PNAS RESEARCH ARTICLE | BIOCHEMISTRY

Structures of NPAS4-ARNT and NPAS4-ARNT2 heterodimers reveal new dimerization modalities in the bHLH-PAS transcription factor family

Xiangnan Sun (孙向楠)[a], Linqian Jing (荆林倩)[a], Fengwei Li (李峰伟)[a], Meina Zhang (张美娜)[a], Xiaotong Diao (刁晓彤)[a], Jingjing Zhuang (庄静静)[a,b], Fraydoon Rastinejad[c], and Dalei Wu (武大雷)[a,d,1]

Edited by Joseph Takahashi, The University of Texas Southwestern Medical Center, Dallas, TX; received May 25, 2022; accepted October 11, 2022

Neuronal PER-ARNT-SIM (PAS) domain protein 4 (NPAS4) is a protective transcriptional regulator whose dysfunction has been linked to a variety of neuropsychiatric and metabolic diseases. As a member of the basic helix–loop–helix PER-ARNT-SIM (bHLH-PAS) transcription factor family, NPAS4 is distinguished by an ability to form functional heterodimers with aryl hydrocarbon receptor nuclear translocator (ARNT) and ARNT2, both of which are also bHLH-PAS family members. Here, we describe the quaternary architectures of NPAS4-ARNT and NPAS4-ARNT2 heterodimers in complexes involving DNA response elements. Our crystallographic studies reveal a uniquely interconnected domain conformation for the NPAS4 protein itself, as well as its differentially configured heterodimeric arrangements with both ARNT and ARNT2. Notably, the PAS-A domains of ARNT and ARNT2 exhibit variable conformations within these two heterodimers. The ARNT PAS-A domain also forms a set of interfaces with the PAS-A and PAS-B domains of NPAS4, different from those previously noted in ARNT heterodimers formed with other class I bHLH-PAS family proteins. Our structural observations together with biochemical and cell-based interrogations of these NPAS4 heterodimers provide molecular glimpses of the NPAS4 protein architecture and extend the known repertoire of heterodimerization patterns within the bHLH-PAS family. The PAS-B domains of NPAS4, ARNT, and ARNT2 all contain ligand-accessible pockets with appropriate volumes required for small-molecule binding. Given NPAS4's linkage to human diseases, the direct visualization of these PAS domains and the further understanding of their relative positioning and interconnections within the NPAS4-ARNT and NPAS4-ARNT2 heterodimers may provide a road map for therapeutic discovery targeting these complexes.

NPAS4 | bHLH-PAS family | heterodimerization | structure | transcription factor

Significance

Neuronal activities can regulate gene expression through the action of certain transcription factors (TFs). One such TF is the neuronal PER-ARNT-SIM (PAS) domain protein 4 (NPAS4), believed to modulate the excitatory-inhibitory balance of neural circuits. NPAS4 dysfunction has been linked to a variety of human neuropsychiatric and metabolic diseases. In a stimulus-dependent manner, NPAS4 forms distinct heterodimers with aryl hydrocarbon receptor nuclear translocator (ARNT) or ARNT2. Here, we visualized the molecular architectures of NPAS4-ARNT and NPAS4-ARNT2 heterodimers, finding unique architectural features and dimerization patterns relative to other members of their family. The visualization of three proteins, especially their PAS domains harboring ligand-binding pockets, and their unique inter-connections in each heterodimer will guide future therapeutic drug discovery directed against them.

As a member of the basic helix–loop–helix PER-ARNT-SIM (bHLH-PAS) family of transcription factors, neuronal PAS domain protein 4 (NPAS4), also known as neuronal transcription factor or limbic-enriched PAS, was originally shown to be highly expressed in the brain cortex and hippocampus (1, 2). The expression level of NPAS4 has been shown to elevate rapidly in response to seizure or cerebral ischemia (3, 4), making it a critical neuronal immediate-early gene (IEG). Functions of NPAS4 involve regulating the development and redistribution of inhibitory synapses (5, 6), coordinating the excitatory-inhibitory balance within neural circuits (7), linking neuronal activities to the formation of contextual memory (8, 9), and mediating the response of the suprachiasmatic nucleus to light (10). NPAS4-deficient mice exhibit social anxiety, hyperactivity, deficits in social interactions, and cognitive impairments (11, 12). In addition, dysregulation of NPAS4 and its downstream signaling pathways has been implicated in neurodevelopmental defects of autism and schizophrenia (13, 14). Therefore, NPAS4 could be a potential therapeutic target for these unmet neuropsychiatric conditions (15).

Outside the central nervous system, NPAS4 has been detected in pancreatic β cells, where its expression level increases under specific stimuli (e.g., endoplasmic reticulum stressors) in a calcium-dependent manner (16). NPAS4 binds to the insulin gene promoter to inhibit its transcription. Moreover, overexpression of NPAS4 prevents β-cell death induced by the immunosuppressant tacrolimus (17), suggesting NPAS4 can function as a cytoprotective IEG factor for enhancing β-cell efficiency under stress. NPAS4 is also expressed in vascular endothelial cells, where it promotes angiogenic sprouting and branch formation by regulating the transcription of vein endiothelial-cadherin (18). Studies using xenograft mice revealed that transplantation of breast cancer cells leads to the reduction of NPAS4 levels and its downstream target gene, brain-derived neurotrophic factor (*BDNF*), within the hippocampus (19). As such, NPAS4 may be responsible for neurocognitive deficits associated with tumor presence (a phenomenon also known as "tumor brain"). Although

Author contributions: X.S. and D.W. designed research; X.S., L.J., M.Z., and X.D. performed research; F.L., J.Z., and F.R. contributed new reagents/analytic tools; X.S., L.J., F.L., and D.W. analyzed data; and X.S., F.R., and D.W. wrote the paper.

Competing interest statement: F.R. is a founder of and consultant for Flare Therapeutics.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: dlwu@sdu.edu.cn.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2208804119/-/DCSupplemental.

Published November 7, 2022.

NPAS4 has not been actively pursued as a therapeutic target in disease settings, the two tandem PAS domains (PAS-A and PAS-B) within its polypeptide provide ligand-binding pockets, which small-molecule drugs can potentially target (20).

Like other class I members of the bHLH-PAS family (*SI Appendix*, Fig. S1*A*), NPAS4 protein requires heterodimerization with corresponding class II members (21) (i.e., aryl hydrocarbon receptor nuclear translocator [ARNT] or ARNT2) to gain the ability to bind DNA response elements and function as a regulator of target genes. Originally, ARNT2 was considered the preferred NPAS4 dimerization partner (1), as ARNT2 expression is largely restricted in the brain (where NPAS4 mainly exists) and kidney, while ARNT is ubiquitously expressed (22). However, in the hippocampal neurons NPAS4-ARNT and NPAS4-ARNT2 heterodimers are both functional and have been shown to have distinctive DNA-binding patterns in response to different excitatory stimuli (23), suggesting that these two NPAS4 heterodimers differentially control gene expression patterns. Besides NPAS4, ARNT2 can heterodimerize with other class I members enriched in neurons, including single-minded homolog 1 (SIM1) (24) and hypoxia-inducible factor (HIF)-1α (25). Although ARNT2 and ARNT share highly conserved protein sequences (*SI Appendix*, Fig. S2), their functional differences in terms of gene regulation have also been described (23, 26, 27). Unlike ARNT, whose single- and multidomain protein structures have been determined in complex with several dimerization partners such as HIF-1α, HIF-2α, HIF-3α, NPAS1, and NPAS3 (20, 28–31), no structural information is available for either ARNT2 or NPAS4 to date.

We previously found that NPAS4 likely heterodimerizes with ARNT in a different way compared with HIF-1α, HIF-2α, NPAS1, and NPAS3 based on mutagenesis and coimmunoprecipitation (Co-IP) studies (21). Specifically, we showed that certain surface residue mutations of ARNT that abolished its heterodimerization with NPAS1, NPAS3, or other class I members did not impact heterodimerization with NPAS4 (20). However, without structural understanding of NPAS4-ARNT, we were unable to discern the true dimerization modality of this heterodimer. Here, we pursued a variety of strategies to directly visualize NPAS4-ARNT2 and NPAS4-ARNT heterodimers in their DNA-bound complexes and further examined their molecular assemblies through a series of biochemical and cell-based studies.

Results

Quaternary Structure of the NPAS4-ARNT2-DNA Complex. Sequence analysis of the C-terminal portions of NPAS4 and ARNT2 proteins (beyond their PAS-B domains) suggested these regions are intrinsically disordered, similar to their counterparts in other bHLH-PAS family members (*SI Appendix*, Fig. S1*A*). Therefore, we coexpressed and purified the NPAS4 (1-348)–ARNT2 (50-439) heterodimers consisting of the bHLH, PAS-A, and PAS-B domains of each protein (Fig. 1*A*). We were unable to crystallize the heterodimeric complexes in the absence of DNA despite numerous attempts and therefore focused on the DNA-bound forms. We tested various double-stranded DNA samples with different lengths, endings, and flanking sequences, all of which contained a central E-box element. Using a 16-mer 5′ overhang DNA containing the core binding motif "GTCGTG" derived from the promoter region of *BDNF*, we successfully obtained NPAS4-ARNT2-DNA crystals that diffracted to 4.3-Å resolution at synchrotron (*SI Appendix*, Table S1). This resolution was in a similar range to those we and others previously reported for crystal structures of other DNA-bound multidomain ARNT heterodimers, such as the NPAS3-ARNT (4.2 Å) (20), HIF-1α-ARNT (3.9 Å), HIF-2α-ARNT (3.6 Å) (30), and aryl hydrocarbon receptor (AHR)–ARNT (4.0 Å) (32) and was lower than the 2.3 to 3.2-Å resolutions that can be obtained when DNA is not included with these proteins (20, 30, 33).

The NPAS4-ARNT2-DNA complex structure was solved in the space group $P4_32_12$, with one complex in the asymmetric unit. The backbone chains for all six domains plus DNA could be clearly interpreted in our electron density maps (*SI Appendix*, Fig. S3*A*). The overall structure of this complex revealed that the bHLH, PAS-A, and PAS-B domains of NPAS4 were in mutual contact, whereas the three corresponding domains of ARNT2 were spatially separated (Fig. 1*B*). The spatial separation of the bHLH, PAS-A, and PAS-B domains of ARNT was observed before in other ARNT heterodimers involving NPAS1, NPAS3, and three HIF-α subunits (21, 31). However, when superimposing the DNA-bound bHLH domains of NPAS4-ARNT2 with the DNA-bound NPAS3-ARNT structure (Fig. 1*C*), we unexpectedly found that the PAS-A domain of ARNT2 had a dramatic positional shift (twisted about 30°) relative to its counterpart in ARNT (Fig. 1*D*). This PAS-A domain positioning creates a unique junctional interface between NPAS4 and ARNT2 (detailed dissection of the domain–domain interfaces is described in the following sections).

Quaternary Structure of the NPAS4-ARNT-DNA Complex. To learn if this distinctive positioning of the ARNT2 PAS-A domain is also shared by the ARNT protein when heterodimerized with NPAS4, we attempted to obtain a crystal structure of NPAS4 (1-348)–ARNT (82-465) using similar protein constructs that contained the three domains from each subunit. Despite our expanded strategy for the growth of NPAS4-ARNT-DNA crystals, which included crystal seeding, dehydration, annealing, changing the reservoir contents (e.g., precipitant concentrations, additives), changing the length and endings of DNA (e.g., blunt, sticky G/C or A/T), and changing the length of protein constructs at their N- or C-termini, we only obtained those diffracted to 10-Å resolution. Inspired by the crystallization of aryl hydrocarbon receptor repressor (AHRR)-ARNT heterodimers that introduced multiple loop truncations in the ARNT protein to improve structure resolution (34), we modified the dimeric proteins by making various truncations at several loop regions of the ARNT PAS-A domain, as these loops were typically invisible (lacking clear electron densities) in our previous structures of ARNT heterodimers (20, 30). These efforts led us to protein-DNA crystals that diffracted to 4.7 Å (*SI Appendix*, Table S1), using a triple-loop truncated (Δ232 to 256, Δ274 to 298, and Δ320 to 330) ARNT construct (*SI Appendix*, Table S2). We could then solve the complex structure of NPAS4-ARNT-DNA with five of the six domains being clearly interpretable in the electron density maps (*SI Appendix*, Fig. S3*B*). However, the electron density for the ARNT PAS-A domain within this complex did not show sufficient quality to be fully traceable. This feature likely reflects the ARNT PAS-A domain's structural flexibility within the complex.

Because we did not encounter such a problem in our previous work on other ARNT heterodimers, we carefully examined all these complex structures and found that besides its intrinsic interactions with other domains within the heterodimers, the ARNT PAS-A domain could also form interactions with symmetric molecules in the crystal lattice. These additional interactions more or less exist in the ARNT heterodimers with three HIF-α subunits, NPAS1, and NPAS3 (20, 30, 31). They not only might help crystal packing but also might help stabilize the conformation of the ARNT PAS-A domain, at least to some extent. In contrast, in the structure of the AHRR-ARNT heterodimer, the ARNT PAS-A domain only

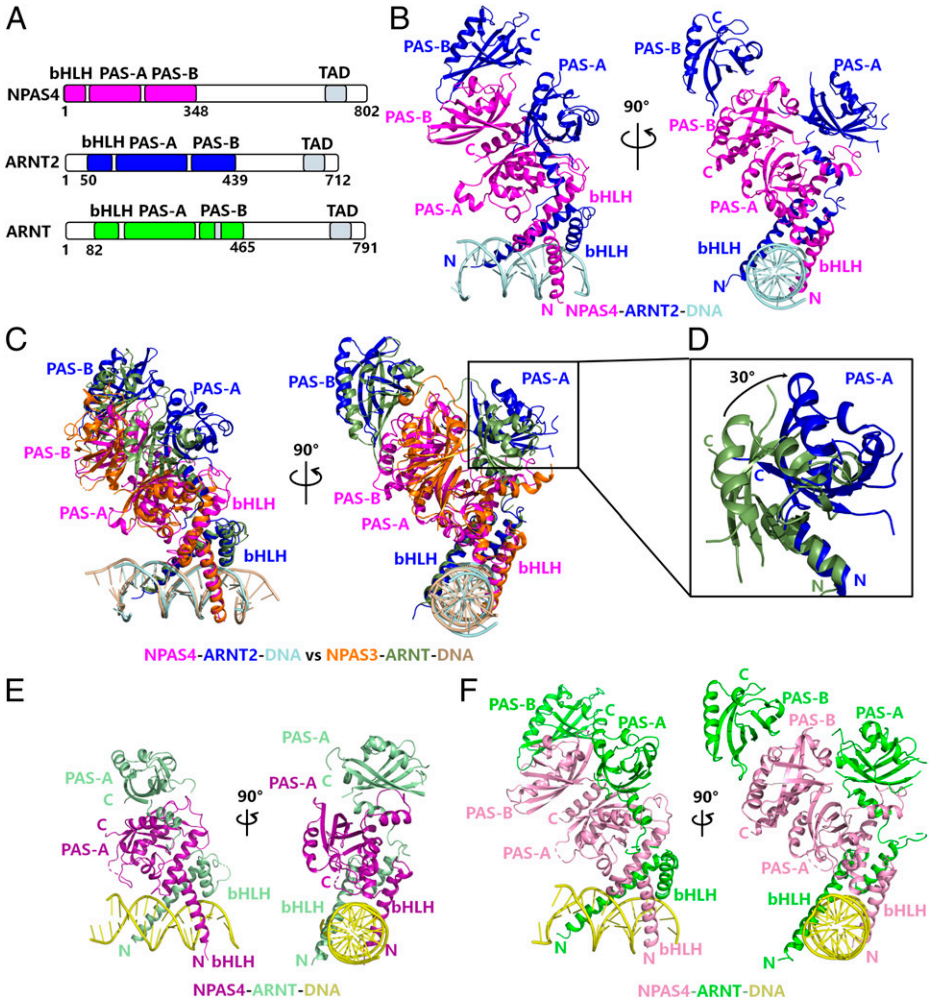

Fig. 1. Overall structures of the NPAS4 heterodimers. (*A*) Schematic representation of the domain arrangements of NPAS4, ARNT2, and ARNT proteins. TAD, transactivation domain. The locations of TADs are only shown as indicative. (*B*) Overall structure of the NPAS4-ARNT2-DNA complex in two views. (*C*) Superimposition of NPAS4-ARNT2-DNA and NPAS3-ARNT-DNA (PDB: 5SY7) structures. (*D*) An enlarged view of the superimposed PAS-A domains of ARNT2 and ARNT from the two heterodimers. (*E* and *F*) Overall structures of four-domain (*E*) and six-domain (*F*) NPAS4-ARNT heterodimers both in complex with DNA. Colors used in labels match those used in figures for the same components.

interacts with AHRR PAS-A, not with any symmetric molecules. This might explain why the electron density of the ARNT PAS-A domain was relatively poor in the AHRR-ARNT heterodimer, even when their overall resolution reached 2.4 Å (34). Therefore, we realized that although the diffraction of NPAS4-ARNT-DNA crystals could be slightly improved by the loop truncations of the ARNT PAS-A domain, the pattern of crystal packing involving this domain had to be modified to clearly visualize its conformation.

To potentially alter the way of protein packing in crystallization, we took an alternative strategy of preparing a four-domain heterodimeric complex consisting of the bHLH and PAS-A domains of both ARNT (82-345) and NPAS4 (1-205). This strategy successfully led to crystals of DNA-bound complexes diffracted to around 4-Å resolution (*SI Appendix*, Table S1). It is noteworthy that in this complex we did not introduce any truncations within the ARNT PAS-A domain. We now found complete electron densities for all of the protein segments in this structure, including the PAS-A domain of ARNT (Fig. 1*E* and *SI Appendix*, Fig. S3*C*). More importantly, when comparing this four-domain structure with the above-mentioned partially solved six-domain NPAS4-ARNT structure, we noticed that their DNA, both bHLH domains, and more importantly NPAS4 PAS-A domain could be easily superposed between the two structures (*SI Appendix*, Fig. S3*D*). These results suggest that the missing ARNT PAS-A domain in the six-domain structure likely would have the same overall structure that we observed in the four-domain complex and guided us in finally solving the six-domain NPAS4-ARNT-DNA structure (Fig. 1*F* and *SI Appendix*, Fig. S3*E* and Table S1).

A Side-by-Side Comparison between Two NPAS4 Heterodimers. With an understanding of the DNA-bound six-domain structures, we were finally able to make the comparison between two NPAS4 heterodimers (NPAS4-ARNT2 versus NPAS4-ARNT). As shown in Fig. 2*A*, the NPAS4 subunit adopted a similar overall confirmation within two complexes, with a root mean squared deviation (RMSD) value of 1.8 Å for all the Cα atoms. Meanwhile, we noticed that ARNT surface residues directly involved in the heterodimerization seen in our previous ARNT heterodimers (20, 30) were fully conserved in ARNT2 (*SI Appendix*, Fig. S2). However, despite this conservation, ARNT2 and ARNT surprisingly did not interact with NPAS4 in an identical manner (Fig. 2*A*). One important difference is that their PAS-A domains occupy notably different positions and orientations in these two NPAS4 heterodimers (Fig. 2*B*). However, since ARNT2 and ARNT proteins have a very high sequence identity at their N-terminal segments (91% for bHLH, 78% for PAS-A, and 79% for PAS-B; *SI Appendix*, Fig. S2), it was not surprising that each of these single

domains showed little structural differences in the two complexes, except for certain loop regions (RMSD values of 1.0 Å, 0.7 Å, and 0.8 Å for Cα atoms of bHLH, PAS-A, and PAS-B domains, respectively; Fig. 2*C*).

To examine the domain–domain interactions in two NPAS4 heterodimers, we numbered their six domain–domain interfaces and calculated the buried surface areas in each case (Fig. 2*D* and *SI Appendix*, Fig. S4*A*). Using the program PISA (35), the overall buried heterodimer interface areas of NPAS4-ARNT2 and NPAS4-ARNT were about 3,030 Å^2 and 3,400 Å^2, respectively (with their loop regions near the interface equally trimmed), suggesting a slightly smaller interaction area in NPAS4-ARNT2 than in NPAS4-ARNT. Since the heterodimer interfaces mainly consist of four domain–domain interfaces (i.e., Interfaces 1∼4, as indicated in Fig. 2*D*), we examined each of them closely to understand their individual contributions to the buried area difference. We found that the buried areas of Interface 2 (1,054 Å^2 versus 1,160 Å^2) and Interface 3 (128 Å^2 versus 268 Å^2) were notably smaller in NPAS4-ARNT2 than in NPAS4-ARNT (*SI Appendix*, Fig. S4*A*). The formation of both of these interfaces depends on the PAS-A domains of ARNT2 or ARNT (Fig. 2*D*), and the differential positioning of their PAS-A domains appears to account for the different buried surfaces in these two heterodimers (Fig. 2*B*).

We next conducted cell-based studies to examine how each of the domain–domain interfaces observed in NPAS4-ARNT2 and NPAS4-ARNT contributed to the stability of their full-length heterodimers in the cellular setting. Co-IP assays were performed in HEK293 cells expressing NPAS4 with a series of single mutations positioned at its domain interfaces (Fig. 2*E*). Among the four heterodimeric domain interfaces (i.e., Interfaces 1∼4), some NPAS4 mutations at Interfaces 1, 2, and 3 destabilized the heterodimerization to different extents, whereas the two mutations at Interface 4 did not show many discernable effects (Fig. 2*E*). These results confirmed our previous findings that ARNT2 or ARNT point mutations at Interfaces 3 and 4 could not disrupt their dimerization with NPAS4 as efficiently as mutations at other interfaces (20). We also found that NPAS4 mutation L74D at Interface 2 exhibited varied effects for ARNT2 and ARNT, likely due to the distinct positioning and conformation of their PAS-A domains (Fig. 2*B* and *SI Appendix*, Fig. S4*B*).

We then tested the effects of the above NPAS4 interface mutations on the transcriptional activities of both NPAS4-ARNT2 and NPAS4-ARNT (Fig. 2*F*). NPAS4 has been shown to act as a transcriptional activator on its target gene *BDNF*. We constructed a luciferase reporter based on a short DNA sequence containing the NPAS4-binding site from the *BDNF* promoter region and conducted assays in HEK293 cells, which showed no background luciferase activity in the absence of NPAS4 overexpression. As shown in Fig. 2*F*, the two NPAS4 mutations L23D and L40D at Interface 1 dramatically decreased the transcriptional activities of both NPAS4 heterodimers. Unlike mutation L23D, which clearly abolished heterodimerization in Co-IP, L40D did not fully disrupt the interactions between NPAS4 and its partners (Fig. 2*E*). Based on the structural information, the NPAS4 residue L40 is adjacent to the ARNT2 residues R75 and K78 (both conserved in ARNT as R101 and K104; *SI Appendix*, Fig. S4*C*). We speculate that after mutation of L40D, this site may form favorable salt bridges with R75 or K78, which could interfere with DNA binding but not dimerization. However, mutations at Interfaces 2∼4 were not very effective in the reduction of transcriptional activities, although some of the mutations were indeed capable of disrupting heterodimerization to variable extents (Fig. 2 *E* and *F*). As both ARNT and ARNT2 proteins are expressed in HEK293 cells, we were not able to accurately discriminate the varied effects of NPAS4 mutants on transcriptional activities of two heterodimers even when we overexpressed ARNT and ARNT2, respectively.

Since we crystallized the two NPAS4 heterodimers in complexes with the same double-stranded DNA, we wanted to measure and compare the DNA-binding affinities of these two complexes using a fluorescence polarization assay. ARNT heterodimers are known to recognize the canonical E-box motif "XXXGTG," with the bHLH domain of ARNT specifically interacting with the DNA base pairs of "GTG" through the H94, E98, and R102 residues (30). Although ARNT2 possesses a nearly identical sequence as ARNT at the bHLH domain (*SI Appendix*, Fig. S2), including the above nucleotide-interacting residues, dissociation constant (K_d) values for NPAS4 (1-348)–ARNT2 (50-439) and NPAS4 (1-348)–ARNT (82-465) were measured as 46.8 ± 3.9 nM and 13.1 ± 0.4 nM, respectively (Fig. 2*G*), suggesting a lower affinity for the NPAS4-ARNT2 heterodimer. Taking these findings into account, we speculate that the protein sequence of bHLH domains may not be the only factor determining the binding ability of NPAS4 heterodimers to DNA and that other domains outside the bHLH may further contribute to DNA-binding affinity. Consistent with this expectation, we previously showed that the HIF-2α PAS-A domain within the HIF-2α–ARNT heterodimer also contributes to DNA binding (30). In our two NPAS4 heterodimers, the different PAS-A sequences and conformations of ARNT2 versus ARNT might potentially influence overall DNA binding. Meanwhile, whether and how the structural discrepancies between two NPAS4 heterodimers further contribute to the functional differences in their DNA binding site recognition, coactivator recruitment, and/or transcriptional signatures remain fascinating questions that warrant deeper investigations in the future.

Unique Tertiary Architecture of the NPAS4 Protein. Next, we studied the NPAS4 subunit within the heterodimers to identify its unique structural features. As shown in Fig. 3*A* (and *SI Appendix*, Fig. S5), the NPAS4 PAS-A domain forms direct interfaces not only with the β-sheet face of its PAS-B domain but also with the α2 helix of the bHLH domain (i.e., Interfaces 5 and 6, respectively, in Fig. 2*D*). Consequently, the NPAS4 subunit adopts an overall compact structure constrained by these internal domain–domain interfaces. Among the class I bHLH-PAS proteins that can heterodimerize with ARNT or ARNT2 (*SI Appendix*, Fig. S1*A*), such as HIF-1α, HIF-2α, HIF-3α, NPAS1, and NPAS3, all of whose structures we previously analyzed as six-domain ARNT heterodimers (20, 30, 31), there is a significant similarity between the three HIF-α isoforms and the NPAS1 and NPAS3 protein structures with respect to their intramolecular domain–domain junctions. For simplicity, we chose HIF-2α and NPAS3 to directly compare with NPAS4 by superimposing their structures at the PAS-A domains (Fig. 3*B*). A short α-helix (Dα) within the NPAS4 PAS-A domain and its following loop flipped out toward the beginning of the α2 helix of the bHLH domain in NPAS4 (Interface 6 in Fig. 3*C*). Moreover, the PAS-A/PAS-B interface of NPAS4 was clearly different from those of HIF-2α and NPAS3 due to the presence of an additional α-helix (Jα) at the tail end of the NPAS4 PAS-B domain (Interface 5 in Fig. 3*C* and *SI Appendix*, Fig. S5). This α-helix inserted in between the PAS-A and PAS-B domains of NPAS4 and created an interface by bridging an otherwise large gap between the two domains. When superimposing representative PAS-B structures of class I bHLH-PAS proteins known to date, we found that NPAS4 uniquely contained this long α-helix at the C terminus of the PAS-B domain (Fig. 3*D*). It should be noted that at this location

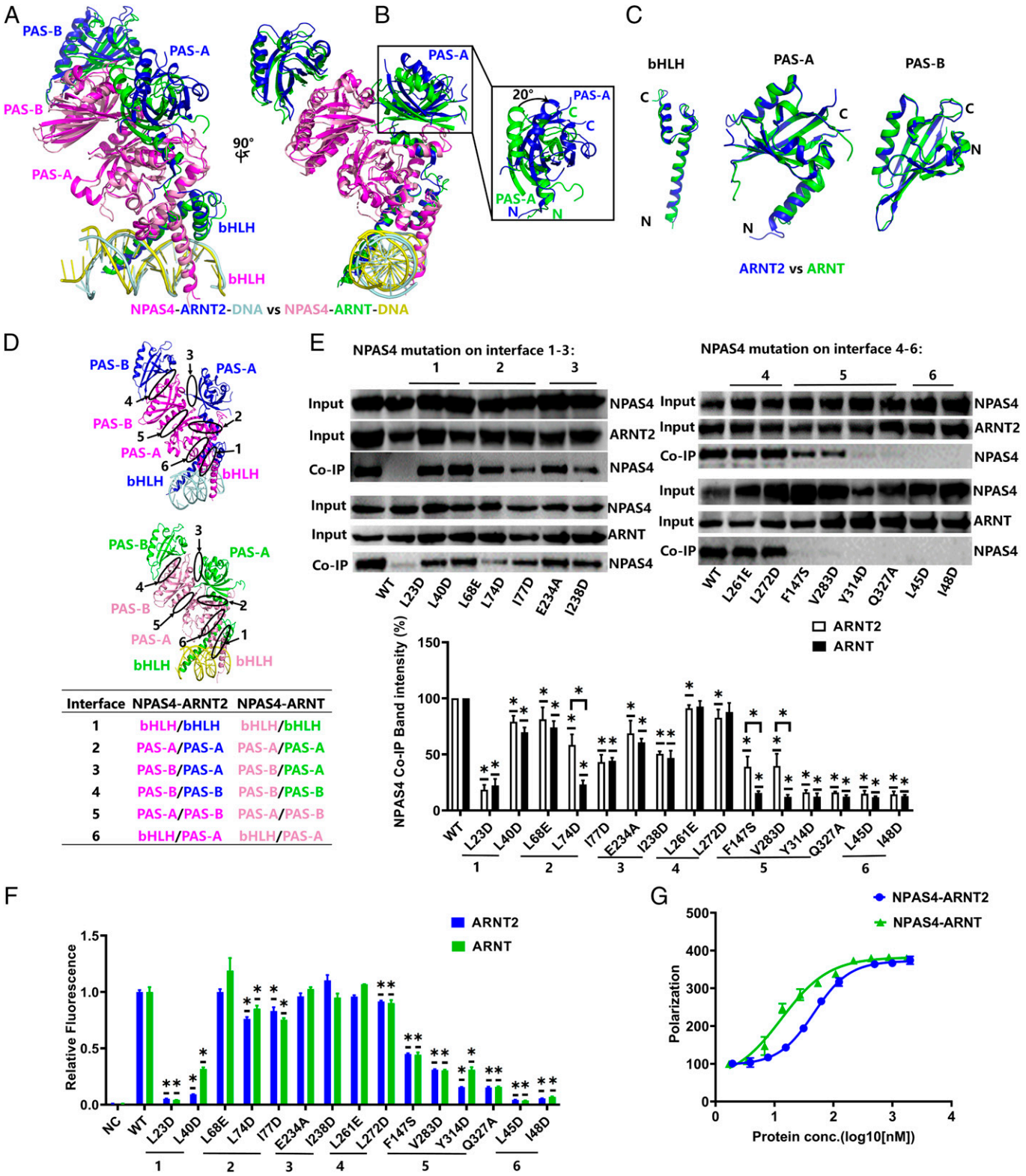

Fig. 2. Comparative analysis of two NPAS4 heterodimers. (*A*) Superimposition of NPAS4-ARNT2-DNA and NPAS4-ARNT-DNA structures. The NPAS4-ARNT-DNA structure presented here is a merged model from both the six- and four-domain complexes. Colors used in labels match those used in figures for the same components. (*B*) An enlarged view of the superimposed PAS-A domains of ARNT2 and ARNT in two NPAS4 heterodimers. (*C*) Direct comparison of single-domain structures of ARNT2 (blue) and ARNT (green). (*D*) The spatial arrangements for each of the six domain–domain interfaces in NPAS4-ARNT2 (*Top*) and NPAS4-ARNT (*Bottom*) heterodimers. Each interface is further noted in the table below. (*E*) Co-IP experiments showing the effects of NPAS4 mutations at Interfaces 1~6 on the stabilities of heterodimers formed with ARNT2 or ARNT in cells. The band densities of precipitated proteins are quantified based on experimental results in triplicate (also shown in *SI Appendix*, Fig. S8). $^*P < 0.05$. (*F*) Luciferase reporter assays evaluating the effects of NPAS4 interface mutations on transactivation of NPAS4-ARNT2 and NPAS4-ARNT heterodimers. NC, negative control (transfection of an empty vector instead of those harboring NPAS4); WT, WT NPAS4. Data are an average of three independent experiments, $^*P < 0.05$. (*G*) DNA-binding affinities of two NPAS4 heterodimers measured using fluorescence polarization. The K_d values of NPAS4-ARNT2 and NPAS4-ARNT binding to the same NPAS4 response element were 46.8 ± 3.9 nM and 13.1 ± 0.4 nM, respectively. Error bars, mean ± SD; $n = 3$ (biological replicates); conc., concentration.

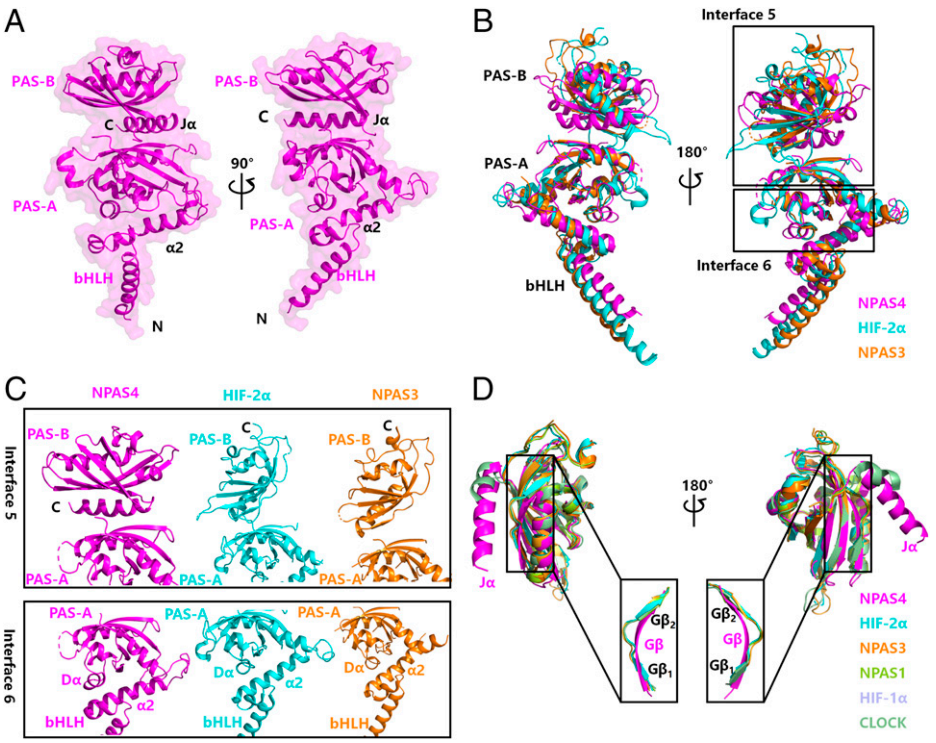

Fig. 3. The unique structure of NPAS4 subunit alone. (*A*) Overall structure of NPAS4 subunit within the NPAS4-ARNT2 heterodimer, shown in two views and with each domain labeled. (*B*) Superimposition of the NPAS4 (magenta), HIF-2α (cyan), and NPAS3 (orange) structures, each from their DNA-bound heterodimers. (*C*) Conformation comparisons of Interface 5 (PAS-A/PAS-B) and Interface 6 (bHLH/PAS-A) in NPAS4, HIF-2α, and NPAS3 subunits in the enlarged views. (*D*) Superimposition of PAS-B domain structures of class I bHLH-PAS proteins, including NPAS4 (magenta), HIF-2α (cyan), NPAS3 (orange), NPAS1 (limon), HIF-1α (light blue), and CLOCK (forest). The unique single-stranded conformation of Gβ in the NPAS4 PAS-B domain (magenta) is highlighted in two enlarged views at the bottom.

the CLOCK protein also possesses an α-helix, which is much shorter and adopts a different orientation (33). In addition, while the Gβ strand of the PAS-B domain in all other class I proteins breaks into two short strands ($Gβ_1$ and $Gβ_2$), NPAS4 is also the only one having this Gβ strand as an undivided structural element (Fig. 3*D*), due to the unique presence of an extra residue at this position in NPAS4 (*SI Appendix*, Fig. S6).

Since both ARNT2 and ARNT wrap around NPAS4, we wanted to know how mutations at intra-NPAS4 domain interfaces would influence the heterodimers as a whole. Therefore, we mutated several residues of NPAS4 at Interfaces 5 and 6 (Fig. 2*D* and *SI Appendix*, Fig. S4 *D* and *E*) and tested their effects using the above-mentioned Co-IP and luciferase reporter assays (Fig. 2 *E* and *F*). Interestingly, these point mutations profoundly disrupted the heterodimerization of both NPAS4-ARNT2 and NPAS4-ARNT (Fig. 2*E*), suggesting that a proper domain integration within the NPAS4 protein might be a precondition for its ability to heterodimerize with either ARNT2 or ARNT. It is noteworthy that two mutations near Interface 5 (F147S and V283D) showed less impactful effects on NPAS4's binding to ARNT2 than to ARNT (Fig. 2*E*), implying a possible better accommodation of ARNT2 in heterodimerization. The buried surface area of Interface 5 in NPAS4-ARNT2 (756 $Å^2$) is relatively larger than that of NPAS4-ARNT (681 $Å^2$) (*SI Appendix*, Fig. S4*A*), which we speculate might contribute to the more drastic effects of F174S and V283D mutations on the disruption of the NPAS4-ARNT complex than that of NPAS4-ARNT2. This discrepancy may also relate to the smaller heterodimer interface area in NPAS4-ARNT2 than in NPAS4-ARNT (*SI Appendix*, Fig. S4*A*). In the luciferase reporter assays, we found that NPAS4 mutations at Interface 5 reduced more than half and those at Interface 6 reduced more than 90% of transcriptional activities for both heterodimers compared with the wild-type (WT) NPAS4 (Fig. 2*F*). These results confirmed that the intraprotein domain arrangement of NPAS4 is indispensable for its heterodimerization and transcriptional activity.

The F147S variant has been highlighted in a human genome database, and a previous study confirmed that this mutation could reduce both the transcriptional activity of NPAS4 and its dimerization with the ARNT2 subunit (36). A location at the dimeric interface was proposed for NPAS4 residue F147 based on homology modeling in that study (36). However, we found that in the structure of the NPAS4-ARNT2 heterodimer, F147 located at the Gβ strand of NPAS4 PAS-A domain, with its side chain residing at the inner face of the β-sheet (*SI Appendix*, Fig. S4*D*). Since F147 is very close to the tail Jα of the PAS-B domain mediating the intraprotein PAS-A/PAS-B interface, it is not surprising that this F147S mutation would compromise the proper domain arrangement of NPAS4.

Comparison of NPAS4-ARNT with Other ARNT Heterodimers. In our prior mutagenesis study, NPAS4, one of the least conserved members in the phylogenetic tree of the bHLH-PAS family (*SI Appendix*, Fig. S1*B*), showed more resistance to heterodimerization disruption when ARNT mutations were introduced at the dimer interfaces relative to other class I members (20). To unveil the possible reason behind this phenomenon, we directly compared NPAS4-ARNT with the other ARNT heterodimeric structures solved previously. We chose NPAS3-ARNT and HIF-2α–ARNT structures as representatives for the comparison, since these two were also solved when DNA elements were bound. When we superimposed NPAS3-ARNT with HIF-2α–ARNT by the bHLH domains and DNA (Fig. 4*A* and *SI Appendix*, Fig. S7*A*), the RMSD value was only 1.9 Å for all the Cα atoms. When we

compared each of these structures with NPAS4-ARNT, the differences were substantially larger, with RMSD values of 4.2 Å and 3.4 Å, respectively. These values and their different domain arrangements as shown in Fig. 4*A* (and *SI Appendix*, Fig. S7*A*) all suggest that NPAS4-ARNT is notably more divergent from the previous ARNT heterodimers.

One key feature of ARNT in all the heterodimeric structures is that its bHLH, PAS-A, and PAS-B domains are spatially displaced from each other (21), providing ARNT with a high degree of flexibility to accommodate its 10 different dimerization partners (*SI Appendix*, Fig. S1*A*). As described in the previous section, NPAS4 protein exhibits a very unique domain arrangement, especially at its PAS-A/PAS-B domain interface (Fig. 3 *A* and *C*). Presumably as a result, the three domains of ARNT wrapping around NPAS4 take on distinct conformations in the accommodation. Through a direct comparison of the dissected heterodimeric domain interfaces (i.e., Interfaces 1∼4 as shown in Fig. 4*B* and *SI Appendix*, Fig. S7*B*), we found that among them, Interfaces 2 and 3 of NPAS4-ARNT are each quite different from those of NPAS3-ARNT and HIF-2α–ARNT. When taking a closer look, we noticed that ARNT utilizes a different set of residues to form Interfaces 2 and 3 with NPAS4 compared with those for NPAS3 or HIF-2α (*SI Appendix*, Fig. S2). Interestingly, some of these NPAS4-contacting ARNT residues (e.g., I178, F303, and V304) are not fully conserved in ARNT2 (corresponding to residues V152, Y277, and A278), which may account for the variations of these two interfaces between NPAS4-ARNT and NPAS4-ARNT2.

To experimentally confirm the observation that NPAS4 interacts with ARNT differently in comparison with other class I members, we mutated several ARNT interface residues and checked their influences on the dimerization of NPAS4-ARNT and NPAS3-ARNT using Co-IP. As shown in Fig. 4*C*, these mutations at Interfaces 1∼3 caused contrasting effects on ARNT's dimerization with NPAS4 and NPAS3. These results all indicate that NPAS4 heterodimerizes with ARNT in a distinct manner compared to other class I members of this family.

Discussion

The bHLH-PAS proteins are heterodimeric transcription factors with distinct cellular enrichments and unique target genes, potentially acting as sensitive monitors and transducers of environmental or physiological stimuli, including day-night cycles, hypoxia, xenobiotics, cellular metabolites, and neuron-specific signals (37–39). The unifying architectural feature of this family is that at the N-terminal segments they all possess the same three domains (i.e., bHLH, PAS-A, and PAS-B), where the PAS domains have the capacity to harbor small molecules and the entirety of the three domains are involved in dimerization (21). While the corresponding domain–domain interfaces do form (i.e., the bHLH/bHLH, PAS-A/PAS-A, and PAS-B/PAS-B interfaces numbered as 1, 2, and 4 in Fig. 2*D*) in the dimerization, there are additional and variable interprotein or intraprotein domain interfaces (e.g., Interfaces 3, 5, and 6 in Fig. 2*D*). This interesting phenomenon that not all heterodimers are identically assembled was first noticed when we reported the six-domain structure of HIF-2α–ARNT and compared it with the six-domain structure of circadian regulator CLOCK-BMAL1 (brain and muscle ARNT-like protein 1, also called ARNTL) (30, 33). The dramatically different quaternary architectures between HIF-2α–ARNT and CLOCK-BMAL1 indicate how the class II members ARNT and BMAL1 each specifically choose different class I members for heterodimerization, thus dividing the whole family into two major groups (*SI Appendix*, Fig. S1*A*). However, recent studies have also revealed the cross-talk and even direct interactions of BMAL1 with AHR or HIF-1α (40–42), which could potentially drive a multitude of signal-responsive functions, with further dimerization modalities to be explored among the bHLH-PAS proteins.

Within the ARNT group of heterodimers (*SI Appendix*, Fig. S1*A*), we previously found initial evidence that while several class I members heterodimerized in a similar way (such as HIF-1α, HIF-2α, NPAS1, NPAS3, and SIM1), there were clear distinctions for NPAS4 and AHR (20). Here, by directly comparing the structure of NPAS4-ARNT with other ARNT heterodimers, we found that indeed the NPAS4-ARNT heterodimer contains distinctive properties at both interprotein and intraprotein levels (Figs. 3 and 4*B*). One unique feature of NPAS4 is the long α-helix Jα at the C terminus of its PAS-B domain, which inserts into the large gap between NPAS4's PAS-A and PAS-B domains (Fig. 3 *A* and *C*). Interestingly, we found that this Jα is also present in the predicted NPAS4 structure by the new AlphaFold2 algorithm (43), suggesting that the formation of this α-helix may be an intrinsic and preexisting feature of NPAS4 that does not simply form upon dimerization. We found that the proper domain arrangement internal to NPAS4 is a key prerequisite for heterodimerization with ARNT (Fig. 2*E*) and may actually define their distinct interprotein domain interfaces, especially as involving the ARNT PAS-A domain. While ARNT is a common partner for many class I members, we found that the precise mode of its heterodimerization to different partners is not identical.

In addition to ARNT, ARNT2 also forms functional heterodimers with NPAS4, particularly in neuronal cells and tissues. Distinct stimulus-responding NPAS4 heterodimers have been reported in hippocampal neurons (23). Action potentials induce NPAS4-ARNT2 to preferentially bind to promoter regions, whereas excitatory postsynaptic potentials induce NPAS4-ARNT to bind at enhancers. Despite their highly conserved protein sequences (*SI Appendix*, Fig. S2), ARNT2 and ARNT exhibit variable DNA-binding abilities in their complexes with NPAS4 (Fig. 2*G*), and their distinct quaternary architectures could be the reason behind these differences (Fig. 2*A*). ARNT is known to recruit coactivators by its PAS-B domain (44, 45). Although there is no evidence as to whether the PAS-A domain is also able to do so, we speculate that the distinct manners in which ARNT and ARNT2 heterodimerize with NPAS4 may modulate the transcriptional output of these complexes differentially, especially as their PAS-A domains employ distinct conformations (Fig. 2*B*). Future studies including the exploration and identification of interacting proteins for these two complexes are warranted to better reveal the physiological roles of the two NPAS4 heterodimers.

As mentioned earlier, due to the presence of two ligand-binding domains in NPAS4's PAS domains and its linkage to neuropsychiatric and metabolic diseases, this protein could be viewed as a novel and actionable drug target. Like other members of the bHLH-PAS family (20), the PAS-B domains of NPAS4 and ARNT2 contain sizable cavities [146 $Å^3$ and 160 $Å^3$, respectively, as calculated using PyVOL (46)] for the binding of small molecules. Since NPAS4 plays a protective role in different cells (including neurons, pancreatic β cells, and endothelial cells) and its down-regulation is linked to the progression of diseases, agonists that enhance NPAS4 activity may be beneficial in treating unmet psychiatric and other diseases. We previously reported agonists for HIF-2α and more recently identified a physiological agonist for HIF-3α, all of which bind to the PAS-B domains of HIF-α proteins and enhance transcriptional activities by promoting the stability of their heterodimerizations with ARNT (31, 47). Besides the intraprotein cavities, the quaternary structures of two

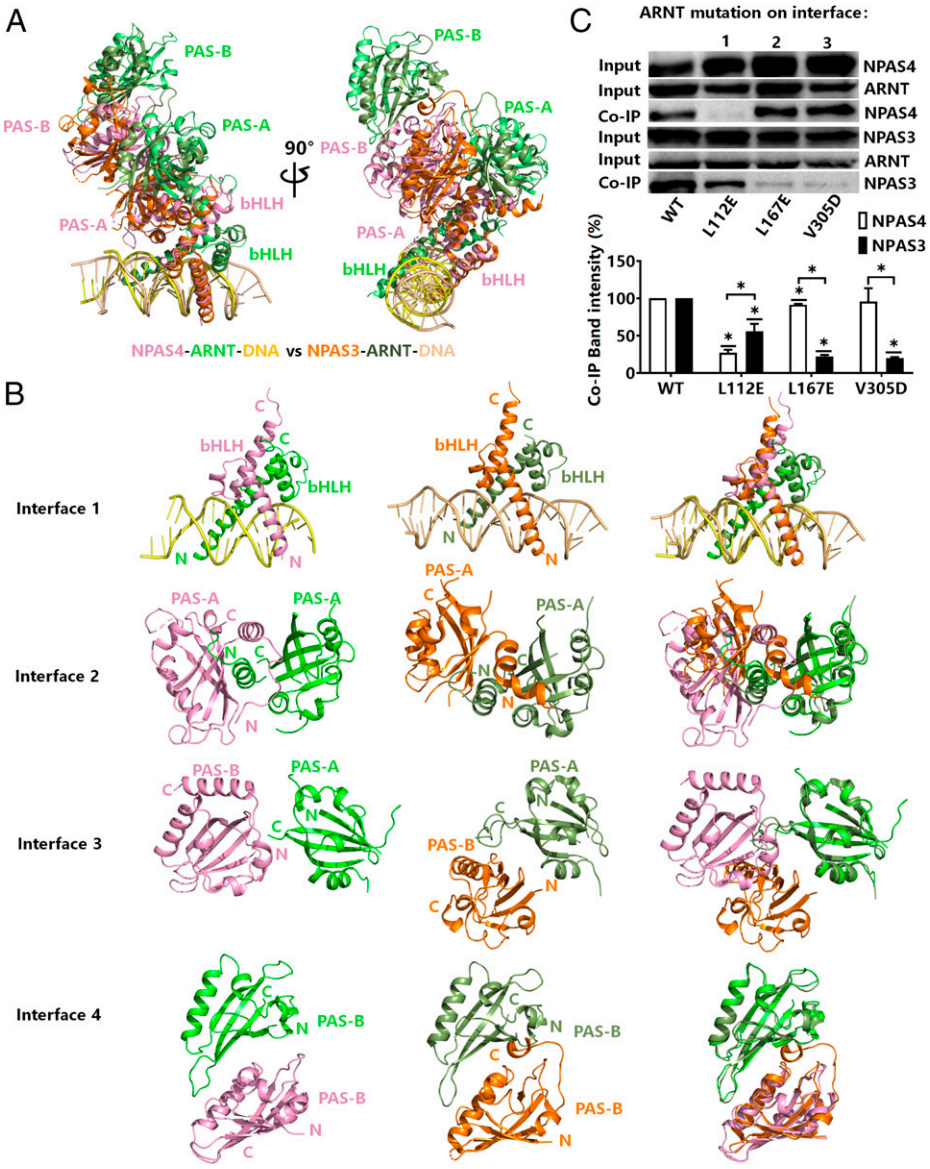

Fig. 4. Comparative analysis of the NPAS4-ARNT and NPAS3-ARNT heterodimers. (*A*) Superimposition of NPAS4-ARNT-DNA and NPAS3-ARNT-DNA structures. The NPAS4-ARNT-DNA structure presented here is a merged model from both the six- and four-domain complexes. (*B*) Direct conformation comparison of dimeric interfaces (Interfaces 1∼4) in two heterodimers. (*C*) Co-IP experiments showing the effects of ARNT mutations at Interfaces 1∼3 on the stabilities of its heterodimers formed with NPAS4 and NPAS3. Colors used in labels match those used in figures for the same components. The band densities of precipitated proteins are quantified based on experimental results in triplicate (also shown in *SI Appendix*, Fig. S9). *$P < 0.05$. Error bars, mean ± SD.; n = 3 (biological replicates).

NPAS4 heterodimers provide opportunities to target the protein–protein interfaces (PPIs). For example, surface cavities formed near these PPIs can be potentially targeted in the rational design of small molecules changing heterodimer stabilities. In addition, the different dimeric interfaces, especially those involving the PAS-A domains of ARNT or ARNT2, could become good target PPIs to discover selective small molecules for these NPAS4 heterodimers. The structural studies presented here provide the initial framework to help drive future discoveries of therapeutic NPAS4-targeting drugs.

Materials and Methods

Plasmid Construction and Site-Directed Mutagenesis. For protein overexpression in *Escherichia coli*, DNA fragment encoding mouse NPAS4 (Uniprot accession number Q8BGD7) was cloned into the pSJ2 (C-terminal His-tagged) vector; those encoding ARNT (Uniprot accession number P53742) and ARNT2 (GenBank accession number AAH54546.1) were cloned into pMKH (with no tag) vector. For Co-IP studies, full-length NPAS4 and NPAS3 were cloned into the pCMV-Tag1 vector (C-terminal HA-tagged), whereas ARNT and ARNT2 were cloned into the pCMV-Tag4 vector (C-terminal Flag-tagged). Site-directed mutagenesis was conducted by PCR using paired primers with desired mutations and confirmed by DNA sequencing. To construct a luciferase reporter containing the NPAS4 response element, three tandem repeats of the *BDNF* promoter core region (48) were subcloned into the pGL3 vector. More details about the expression vectors and protein sequences are provided in *SI Appendix*, Table S2.

Protein Expression and Purification. The plasmid pSJ2-NPAS4 was cotransformed with pMKH-ARNT or pMKH-ARNT2 into *Transetta* (DE3) competent cells (TransGen Biotech). Similar to the procedure previously described (30), recombinant protein complexes were expressed at 16 °C overnight and then purified with a three-step chromatography protocol using Ni Bestarose FF (Bestchrom), SP Sepharose (Solarbio), and a Superdex 200-pg gel-filtration column (Cytiva). To prepare the DNA-bound NPAS4-ARNT and NPAS4-ARNT2 complexes, the synthetic 16-mer double-stranded DNA (F: 5′-GGAGGTCGTGAGTGAT-3′ and R: 5′-CCATCACTCACGACCT-3′) was mixed with the heterodimeric proteins at a molar ratio of 1.2:1. Subsequently, the mixture of protein-DNA complex was further purified by gel filtration in running buffer containing 20 mM Tris (pH 8.0) and 150 mM NaCl.

Crystallization and X-ray Data Collection. NPAS4-ARNT2-DNA crystals were grown at 16 °C in sitting drops formed by an equal volume of complex (4 mg/mL) and a reservoir consisting of 200 mM sodium citrate tribasic, 12% PEG3350. Crystallization of the DNA-bound six-domain NPAS4-ARNT heterodimer was carried out similarly by mixing an equal volume of complex (4 mg/mL) and a reservoir solution containing 120 mM ammonium tartrate dibasic, 9% PEG3350. Finally, the DNA-bound four-domain NPAS4-ARNT crystals were grown in drops formed by an equal volume of complex (2 mg/mL) and a reservoir consisting of 50 mM sodium citrate tribasic, 8% PEG3350. Before being flash frozen in liquid nitrogen, all crystals were soaked in the reservoir plus 25% glycerol as the cryoprotectant. Diffraction data for the NPAS4-ARNT2-DNA and NPAS4-ARNT-DNA crystals were collected at the Shanghai Synchrotron Radiation Facility (SSRF) on the BL02U1 and BL19U1 beamlines (49). The data were subsequently processed using the HKL3000 program (50).

Structure Determination and Refinement. The structures of the NPAS4-ARNT2-DNA and NPAS4-ARNT-DNA complexes were all solved by the molecular replacement method with Phaser (51), using the HIF-2α–ARNT-DNA structure (Protein Data Bank [PDB] accession number 4ZPK) as a search model (30). Further manual model building was facilitated using Coot (52), combined with structure refinement using Phenix (53). The diffraction data and final statistics are summarized in *SI Appendix*, Table S1. The Ramachandran statistics are 94.9%/0%, 93.9%/0.2%, and 94.1%/0% (favored/outliers) for the structures of NPAS4-ARNT2-DNA and the six-domain and four-domain NPAS4-ARNT-DNA complexes, respectively. All the structural figures were prepared using PyMol (Schrödinger, LLC).

Fluorescence Polarization DNA-Binding Assay. The 21-mer fluoresceinated double-stranded DNA was prepared by combining a 6-FAM–labeled forward strand (5′-GGATGAGGTCGTGAGTGATGA-3′) with the unlabeled reverse strand (5′-TCATCACTCACGACCTCATCC-3′) in the annealing buffer of 10 mM Tris (pH 7.5), 1 mM ethylene diamine tetraacetic acid (EDTA), and 2 mM $MgCl_2$. In the binding assay, 2 nM DNA was incubated with purified protein complexes for 30 min, whereas the final protein concentrations were varied by serial dilution in the binding buffer consisting of 20 mM Tris (pH 8.0), 50 mM NaCl, and 10 mM dithiothreitol. The fluorescence polarization signals were recorded using a Spark microplate reader (Tecan) and processed as previously described (30).

Co-IP. HEK293 cells (Procell, CL-0001) were seeded in 6-cm dishes and cultured in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum (Thermo Fisher Scientific) at 37 °C with 5% CO_2. One day later, cells were cotransfected with 800 ng pCMV-Tag1-NPAS4 or pCMV-Tag1-NPAS3 (WT or mutants) as well as 800 ng pCMV-Tag4-ARNT (WT or mutants) or pCMV-Tag4-ARNT2 plasmids using 3.2 μL jetPRIME reagent (Polyplus-transfection). After a 4-h incubation time, the medium was refreshed. Another 24 h later, cells were harvested and immunoprecipitation was performed similarly to our previous work (20). After the protein concentration measurement for each sample, a fraction of supernatant was saved as input for Western blot examinations using anti-Flag polyclonal antibody (Proteintech, 20543–1-AP) and anti-HA polyclonal antibody (Proteintech, 51064–2-AP). Immunoprecipitation was performed with the supernatant and anti-Flag affinity gel suspension (Beyotime, P2271) according to the manufacturer's instructions, followed by Western blot using anti-HA polyclonal antibody (Proteintech, 51064–2-AP). Horseradish peroxidase–conjugated goat anti-rabbit immunoglobulin G antibody (D110058) was purchased from Sangong Biotech. The primary and secondary antibodies were used as 1:1,000 and 1:6,000 dilutions, respectively. All Co-IP experiments were performed independently at least three times. Band quantification was conducted with Image J software (https://imagej.nih.gov/ij/), and further statistical analysis was performed using the paired Student's *t*-test, with respect to the corresponding control sample in each experiment.

Luciferase Reporter Assay. HEK293 cells were seeded in 48-well plates and 1 d later transfected with 40 ng of pCMV-Tag1-NPAS4 (WT, mutants, or empty vector) and 80 ng pCMV-Tag4-ARNT or ARNT2 plasmids, 1 ng pRL (control Renilla luciferase), and 40 ng BDNF promoter-Luc reporter using 0.32 μL jetPRIME reagent for each well. Medium was refreshed 4 h after the transfection, and luciferase activity was measured another 24 h later using the Dual-Glo Luciferase Assay System (Promega). Final data were normalized by the relative ratio of firefly and Renilla luciferase activity.

Data, Materials, and Software Availability. Coordinates and structure factors have been deposited in the PDB (https://www.rcsb.org/) under accession numbers 7XI3 (NPAS4-ARNT2-DNA) (54), 7XI4 (six-domain NPAS4-ARNT-DNA) (55), and 7XHV (four-domain NPAS4-ARNT-DNA) (56).

ACKNOWLEDGMENTS. We thank Dr. Xiaoju Li from the State Key Laboratory of Microbial Technology of Shandong University for her help with experiments of X-ray diffraction. We also thank the staff of the BL19U1 beamline of the National Facility for Protein Science in Shanghai at the SSRF, as well as the staff of the BL02U1 beamline at the SSRF for their assistance during data collection. This work was supported by grants from the Shandong Provincial Natural Science Foundation, China (ZR2021JQ30), the National Natural Science Foundation of China (22177063), the Interdisciplinary Innovative Research Group Program of Shandong University (2020QNQT009), the Natural Science Foundation of Jiangsu Province (BK20170399), and the Taishan Scholars Project of Shandong (tsqn201909004) to D.W.

Author affiliations: [a]Helmholtz International Lab, State Key Laboratory of Microbial Technology, Shandong University, Qingdao 266237, China; [b]Marine College, Shandong University, Weihai 264209, China; [c]Target Discovery Institute, University of Oxford, Oxford OX3 7FZ, United Kingdom; and [d]Suzhou Research Institute, Shandong University, Suzhou 215123, China

1. N. Ooe, K. Saito, N. Mikami, I. Nakatuka, H. Kaneko, Identification of a novel basic helix-loop-helix-PAS factor, NXF, reveals a Sim2 competitive, positive regulatory role in dendritic-cytoskeleton modulator Drebrin gene expression. *Mol. Cell. Biol.* **24**, 608–616 (2004).
2. M. Moser, R. Knoth, C. Bode, C. Patterson, LE-PAS, a novel Arnt-dependent HLH-PAS protein, is expressed in limbic tissues and transactivates the CNS midline enhancer element. *Brain Res. Mol. Brain Res.* **128**, 141–149 (2004).
3. W. D. Flood, R. W. Moyer, A. Tsykin, G. R. Sutherland, S. A. Koblar, Nxf and Fbxo33: Novel seizure-responsive genes in mice. *Eur. J. Neurosci.* **20**, 1819–1826 (2004).
4. M. Shamloo *et al.*, Npas4, a novel helix-loop-helix PAS domain protein, is regulated in response to cerebral ischemia. *Eur. J. Neurosci.* **24**, 2705–2720 (2006).
5. Y. Lin *et al.*, Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455**, 1198–1204 (2008).
6. B. L. Bloodgood, N. Sharma, H. A. Browne, A. Z. Trepman, M. E. Greenberg, The activity-dependent transcription factor NPAS4 regulates domain-specific inhibition. *Nature* **503**, 121–125 (2013).
7. I. Spiegel *et al.*, Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs. *Cell* **157**, 1216–1229 (2014).
8. K. Ramamoorthi *et al.*, Npas4 regulates a transcriptional program in CA3 required for contextual memory formation. *Science* **334**, 1669–1675 (2011).
9. X. Sun, Y. Lin, Npas4: Linking neuronal activity to memory. *Trends Neurosci.* **39**, 264–275 (2016).
10. P. Xu *et al.*, NPAS4 regulates the transcriptional response of the suprachiasmatic nucleus to light and circadian behavior. *Neuron* **109**, 3268–3282.e6 (2021).
11. L. Coutellier, S. Beraki, P. M. Ardestani, N. L. Saw, M. Shamloo, Npas4: A neuronal transcription factor with a key role in social and cognitive functions relevant to developmental disorders. *PLoS One* **7**, e46604 (2012).
12. D. Ibi *et al.*, Social isolation rearing-induced impairment of the hippocampal neurogenesis is associated with deficits in spatial memory and emotion-related behaviors in juvenile mice. *J. Neurochem.* **105**, 921–932 (2008).
13. A. S. C. Louis Sam Titus, D. Sharma, M. S. Kim, S. R. D'Mello, The Bdnf and Npas4 genes are targets of HDAC3-mediated transcriptional repression. *BMC Neurosci.* **20**, 65 (2019).
14. J. J. Rossi *et al.*, Molecular characterisation of rare loss-of-function NPAS3 and NPAS4 variants identified in individuals with neurodevelopmental disorders. *Sci. Rep.* **11**, 6602 (2021).
15. J. Fu, O. Guo, Z. Zhen, J. Zhen, Essential functions of the transcription factor Npas4 in neural circuit development, plasticity, and diseases. *Front. Neurosci.* **14**, 603373 (2020).
16. P. V. Sabatini *et al.*, Npas4 is a novel activity-regulated cytoprotective factor in pancreatic β-cells. *Diabetes* **62**, 2808–2820 (2013).
17. T. Speckmann, P. V. Sabatini, C. Nian, R. G. Smith, F. C. Lynn, Npas4 transcription factor expression is regulated by calcium signaling pathways and prevents tacrolimus-induced cytotoxicity in pancreatic beta cells. *J. Biol. Chem.* **291**, 2682–2695 (2016).
18. J. S. Esser *et al.*, The neuronal transcription factor NPAS4 is a strong inducer of sprouting angiogenesis and tip cell formation. *Cardiovasc. Res.* **113**, 222–223 (2017).
19. A. Kovalchuk *et al.*, Growth of triple negative and progesterone positive breast cancer causes oxidative stress and down-regulates neuroprotective transcription factor NPAS4 and NPAS4-regulated genes in hippocampal tissues of tumorgraft mice–An aging connection. *Front. Genet.* **9**, 58 (2018).
20. D. Wu, X. Su, N. Potluri, Y. Kim, F. Rastinejad, NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors. *eLife* **5**, e18790 (2016).
21. D. Wu, F. Rastinejad, Structural characterization of mammalian bHLH-PAS transcription factors. *Curr. Opin. Struct. Biol.* **43**, 1–9 (2017).
22. K. Hirose *et al.*, cDNA cloning and tissue-specific expression of a novel basic helix-loop-helix/PAS factor (Arnt2) with close sequence similarity to the aryl hydrocarbon receptor nuclear translocator (Arnt). *Mol. Cell. Biol.* **16**, 1706–1713 (1996).
23. G. S. Brigidi *et al.*, Genomic decoding of neuronal depolarization by stimulus-specific NPAS4 heterodimers. *Cell* **179**, 373–391.e27 (2019).

24. J. L. Michaud, C. DeRossi, N. R. May, B. C. Holdener, C. M. Fan, ARNT2 acts as the dimerization partner of SIM1 for the development of the hypothalamus. *Mech. Dev.* **90**, 253–261 (2000).
25. E. Maltepe, B. Keith, A. M. Arsham, J. R. Brorson, M. C. Simon, The role of ARNT2 in tumor angiogenesis and the neural response to hypoxia. *Biochem. Biophys. Res. Commun.* **273**, 231–238 (2000).
26. N. Hao, V. L. Bhakti, D. J. Peet, M. L. Whitelaw, Reciprocal regulation of the basic helix-loop-helix/Per-Arnt-Sim partner proteins, Arnt and Arnt2, during neuronal differentiation. *Nucleic Acids Res.* **41**, 5626–5638 (2013).
27. O. Hankinson, Why does ARNT2 behave differently from ARNT? *Toxicol. Sci.* **103**, 1–3 (2008).
28. T. H. Scheuermann *et al.*, Artificial ligand binding within the HIF2α PAS-B domain of the HIF2 transcription factor. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 450–455 (2009).
29. R. Cardoso *et al.*, Identification of Cys255 in HIF-1α as a novel site for development of covalent inhibitors of HIF-1α/ARNT PasB domain protein-protein interaction. *Protein Sci.* **21**, 1885–1896 (2012).
30. D. Wu, N. Potluri, J. Lu, Y. Kim, F. Rastinejad, Structural integration in hypoxia-inducible factors. *Nature* **524**, 303–308 (2015).
31. X. Diao *et al.*, Identification of oleoylethanolamide as an endogenous ligand for HIF-3α. *Nat. Commun.* **13**, 2529 (2022).
32. S. H. Seok *et al.*, Structural hierarchy controlling dimerization and target DNA recognition in the AHR transcriptional complex. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5431–5436 (2017).
33. N. Huang *et al.*, Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. *Science* **337**, 189–194 (2012).
34. S. Sakurai, T. Shimizu, U. Ohto, The crystal structure of the AhRR-ARNT heterodimer reveals the structural basis of the repression of AhR-mediated transcription. *J. Biol. Chem.* **292**, 17609–17616 (2017).
35. E. Krissinel, Stock-based detection of protein oligomeric states in jsPISA. *Nucleic Acids Res.* **43** (W1), W314–W319 (2015).
36. D. C. Bersten, J. B. Bruning, D. J. Peet, M. L. Whitelaw, Human variants in the neuronal basic helix-loop-helix/Per-Arnt-Sim (bHLH/PAS) transcription factor complex NPAS4/ARNT2 disrupt function. *PLoS One* **9**, e85768 (2014).
37. B. E. McIntosh, J. B. Hogenesch, C. A. Bradfield, Mammalian Per-Arnt-Sim proteins in environmental adaptation. *Annu. Rev. Physiol.* **72**, 625–645 (2010).
38. D. C. Bersten, A. E. Sullivan, D. J. Peet, M. L. Whitelaw, bHLH-PAS proteins in cancer. *Nat. Rev. Cancer* **13**, 827–841 (2013).
39. B. Greb-Markiewicz, M. Kolonko, Subcellular localization signals of bHLH-PAS proteins: Their significance, current state of knowledge and future perspectives. *Int. J. Mol. Sci.* **20**, 4746 (2019).
40. C. X. Xu, S. L. Krager, D. F. Liao, S. A. Tischkau, Disruption of CLOCK-BMAL1 transcriptional activity is responsible for aryl hydrocarbon receptor-mediated regulation of Period1 gene. *Toxicol. Sci.* **115**, 98–108 (2010).
41. J. L. Fribourgh, C. L. Partch, Assembly and function of bHLH-PAS complexes. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5330–5332 (2017).
42. M. E. Vaughan *et al.*, Cryptochromes suppress HIF1α in muscles. *iScience* **23**, 101338 (2020).
43. M. Varadi *et al.*, AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50** (D1), D439–D444 (2022).
44. C. L. Partch, P. B. Card, C. A. Amezcua, K. H. Gardner, Molecular basis of coiled coil coactivator recruitment by the aryl hydrocarbon receptor nuclear translocator (ARNT). *J. Biol. Chem.* **284**, 15184–15192 (2009).
45. Y. Guo *et al.*, Regulating the ARNT/TACC3 axis: Multiple approaches to manipulating protein/protein interactions with small molecules. *ACS Chem. Biol.* **8**, 626–635 (2013).
46. R. H. B. Smith, A. C. Dar, A. Schlessinger, PyVOL: A PyMOL plugin for visualization, comparison, and volume calculation of drug-binding sites. *bioRxiv* [Preprint]. https://doi.org/10.1101/816702 (Accessed 24 October 2019).
47. D. Wu *et al.*, Bidirectional modulation of HIF-2 activity through chemical ligands. *Nat. Chem. Biol.* **15**, 367–376 (2019).
48. Y. Funahashi *et al.*, Phosphorylation of Npas4 by MAPK regulates reward-related gene expression and behaviors. *Cell Rep.* **29**, 3235–3252.e9 (2019).
49. W. Z. Zhang *et al.*, The protein complex crystallography beamline (BL19U1) at the Shanghai synchrotron radiation facility. *Nucl. Sci. Tech.* **30**, 170 (2019).
50. W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, HKL-3000: The integration of data reduction and structure solution–From diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859–866 (2006).
51. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
52. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
53. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
54. X. Sun *et al.*, Structures of NPAS4-ARNT and NPAS4-ARNT2 heterodimers reveal new dimerization modalities in the bHLH-PAS transcription factor family. PDB. https://www.rcsb.org/structure/unreleased/7XI3. Deposited 11 April 2022.
55. X. Sun *et al.*, Structures of NPAS4-ARNT and NPAS4-ARNT2 heterodimers reveal new dimerization modalities in the bHLH-PAS transcription factor family. PDB. https://www.rcsb.org/structure/unreleased/7XI4. Deposited 12 April 2022.
56. X. Sun *et al.*, Structures of NPAS4-ARNT and NPAS4-ARNT2 heterodimers reveal new dimerization modalities in the bHLH-PAS transcription factor family. PDB. https://www.rcsb.org/structure/unreleased/7XHV. Deposited 10 April 2022.