaction is not bridged or stabilized by RNA. We also found that Flag-PGC-1β(WT), but not Flag-PGC-1β(ΔCBM), coimmunoprecipitates with CBP20 in an RNase I-sensitive manner, consistent with our finding that PGC-1β binding to CBP20 is not direct and is mediated by CBP80.

We next performed anti-Flag immunoprecipitations of lysates of MCF-7 cells transiently expressing at the physiological level either 3x-Flag-tagged full-length human CBP80 [i.e., Flag-CBP80(WT)] or one of two 3x-Flag-tagged human CBP80 variants (Fig. 1F; Supplemental Fig. S1J). One variant, Flag-CBP80(Δ1–44), lacks the first N-terminal α helix that interacts with the PGC-1β CBM, and the other variant, Flag-CBP80(9Mut), harbors an alanine at each of the nine amino acids that directly contact the PGC-1β CBM. While PGC-1β coimmunoprecipitated with Flag-CBP80(WT) in an RNase I-insensitive manner, PGC-1β failed to coimmunoprecipitate with either Flag-CBP80 variant (Fig. 1G). Notably, CBP20 binding to

Flag-CBP80 was unaffected by the mutations in either variant (Fig. 1G), reinforcing the idea that PGC-1β binding to CBP80 does not destabilize the CBC. Taken together, these co-IP experiments confirm that cellular PGC-1β and CBC bind via the C-terminal α helix of PGC-1β and the N-terminal α helices of CBP80, which we call the PGC-1-dependent AD (see below).

The CBM is conserved in PGC-1α and mediates binding to CBP80 in cells

To investigate the biological significance of the CBM, we turned to human PGC-1α and, ultimately, mouse PGC-1α, the latter of which has a better-characterized network of coactivated transcripts than does either human PGC-1β or human PGC-1α. Pull-downs of GST-tagged human PGC-1α deletion variants demonstrated that C-terminal amino acids 768–798 of PGC-1α are necessary for binding to CBP80 (Fig. 2A,B). From these data and the observations that the nine-amino-acid sequence defining the CBM of PGC-1β is conserved in PGC-1α (Supplemental Fig. S1B) and that the C-terminal end of PGC-1α, like that of PGC-1β, forms an α helix in solution (Supplemental Fig. S1C)

and binds the CBC (Supplemental information), we conclude that the human PGC-1α CBM consists of amino acids 785–793. Important to subsequent experiments in this study, human PGC-1α and mouse PGC-1α are 95% identical, and there are no amino acid differences spanning the last 158 C-terminal amino acids between the two proteins.

Corroborating our in vitro results, immunoprecipitations of cellular CBP80 coimmunoprecipitated the longest and CBM-containing PGC-1α isoform (Supplemental Fig. S2A) and, as a positive control, CBP20 from lysates of mouse C2C12 myoblasts (MBs) either without or with formaldehyde cross-linking prior to lysis and in an RNase I-resistant manner (Supplemental Fig. S2B). Thus, binding of this CBM-containing PGC-1α isoform to CBP80 is robust, occurs in intact cells, and is stable after cellular RNA degradation. As a control, RNase I was shown to be active by the failure of poly(A)-binding protein C1 (PABPC1) to coimmunoprecipitate with CBP80 in the presence of RNase I. Furthermore, PGC-1α coimmunoprecipitated with CBP80(WT) but not CBP80(Δ1–44) or CBP80(9Mut) in MCF-7 cell lysates (Fig. 1G), confirming that CBP80 binds PGC-1α and PGC-1β through the same interface.
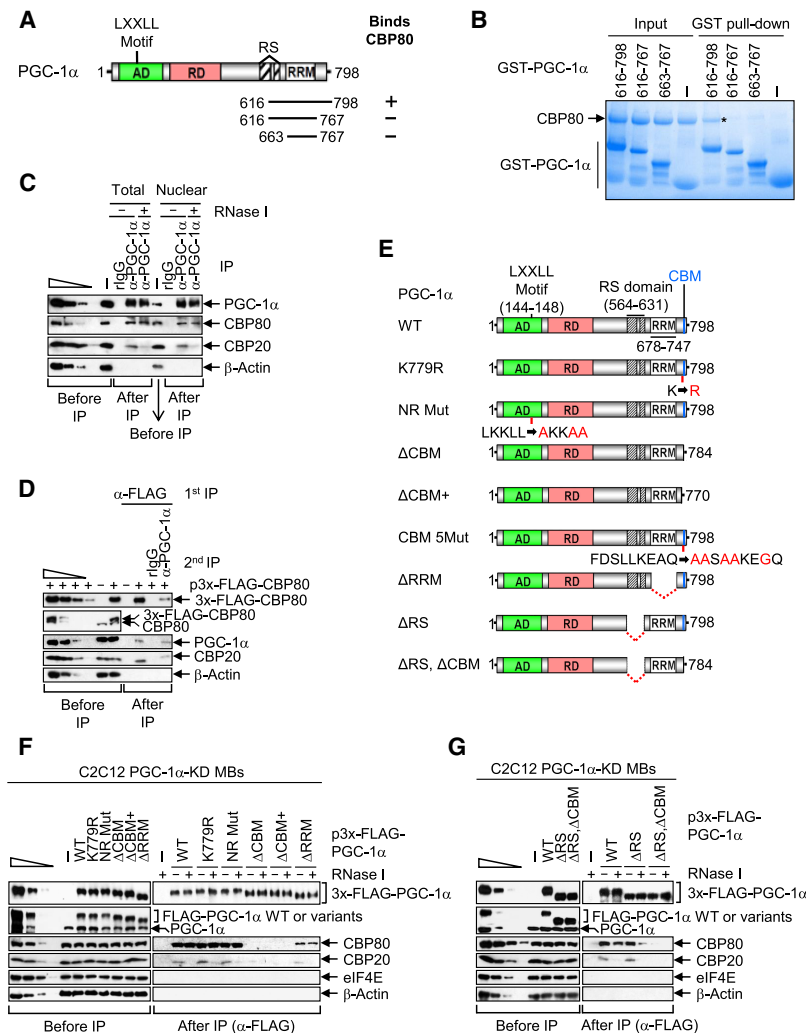

Figure 2. PGC-1α interacts with the CBC largely via its CBM. (A) Diagram of human PGC-1α and its regions tested for the ability to pull down CBP80, as in Figure 1A. (LXXLL motif) α-Helical sequence that binds NRs, where XX represent two lysines in PGC-1α. (B) GST pull-down analyses of CBP80 with GST alone or GST-tagged regions of human PGC-1α, as in Figure 1B. (C) Western blots of total cell lysates or nuclear lysates of C2C12 myoblasts (MBs) before or after immunoprecipitation, the latter in the presence (+) or absence (−) of RNase I, using anti-PGC-1α or, as a control, rabbit IgG (rIgG). (D) Western blots of lysates of C2C12 MBs transiently transfected with plasmid producing Flag-CBP80(WT) (+) or Flag alone (−) either before immunoprecipitation, after a first anti-Flag immunoprecipitation, or after a second immunoprecipitation using anti-PGC-1α or, as a control, rIgG. Samples from the second immunoprecipitation were loaded so that the amounts of PGC-1α in the first and second immunoprecipitations were equivalent. (E) Diagrams (as in A) of human PGC-1α(WT) and variants. (F,G) Western blots of lysates of C2C12 MBs transiently transfected with plasmid encoding the specified Flag-tagged PGC-1α variant or Flag alone (−) before or after anti-Flag immunoprecipitation, the latter in the presence (+) or absence (−) of RNase I. (KD) Knockdown.

Additional immunoprecipitation experiments revealed that nuclear PGC-1α coimmunoprecipitated levels of CBP80 and CBP20 nearly similar to those coimmunoprecipitated by total-cell PGC-1α, indicating that PGC-1α and the CBC interact largely in the nucleus (Fig. 2C; Supplemental Fig. S2C). Consistent with our previous findings with PGC1β, anti-PGC-1α immunoprecipitated CBP80 in a RNase I-insensitive manner and CBP20 in an RNase I-sensitive manner (Fig. 2C). Furthermore, tandem anti-Flag immunoprecipitation followed by anti-PGC-1α immunoprecipitation using lysates of C2C12 MBs expressing Flag-CBP80(WT) demonstrated that PGC-1α associates with CBP80 and CBP20 in the same complex (Fig. 2D).

To determine whether the CBM or other functional domains of PGC-1α are responsible for its binding to the CBC in cells, we transiently expressed a series of 3x-Flag-tagged PGC-1α variants in C2C12 MBs in which PGC-1α was knocked down. To generate PGC-1α knockdown C2C12 MBs, MBs were stably transduced with a lentivirus expressing *PGC-1α* shRNA so as to down-regulate cellular PGC-1α to 30% of its normal level (Supplemental Fig. S2A). We investigated the following PGC-1α variants of the major PGC-1α isoform in C2C12 MBs (Fig. 2E): (1) Flag-PGC-1α(K779R), in which the K779 methylation site is replaced by an arginine, which has been shown to limit PGC-1α binding to eRNAs in C2C12 myotubes (MTs) (Aguilo et al. 2016); (2) Flag-PGC-1α(NR Mut), in which the N-terminal LKKLL NR-binding motif is mutated to AKKAA so as to block PGC-1α binding to NRs and the recruitment of PGC-1α to NR-responsive promoters (Kressler et al. 2002); (3) Flag-PGC-1α(ΔCBM), which lacks the CBM and remaining five C-terminal residues; (4) Flag-PGC-1α(ΔCBM⁺), which lacks the CBM and 14 residues upstream of the CBM, including K779; (5) Flag-PGC-1α (CBM 5Mut), which harbors five amino acid changes in the CBM that we found alter binding to CBP80; (6) Flag-PGC-1α(ΔRRM), which lacks the RRM; (7) Flag-PGC-1α (ΔRS), which lacks the RS domain; and (8) Flag-PGC-1α (ΔRS, ΔCBM), which lacks the RS domain, the CBM, and the remaining five C-terminal residues. Importantly, Flag-PGC-1α(WT) and each variant were expressed in PGC-1α knockdown C2C12 MBs at the physiological level (Fig. 2F,G; Supplemental Fig. S2D). Anti-Flag immunoprecipitations of the K779R or NR Mut variant coimmunoprecipitated CBP80 and CBP20 with the same efficiency and RNase I (in)sensitivity as PGC-1α(WT) (Fig. 2F), indicating that blocking K779 methylation or disrupting the interaction of PGC-1α with NRs does not affect the binding of PGC-1α to the CBC in C2C12 MBs. Anti-Flag immunoprecipitations of the four variants in which the CBM is removed (i.e., ΔCBM, ΔCBM⁺, or ΔRS, ΔCBM) or mutated (CBM 5Mut) failed to coimmunoprecipitate CBP80 and CBP20, as expected (Fig. 2F,G; Supplemental Fig. S2D). In control experiments, PGC-1α(WT), PGC-1α(ΔCBM), and PGC-1α(CBM 5Mut) manifested a similar subcellular localization (Supplemental Fig. S2E, top).

Relative to the immunoprecipitation of PGC-1α(WT), immunoprecipitation of the ΔRRM variant significantly reduced the co-IP of CBP80 and CBP20 (Fig. 2F), and, in the case of the ΔRS variant, the co-IP of CBP80 became RNase I-sensitive (Fig. 2G). Importantly, PGC-1α(ΔRRM) and PGC-1α(ΔRS) were enriched in and depleted from, respectively, the chromatin fraction; i.e., where PGC-1α functions in transcription (Supplemental Fig. S2E, bottom). Taken together, these results indicate that, in addition to the CBM directly binding CBP80, the RRM and RS domain may involve other protein–protein and/or protein–RNA interactions that control the subcellular localization of PGC-1α and/or directly stabilize its association with the CBC in cells. Notably, neither Flag-PGC-1α(WT) nor any of the variants coimmunoprecipitated with eukaryotic translation initiation factor 4E (eIF4E) (Fig. 2F, G), indicating that PGC-1α is removed from newly synthesized transcripts by the time the CBC is replaced by eIF4E and thus is not associated with the bulk of translationally active cytoplasmic mRNAs (Lejeune et al. 2002; Sato and Maquat 2009). These results are consistent with our previous observation that PGC-1α and CBP80 largely associate in the nucleus (Fig. 2C).

PGC-1α associates with the cap of nascent transcripts via its CBM

Since PGC-1α interacts with the CBC in cells, we tested whether PGC-1α physically associates with transcripts whose synthesis it directly regulates. Chromatin immunoprecipitation (ChIP) coupled to quantitative PCR (ChIP-qPCR) using C2C12 MTs revealed that PGC-1α binds to the promoters of genes encoding the metabolic enzymes isocitrate dehydrogenase 3b (IDH3B), 6-phosphofructokinase (PFKL), and sirtuin 5 (SIRT5) as well as to regions situated ≥4 kb upstream of each gene that produce eRNAs (Supplemental Fig. S3A; Aguilo et al. 2016). Consistent with these results, RT-qPCR revealed that anti-PGC-1α immunoprecipitates *Idh3b*, *Pfkl*, and *Sirt5* pre-mRNAs, mRNAs, and their corresponding eRNAs but not *Hprt* or *β-Actin* pre-mRNAs and mRNAs, which are not regulated by PGC-1α [Fig. 3A,B; Supplemental Fig. S3B–F; Aguilo et al. 2016; data not shown for poly (A)⁺ RNA]. Importantly, tandem anti-Flag-CBP80 immunoprecipitation followed by anti-PGC-1α immunoprecipitation using lysates of C2C12 MBs expressing Flag-CBP80 (WT) demonstrated that PGC-1α associates with its responsive transcripts but not nonresponsive transcripts in complex with CBP80 (Fig. 3C,D).

Using the same transfection conditions that were used to test the importance of PGC-1α domains for the interaction of PGC-1α with CBP80 and with a focus on potential RNP-interacting domains—i.e., the CBM, the RS domain, and the RRM—that contribute to the association of PGC-1α with CBP80 in cells (Fig. 2F,G), RT-qPCR after anti-Flag immunoprecipitation demonstrated that the CBM is required for most, if not all, of PGC-1α binding to its target pre-mRNAs, mRNAs (Fig. 3E,F), and eRNAs (Supplemental Fig. S3G). In general, deleting the PGC-1α RRM or RS domain moderately limited PGC-1α binding to target transcripts (Fig. 3E,F; Supplemental Fig. S3G), reflecting roles for these domains in the association with CBP80 in cells. Deleting both the CBM and the RS domain
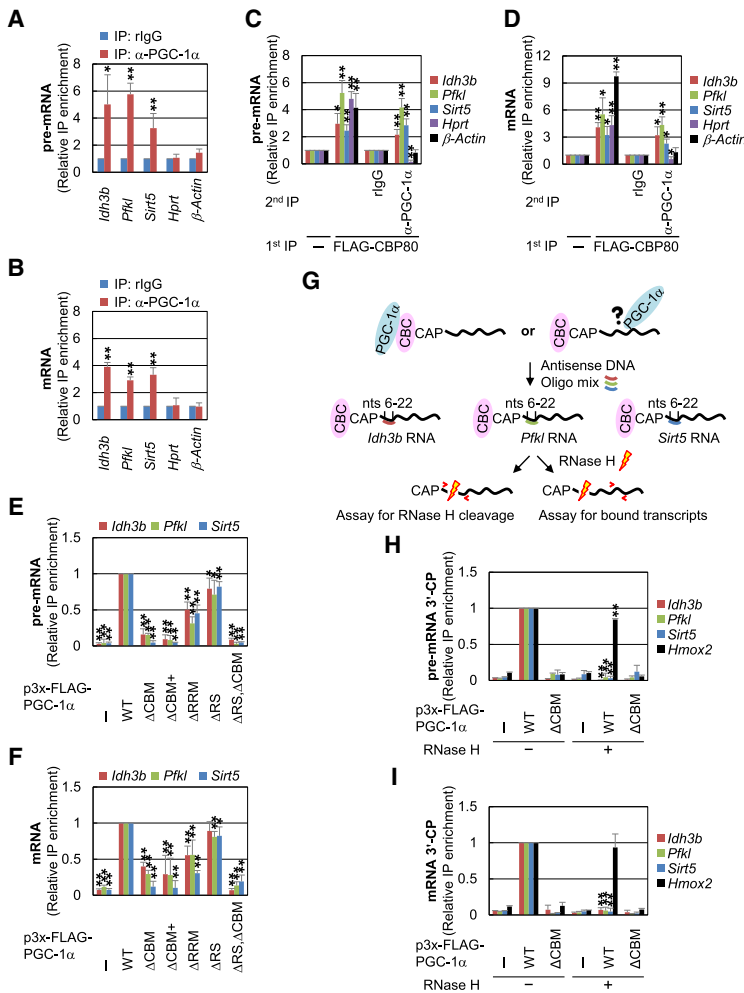
Figure 3. PGC-1α interacts with the cap of newly synthesized target transcripts largely via its CBM. (*A,B*) Histogram representation of RT-qPCR quantitations of pre-mRNA (*A*) or mRNA (*B*) from three PGC-1α-responsive genes (*Idh3b*, *Pfk1*, and *Sirt5*) and, as negative controls, two unresponsive genes (*Hprt* and *β-Actin*) after immunoprecipitation relative to before immunoprecipitation using anti-PGC-1α (red) or, as a control, rIgG (blue), where values after rIgG immunoprecipitation relative to before immunoprecipitation are set to 1. (*C,D*) As in *A* and *B* but using lysates of C2C12 cells treated as in Figure 2D and after a first anti-Flag immunoprecipitation or a second immunoprecipitation using anti-PGC-1α or, as a control, rIgG, where values in the first anti-Flag immunoprecipitation in the presence of Flag relative to before immunoprecipitation and values after the second immunoprecipitation using rIgG relative to before the second immunoprecipitation are set to 1. (*E,F*) As in *A* and *B* but after anti-Flag immunoprecipitation using lysates of PGC-1α knockdown C2C12 MBs transiently transfected with plasmid encoding the specified Flag-PGC-1α variant or Flag alone (−), where values after anti-Flag-PGC-1α(WT) immunoprecipitation relative to before immunoprecipitation were set to 1. (*G*) Schematic for antisense DNA oligonucleotide-directed RNase H-mediated cleavage of cellular pre-mRNP and mRNP. The mixture of three oligonucleotides, each complementary to nucleotides 6–22 of the denoted transcript, directs RNase H-mediated cleavage so as to remove the 5′ cap structure (CAP), including the associated proteins, from the body of the transcript. Assays for RNase H cleavage are presented (Supplemental Fig. S3I). (*H,I*) As in *E* and *F* but after anti-Flag immunoprecipitation, the latter in the presence of the oligonucleotide mix and either RNase H (+) or no RNase H (−), where values after anti-Flag-PGC-1α(WT) immunoprecipitation in the absence of RNase H relative to before immunoprecipitation were set to 1. For all histograms, results are means ± SD. $n ≥ 3$; $n = 2$ for *H* and *I*. (*) $P < 0.05$; (**) $P < 0.01$; (no asterisks) $P > 0.05$ (i.e., not significant) compared with control immunoprecipitation (*A–D*), compared with PGC-1α knockdown C2C12 MBs expressing Flag-PGC-1α(WT) (*E,F*), or compared with minus RNase H treatment (*H,I*) by a two-tailed unpaired Student's *t*-test.

completely abolished binding (Fig. 3E,F; Supplemental Fig. S3G), confirming that the CBM and the RS domain cooperate in PGC-1α binding to target transcripts.

Given that the CBM of PGC-1α associates with the cap-binding protein CBP80 in vitro and in vivo and that PGC-1α binds responsive transcripts in complex with CBC, we hypothesized that PGC-1α associates in cells with the cap of transcripts synthesized from PGC-1α target genes. To test this, RNase H was used to cleave nucleotides 6–22 of *Idh3b*, *Pfkl*, and *Sirt5* pre-mRNPs and mRNPs (Fig. 3G) in anti-Flag immunoprecipitates using lysates of PGC-1α knockdown MBs expressing Flag, Flag-PGC-1α (WT), or Flag-PGC-1α(ΔCBM). Site-specific RNase H cleavage (Kurosaki and Maquat 2013), which was directed using a mixture of three antisense DNA oligonucleotides while RNPs were associated with anti-Flag-bound beads, had no effect on immunoprecipitation efficiencies (Supplemental Fig. S3H) and was deemed complete, as evidenced by the loss of RT-qPCR products that spanned each cleavage site in anti-Flag-PGC-1α(WT) immunoprecipitations (Supplemental Fig. S3I). As controls for cleav-

age specificity, the mixture of the three oligonucleotides failed to promote RNase H-mediated cleavage of pre-mRNA or mRNA from the heme oxygenase 2 (*Hmox2*) gene, which is also PGC-1α-responsive (Aguilo et al. 2016) and thus encodes transcripts that would be retained on beads along with the other PGC-1α target transcripts (Supplemental Fig. S3I).

Results demonstrated that co-IP of the bodies of the three PGC-1α-responsive pre-mRNAs and their corresponding mRNAs, but not uncleaved control *Hmox2* pre-mRNA and mRNA, with Flag-PGC-1α(WT) is abolished after RNase H treatment (Fig. 3H,I; data not shown using additional pairs of primers that amplify the 5′ untranslated region [UTR], the coding sequence, or the 3′ UTR of *Hmox2* mRNA). Importantly, the three antisense DNA oligonucleotides were added to samples that were untreated or treated with RNase H, ruling out the possibility that the interaction of PGC-1α with its target transcripts is disrupted by modification of their 5′ UTR structures upon antisense DNA oligonucleotide annealing. We conclude that PGC-1α does not bind throughout

the bodies of target transcripts but at their 5′ ends, where the cap and the CBC reside. However, we do not rule out that the RRM and/or the RS domain of PGC-1α, the latter of which was found recently to bind RNA (Long et al. 2016), may facilitate or stabilize the PGC-1α CBM–CBP80 interaction by binding to the 5′-most ~5 nucleotides (nt) of target transcripts (which were not removed by RNase H treatment) or transiently binding the bodies of the nascent pre-mRNAs as they are synthesized.

The CBM controls the transcriptional activity of PGC-1α

Considering the critical role of the CBM in PGC-1α binding to the 5′ cap of pre-mRNAs synthesized from PGC-1α-responsive genes, we tested whether this motif is also key to PGC-1α-mediated transcriptional activation. In control experiments we demonstrated using RT-qPCR that expressing Flag-PGC-1α(WT) in PGC-1α knockdown C2C12 MBs restores the levels of target pre-mRNAs, mRNAs, and eRNAs to those in wild-type C2C12 MBs (Fig. 4A,B; Supplemental Fig. S4A); i.e., Flag-PGC-1α (WT) can functionally replace cellular PGC-1α. Similar analyses of PGC-1α(ΔCBM) and PGC-1α(CBM 5Mut) revealed that each failed to support the expression of these transcripts (Fig. 4A,B; Supplemental Fig. S4A). PGC-1α knockdown and replacement with Flag-PGC-1α(ΔCBM) or PGC-1α(CBM 5Mut) did not decrease the level of mRNAs more than the levels of pre-mRNAs, indicating that the PGC-1α CBM promotes gene expression at the transcriptional level. To validate this hypothesis, we assessed the decay rates of the PGC-1α-responsive mRNAs under conditions where transcription was blocked using actinomycin D (ActD). We found that PGC-1α knockdown or its replacement with Flag-PGC-1α(ΔCBM) or PGC-1α(CBM 5Mut) did not detectably decrease the half-lives of the relatively stable *Idh3b* and *Pfkl* mRNAs, which manifested half-lives of >10 h under all conditions tested, or the half-life of the more labile *Sirt5* mRNA (Supplemental Fig. S4B).

To further confirm that PGC-1α(CBM)-dependent gene expression relies on the ability of PGC-1α to bind CBP80, we analyzed C2C12 cells in which endogenous CBP80 was down-regulated to 10% of normal and replaced with either Flag, Flag-CBP80(WT), or one of Flag-CBP80 variants (Supplemental Fig. S4C). Compared with Flag-CBP80(WT), neither Flag-CBP80(Δ1–44) nor Flag-CBP80 (9Mut) supported the expression of PGC-1α-responsive genes (Fig. 4C,D; Supplemental Fig. S4D). In these experiments, RNA levels were normalized to the level of U1 snRNA, which is unaltered by transient CBP80 down-regulation (Pabis et al. 2013). These results validate the functional relevance of the interaction between PGC-1α and CBP80 in the transcription of PGC-1α-responsive genes.

To extend our studies beyond the three PGC-1α-responsive gene loci, we aimed to determine the extent to which the PGC-1α CBM globally regulates the C2C12 MB transcriptome. Thus, we used RNA sequencing (RNA-seq) to compare transcripts in wild-type C2C12 MBs and PGC-1α knockdown C2C12 MBs expressing Flag alone, Flag-
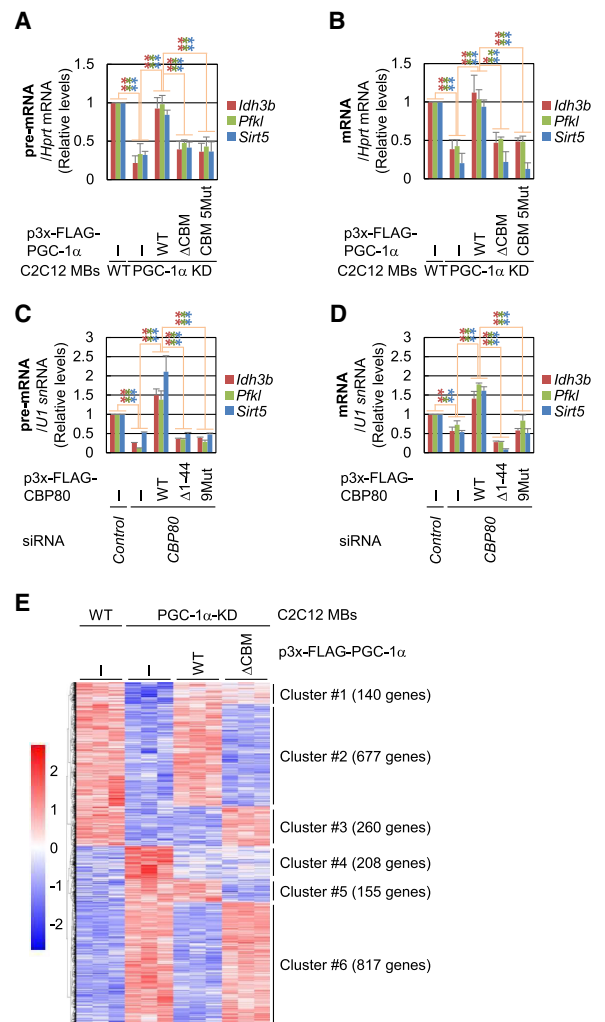
Figure 4. The CBM controls the transcriptional activity of PGC-1α. (*A,B*) Histogram representation of RT-qPCR quantitations of pre-mRNA (*A*) or mRNA (*B*) from the three PGC-1α-responsive genes normalized to the level of *Hprt* mRNA using lysates of PGC-1α knockdown C2C12 MBs treated as in Figure 2, F and G, or Supplemental Figure S2D. (*C,D*) As in *A* and *B* but using lysates of C2C12 MBs transiently cotransfected with either control or *CBP80* siRNA and with plasmid encoding the specified siRNA-resistant Flag-CBP80 variant or Flag alone (−). pre-mRNA and mRNA levels were normalized to the level of U1 snRNA. (*E*) Heat map of hierarchically clustered mRNA expression from RNA sequencing (RNA-seq) analyses of the specified cells: wild-type C2C12 MBs expressing either Flag (−) or PGC-1α knockdown C2C12 MBs expressing Flag (−), Flag-PGC-1α(WT), or Flag-PGC-1α(ΔCBM). Shown are genes whose expression was significantly affected by PGC-1α knockdown and significantly rescued by Flag-PGC-1α(WT) and/or Flag-PGC-1α(ΔCBM). The color key represents row-scaled expression values. For all histograms, results are means ± SD. n = 3. (*) P < 0.05; (**) P < 0.01 by a two-tailed unpaired Student's t-test.

PGC-1α(WT), or Flag-PGC-1α(ΔCBM). Hierarchal clustering of genes whose expression was affected by PGC-1α knockdown and rescued by Flag-PGC-1α(WT) and/or Flag-PGC-1α(ΔCBM) identified six groups of genes that

are differentially regulated (Fig. 4E). Clusters 1–3 and clusters 4–6 comprise genes whose expression is downregulated or up-regulated, respectively, by PGC-1α knockdown. We focused our attention on clusters 1–3 for two reasons: The literature describes PGC-1α largely as a transcriptional coactivator, and we characterized the binding of PGC-1α to pre-mRNAs and mRNAs that are up-regulated by PGC-1α. While the expression of most PGC-1α target genes in clusters 1–3 was restored in C2C12 PGC-1α knockdown MBs expressing Flag-PGC-1α(WT) (Fig. 4E, cf. clusters 1–2 and clusters 1–3), our finding that the expression of <25% of PGC-1α target genes was restored by Flag-PGC-1α(ΔCBM) but not by Flag-PGC-1α(WT) (Fig. 4E, cf. cluster 3 and clusters 1–3) indicates that some (i.e., <25%) of the PGC-1α target genes can be specifically regulated by PGC-1α isoforms that lack the CBM. Nevertheless, our finding that the expression of >75% of PGC-1α target genes that was restored by Flag-PGC-1α(WT) was not restored by Flag-PGC-1α(ΔCBM) (Fig. 4A, cf. cluster 2 and clusters 1–2) validates that the CBM is critical for the activation of the majority of genes that are responsive to the major and longest CBM-containing PGC-1α isoform.

PGC-1α promotes the differentiation of C2C12 MBs to MTs via its CBM

Gene ontology (GO) analyses of genes that are transcriptionally induced by PGC-1α via its CBM (Supplemental Table S3) revealed a significant enrichment in genes associated with "muscle structure development" (GO: 0061061) and, more particularly, "regulation of striated muscle differentiation" (GO: 0051146) (Fig. 5A; Supplemental Fig. S5A). Thus, we hypothesized that the previously observed dependence of C2C12 MB differentiation into MTs on PGC-1α (Baldelli et al. 2014) requires the PGC-1α CBM and that differentiation would be impaired if PGC-1α were replaced by PGC-1α(ΔCBM).

To test this hypothesis, we first used RT-qPCR to validate RNA-seq data indicating that the expression of pre-mRNAs and mRNAs encoded by four PGC-1α CBM-reliant target genes, which have also been shown to be bound by Flag-PGC-1α(WT) in MTs (Baresic et al. 2014), are up-regulated in a CBM-dependent manner. These genes encode proteins called the master early differentiation factor paired-box 7 (PAX7), salt-inducible kinase 1 (SIK1), SET and MYND domain-containing 1 (SMYD1), and growth
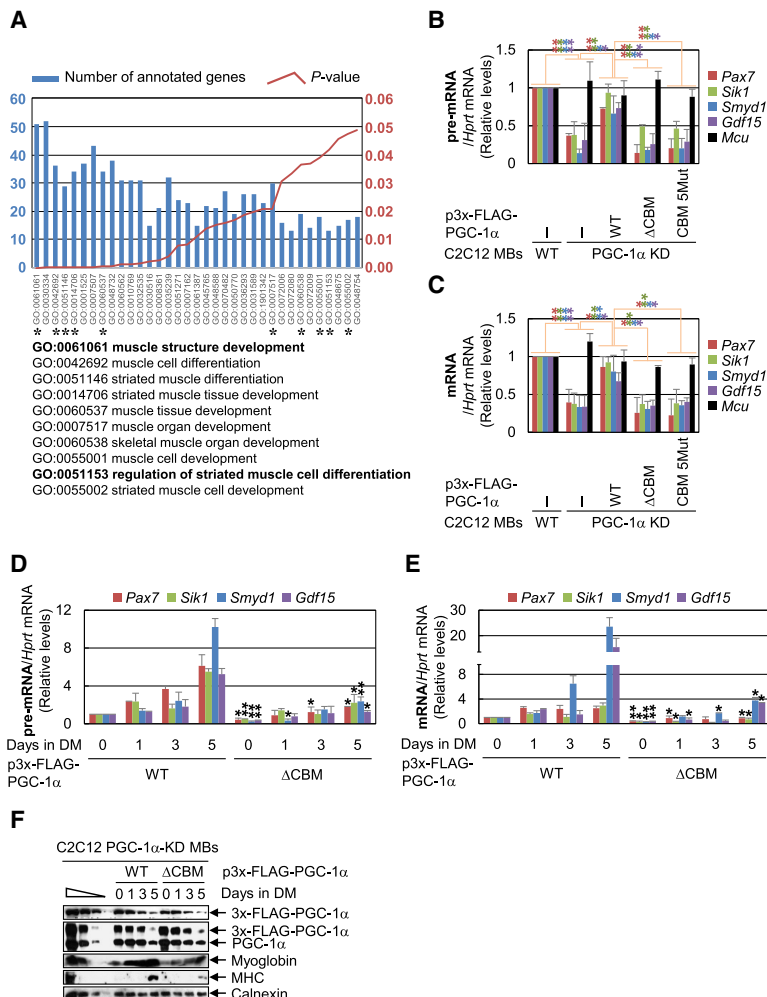

Figure 5. PGC-1α CBM promotes myogenesis by inducing myogenic gene transcription. (A) GO analysis of C2C12-MB genes whose up-regulation is dependent on the PGC-1α CBM (Supplemental Table S3). Asterisks specify myogenesis-related GO terms. (B,C) Histogram representation of RT-qPCR quantitations of PGC-1α CBM-dependent myogenic pre-mRNAs (B) or mRNAs (C) identified in Supplemental Figure S5A in cells treated as in Figure 4, A and B. (D,E) Histogram representation of RT-qPCR quantitations of PGC-1α CBM-dependent promyogenic pre-mRNAs (D) or mRNAs (E) using lysates of PGC-1α knockdown MBs expressing Flag-PGC-1α(WT) or Flag-PGC-1α (ΔCBM) before (0 d) and after (1, 3, and 5 d) culturing in differentiation medium (DM). (F) Western blots showing the expression of exogenous Flag-PGC-1α (WT) or Flag-PGC-1α(ΔCBM) and endogenous PGC-1α, CBP80, and myogenic markers in C2C12 MBs treated as in D and E. For all RT-qPCR analyses, results are means ± SD. n = 3 for B and C; n = 2 for D and E. (*) P < 0.05; (**) P < 0.01; (no asterisks) P > 0.05 between the indicated conditions (B,C) or comparing PCG-1α knockdown C2C12 MBs expressing Flag-PGC-1α(ΔCBM) with Flag-PGC-1α(WT) (D,E) by a two-tailed unpaired Student's t-test.

differentiation factor 15 (GDF15). Results indicated that the pre-mRNAs and mRNAs encoding all four promyogenic proteins are down-regulated in PGC-1α knockdown MBs expressing Flag alone, Flag-PGC-1α(ΔCBM), or Flag-PGC-1α(CBM 5Mut) relative to either wild-type MBs or PGC-1α knockdown MBs expressing Flag-PGC-1α(WT) (Fig. 5B,C, where pre-mRNA and mRNA encoded by the mitochondrial uniporter [Mcu] gene served as negative controls). Of note, the stability of these short-lived promyogenic mRNAs was generally not affected under the different conditions tested (Supplemental Fig. S5B). Additionally, expression of the corresponding genes was not supported by Flag-CBP80(Δ1–44) or Flag-CBP80(9Mut) (Supplemental Fig. S5C,D), further confirming that their transcription relies on the direct interaction between PGC-1α and CBP80.

Next, we cultured PGC-1α knockdown MBs in differentiation medium to induce their maturation into MTs. Results using RT-qPCR revealed that the differentiation-induced expression of the four selected PGC-1α CBM-dependent promyogenic pre-mRNAs and mRNAs was strongly inhibited in PGC-1α knockdown MBs expressing Flag-PGC-1α(ΔCBM) relative to PGC-1α knockdown MBs expressing Flag-PGC-1α(WT) (Fig. 5D,E). Moreover, MBs expressing PGC-1α(ΔCBM) were characterized by a slower differentiation rate relative to MBs expressing Flag-PGC-1α(WT), as evidenced by delayed gain of the MT protein markers myoglobin and myosin heavy chain (MHC) (Fig. 5F) and delayed formation of multinucleated and elongated MTs (Supplemental Fig. S5E). We conclude that PGC-1α promotes the differentiation of C2C12 MBs into MTs by up-regulating a specific set of promyogenic genes in a mechanism that involves its CBM binding to the CBC of nascent transcripts that derive from PGC-1α-responsive genes.

Discussion

Here, we demonstrated that PGC-1α promotes the transcription of the large majority of its target genes by interacting with the CBC of nascent transcripts that derive from these genes. This interaction is mediated by a C-terminal motif in PGC-1α that we call the CBM and the N-terminal α helices of CBP80 that we call the PGC-1-dependent AD. Notably, the CBM (which we define based on its primary sequence, its α-helical structure, and its C-terminal position) is also present in PGC-1β as well as the more divergent PGC-1 family member PRC (Supplemental Fig. S1B,C; Supplemental information). We propose that PGC-1α CBM binding to the 5′ end cap of nascent pre-mRNAs constitutes a type of quality control; i.e., a pre-mRNP surveillance mechanism. This mechanism requires that PGC-1α-responsive genes produce nascent transcripts that not only harbor a 5′ end cap but also have CBC bound to the cap in order for PGC-1α to promote gene expression.

We found that the PGC-1α CBM is fully conserved in all examined species ranging from African clawed toads to humans and that only one residue of the CBM that is not required for binding to CBP80 is different in zebrafish,

which is the most evolutionarily distant organism that we found harboring a PGC-1α ortholog. This suggests that PGC-1α CBM-mediated gene transcription is conserved in a wide range of animal species. Interestingly, a blastp search identified several proteins bearing the FDXLLXXA sequence, where X can be any residue, and non-X residues are at the CBM–CBP80 interface and mediate the interaction. However, whereas the PGC-1 CBM forms an extended C-terminal α helix whose ability to interact with CBP80 is likely conferred by an upstream flexible linker, the vast majority of BLASTp matches is to different structures and/or positions within the host protein. This is illustrated by SLC3A2, in which 84-FDSLLQSA-91 is included in an internal α helix (PDB ID 2DH2), and APAF1, in which 1120-FDLLL-1124 is part of a loop between β strands (PDB ID 3JBT). Thus, while future studies of proteins harboring this motif as part of a C-terminal α helix (such as in the uncharacterized Bacillus cereus protein Q739P5; PDB ID 5F2H) might elucidate other CBP80-binding proteins, here we limited our analyses to PGC-1 family members.

While our model was initially built on observations made using the pre-mRNAs, mRNAs, and eRNAs of three well-characterized PGC-1α-responsive gene loci in C2C12 MTs (Aguilo et al. 2016), our RNA-seq analyses enable us to expand this model to >650 PGC-1α(CBM)-responsive genes, including a cohort of promyogenic genes, providing physiological relevance for the CBM-dependent control of transcription. Notably, we found that the CBM regulates most, but not all, PGC-1α-responsive genes in C2C12 MBs. Those genes whose expression level in PGC-1α knockdown cells was restored by PGC-1α(ΔCBM) likely correspond to genes that are specifically regulated by smaller PGC-1α isoforms lacking the CBM. Distinct PGC-1α isoforms may be recruited to distinct promoters (Martínez-Redondo et al. 2016), and those isoforms lacking the CBM may possibly retain transcriptional activity by their ability to recruit HATs via their N-terminal regions. These smaller isoforms include PGC-1α4, whose overexpression in skeletal muscle has been shown to promote skeletal muscle hypertrophy (Ruas et al. 2012). In keeping with data indicating that the CBM functions comparably in PGC-1β, human PGC-1β pre-mRNA undergoes alternative splicing that results in the production of a variant of PGC-1β, PGC-1β-b, in which the CBM-containing C-terminal 33 amino acids of the major isoform are replaced by 27 different residues (Meirhaeghe et al. 2003). While yet to be characterized, the existence of this intriguing variant suggests that the physiological relevance of the CBM is not limited to PGC-1α but also typifies PGC-1β.

In view of the recent development of low-molecular-weight α-helix proteomimetics as potent pharmaceuticals (Chang et al. 2013), our finding that the CBM forms an α-helical interface with CBP80 offers the possibility of developing new therapeutics that target PGC-1α, PGC-1β, and/or PRC for disorders that include not only cardiovascular diseases but also cancer and metabolic and neurodegenerative pathologies (Villena 2015; Kadlec et al. 2016).

Materials and methods

Production and purification of GST, GST-PGC-1β, and GST-PGC-1α proteins

BL21-*CodonPlus(DE3)*-RIPL cells (Agilent Technologies) were transformed using pGEX-6p-1 *Escherichia coli* protein expression vectors producing GST, GST-PGC-1β, or GST-PGC-1α variants and grown at 37°C with shaking in 250 mL of Luria broth containing 30 μg/mL chloramphenicol and 100 μg/mL ampicillin to an OD_{600} of ~0.5. Cells were then transferred to 18°C with shaking, and plasmid expression was induced by adding 200 μL of 1 M IPTG. Cells were pelleted and stored at −80°C until use. Pellets were resuspended in 1 M NaCl, 25 mM HEPES (pH 7.4), and 10 mM DTT. Each sample was sonicated on ice using a microtip (45 sec on and 45 sec off, three times, using a Bronson Sonifier 250) and centrifuged at 17,000*g* for 90 min at 4°C. Cleared lysates were divided into halves, and each half was added to two tandem 5-mL GST-Trap HP columns (GE Healthcare) that had been pre-equilibrated with equilibration buffer (1 M NaCl, 25 mM HEPES at pH 7.4). After washing each column with 25 mL of equilibration buffer, GST-PGC-1β and GST-PGC-1α were eluted with 20 mL of elution buffer (0.33 g of reduced glutathione dissolved in 100 mL of 100 mM NaCl, 10 mM Tris-HCl at pH 8.0). Purified samples were dialyzed overnight to remove glutathione and subsequently concentrated in 150 mM NaCl, 50 mM HEPES (pH 7.5), and 5 mM 2-mercaptoethanol using a Vivaspin 20 centrifugal concentrator (Fisher Scientific). Concentrations of recombinant proteins were estimated after electrophoresis in SDS–polyacrylamide gels using Coomassie blue (Sigma), where serial dilutions of bovine serum albumin (Rockland) provided concentration standards.

Production and purification of CBP20 and CBP80 proteins and formation and purification of the CBC

To generate human CBP20, BL21 DE3 *E. coli* harboring pGEX-6p-1-CBP20 (Kim et al. 2009) were propagated at 37°C with shaking in 1 L of Luria broth that had been supplemented with 50 μg/mL ampicillin. When cells reached an OD_{600} of 0.4–0.8, 300 μL of 1 M IPTG was added, and cells were subsequently cultured overnight at 18°C. Cells (from 3 to 6 L) were lysed by sonication in 30 mL of equilibration buffer supplemented with 1 mM DTT, 0.5 mM EDTA (pH 8.0), and a cOmplete Mini Tablet EDTA-free protease inhibitor (Roche). Cells were sonicated on ice, and lysates were centrifuged at 17,000*g* for 30 min. Cleared lysate was purified as mentioned above. After adding 1 mg of purified PreScission protease (GE Healthcare) to cleave off the GST tag, eluted samples were dialyzed (molecular weight cutoff [MWCO] 12–14,000 kDa; Spectrapor) overnight at 4°C in 2 L of dialysis buffer (50 mM Tris-HCl at pH 8.0, 200 mM NaCl, 2 mM DTT). GST was removed by passage through tandem GST columns. Samples were concentrated in 3-kDa MWCO concentrators (Sartorius) to a concentration <5 mL and then loaded onto a Superdex 200 (GE Healthcare) gel filtration column using a Bio-Rad BioLogic Duo-Flow fast protein liquid chromatographer that had been equilibrated with gel column buffer (3 mM DTT, 100 mM NaCl, 25 mM Tris-HCl at pH 8.0). Glycerol was added to a final concentration of 15% to CBP20 fractions that were then flash-frozen in liquid N_2 and stored at −80°C.

To form the CBC, the pellet of SF9 insect cells infected with baculovirus expressing HIS-tagged CBP80(Δ2–23) (Genscript) was resuspended in 40 mL of Q column buffer A (200 mM NaCl, 50 mM HEPES at pH 7.5, 10 mM 2-mercaptoethanol) and sonicated on ice, and lysates were centrifuged at 17,000*g* for 30 min. Notably, amino acids 2–23 were removed from CBP80 for improved stability (Mazza et al. 2001, 2002a). Cleared lysate was loaded

onto a series of three 5-mL Q columns (HiTrap Q HP, GE Healthcare) that had been equilibrated in Q column buffer A and subjected to a gradient from Q column buffer A to Q column buffer B (800 mM NaCl, 50 mM HEPES at pH 7.5, 10 mM 2-mercaptoethanol) using fast protein liquid chromatography (FPLC). Fractions were concentrated as was performed for CBP20 and then separated using a Superdex 200 gel filtration column (GE Healthcare) that had been equilibrated with 50 mM NaCl, 10 mM HEPES (pH 7.5), and 5 mM 2-mercaptoethanol. Fractions containing CBP80 were identified after separation in 4%–20% polyacrylamide (Bio-Rad) followed by Coomassie blue staining.

To form the CBC, purified CBP20 was added to the lysate of SF9 insect cells infected with baculovirus expressing HIS-tagged CBP80(Δ2–23), and the complex was processed as described for CBP80.

GST pull-downs

Equal amounts (298 pmol) of GST, GST-tagged PGC-1α, or GST-tagged PGC-1α fragments and HIS-TEV-tagged CBP80(Δ2–23) were added to 1 mL of a 50:50 slurry of glutathione Sepharose 4B beads (GE Healthcare). Bound material was washed three times using 150 mM NaCl, 25 mM HEPES (pH 7.5), and 5 mM 2-mercaptoethanol and subsequently eluted by boiling in 4× SDS loading buffer (250 mM Tris-HCl at pH 6.8, 400 mM DTT, 8% SDS, 40% glycerol, 0.02% Bromophenol blue).

X-ray crystal structure determination

Crystals of the CBC-PGC-1β m^7GpppA complex were identified using hanging drop vapor diffusion and condition number 20 (0.2 M magnesium chloride, 0.1 M Tris at pH 7.0, 10% [w/v] polyethylene glycol 8000) of the JCSG+ crystallization screen and incubated at 20°C. Crystals were formed by adding 1 μL of complex to 1 μL of reservoir solution to form a drop incubated over 0.8 mL of reservoir solution and in a 24-well crystallization tray. The complex was formed by adding 25 μL of 9.8 mg/mL CBC, 5 μL of 20 μg/mL PGC-1β peptide spanning amino acids 994–1023 (Ac-SEEALPASGKSKYEAMDFDSLLKEAQQSLH, where p-iodo-tyrosine [L-3-I-Tyr] was used in place of Y1006 for possible phasing applications; GenScript) (Xie et al. 2004), and 1 μL of 10 mM m^7GpppA (New England Biolabs) in gel filtration buffer (10 mM Tris-HCl at pH 7.5, 50 mM NaCl, 5 mM 2-mercaptoethanol). Crystals grew to as much as 900 μm in length. Crystals were cryoprotected in 50/50 (v/v) paratone N (Hampton Research) and silicon oil (Hampton Research), snap-frozen, and stored in liquid N_2 until used for X-ray diffraction studies. X-ray data collection was performed at Stanford Synchrotron Radiation Lightsource Beamline 12-2 using a wavelength of 0.9553 Å and a PILATUS detector. Diffraction data were processed using autoxds (http://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script). Phenix (Adams et al. 2010) was used for molecular replacement and refinement, Coot (Emsley et al. 2010) was used for modeling, and figures were generated using University of California at San Francisco (UCSF) Chimera (Pettersen et al. 2004). Sigma-A weighted 2mFo-DFc maps were generated using Phenix software (phenix.maps), and electron density was shown using UCSF Chimera 1.01 Å and within 2 Å around the residues and regions shown.

Cell culture, transient transfections, and generation of stable PGC-1α knockdown cells

Mouse C2C12 MBs and HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco) containing 15%

fetal bovine serum (Gibco) and 1% penicillin/streptomycin (Gibco). Human MCF-7 breast adenocarcinoma cells were cultured in DMEM (Gibco) containing nutrient mixture F-12 (lacking phenol red), 10% fetal bovine serum (Gibco), and 1% penicillin/streptomycin (Gibco). To generate C2C12 MBs in which PGC-1α or, as a negative control, GFP was stably knocked down, HEK293T cells were transiently transfected with VSV-G plasmid (Addgene, 8454), pCG-gag-pol (Ulm et al. 2007), and MISSION shRNA pLKO.1-puro vector for PGC-1α (TRCN0000234017: CCGGTCCAGTAAGCACACGTTTATTC TCGAGAATAAACGTGTGCTTACTGGATTTTTG; Sigma) or pRetroSuper-GFP-shRNA, respectively (Brummelkamp et al. 2002). C2C12 MBs were transduced using the resulting viral supernatants (multiplicity of infection [MOI] of 1) supplemented with 2 µg/mL hexadimethrine bromide (Sigma). After 24 h, cells were selected for resistance to 2 mg/mL puromycin (Gibco), and resistant cells were pooled.

Cells were transiently transfected with the indicated plasmids using Lipofectamine 2000 (Life Technologies) and harvested 32 h later, as described previously (Kurosaki et al. 2014). When cotransfecting siRNA and plasmid, cells were transiently cotransfected with 50 nM in vitro synthesized siRNA (Dharmacon) and the indicated plasmids using jetPRIME (Polyplus-Transfection) and harvested 48 h later. The following siRNA sequence was used for mouse CBP80: 5′-GGGAGGAGAGUGAG GAAUAUU-3′.

Experiments in which MBs were differentiated into MTs were essentially as described (Wang et al. 2013). Briefly, MBs were cultured to 95%–100% confluency before the medium was changed to DMEM supplemented with 2% horse serum (Gibco).

Cell lysis and fractionation

Total-cell lysates were prepared using hypotonic gentle lysis buffer (10 mM Tris at pH 7.4, 10 mM NaCl, 10 mM EDTA, 1% [w/w] Triton X-100) (Cho et al. 2009; Kurosaki and Maquat 2013) containing protease inhibitor cocktail (Roche). Cytoplasmic, nuclear, and chromatin fractions were purified as described previously (Nojima et al. 2016) except that equal volumes of NUN1 and NUN2 buffers were used, and nuclei were lysed by five 1-sec vortex pulses every 2 min for 10 min on ice.

Western blotting

Cell lysates were electrophoresed in 8%, 10%, or 12% polyacrylamide gels, and proteins were transferred to a nitrocellulose membrane (Amersham Biosciences) by electroblotting. Proteins were detected using the following antibodies: anti-PGC-1β (Bethyl Laboratories, A302-273A), anti-PGC-1α (Novus, NBP1-04676), anti-Flag (Sigma-Aldrich, A8592), anti-β-actin (Sigma-Aldrich, A2228), anti-CBP80 (Bethyl Laboratories, A301-793A or, alternatively, Choe et al. 2010), anti-CBP20 (Santa Cruz Biotechnology, sc-48793), anti-PABPC1 (Abcam, ab21060), anti-eIF4E (Bethyl Laboratories, A301-153A), anti-α-Tubulin (Santa Cruz Biotechnology, sc-58666), anti-P62 (BabCO, mAb414), anti-vimentin (Santa Cruz Biotechnology, sc-6260), anti-pSer2-RNA-PII (Abcam, ab5095), anti-GAPDH (Cell Signaling Technology, 2118), anti-NONO (Bethyl Laboratories, A300-582A), anti-myoglobin (Santa Cruz Biotechnology, sc25607), anti-MHC (Abcam, ab91506), or anti-calnexin (Enzo, SPA-865F).

Immunoprecipitations

Immunoprecipitations were performed after the addition of NaCl to 150 mM as described previously (Kim et al. 2005), except that

RNase I (Invitrogen) was used in place of RNase A, and antibodies were as described for Western blotting (see below). For tandem immunoprecipitations, C2C12 cells expressing Flag-CBP80 were lysed and immunoprecipitated using anti-Flag M2 affinity beads (Sigma). Beads were washed twice using ice-cold NET-2 buffer (50 mM Tris-HCl at pH 7.4, 150 mM NaCl, 0.05% NP-40, Complete mini, EDTA-free protease inhibitor cocktail [Roche]). After an additional five washes using NET2 buffer, Flag-CBP80-containing complexes were affinity-eluted from the beads using elution buffer (200 µg/mL Flag peptide [Sigma], 5% glycerol in NET2 buffer) overnight at 4°C with rocking. The second immunoprecipitation was performed for 2 h at 4°C using anti-PGC-1α that had been coupled to protein A (Roche) according to the manufacturer's instructions. Beads were washed five times using NET2 buffer, and complexes were eluted using 2× SDS sample buffer.

RT-qPCR

RT-qPCR was undertaken using the 7500 Fast real-time PCR system (Applied Biosystems), Fast SYBR Green Master mix (Applied Biosystems), and PCR primers specified in Supplemental Table S4. RT was performed using total-cell RNA (TRIzol reagent, Life Technologies) and SuperScript III reverse transcriptase (Invitrogen) to quantify levels of pre-mRNA, mRNA, or eRNA.

DNA oligonucleotide-directed RNase H-mediated cleavage of cellular RNPs

Cleavage was performed essentially as described (Kurosaki and Maquat 2013). DNA oligonucleotides (Integrated DNA Technologies) were as follows: *Idh3b* pre-mRNA + mRNA (5′-ACTTCC CCCGCGACTTC-3′), *Pfkl* pre-mRNA + mRNA (5′-GGGACA AACCGGGTACA-3′), *Sirt5* pre-mRNA (5′-TGCAGGTGCGTG TCTGG-3′), and *Sirt5* mRNA (5′-TGCAGCCTGCCTGGGGA-3′).

RNase H cleavage efficiency was assessed using RT-qPCR and primers spanning the cleavage site as specified in Supplemental Table S4.

RNA-seq library preparation and sequencing

The total RNA concentration was determined using the Nano-Drop 1000 spectrophotometer (NanoDrop), and RNA quality was assessed using the Agilent Bioanalyzer (Agilent). The TruSeq stranded mRNA sample preparation kit (Illumina) was used for next-generation sequencing library construction following the manufacturer's protocols. Briefly, poly(A)+ RNA was purified from 200 ng of total C2C12 cell RNA using oligo-dT magnetic beads and fragmented. First strand cDNA synthesis was performed using random hexamer priming followed by second strand cDNA synthesis using dUTP incorporation for strand marking. The resulting double-stranded cDNA was then subjected to end repair and 3′ end adenylation. Illumina adaptors were ligated to both cDNA ends. Adapted cDNA was purified using Ampure beads and PCR-amplified using primers specific to the adaptor sequences to generate cDNA amplicons of ~200–500 base pairs. Amplified libraries were hybridized to the Illumina single-end flow cell and further amplified using the cBot (Illumina). Single-end reads of 100 nt were generated for each sample using HiSeq 2500v4 (Illumina).

GO analysis

GO term enrichment analyses were performed using PGC-1α CBM-dependent genes (Supplemental Table S3), the online

ToppGene suite (Chen et al. 2009), and a Bonferroni-corrected *P*-value.

References

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221.

Aguilo F, Li S, Balasubramaniyan N, Sancho A, Benko S, Zhang F, Vashisht A, Rengasamy M, Andino B, Chen CH, et al. 2016. Deposition of 5-methylcytosine on enhancer RNAs enables the coactivator function of PGC-1α. *Cell Rep* **14:** 479–492.

Baldelli S, Aquilano K, Ciriolo MR. 2014. PGC-1α buffers ROS-mediated removal of mitochondria during myogenesis. *Cell Death Dis* **5:** e1515.

Baresic M, Salatino S, Kupr B, van Nimwegen E, Handschin C. 2014. Transcriptional network analysis in muscle reveals AP-1 as a partner of PGC-1α in the regulation of the hypoxic gene program. *Mol Cell Biol* **34:** 2996–3012.

Brummelkamp TR, Bernards R, Agami R. 2002. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296:** 550–553.

Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. 2013. Scalable Web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res* **41:** W349–W357.

Calero G, Wilson KF, Ly T, Rios-Steiner JL, Clardy JC, Cerione RA. 2002. Structural basis of m⁷GpppG binding to the nuclear cap-binding protein complex. *Nat Struct Biol* **9:** 912–917.

Calvo JA, Daniels TG, Wang X, Paul A, Lin J, Spiegelman BM, Stevenson SC, Rangwala SM. 2008. Muscle-specific expression of PPARγ coactivator-1α improves exercise performance and increases peak oxygen uptake. *J Appl Physiol* **104:** 1304–1312.

Chang YS, Graves B, Guerlavais V, Tovar C, Packman K, To KH, Olson KA, Kesavan K, Gangurde P, Mukherjee A, et al. 2013. Stapled α-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc Natl Acad Sci* **110:** E3445–E3454.

Chen J, Bardes EE, Aronow BJ, Jegga AG. 2009. ToppGene suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* **37:** W305–W311.

Cho H, Kim KM, Kim YK. 2009. Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex. *Mol Cell* **33:** 75–86.

Choe J, Cho H, Lee HC, Kim YK. 2010. MicroRNA/Argonaute 2 regulates nonsense-mediated messenger RNA decay. *EMBO Rep* **11:** 380–386.

Correia JC, Ferreira DM, Ruas JL. 2015. Intercellular: local and systemic actions of skeletal muscle PGC-1s. *Trends Endocrinol Metab* **26:** 305–314.

Dinulovic I, Furrer R, Beer M, Ferry A, Cardel B, Handschin C. 2016. Muscle PGC-1α modulates satellite cell number and proliferation by remodeling the stem cell niche. *Skelet Muscle* **6:** 39.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501.

Gonatopoulos-Pournatzis T, Cowling VH. 2014. Cap-binding complex (CBC). *Biochem J* **457:** 231–242.

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458:** 223–227.

Kadlec AO, Chabowski D, Ait-Aissa K, Gutterman DD. 2016. Role of PGC-1α in vascular regulation: implications for atherosclerosis. *Arterioscler Thromb Vasc Biol* **36:** 1467–1474.

Kataoka N, Ohno M, Moda I, Shimura Y. 1995. Identification of the factors that interact with NCBP, an 80 kDa nuclear cap binding protein. *Nucleic Acids Res* **23:** 3638–3641.

Kim YK, Furic L, Desgroseillers L, Maquat LE. 2005. Mammalian Staufen1 recruits Upf1 to specific mRNA 3′UTRs so as to elicit mRNA decay. *Cell* **120:** 195–208.

Kim KM, Cho H, Choi K, Kim J, Kim BW, Ko YG, Jang SK, Kim YK. 2009. A new MIF4G domain-containing protein, CTIF, directs nuclear cap-binding protein CBP80/20-dependent translation. *Genes Dev* **23:** 2033–2045.

Kressler D, Schreiber SN, Knutti D, Kralli A. 2002. The PGC-1-related protein PERC is a selective coactivator of estrogen receptor α. *J Biol Chem* **277:** 13918–13925.

Kurosaki T, Maquat LE. 2013. Rules that govern UPF1 binding to mRNA 3′ UTRs. *Proc Natl Acad Sci* **110:** 3357–3362.

Kurosaki T, Li W, Hoque M, Popp MW, Ermolenko DN, Tian B, Maquat LE. 2014. A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. *Genes Dev* **28:** 1900–1916.

Lejeune F, Ishigaki Y, Li X, Maquat LE. 2002. The exon junction complex is detected on CBP80-bound but not eIF4E-bound

mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J* **21:** 3536–3545.

Lin J, Wu H, Tarr PT, Zhang CY, Wu Z, Boss O, Michael LF, Puigserver P, Isotani E, Olson EN, et al. 2002. Transcriptional co-activator PGC-1α drives the formation of slow-twitch muscle fibres. *Nature* **418:** 797–801.

Long J, Badal SS, Ye Z, Wang Y, Ayanga BA, Galvan DL, Green NH, Chang BH, Overbeek PA, Danesh FR. 2016. Long noncoding RNA Tug1 regulates mitochondrial bioenergetics in diabetic nephropathy. *J Clin Invest* **126:** 4205–4218.

Martínez-Redondo V, Pettersson AT, Ruas JL. 2015. The hitchhiker's guide to PGC-1α isoform structure and biological functions. *Diabetologia* **58:** 1969–1977.

Martínez-Redondo V, Jannig PR, Correia JC, Ferreira DM, Cervenka I, Lindvall JM, Sinha I, Izadi M, Pettersson-Klein AT, Agudelo LZ, et al. 2016. Peroxisome proliferator-activated receptor γ coactivator-1α isoforms selectively regulate multiple splicing events on target genes. *J Biol Chem* **291:** 15169–15184.

Mazza C, Ohno M, Segref A, Mattaj IW, Cusack S. 2001. Crystal structure of the human nuclear cap binding complex. *Mol Cell* **8:** 383–396.

Mazza C, Segref A, Mattaj IW, Cusack S. 2002a. Co-crystallization of the human nuclear cap-binding complex with a m⁷GpppG cap analogue using protein engineering. *Acta Crystallogr D Biol Crystallogr* **58:** 2194–2197.

Mazza C, Segref A, Mattaj IW, Cusack S. 2002b. Large-scale induced fit recognition of an m⁷GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J* **21:** 5548–5557.

Meirhaeghe A, Crowley V, Lenaghan C, Lelliott C, Green K, Stewart A, Hart K, Schinner S, Sethi JK, Yeo G, et al. 2003. Characterization of the human, mouse and rat PGC1β (peroxisome-proliferator-activated receptor-γ co-activator 1β) gene in vitro and in vivo. *Biochem J* **373:** 155–165.

Michael LF, Wu Z, Cheatham RB, Puigserver P, Adelmant G, Lehman JJ, Kelly DP, Spiegelman BM. 2001. Restoration of insulin-sensitive glucose transporter (GLUT4) gene expression in muscle cells by the transcriptional coactivator PGC-1. *Proc Natl Acad Sci* **98:** 3820–3825.

Nelson JW, Kallenbach NR. 1986. Stabilization of the ribonuclease S-peptide α-helix by trifluoroethanol. *Proteins* **1:** 211–217.

Nojima T, Gomes T, Carmo-Fonseca M, Proudfoot NJ. 2016. Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. *Nat Protoc* **11:** 413–428.

Pabis M, Neufeld N, Steiner MC, Bojic T, Shav-Tal Y, Neugebauer KM. 2013. The nuclear cap-binding complex interacts with the U4/U6.U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA* **19:** 1054–1063.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25:** 1605–1612.

Ruas JL, White JP, Rao RR, Kleiner S, Brannan KT, Harrison BC, Greene NP, Wu J, Estall JL, Irving BA, et al. 2012. A PGC-1α isoform induced by resistance training regulates skeletal muscle hypertrophy. *Cell* **151:** 1319–1331.

Sato H, Maquat LE. 2009. Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin β. *Genes Dev* **23:** 2537–2550.

Shatsky IN, Dmitriev SE, Andreev DE, Terenin IM. 2014. Transcriptome-wide studies uncover the diversity of modes of mRNA recruitment to eukaryotic ribosomes. *Crit Rev Biochem Mol Biol* **49:** 164–177.

Ulm JW, Perron M, Sodroski J, Mulligan RC. 2007. Complex determinants within the Moloney murine leukemia virus capsid modulate susceptibility of the virus to Fv1 and Ref1-mediated restriction. *Virology* **363:** 245–255.

Villena JA. 2015. New insights into PGC-1 coactivators: redefining their role in the regulation of mitochondrial function and beyond. *FEBS J* **282:** 647–672.

Wang J, Gong C, Maquat LE. 2013. Control of myogenesis by rodent SINE-containing lncRNAs. *Genes Dev* **27:** 793–804.

Xie J, Wang L, Wu N, Brock A, Spraggon G, Schultz PG. 2004. The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination. *Nat Biotech* **22:** 1297–1301.