Corepressor inactivation through tandem linear motif-induced polymerization

Nishit Goradia[1,#], Stefan Werner[2,3,#], Edukondalu Mullapudi[1], Sarah Greimeier[2], Lina Merkens[2], Andras Lang[4], Haydyn Mertens[1], Aleksandra Węglarz[2], Simon Sander[2], Grzegorz Chojnowski[1], Harriet Wikman[2], Oliver Ohlenschläger[4], Gunhild von Amsberg[,5,6], Klaus Pantel[2,*], Matthias Wilmanns[1,7,*]

[#] Equal contributions

[1] European Molecular Biology Laboratory, Hamburg Unit, Notkestrasse 85, 22607 Hamburg, Germany

[2] University Medical Center Hamburg-Eppendorf, Department of Tumor Biology, University Cancer Center Hamburg, Martinistrasse 52, 20246 Hamburg, Germany

[3] University Medical Center Hamburg-Eppendorf, Mildred Scheel Cancer Career Center HaTriCS4, Martinistrasse 52, 20246 Hamburg, Germany

[4] Leibniz Institute on Aging, Fritz-Lipmann-Institute, Beutenbergstraße 11, 07745 Jena, Germany

[5] Martini Clinic, Martinistrasse 52, 20246 Hamburg, Germany

[6] University Medical Center Hamburg-Eppendorf, Department of Hematology and Oncology, Hamburg, Germany

[7] University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany

Correspondence:

Klaus Pantel, email: pantel@uke.de

Matthias Wilmanns, email: matthias.wilmanns@embl-hamburg.de

Abstract

The transcriptional corepressor C-terminal binding proteins (CtBPs) form redox-dependent homo-tetrameric assemblies that regulate the expression of various epithelial, pro-apoptotic and pro-oncogenic genes through their involvement in the multi-component Polycomb repressive complex 2 (PRC2) [1]. As CtBPs bind to short linear PxDLS-like sequence motifs [2], the presence of a tandem ALDLS motif on the putative metastasis suppressor Retinoic acid-induced 2 (RAI2) protein suggested a possible dual, non-canonical interaction with CtBPs [3]. Here, we show that RAI2 induces CtBP polymerization through well-ordered filaments of stacked tetrameric CtBP layers. Consistently, we observed RAI2-mediated CtBP nuclear foci in RAI2-expressing cancer cells. RAI2-mediated CtBP polymerization synchronized with the relief of CtBP-mediated repression of the tumor suppressor gene *CDKN1A*. Conversely, we found diminishment of *RAI2* gene expression during the progression of prostate cancer to treatment-resistant subtypes, concomitant with a marked increase of neuroendocrine markers in a *RAI2* depleted prostate cancer cell line. Taken together, our data demonstrate a previously unknown mechanism of multivalent short linear sequence motif-induced polymerization, tipping the balance of opposing oncogenic and tumor suppressive activities. Our findings serve as a model for a broad spectrum of related multivalent protein/protein interactions with diverse functional outcomes and provide a blueprint for the design of tandem motif-based inhibitor compounds targeting oncogenic CtBPs.

Introduction

Imbalanced transcription of a large number of genes can be caused at various levels, ranging from changes in individual transcription factors to malfunctional corepressor and coactivator complexes [4]. Interferences of these regulatory networks are crucial in cancer etiology, especially when genes with established oncogenic and oncogene-suppressing readouts are involved [5-7]. C-terminal binding proteins (CtBPs) are ubiquitous master transcriptional coregulators associated with the Polycomb repressive complex 2 (PRC2) [8,9], with essential functions in the development and oncogenesis of various tumor entities, including breast cancer (BC) and prostate cancer (PC) [10,11]. CtBPs can affect cancer progression and promote cell survival, proliferation, motility and the epithelial-mesenchymal transition by repressing many tumor suppressor genes such as *CDKN1A*, *CDH1* or *BRCA1* [12-15].

In humans, two closely related proteins CtBP1 and CtBP2 are encoded by two paralogous genes *CTBP1* and *CTBP2*, respectively [16]. Both CtBPs have a conserved three-domain arrangement in common that includes a split binding domain interacting with PxDLS sequence motif-containing proteins, a central dehydrogenase domain and a C-terminal domain, which is less conserved in the two CtBP paralogues. Binding of NAD^+/NADH to each dehydrogenase domain is required for CtBP dimerization, which allows further NAD^+/NADH -independent assembly into active CtBP tetramers as part of large corepressor complexes for recruitment to multiple chromatin-binding sites [12,17-19].

Although a single PxDLS-like short linear motif (SLiM) on a target protein sequence is sufficient for CtBP binding [20-22], it was intriguing to detect the presence of two such repeated non-canonical motifs with identical sequences (ALDLS) on the Retinoic Acid-induced 2 (RAI2) protein [3]. RAI2 has been characterized as a potential metastasis suppressor protein that inhibits the dissemination of tumor cells to the bone marrow and when absent, induces an

aggressive tumor phenotype in BC cells [3,23]. Hence, the antagonistic roles of RAI2 and CtBP as tumor suppressor and oncogene, respectively, could be explained through their ability to direct binding each other. Indeed, owing to the tetrameric assembly of CtBP, there are multiple PxDLS-binding sites accessible for interactions with RAI2. We therefore hypothesized that multivalent protein/protein interactions may lead to higher oligomeric assemblies beyond the established active tetrameric state of CtBP and potentially interfere with its function as a master gene coregulator.

Here, we demonstrate that the presence of a two-fold repeated RAI2 ALDLS motif sequence segment leads to well-defined CtBP polymerization at the molecular level and nuclear foci at the cellular level. We further show that the presence of RAI2 leads to epigenetically regulated relief of CtBP-mediated repression of the key tumor suppressor gene *CDKN1A* through a reduction of histone 3 tri-methylation (H3K27me3). In agreement with these findings, we reveal that the appearance of neuroendocrine traits in a relevant prostate tumor cell line model and PC progression to treatment-resistant stages in metastatic PC patient samples are both associated with the loss of RAI2, thus abolishing the molecular basis of CtBP polymerization and subsequent functional inactivation. Ultimately, our data may serve as a paradigm for the inactivation of other protein factors by polymerization and could provide molecular insights into the development of therapy resistance, thus opening novel directions in cancer therapy.

Results

CtBP binding to the tandem motif-containing RAI2 is autonomous. As RAI2 remarkably contains two CtBP-binding motifs with identical PxDLS-like sequences (ALDLS) in close proximity, separated by a 21-residue linker, we first investigated whether there is preferential binding by either of the ALDLS motifs through a specific protein environment

and whether there is any interference between the two motifs. As the PxDLS-binding domain is conserved between CtBP1 and CtBP2 sequences, we treated them analogously in this study. First, we analyzed the molecular properties of the two binding partners RAI2 and CtBP1/2 individually. Whereas the structural basis of the tetrameric CtBP arrangement has already been established [24], there are no experimental structural data on RAI2 owing to the lack of any defined fold predictions **(Extended Data Fig. 1)**. Although we were unable to obtain pure, full-length (FL) RAI2 protein due to extensive degradation, we established purification protocols for three truncated versions of RAI2 constructs with both ALDLS motifs intact, referred to as RAI2(303-530), RAI2(303-465) and RAI2(303-362) **(Fig. 1A, Extended Data Fig. 2A).** Circular dichroism (CD) spectroscopy and small angle X-ray scattering (SAXS) data confirmed that RAI2 is mostly intrinsically disordered **(Fig. 1B-C, Extended Data Fig. 2B, 3 and 4).** In addition, nuclear magnetic resonance (NMR) spectroscopy chemical shift analysis of RAI2(303-362) showed typical random coil values over the entire sequence of this protein fragment (**Fig. 1E-F, Extended Data Fig. 5, Extended Data Table 1).**

For quantitative characterization of the CtBP/RAI2 interaction, we first investigated the binding effects of the ALDLS motifs embedded in the RAI2 protein sequence. Both RAI2(303-362) variants with one of the two ALDLS sites impaired (M1 and M2) bind with an approximately 5 µM dissociation constant **(Fig. 1D, Extended Data Table 2)** and displayed no gain in binding affinity when compared to previous data on synthetic peptides with one ALDLS motif each [3]. RAI2 with both CtBP-binding sites intact showed only a modest five-fold increase in binding affinity, rendering any cooperative effect of the two binding sites acting in tandem to be unlikely **(Fig. 1D, Extended Data Table 2)**. We also detected an approximately two-fold change in binding stoichiometry, demonstrating that RAI2 ALDLS sites are accessible to CtBP1 in parallel. The binding affinity did not increase further when

using longer RAI2 constructs (WT, 303-465, 303-530), which supports a protein environment-independent nature of the RAI2/CtBP1 interaction as separate linear sequence motifs **(Extended Data Table 2).** Based on these findings, we used either RAI2(WT, 303-362) or RAI2(WT, 303-465) in all subsequent biophysical and structural experiments. CD and SAXS data of the respective CtBP1/RAI2 complexes reveal partly folded protein content, which is expected for complexes of folded and unfolded protein partners (**Fig. 1B-C, Extended Data Fig. 3, 4 and 6A).**

To gain insight into the molecular basis of this interaction, we titrated purified CtBP1 [25] into RAI2(WT, 303-362) as well as the two RAI2 variants M1 and M2 containing a single ALDLS motif and characterized the interactions on a single residue basis by NMR spectroscopy (**Fig. 1F-H)**. The marked decrease in chemical shift intensities of the RAI2 M1 and M2 variants upon CtBP1 complex formation of the remaining intact ALDLS sequence demonstrated related CtBP-binding properties of the two RAI2 ALDLS motifs (**Fig. 1H middle and bottom panel, Extended Data Table 3)**. In contrast to the single ALDLS-containing RAI2 variants, we found a bilobal intensity decrease distribution around the ALDLS motifs in RAI2(WT, 303-362), confirming that CtBP may bind simultaneously to both RAI2 ALDLS sites **(Fig. 1H top panel, Extended Data Table 3).**

Tandem ALDLS motif-containing RAI2 induces CtBP oligomerization. Given that both protein binding partners have multivalent binding sites, four in CtBP and two in RAI2, we wondered about potential assembly-mediated oligomerization [3,26]. As evident from analytical size exclusion chromatography (SEC) profiles, titration of RAI2(WT, 303-465) to CtBP1 leads to the formation of a high molecular weight complex indicating higher order oligomerization **(Fig. 2A, Extended Data Fig. 6B, Extended Data Table 4)**. Addition of the M1 and M2 variants to CtBP1 led to an increase of the retention volume, which is suggestive of

non-oligomerized, lower molecular weight CtBP1/RAI2 complexes **(Fig. 2A, Extended Data Table 4)**. As expected, there was no change in the CtBP1 elution profile when titrated with the RAI2 M1+M2 variant **(Extended Data Fig. 6C).**

To further investigate the nature of this high molecular weight CtBP1/RAI2 complex with both ALDLS binding sites intact, we used SAXS to assess its maximum particle dimension (D_{max}) in solution. When RAI2(303-465) M1 or M2 variants were added to CtBP1, we observed a moderate increase of the CtBP1 particle size from 14 nm to about 18 nm **(Fig. 2B, Extended Data Table 5)**. However, when RAI2(WT, 303-465) was added to CtBP1, the average particle size doubled from 14 nm to 28 nm, consistent with an elongated assembly of both protein components. These observations also hold true for the titration of shorter RAI2(303-362) constructs to CtBP1 **(Extended Data Fig. 3, 4 and 6D-F, Extended data Table 5).**

To understand the implications of this interaction in a cellular environment, we analyzed CtBP1/RAI2 assembly in KPL-1 BC cells [3] by confocal laser-scanning microscopy**.** In the presence of RAI2(WT), we detected distinct CtBP1/RAI2 nuclear foci. The number and volume of these foci significantly dropped when RAI2 variants (M1, M2, M1+M2) were used in which one or both ALDLS CtBP-binding sites were mutated **(Fig. 2C-D).** Our findings for CtBP2/RAI2 assembly are consistent with those observed for CtBP1 **(Extended Data Fig. 7).**

Next, we used high-resolution structural biology approaches to unravel the molecular basis of RAI2-mediated CtBP polymerization. To define a reference module for RAI2-mediated oligomerization, we first determined the high-resolution crystal structure of CtBP2 in the presence of a single CtBP binding site RAI2(M2, 303-465) variant **(Fig. 2E, Extended Data Table 6).** The structure comprises a CtBP/RAI2 complex with 4:4 stoichiometry that forms a flat X-shaped arrangement with overall dimensions of 9.5 x 9.5 x 5 nm **(Fig. 2E, left and**

middle panel). Apart from RAI2 residues 315-321 directly interacting with CtBP2 (**Fig. 2E, right panel, Extended Fig. 10B, right panel**), there was no interpretable electron density for the remaining RAI2 sequence, confirming that it remains unfolded upon assembly with CtBP, as expected from our CD and SAXS data **(Fig. 1B-C).**

We then investigated a molecular mechanism for RAI2-induced CtBP polymerization by cryogenic-electron microscopy (EM), demonstrating that the interaction of CtBP1 with RAI2 gives rise to extended polymeric structures with repeating structural entities at 5 nm intervals **(Fig. 2F-G).** These distances coincide with the layer thickness of individual CtBP1 tetramers **(Fig. 2E).** However, any further high-resolution structural analysis was not feasible due to the inherent flexibility and structural disorder of these fibers.

Molecular Basis of RAI2-induced CtBP oligomerization. To obtain structural insight into RAI2-mediated CtBP1 polymerization at high resolution, we shortened the RAI2 linker connecting the two ALDLS CtBP-binding sites by 11 residues (331-341) **(Fig. 1E)**. The resulting RAI2 variant (Δ11, 303-362) retained the ability to bind CtBP1 with comparable affinity, stoichiometry and particle size and the ability to form nuclei foci **(Extended Data Tables 2 and 4**, **Extended Data Fig. 8A-D)**. As evident from cryogenic-EM 2D classes, we found a well-defined filament that lacked the high flexibility of the CtBP1(WT)/RAI2(WT, 303-362) complex **(Fig. 2F-G),** thus allowing determination of its structure to 3.0 Å resolution (**Fig. 3, Extended Data Fig. 9 and 10A, Extended Data Table 7).** The structure revealed a RAI2-mediated CtBP1 filament with stacked tetrameric CtBP1 layers, confirming the identity of the layers observed at 5 nm intervals for both RAI2(WT) and RAI2(Δ11)-mediated CtBP1 polymers (**Fig. 2F-G)**. For structural refinement, we restricted our analysis to a fiber comprising six CtBP1 layers, with overall dimensions of 6 x 5 nm = 30 nm in length and 9.5 nm in thickness **(Fig. 3A)**, which is consistent with the overall dimensions of tetrameric

CtBP2 determined by X-ray crystallography **(Fig. 2E)**. In this structure, the CtBP1/RAI2 fiber is assembled by 322 symmetry, in which the 3-fold axis defines the longitudinal fiber axis and coincides with one of the two-fold axes of the CtBP1 tetrameric assembly **(Fig. 3A)**. This arrangement is rotated by 120 degrees and translated by 5 nm for neighboring CtBP1 tetrameric layers. Hence, the minimal unique entity of the overall RAI2-mediated filament consists of three tetrameric CtBP1 layers, which is two-fold repeated in the refined overall filament structure with a total of 6 x 4 = 24 CtBP1 molecules. Despite the substantial size of the surface interfaces of neighboring CtBP tetrameric layers in the order of 1600 $Å^2$, they are almost void of any specific interactions **(Fig. 3B-C)**, which explains why their formation requires the presence of RAI2 as a polymerization mediator.

Within this arrangement, we observed two RAI2 peptides per CtBP1/CtBP1 layer interface that directly connect the RAI2-binding sites of neighboring CtBP1 tetrameric layers via their two ALDLS motifs 1 and 2 (**Fig. 3B, 3D, Extended Data Fig. 10B, left panel**). This arrangement with four CtBP1 molecules per CtBP1 layer and two RAI2 peptides per CtBP1 layer interface is in agreement with the N=0.5 stoichiometry observed in our isothermal titration calorimetry (ITC) data **(Fig. 1D, right panel)**. Both RAI2 ALDLS CtBP-binding sites are characterized by a parallel β-sheet interaction involving β-strand 2 of each CtBP1 ALDLS-binding domain and the central LDL sequence of each of the two RAI2 ALDLS motifs 1 and 2 **(Fig. 3D, Extended Data Fig. 10B, left panel, and 11)**. This finding agrees with our X-ray crystallography data of tetrameric CtBP2(31-364)/RAI2(M2, 303-465) variant assembly, in which the CtBP/RAI2 interaction is limited to only the first RAI2 ALDLS motif **(Fig. 2E right panel, Extended Data Figure 10B, right panel)**. We did not detect any further specific interactions of any LDL-motif flanking residues, explaining the relatively moderate binding affinity of the RAI2 interaction with CtBP1 **(Fig. 1D)**.

To probe whether our structural findings on stacked tetrameric CtBP1 layers also apply to RAI2-mediated CtBP polymerization, we mutated one of the few CtBP1 residues (K46) involved in the limited CtBP1/CtBP1 layer interactions **(Fig. 3C)** to tryptophan. The binding affinity of CtBP1(K46W) with RAI2 was substantially lowered, regardless of whether the WT or Δ11 RAI2 variants were used. In addition, we observed a drop of RAI2/CtBP1 stoichiometry upon binding to values below 0.3 **(Extended Data Table 2),** which could indicate reduced polymerization and fragmented CtBP/RAI2 fiber formation. In agreement with these observations, SEC and negative stain EM data of CtBP1(K46W) with these RAI2 variants showed increased retention volume and a significant reduction of ordered filament formation, respectively (**Extended Data Fig. 12, Extended Data Table 4).** Taken together, our data demonstrate that ordered CtBP1/RAI2 filament formation depends on proper interface formation of neighboring CtBP1 tetramer layers, which are not sufficient to generate CtBP1 filaments in the absence of RAI2.

RAI2-induced polymerization relieves CtBP repressor activity. To study the implications of RAI2-induced CtBP-polymerization in a relevant tumor cell model, we chose the VCaP PC cell line because of its high endogenous RAI2 expression and hence the presence of distinct CtBP/RAI2 foci **(Fig. 4A-B, Extended Data Fig. 13, 14A-B)**. By contrast, there were no related foci in VCaP cells without RAI2 (VCaP KO) **(Fig. 4B, Extended Data Fig. 13, 14B)**. To probe the implications of RAI2-mediated CtBP polymerization on transcriptional regulation of the tumor suppressor gene *CDKN1A* [15,27], we first applied an *in vitro* transactivation assay in Human Embryonic Kidney (HEK) 293T cells. We found that concomitant expression of RAI2 and CtBP1 caused a 2.5-fold increase in *CDKN1A* promotor activity**,** implying relief of CtBP corepressor activity in the presence of RAI2 **(Extended Data Fig. 15)**. In line with this, levels of the *CDKN1A*-encoded protein p21 (CDKN1A) were low in VCaP KO cells but high in

parental (PAR) VCaP cells with concomitant RAI2 expression **(Fig. 4D)**. Under genotoxic stress conditions using hydrogen peroxide or Camptothecin treatment [28], both *CDKNA1* gene expression and CDKN1A protein levels were significantly increased in VCaP PAR cells **(Fig. 4C-D)**. This was further accompanied by increased number and volume of CtBP/RAI2 foci (**Extended Data Fig. 13, 14C**). In VCaP KO cells, we did not observe any of these effects, demonstrating that RAI2 relieves the repression of direct targets of CtBPs, including *CDKNA1* (**Fig. 4C-D)**.

RAI2-induced CtBP polymerization associates with the inhibition of EZH2 activity. As CtBP repressor function is connected to PRC2 activity [9], we next assessed the effects of CtBP in the absence and presence of RAI2 on key PRC2 markers, including EZH2 and EZH2-induced posttranslational modification H3K27me3, in VCaP cells by chromatin-immune precipitation (ChIP) (**Fig. 4E**). As expected, we found impaired binding of CtBP to the *CDKN1A* promoter in PAR cells compared with RAI2 KO cells. Whereas the level of EZH2 binding to chromatin remained unaltered irrespective of the presence of RAI2, we found a strong increase of H3K27me3 at the target promotor in VCaP KO cells, indicating catalytic activity of EZH2 in line with CtBP-mediated repression of the *CDKN1A* gene **(Fig. 4D**)**.** Since the level of histone H3K27me3 modification is low in VCaP PAR cells, our data suggest that RAI2-mediated CtBP polymerization downregulates EZH2 activity. These data are supported by colocalization of both EZH2 and H3K27me3 with CtBP/RAI2 foci in VCaP PAR cells **(Fig. 4F-G).** In contrast, in the absence of CtBP1/RAI2 foci we found only a diffuse distribution of these components within VCaP KO cell nuclei (**Extended Data Fig. 16A-B**). CtBP2 showed the same pattern of colocalization with EZH2 and H3K27me3 as CtBP1 **(Extended Data Fig. 16C-F).** Interestingly, these findings are in agreement with a highly significant anti-correlation of gene expression levels of *RAI2* and *EZH2* in a primary prostate adenocarcinoma cohort deposited in the

Cancer Genome Atlas (TCGA) PanCancer Atlas, indicating that both genes might have opposing effects in primary prostate tumors [29] **(Extended Data Fig. 17)**.

Relevance of *RAI2* gene expression in prostate cancer progression. To demonstrate the potential clinical relevance of RAI2-induced CtBP polymerization, we probed its role in stepwise PC progression [11,30]. Under androgen deprivation therapy (ADT), PC frequently progresses from the hormone-sensitive (HSPC) to castration-resistant (CRPC) stage and can further transdifferentiate to treatment-refractory variants including tumor lesions with neuroendocrine traits [30,31]. Probing the antagonistic roles of CtBP and RAI2 in PC progression is relevant, as both proteins have independently been shown to be involved in PC-associated androgen signaling processes [3,32,33]. To this end, we performed gene expression analyses of circulating tumor cells (CTCs) enriched from blood samples of metastatic PC patients from metastatic HSPC, CRPC, aggressive variant PC (AVPC) and histological evident neuroendocrine prostate cancer (NEPC) [34,35]. From more than 100 blood samples analyzed **(Extended Data Table 8)**, the percentage of patients with detectable CTCs increased from 67-70% in HSPC/CRPC patients to 89-91% in treatment-resistant AVPC/NEPC disease stages **(Fig. 5A).** Remarkably, whereas *RAI2* gene expression was detected in 89% of CTC-positive blood samples in CRPC patients, this rate dropped to 35% in AVPC and 19% in NEPC patients (**Fig. 5B**), suggesting that loss of RAI2 could be associated with PC progression to disease stages that are refractory to treatment. Because EZH2 is a known driver of the development of CRPC and trans-differentiation to NEPC [36,37], we tested whether RAI2 loss alone could induce neuroendocrine traits. Indeed, we detected a strong increase in the expression levels of the neuroendocrine markers Synaptophysin (*SYP*) and Chromogranin (*CHGA*) in VCaP KO cells when compared to VCaP PAR cells (**Fig. 5C-D**). In VCaP KO cells, we also observed a higher nucleus/cytoplasm ratio and smaller cell size compared with the VCaP PAR cells,

which is also indicative of NEPC [35] (**Fig. 5E**). In summary, our data demonstrate that loss of RAI2 in prostate tumor cells matches our observations in the development of treatment-resistant neuroendocrine traits associated with severe reduction of RAI2 from CTCs of patients at different PC stages. This is in agreement with our molecular data, demonstrating the requirement of RAI2-induced CtBP polymerization for relief of repression of crucial tumor suppressors such as *CDKN1A*, ultimately revealing how the balance of CtBP and RAI2 has critical implications in cancer etiology.

Discussion

Unraveling the principles and mechanisms of how biological activities are regulated remains one of most fascinating topics in the life sciences. Those involving mostly folded proteins such as allosteric processes, changes of assembly states and post-translational modifications are well established, but insight into molecular mechanisms involving unfolded proteins is lacking, in part because they are less amenable to structural characterization. This is particularly true for processes involving SLiMs from unfolded proteins [38,39].

Here, we reveal that the RAI2 protein that lacks any folded domain structure has a unique ability to induce CtBP polymerization via multivalent interaction sites from both binding partners: two in RAI2 through a tandem arrangement of two identical ALDLS sequence motifs and four in CtBP because of its tetrameric assembly [17-19]. The RAI2 ALDLS motifs belong to a larger group of SLiMs with a PxDLS consensus sequence, which occurs in one or more copies in a variety of CtBP-binding proteins [22,40]. Our data thus add a new dimension to CtBP activity regulation through a previously unknown polymerization ability mediated by multiple SLiMs. This inactivation mechanism complements its proven loss of activity in the absence of NAD^+/NADH, which is required for CtBP tetrameric assembly and the potential induction of its metabolic- and redox-sensing properties **(Fig. 6)** [41-44]. As CtBP binding of

PxDLS protein targets does not require oligomerization [22,45], CtBP tetramerization has been proposed as a prerequisite for simultaneous recruitment of multiple chromatin-modifying enzymes and/or proper nuclear localization [19,26].

There is only a limited understanding of protein polymerization mediated by multivalent interactions, especially involving intrinsically disordered SLiM motif-containing protein partners that match the overall structural features of RAI2 [46,47]. Nevertheless, the available data are suggestive of models proposing molecular condensates, which could generate membrane-less compartments with enhanced local concentrations of target proteins [48-50]. Interestingly, both PRC1 and PRC2 machineries are known to generate phase separated condensates known as "polycomb bodies" for efficient chromatin compaction, induced by interactions with disordered protein components [51]. Our observation that EZH2 colocalizes with CtBP/RAI2 foci **(Fig. 4)** suggest that RAI2-mediated CtBP polymerization may have implications for PRC2 compartmentalization.

Bipartite tandem CtBP-binding PxDLS motifs are also part of other proteins, such as the ligand-dependent corepressor (LCoR) protein or the receptor-interacting protein 140 (RIP140) [52,53]. Thus, these proteins may also be capable of inducing CtBP polymerization. In this model, binding of either RAI2 or LCoR to CtBP might be mutually exclusive. Interestingly, in LNCaP PC cells, PRC2-associated LCoR isoform 1 (PALI1) physically interacts with the PRC2 complex and its knockdown causes a drastic reduction of genome-wide EZH2 activity [54], which is in direct opposition to our observation that RAI2 loss causes increased EZH2 activity at a target promotor.

Due to its multifunctionality, well-defined domain organization and the requirement for oligomeric assembly in order to act as a major corepressor, CtBP provides multiple opportunities for therapeutic intervention [55,56]**.** Our data on RAI2-mediated CtBP

polymerization as a powerful mechanism of CtBP repressor activity regulation may inspire new routes for future drug development, by chemically exploiting the recurrence of two tandem PxDLS motifs as an effective way to mimic RAI2-mediated CtBP activity through polymerization. Such polymerization could potentially interfere with crucial transition processes leading to increased cancer malignancy, as demonstrated for the CRPC/NEPC transition based on our findings in this contribution and EMT-like processes in other cancer types. To what extent target deactivation by polymerization through multivalent interactions as a molecular principle is applicable to other molecular processes under healthy and disease conditions provides a fascinating future perspective.

Material and Methods

Secondary and tertiary structure prediction. IUPRED2A was employed for sequence disorder prediction [57,58]. AlphaFold/ColabFold was used for tertiary structure prediction [59-61].

Plasmid Construction. Plasmids pcDNA3.1-hCtBP1 isoform 1 (Uniprot ID: Q13363) and pCMV-SPORT6-CtBP2 (Uniprot ID: P56545) were purchased from Genescript USA Inc (Piscataway, NJ, USA) and Addgene (Watertown, MA, USA), respectively. CtBP1 and CtBP2 were cloned using restriction sites NcoI and XhoI into pETM14 (EMBL, https://grp-pepcore.embl-community.io/vectors/ecoli.html, Germany) in frame with an N-terminal polyhistidine-tag and a 3C protease site sequence LEVLFGP. CtBP1 and CtBP2 mutants were generated using gene blocks (Integrated DNA Technologies, IA, USA) and then subcloned into the original plasmid using the CloneJET PCR cloning kit (Thermo Fischer Scientific, MA, USA). RAI2 protein expression plasmids in mammalian cells were established as described [3]. The cDNA of truncated versions of RAI2 was cloned into the expression vector pETGB1a (EMBL), using the restriction sites NcoI and XhoI in frame with an N-terminal hexahistidine-G-binding protein domain fusion tag and a Tobacco Etch Virus (TEV) protease site with

sequence ENLYFQG. RAI2 variants M1 (L319A, S320A), M2 (L345A, S346A), and M1+2 (L319A, S320A, L345A, S346A) were generated using gene blocks (Integrated DNA Technologies, IA, USA). All vectors were confirmed by DNA sequencing.

Protein expression and purification. CtBP1 and CtBP2 variants were expressed in *Escherichia (E.) coli* strain BL21(DE3) pLysS. Cultures were grown in the Terrific Broth medium at 37 °C to an optical density of 1.2, induced with 0.2 mM isopropyl thiogalactose (IPTG) for 18 hours, pelleted by centrifugation (30 minutes, 5000 x g) and resuspended in 50 ml lysis buffer containing 50 mM HEPES/NaOH (pH 7.5), 300 mM NaCl, 10 mM imidazole, 10% (v/v) glycerol. Cells were lysed three times by using an emulsifier under a constant pressure of 10,000 psi. Cell debris was removed by centrifugation (45 min, 21,000 x g), passed through a 0.45 µM filter, incubated with nickel-charged affinity (Ni-NTA) beads in 10 ml slurry and equilibrated with lysis buffer for 60 minutes and passed through it. The column was then washed with 100 ml lysis buffer, followed by 100 ml wash with lysis buffer containing 30 mM imidazole and 1M NaCl and finally eluted with a 30 ml volume of 200 mM imidazole. The eluted fusion proteins were cleaved using 3C protease in overnight dialysis buffer, which was identical to the lysis buffer. The cleaved CtBP proteins were again passed through Ni-NTA beads equilibrated with lysis buffer, to remove 3C protease and further impurities. The proteins were dialyzed in 4 liters of 50 mM Tris/HCl (pH 8.0), 5% [v/v] glycerol for 6 hours and then passed through a pre-equilibrated HiTrap Q HP 5ml (Cytiva, MA, US) column and eluted with 125 ml of a 0-1,0 M NaCl gradient to obtain pure proteins as confirmed by Sodium Dodecyl Sulfate (SDS)-Polyacrylamide Gel Electrophoresis (PAGE) analysis. Highest purity protein fractions were pooled, dialyzed against SEC buffer (50 mM HEPES (pH 7.2), 150 mM NaCl), concentrated, and loaded onto a Superdex 200 column (Cytiva, MA, USA).

Due to the presence of disordered N- and C-termini, CtBP1(28-353) and CtBP2(31-364) were utilized for NMR spectroscopy and X-ray crystallography experiments, respectively.

RAI2(303-362), RAI2(303-465) and RAI2(303-530) variants were transformed in BL21 DE3 LOBSTR *E. coli* cells and plated onto kanamycin plates. A single colony from the plates was inoculated overnight into 100 ml of Luria Bertani (LB) medium containing 50 µg/ml kanamycin. This primary culture was inoculated into 3 liters of Terrific Broth Medium until an OD_{600nm} of 1.2, protein expression induced with 0.5 mM IPTG for 18 hours and pelleted by centrifugation (30 min, 5000 x g). For NMR spectroscopy, primary culture was inoculated into 800 ml LB medium until an OD_{600nm} of 0.7, spun down at 6000 × *g* for 20 minutes and then inoculated into 2 liters of M9 medium containing labeled with 3 g/l $^{15}NH_4Cl$ (1.5 g/l) and $^{13}C_6$-glucose. Protein expression was induced at an OD_{600nm} of 0.7 by adding 0.5 mM IPTG at 18 °C for 18 hours and then pelleted by centrifugation (30 minutes, 5000 x g).

The pellets were resuspended in 50 ml lysis buffer containing 50 mM HEPES/NaOH (pH 7.5), 300 mM NaCl. Cells were lysed three times by using an emulsifier under a constant pressure of 10,000 psi and cell debris were removed by centrifugation (40 minutes, 21,000 x g), passed through a 0.45 µM filter and loaded onto a pre-equilibrated HisTrap HP 5ml column (GE Healthcare, IL, US). The column was washed with 50 ml lysis buffer and eluted with buffer containing 200 mM imidazole. The eluted fusion proteins were cleaved using TEV protease in the overnight dialysis buffer. The cleaved proteins were again passed through the same HisTrapHP 5ml column equilibrated with lysis buffer to remove TEV protease and fusion tags. The flowthrough was then concentrated using a 3.5 kDa concentrator (Merck KGaA, Germany) and loaded on a Superdex 75 column SEC column (Cytiva, MA, USA), pre-equilibrated with 50 mM HEPES (pH 7.2), 150 mM NaCl buffer. Each purification step was analyzed on SDS-PAGE before proceeding to the next. Purified proteins were concentrated,

aliquoted and stored at -80 °C or utilized for further biophysical and high-resolution structural biology experiments.

Quantitative determination of binding affinities. Isothermal titration calorimetry (ITC) experiments were carried out with a VP-ITC system (Malvern Pananalytical, UK). Experiments were performed at 25 °C in 50 mM HEPES (pH 7.2), 150 mM NaCl. Proteins were dialyzed against this buffer overnight. Purified CtBP1 and CtBP2 were placed in the reaction cell while the RAI2 constructs were loaded in the injection syringe. The CtBP concentration used was 10 µM for RAI2(M1, 303-362) or RAI2(M2, 303-362) in the syringe or 20 µM for the wild type RAI2 constructs. The concentration of the RAI2(303-362) variants in the syringe was 100 µM. Injections of 10 µl of RAI2(303-362) variants were performed at 4-minute intervals. ITC data were processed using NITPIC, SEDPHAT and GUSSI [55,62]. The values reported are the means of three independent measurements, and the ± errors represent standard deviations. Significance between CtBP1(WT)/RAI2(WT, 303-362) and other CtBP/RAI2 complexes was evaluated by two-sided Student's *t*-test and the respective p values are shown.

Analytical Size Exclusion Chromatography. The complex formation between CtBP1 and CtBP2 variants and RAI2 variants was investigated on an analytical Superose 6 3.2/300 GL column (Cytiva, MA, USA) coupled to a 1260 Infinity HPLC system (Agilent, CA, USA). Proteins were dialyzed against the SEC buffer and injected onto a pre-equilibrated S6 column using an autosampler. The concentration of CtBP1 and CtBP2 variants was kept constant at 100 µM and RAI2 variants were titrated at stoichiometric ratios of 1:0.1, 1:0.2, 1:0.5, 1:1 and 1:2. The gel filtration protein standard containing thyroglobulin (670 kDa), gammaglobulin (158 kDa), ovalbumin (44 kDa), myoglobin (17.5 kDa) and vitamin B12 (1.35 kDa) was used to calibrate the column (Bio-Rad, CA, USA). The curves were plotted using GraphPad Prism (Dotmatics, MA, USA).

Circular dichroism (CD) spectropolarimetry measurements. Prior to each measurement, protein samples were dialyzed against 10 mM K_2PO_4 (pH 7.2), 100mM NaF and diluted to a concentration of 0.25 mg/ml. Spectra were recorded at 20 °C on a Chirascan CD Spectrometer (Applied Photophysics, UK), between 180 and 260 nm in a 0.1 cm cuvette. The number of scans per measurement was 10. Machine settings were as follows: 1 nm bandwidth, 0.5 s response, and 1 nm data pitch. Spectra were background subtracted, converted to mean residue ellipticity [Θ] and plotted using GraphPad Prism (Dotmatics, MA, USA). The secondary structure content of proteins was calculated using the neural network CDNN [63].

Small Angle X-ray Scattering (SAXS). SAXS data were collected on beamline P12 at PETRA III (EMBL/DESY, Hamburg, Germany), using a PILATUS 6M pixel detector (DECTRIS, Switzerland). A total of 40 µl sample volumes were exposed to X-rays while flowing through a temperature-controlled 1.2 mm width quartz capillary at 20 °C. A total of 40 image frames of 0.2 s exposure time each were collected. Data were normalized to the transmitted X-ray beam intensity, averaged, buffer contribution subtracted and placed on absolute scale relative to water using SASFLOW [64,65]. Data correction was performed using PRIMUS*qt* and the ATSAS 3.0.1 software package [66]. The forward scattering intensity *I(0)* and radius of gyration R_g were determined from Guinier analysis assuming that at angles $s \leq 1.3/R_g$ the intensity is represented as $I(s)=I(0)\exp(-(sR_g)2/3))$. These parameters were also estimated from the full SAXS curves using the indirect Fourier transform method implemented in GNOM [67], along with the distance distribution function *p(r)* and maximum particle dimensions D_{max} estimates. Molecular masses of solutes were estimated from *I(0)* by computing the partial specific volume and contrast between the protein sequence and

buffer components. Theoretical scattering intensities from structural coordinates were computed with CRYSOL [68].

Low-resolution shapes were reconstructed from SAXS data using DAMMIF [69]. Rigid body models for different CtBP1/RAI2(303-465) and CtBP1/RAI2(303-362) complexes were computed from the experimental data using CORAL [65] followed by the generation of fits against experimental scattering data in solution using CRYSOL. The available high resolution crystal structure of CtBP1 (PDB ID: 6CDF) and CtBP2(WT, 31-364)/RAI2(315-321) were used as input rigid bodies. For the full-length CtBP1 tetramer, a monomer was extracted from 6CDF and P222 symmetry was applied during structure calculation. Additional residues corresponding to the 24 and 83 N-terminal and C-terminal intrinsically disordered residues absent in the crystal structure, respectively, were modelled as dummy residues. This tetramer along with the interface information obtained from our crystal structure was then further used to calculate the stoichiometric 4:4 complexes for RAI2 M1, M2 and M1+M2 mutants. For the large polymeric WT complex formed between CtBP1 and RAI2, tetrameric CtBP1 assemblies were used as rigid bodies and tetramers linked by RAI2 monomers at CtBP1 binding sites identified by consensus between the crystallographic data and NMR chemical shift analysis. For the polymeric Δ11 complex, a theoretical scattering plot was generated for the cryo-EM structure and the fits were generated against the WT and Δ11 experimental scattering data. The goodness of the fits is reflected in the respective χ^2 values. Kratky plots were obtained using ATSAS.

X-ray structure determination. Human CtBP2(31-364)/RAI2(M2, 303-465) complex at a concentration of 10 mg/ml was crystallized from 20% [w/v] PEG-6000, 100 mM Tris (pH 7.5) at 4 °C. Prior to X-ray data collection, crystals were cryoprotected with 25 % [v/v] glycerol, mounted on cryo-loops (Hampton Research, CA, USA) and were flash frozen in liquid

nitrogen. Diffraction X-ray data were collected at beamline P13 at PETRA III (EMBL/DESY, Hamburg, Germany). The data were integrated using XDS [70], scaled and merged using AIMLESS [71] within CCP4 suite [72]. The structure was determined by molecular replacement using Phaser [73], using another CtBP2 structure (PDB ID: 2OME) as a search model. The asymmetric unit of these crystals contained four CtBP2 molecules. The fourth molecule was disordered due to poor crystal packing, as evident by high B-factors of the final structural model. After partial coordinate refinement, the *2Fo-Fc* map resolved distinct electron density for three RAI2 peptides with side-chain features that enabled modeling of seven residues 315-321 (sequence: EALDLSM) of each RAI2 polypeptide chain. Iterative model building and refinement was performed in COOT [74,75], Phenix [76] and Refmac5 [77], using non-crystallographic symmetry (NCS) constraints and the Translation-Libration-Screw-rotation parameters. The final model was refined to R/Rfree of 0.23/0.29 at 2.64 Å resolution. For further structural analysis, a biological tetramer was generated using polypeptide chains A and B and the equivalent chains of a crystallographic symmetry mate.

NMR spectroscopy. NMR data were acquired at 20 °C in 20 mM Na_2PO_4 and 50 mM NaCl (pH 6.8), 10% D_2O and 1.5 mM Tris(2-carboxyethyl)phosphin-hydrochloride on a 600 MHz NMR spectrometer (Bruker, MA, USA) equipped with a cryo-probe and a 700 MHz Avance III spectrometer. 2D [^{1}H,^{15}N] heteronuclear single quantum coherence (HSQC) spectra were recorded using States/TPPI quadrature detection mode in the indirect dimension. Typically, 512 x 2048 complex points with 8 transients were used for recording the 12-14 and 24-28 ppm spectral width in the ^{1}H and ^{15}N dimensions, respectively. The transmitter frequency for the proton channel was placed on-resonance with the water signal at a ^{15}N frequency offset of 119-120 ppm. The ^{15}N–{^{1}H} Nuclear Overhauser Effect (NOE) was recorded in Echo-Antiecho detection mode in the indirect dimension in an interleaved manner with and

without proton saturation using 5 ms separated 120° pulse trains for a total of 5 s as recycle delay. Residue specific heteronuclear-NOEs were calculated as the ratio of ^{15}N–^{1}H cross-peak intensities (I_{sat}/I_{unsat}) using CCPNmr [78].

For NMR backbone and side-chain assignments, [^{1}H,^{15}N]-HSQC, [^{1}H,^{13}C]-HSQC, HNCO, HN(CA)CO, HNCACB, CBCA(CO)NH, CC(CO)NH, TOCSY-HSQC (80 ms mixing time) and NOESY-HSQC (120 and 200 ms mixing times) were used. For triple resonance experiments, the transmitter for the ^{15}N channel was set to 119 ppm with 26 ppm spectral width. The carbon carrier frequencies were set to 38 ppm for HNCACB or 43 ppm for CC(CO)NH in experiments sensitive to backbone and side-chain ^{13}C spins with 76 or 74 ppm width, while it was set to 173 or 173.5 ppm in experiments sensitive only to carbonyl ^{13}C spins using 13 ppm spectral width. Increments in the indirect dimensions were 72 x 128 for HNCACB and 168 for CC(CO)NH, respectively, with 32 x 14 ppm spectral width, and 119 x 176 for HNCO and HN(CA)CO for ^{15}N and ^{13}C. Transients were 8, 16 or 32 depending on sample concentration, measurement type and sensitivity. Traditional or non-uniform 25% sampling acquisition modes were used for 3D measurements. Non-uniform sampling processing was achieved by the multi-dimensional decomposition method to replenish missing data points [79,80]. Proton chemical shifts were referenced relative to the water signal, with indirect referencing in the ^{15}N dimension using a $^{15}N/^{1}H$ frequency ratio value of 0.101329118 according to Biological Magnetic Resonance Bank (BMRB) [81]. Additionally, the spectra of RAI2(WT, 303-362) and its two variants were assigned. Next, CtBP1(28-353) was titrated at concentration ratios of 1:0.4, 1:0.7 and 1:1 stoichiometry onto RAI2 variants at 100 μM concentration, followed by [^{1}H,^{15}N]-HSQC experiments to perform chemical shift perturbation measurements. Data were processed using TOPSPIN v.4.0.6 (Bruker, MA, US) and analyzed with CARA [82], CCPNmr Analysis and SPARKY [78,83].

Negative-stain electron microscopy (EM). Carbon-coated copper grids were glow-discharged for 30 s at 25 mA using a GloQube Plus Glow Discharge System (Electron Microscopy Sciences, PA, USA). Grids were applied with 4 µl protein, incubated for 30 s and then blotted off from the side. Next, the grid was washed with 4 µl of 2% $(NH_4)_2MoO_4$ staining solution, followed by 4 µl staining for 30 s. The stain was blotted off from the side and grids were air-dried. Grids were imaged using a Talos L120C Transmission Electron Microscope equipped with a 4K Ceta CEMOS camera using TIA 4.1.5 (Thermo Fisher Scientific, MA, US).

Cryogenic-EM. Quantifoil R2/2, Cu 200 mesh holey carbon grids (Jena Bioscience, Germany) were glow-discharged at 25 mA for 60 s using the glow-cube device. Purified CtBP1 (WT) and RAI2 (Δ11, 303-362) were mixed at a concentration of 50 µM at 1:1 stoichiometry in a buffer containing 50 mM HEPES (pH 7.2), 50 mM NaCl. Next, 5 µl sample was applied on the glow-discharged grid, incubated for 10 s, blotted for 2 s with a blotting force of -12 at 4 °C and 100 % humidity, followed by plunge freezing in 63:37 propane:ethane ratio, using Vitrobot Mark IV device (Thermo Fisher Scientific, MA, US) set to 100% humidity and 4 °C.

Cryo-EM data were acquired on Titan Krios electron microscope operating at 300 kV equipped with a K3 direct electron detector (Gatan, CA, USA) using a GIF-Quantum energy filter with 20 eV slit width (Thermo Fisher Scientific, MA, USA). Movies were recorded with EPU software in counting mode (Thermo Fisher Scientific, MA, USA). Each micrograph was fractionated over 50 frames, with a total dose of 50 $e^-/Å^2$ for 1.8 s. Image processing was performed using Cryosparc [84]. Movies were drift corrected using patch motion correction. Contrast transfer function was performed using patch estimation, and micrographs exceeding 5 Å resolution were selected for subsequent steps.

Manual picking was performed on a subset of micrographs and picked particles were subjected to 2D classification. Selected classes were used for template-based particle picking on the full data set. After three rounds of 2D classification, particles with different orientations were selected to generate an initial model using *ab initio* reconstruction in Cryosparc [84]. The initial model obtained was used as reference for heterogenous 3D refinement with three or more classes to determine the level of heterogeneity in the dataset. Selected classes were subjected to both homogenous and non-uniform 3D refinement by enabling defocused global CTF refinement. The refined particles were subjected to local motion correction, followed by homogenous 3D refinement with or without imposing D2 symmetry resulting in final density maps of 3.0 and 3.2 Å resolution, respectively.

The D2 symmetrized cryo-EM map was fitted with six copies of CtBP1 tetramer crystal structures (PDB code 6CDF) in Chimera [85]. RAI2 peptides were built manually in COOT [74,75] and were subjected to real space refinement in Phenix [76] using the secondary structure restraints and NCS constraints. Ramachandran and rotamer outliers were fixed in ISOLDE [86] and the model was further refined in Phenix [76]. Validation of the final model and data was carried out using CheckMySequence [87], MolProbity [76,88] and tools from Phenix suite. Interface and solvent accessible areas of the fiber complex was calculated using PDBePISA [89].

Cell Culture. KPL-1, VCaP and HEK 293T cell identity was assured by tandem repeat profiling and were regularly tested for mycoplasma contamination using the Venor®GeM detection kit (Minerva Biolabs, Germany). Cells were grown as monolayers according in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum and 2 mM L-glutamine at 37°C in a humidified atmosphere containing 10% CO_2.

Quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR) analysis. RNA was extracted from cultured cells during exponentially growth using Nucleospin RNA Kit (Macherey Nagel, Germany). A total of 1000 ng RNA of each sample was transcribed using the First Strand cDNA Synthesis Kit (Thermo Scientific, MA, USA) and random hexamers. Each qRT-PCR was run in technical triplicates using SYBR Green (Thermo Scientific, MA, USA) with the following primers: *CDKN1A*: CATGTGGACCTGTCACTGTCTTGTA, GAAGATCAGCCGGCGTTTG; *RPLP0*: TGAGGTCCTCCTTGGTGAACA, CCCAGCTCTGGAGAAACTGC. Mean crossing point-values were used for each gene. Data were analyzed by applying the Comparative CT Method (ΔΔCT)-method using gene expression of *B2M* for normalization. Significance between different cell lines and results of three independent experiments were evaluated by two-sided Student's *t*-test.

Western Blot Analysis. Whole cell extracts from cultured cells were prepared by direct lysis and sonication of cells in 2% SDS sample buffer containing phosphatase and protease inhibitors (Roche, Germany). Cell extracts were separated in denaturing 8% or 15% (solely used for CDKN1A) polyacrylamide gels and blotted onto a nitrocellulose membrane. Proteins were detected by incubating with RAI2 (D4W9P; Cell Signaling Technology, MA, US), CtBP1 (Clone 3/CtBP1; BD Bioscience, NJ, US), CtBP2 (Clone 16/CtBP2; BD Bioscience, NJ, USA) and CDKN1A (Clone 12D1; Cell Signaling Technology, MA, USA) antibodies. HSC70 protein (Clone B-6; Santa Cruz Biotechnology, CA, USA) was used as loading control and HRP-(Cell Signaling Technology, MA, USA) or IRDye® (Li-Cor Biosciences, Germany) coupled secondary antibodies were used for protein detection with either the Curix 60 Processor (Agfa HealthCare, Belgium) on Super RX films (Fujifilm, Japan) or the Odyssey® CLx Imaging System (Li-Cor Biosciences, Germany).

Immunofluorescence Staining. Cells were fixed in 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS) (pH 7.4) for 10 minutes, washed 3 times with PBS and permeabilized with 0.2% Triton X-100 in PBS for 10 minutes. After incubation with 1% bovine serum albumin (BSA) in PBS for 30 minutes, cells were further incubated with primary antibodies RAI2 (Clone D4W9P, 1:250, Cell Signaling Technology, MA, USA), CtBP1 (Clone 3, 1:250; BD Biosciences, NJ, USA), CtBP2 (Clone 16, 1:250; BD Biosciences, NJ, USA), EZH2 (D2C9, 1:125, Cell Signaling Technology, MA, USA) or H3K27me3 (C36B11, 1:250, Cell Signaling Technology, MA, USA) in 1% BSA (w/v) in PBS for 90 minutes. After three washes with PBS, specific antibody binding was visualized with Alexa Fluor 488 goat anti-rabbit IgG (H+L) and Alexa Fluor 546 goat anti-mouse IgG (H+L) (Life Technologies, CA, USA) in 1% BSA in PBS. After three washes with PBS, nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) and mounted in Mowiol (Sigma-Aldrich, MO, USA) according to the manufacturer's instructions. Confocal laser scanning microscopy was performed on a Leica TCS SP5 microscope (Leica Microsystems, Germany). Imaris imaging software (Oxford Instruments, UK) were used to identify and measure the number and size of volumes of CtBP1 and CtBP2 foci. For this purpose, we evenly applied background subtraction and a lower volume cut-off of 0.1 μm^3 in statistical analysis by two-sided Student's *t*-test.

Colocalization of CtBP1/RAI2 and CtBP2/RAI2 foci with EZH2 or H3K27me3 in VCaP PAR cells was verified using fluorescence intensity profiles measured along lines manually drawn on the microscopy images. For this, we first macroscopically identified 100 individual foci of each condition in single image stacks. The lines were selected to intersect foci observed on channels corresponding to CtBP1 or CtBP2 ignoring any other signal. The procedure resulted in 100 intensity profiles with two channels for each of the four conditions. Profiles were aligned by applying a horizontal shift maximizing cross-correlation function in the CtBP1 or

CtBP2 channel with an arbitrarily selected reference profile. The shifts were applied to a second channel of each profile prior to the calculation of mean intensities and standard deviations, which were the result of about 90% colocalization efficiency for CtBP/ H3K27me3 and 50% colocalization efficiency for CtBP/EZH2.

Transactivation Assays. The reporter plasmid (CDKN1A/WAF1 promoter, Addgene plasmid #16451) was kindly provided by Dr. Bert Vogelstein (John Hopkins University, MD, USA) [90]. HEK 293T cells were transiently cotransfected with individual expression plasmids encoding full length RAI2 and CtBP1 and the reporter plasmid using Lipofectamine 2000 (Thermo Fischer Scientific , MA, USA). Cotransfection with the pEGFP-N2 plasmid (Takara Bio, CA, USA) served as a control for different transfection efficiencies. The medium was changed after 24 hours, and measurements were performed after 48 hours. Cells were rinsed with PBS before adding 20 µl PBS for the Green Fluorescent Protein fluorescence intensity measurement on an Infinite 200 Pro plate reader (TECAN, Switzerland), using fluorescence excitation at 465 nm and fluorescence emission at 535 nm. Subsequently, luciferase luminescence was determined by adding 10 µl Luc-pairTM Firefly Luciferase HT Assay Kit solution (Gene Copoeia, MD, USA) in an Infinite 200 Pro plate reader (TECAN, Switzerland). Significance between different cell lines and results of three independent experiments were evaluated by two-sided Student's *t*-test.

Chromatin immunoprecipitation (IP). Cells were crosslinked with 1 % PFA for 10 minutes at room temperature (RT), washed twice with ice-cold 1 x PBS and scraped in low-salt IP buffer (150 mM NaCl, 50 mM Tris/HCl (pH 7.5), 5 mM EDTA, 0.5 % Igepal, 1 % Triton X-100, 1 x proteinase inhibitor cocktail (Roche, Germany). After centrifugation, the nuclear pellet was resuspended in 300 µl IP buffer and sonicated four times for 10 minutes (30 s in "on" mode, 30 s in "off" mode) using the Bioruptor Plus (Diagenode, Belgium). Sheared chromatin was

collected by centrifugation and a 10-µg sample was used for IP. Following preclearing for 2 hours at 4 °C, chromatin was incubated with the specific antibodies for CtBPs (sc-17805, Santa Cruz, CA, US), EZH2 (C15410039, Diagenode, Belgium), H3K27me3 (C15410069, Diagenode, Belgium) mouse IgG (Diagenode, C15400001, Belgium) or rabbit IgG (Diagenode, C15410206, Belgium) overnight by gentle rotation at 4 °C. Recombinant magnetic protein G beads (Invitrogen, MA, USA) were added for 1 hour at 4 °C. Afterwards the beads were washed with low-salt buffer 1 (20 mM Tris/HCl (pH 8.2), 10 mM EDTA, 150 mM NaCl, 0.1 % SDS, 1 % Triton X-100), high-salt buffer 2 (20 mM Tris/HCl (pH 8.2), 10 mM EDTA, 500 mM NaCl, 0.1 % SDS, 1 % Triton X-100), buffer 3 (10 mM Tris/HCl (pH 8.1),10 mM EDTA, 1 % Igepal, 1 % deoxycholate, 250 mM LiCl) and 3 times with Tris-EDTA buffer (10 mM Tris/HCl (pH 8.0), 1 mM EDTA). Each washing step was performed for 10 minutes at 4 °C. Sample was eluted in 100 µl elution buffer containing 0.1 % SDS, 0.1 M $NaHCO_3$, (pH 8.0) for 1 hour at RT. To reverse crosslink, samples were supplemented with 0.2 M NaCl and incubated overnight at 65 °C. Proteinase K was added at a final concentration of 0.5 µg/µl for 2 hours at 42 °C. DNA was purified with a commercial gel and PCR clean-up kit (Macherey-Nagel, Germany) and was directly used for quantitative PCR with SimpleChIP human CDKN1A promoter primers (Cell Signaling Technology, MA, US). ChIP-qPCR data were normalized to input DNA and presented as percentage of input. P-values (Student's *t*-test) were calculated based on at least three biological replicates.

In silico Validation. *RAI2* mRNA expression was analyzed in the prostate adenocarcinoma cohort of the TCGA PanCancer Atlas at the cBIO-portal [29,91]. For correlative analysis with *EZH2* gene expression, the coexpression tab has been applied.

Gene expression analysis in Circulating Tumor Cells (CTC). A total of 117 blood samples were obtained from 92 patients with PC, who were treated at the Center of Oncology,

University Medical Center Hamburg-Eppendorf from June 2018 to July 2022. The study was approved by the ethical commission of the 'Hamburger Ärztekammer' (PV3779, PV5392). Patients provided informed consent for the participation in this study. From each sample, 7.5 ml blood were drawn into EDTA-blood collection tubes and processed within 2 h using the AdnaTest (Qiagen, Hilden, Germany). Patients were divided into HSPC, CRPC, AVPC [34] [34].and NEPC [35], based on their clinical status. Additionally, blood from 10 healthy male donors was obtained as negative control from the Institute of Transfusion Medicine, University Medical Center Hamburg-Eppendorf.

CTC gene expression was analyzed as described [92]. Missing values were replaced by ΔCq max +2 in a gene-specific manner. All ΔCq values higher than the mean ΔCq + 1 SD of the healthy donors were replaced as well. Samples were considered positive for the expression of the respective genes if a ΔCq value below the expression in healthy donors was calculated. Samples positive for AR or KRT19 or for both were defined as CTC positive.

Data Availability. The SAXS data have been deposited at the SASBDB (www.sasbdb.org) with accession codes: **SASDQW5**, **SASDQZ5**, **SASDQ26**, **SASDQ46**, **SASDQ56**, **SASDQ66**, **SASDQ76**, **SASDQ86**, **SASDQ96**, **SASDQA6**, **SASDQB6**, **SASDQC6**, **SASDQD6**, **SASDQE6** and **SASDQF6**. Chemical shifts of RAI2(WT, 303-465) were assigned and deposited in the BMRB under accession number **28085**. The cryo-EM density map has been deposited in the Electron Microscopy Data Bank (EMDB) under the accession code **EMD-15603** while the structure has been deposited in the Protein Data Bank with the code **8ARI**. The X-ray structure has been deposited in the Protein Data Bank with the code **8ATI**.

Acknowledgements: The SAXS and X-ray diffraction data were collected at beamlines P12 and P13 operated by the EMBL Hamburg Unit at the PETRA III storage ring (DESY, Hamburg, Germany), respectively. We also acknowledge access to the EM facility at the

Centre for Structure Systems Biology, Hamburg, for sample screening and collection. We acknowledge technical support by the SPC facility at EMBL Hamburg, University Medical Center Hamburg-Eppendorf Microscopy Imaging Facility (DFG Research Infrastructure Portal: RI_00489) and FACS Sorting Core Unit. We thank Eric Metzger from the University Freiburg Medical Center for helping to set up chromatin binding experiments. We are grateful for technical assistance of Bettina Steinbach and thank Katharina Besler from the University Medical Center Hamburg-Eppendorf for supporting the project administration. We are also grateful for technical and scientific advice at different stages of the project by EMBL colleagues Annabel Parret, Kyung Min Noh, Christian Tischer and Toby Gibson, and Caitlin MacCarthy from the Max Planck Institute for Molecular Biomedicine, Münster, Germany. This project has been supported by Deutsche Forschungsgemeinschaft (DFG 218826742 to HW, KP and MW), (SPP 2084: µBONE to KP and HW), (WE 5844/5-1 to SW and 341/25-1 to KP) and by Deutsche Krebshilfe (70113304) to SW.

Author Contributions

NG, SW, AL, OO, GvA, KP and MW designed the project. NG, SW, EM, SG, LM, AL, AW, SS and GvA performed the experiments. NG, SW, EM, SG, LM, HM, SS, GC, OO, GvA and MW analyzed the data. NG, SW, EM, OO and MW wrote the manuscript. SW, HW, GvA, KP, and MW supported the project.

Competing interests: The authors declare no competing interests.

Reference List

1. Byun, J.S. & Gardner, K. C-Terminal Binding Protein: A Molecular Link between Metabolic Imbalance and Epigenetic Regulation in Breast Cancer. *Int J Cell Biol* **2013**, 647975 (2013).

2. Quinlan, K.G. et al. Role of the C-terminal binding protein PXDLS motif binding cleft in protein interactions and transcriptional repression. *Mol Cell Biol* **26**, 8202-13 (2006).

3. Werner, S. et al. Suppression of early hematogenous dissemination of human breast cancer cells to bone marrow by retinoic Acid-induced 2. *Cancer Discov* **5**, 506-19 (2015).

4. Lee, T.I. & Young, R.A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237-51 (2013).

5. Bradner, J.E., Hnisz, D. & Young, R.A. Transcriptional Addiction in Cancer. *Cell* **168**, 629-643 (2017).

6. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* **12**, 31-46 (2022).

7. Morgan, M.P., Finnegan, E. & Das, S. The role of transcription factors in the acquisition of the four latest proposed hallmarks of cancer and corresponding enabling characteristics. *Semin Cancer Biol* **86**, 1203-1215 (2022).

8. Basu, A. & Atchison, M.L. CtBP levels control intergenic transcripts, PHO/YY1 DNA binding, and PcG recruitment to DNA. *J Cell Biochem* **110**, 62-9 (2010).

9. Yamamoto, M. et al. The PRDM14-CtBP1/2-PRC2 complex regulates transcriptional repression during the transition from primed to naive pluripotency. *J Cell Sci* **133**(2020).

10. Di, L.J. et al. Genome-wide profiles of CtBP link metabolism with genome stability and epithelial reprogramming in breast cancer. *Nat Commun* **4**, 1449 (2013).

11. Wang, R. et al. Role of transcriptional corepressor CtBP1 in prostate cancer progression. *Neoplasia* **14**, 905-14 (2012).

12. Chinnadurai, G. The transcriptional corepressor CtBP: a foe of multiple tumor suppressors. *Cancer Res* **69**, 731-4 (2009).

13. Di, L.J., Fernandez, A.G., De Siervi, A., Longo, D.L. & Gardner, K. Transcriptional regulation of BRCA1 expression by a metabolic switch. *Nat Struct Mol Biol* **17**, 1406-13 (2010).

14. Grooteclaes, M. et al. C-terminal-binding protein corepresses epithelial and proapoptotic gene expression programs. *Proc Natl Acad Sci U S A* **100**, 4568-73 (2003).

15. Madison, D.L. & Lundblad, J.R. C-terminal binding protein and poly(ADP)ribose polymerase 1 contribute to repression of the p21(waf1/cip1) promoter. *Oncogene* **29**, 6027-39 (2010).

16. Katsanis, N. & Fisher, E.M. A novel C-terminal binding protein (CTBP2) is closely related to CTBP1, an adenovirus E1A-binding protein, and maps to human chromosome 21q21.3. *Genomics* **47**, 294-9 (1998).

17. Erlandsen, H., Jecrois, A.M., Nichols, J.C., Cole, J.L. & Royer, W.E., Jr. NADH/NAD(+) binding and linked tetrameric assembly of the oncogenic transcription factors CtBP1 and CtBP2. *FEBS Lett* **596**, 479-490 (2022).

18. Jecrois, A.M. et al. Cryo-EM structure of CtBP2 confirms tetrameric architecture. *Structure* **29**, 310-319 e5 (2021).

19. Madison, D.L., Wirz, J.A., Siess, D. & Lundblad, J.R. Nicotinamide adenine dinucleotide-induced multimerization of the co-repressor CtBP1 relies on a switching tryptophan. *J Biol Chem* **288**, 27836-48 (2013).

20. Kim, T.W. et al. Ctbp2 Modulates NuRD-Mediated Deacetylation of H3K27 and Facilitates PRC2-Mediated H3K27me3 in Active Embryonic Stem Cell Genes During Exit from Pluripotency. *Stem Cells* **33**, 2442-55 (2015).

21. Shalev, M. et al. The PXDLS linear motif regulates circadian rhythmicity through protein-protein interactions. *Nucleic Acids Res* **42**, 11879-90 (2014).

22. Zhao, L.J., Kuppuswamy, M., Vijayalingam, S. & Chinnadurai, G. Interaction of ZEB and histone deacetylase with the PLDLS-binding cleft region of monomeric C-terminal binding protein 2. *BMC Mol Biol* **10**, 89 (2009).

23. Nishikawa, S. et al. Low RAI2 expression is a marker of poor prognosis in breast cancer. *Breast Cancer Res Treat* **187**, 81-93 (2021).

24. Bellesis, A.G., Jecrois, A.M., Hayes, J.A., Schiffer, C.A. & Royer, W.E., Jr. Assembly of human C-terminal binding protein (CtBP) into tetramers. *J Biol Chem* **293**, 9101-9112 (2018).

25. Kumar, V. et al. Transcription corepressor CtBP is an NAD(+)-regulated dehydrogenase. *Mol Cell* **10**, 857-69 (2002).

26. Kuppuswamy, M. et al. Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex. *Mol Cell Biol* **28**, 269-81 (2008).

27. Yu, Y. et al. RBBP8/CtIP suppresses P21 expression by interacting with CtBP and BRCA1 in gastric cancer. *Oncogene* **39**, 1273-1289 (2020).

28. Hsu, C.H., Altschuler, S.J. & Wu, L.F. Patterns of Early p21 Dynamics Determine Proliferation-Senescence Cell Fate after Chemotherapy. *Cell* **178**, 361-373 e12 (2019).

29. Hoadley, K.A. et al. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304 e6 (2018).

30. Davies, A.H., Beltran, H. & Zoubeidi, A. Cellular plasticity and the neuroendocrine phenotype in prostate cancer. *Nat Rev Urol* **15**, 271-286 (2018).

31. Merkens, L. et al. Aggressive variants of prostate cancer: underlying mechanisms of neuroendocrine transdifferentiation. *J Exp Clin Cancer Res* **41**, 46 (2022).

32. Takayama, K. et al. Androgen-responsive long noncoding RNA CTBP1-AS promotes prostate cancer. *EMBO J* **32**, 1665-80 (2013).

33. Takayama, K. et al. CtBP2 modulates the androgen receptor to promote prostate cancer progression. *Cancer Res* **74**, 6542-53 (2014).

34. Aparicio, A.M. et al. Platinum-based chemotherapy for variant castrate-resistant prostate cancer. *Clin Cancer Res* **19**, 3621-30 (2013).

35. Epstein, J.I. et al. Proposed morphologic classification of prostate cancer with neuroendocrine differentiation. *Am J Surg Pathol* **38**, 756-67 (2014).

36. Davies, A. et al. An androgen receptor switch underlies lineage infidelity in treatment-resistant prostate cancer. *Nat Cell Biol* **23**, 1023-1034 (2021).

37. Varambally, S. et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624-9 (2002).

38. van der Lee, R. et al. Classification of intrinsically disordered regions and proteins. *Chem Rev* **114**, 6589-631 (2014).

39. Van Roey, K. & Davey, N.E. Motif co-regulation and co-operativity are common mechanisms in transcriptional, post-transcriptional and post-translational regulation. *Cell Commun Signal* **13**, 45 (2015).

40. Hickabottom, M., Parker, G.A., Freemont, P., Crook, T. & Allday, M.J. Two nonconsensus sites in the Epstein-Barr virus oncoprotein EBNA3A cooperate to bind the co-repressor carboxyl-terminal-binding protein (CtBP). *J Biol Chem* **277**, 47197-204 (2002).

41. Deng, Y. et al. Redox-dependent Brca1 transcriptional regulation by an NADH-sensor CtBP1. *Oncogene* **29**, 6603-8 (2010).

42. Fjeld, C.C., Birdsong, W.T. & Goodman, R.H. Differential binding of NAD+ and NADH allows the transcriptional corepressor carboxyl-terminal binding protein to serve as a metabolic sensor. *Proc Natl Acad Sci U S A* **100**, 9202-7 (2003).

43. Sekiya, M. et al. The transcriptional corepressor CtBP2 serves as a metabolite sensor orchestrating hepatic glucose and lipid homeostasis. *Nat Commun* **12**, 6315 (2021).

44. Zhang, Q. et al. Redox sensor CtBP mediates hypoxia-induced tumor cell migration. *Proc Natl Acad Sci U S A* **103**, 9029-33 (2006).

45. Bhambhani, C., Chang, J.L., Akey, D.L. & Cadigan, K.M. The oligomeric state of CtBP determines its role as a transcriptional co-activator and co-repressor of Wingless targets. *EMBO J* **30**, 2031-43 (2011).

46. Fung, H.Y.J., Birol, M. & Rhoades, E. IDPs in macromolecular complexes: the roles of multivalent interactions in diverse assemblies. *Curr Opin Struct Biol* **49**, 36-43 (2018).

47. Weng, J. & Wang, W. Dynamic multivalent interactions of intrinsically disordered proteins. *Curr Opin Struct Biol* **62**, 9-13 (2020).

48. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419-434 (2019).

49. Banani, S.F., Lee, H.O., Hyman, A.A. & Rosen, M.K. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285-298 (2017).

50. Bienz, M. Head-to-Tail Polymerization in the Assembly of Biomolecular Condensates. *Cell* **182**, 799-811 (2020).

51. Guo, Y., Zhao, S. & Wang, G.G. Polycomb Gene Silencing Mechanisms: PRC2 Chromatin Targeting, H3K27me3 'Readout', and Phase Separation-Based Compaction. *Trends Genet* **37**, 547-565 (2021).

52. Augereau, P. et al. The nuclear receptor transcriptional coregulator RIP140. *Nucl Recept Signal* **4**, e024 (2006).

53. Palijan, A. et al. Ligand-dependent corepressor LCoR is an attenuator of progesterone-regulated gene expression. *J Biol Chem* **284**, 30275-87 (2009).

54. Fong, K.W. et al. PALI1 promotes tumor growth through competitive recruitment of PRC2 to G9A-target chromatin for dual epigenetic silencing. *Mol Cell* **82**, 4611-4626 e7 (2022).

55. Blevins, M.A. et al. Small Molecule, NSC95397, Inhibits the CtBP1-Protein Partner Interaction and CtBP1-Mediated Transcriptional Repression. *J Biomol Screen* **20**, 663-72 (2015).

56. Dcona, M.M., Morris, B.L., Ellis, K.C. & Grossman, S.R. CtBP- an emerging oncogene and novel small molecule drug target: Advances in the understanding of its oncogenic action and identification of therapeutic inhibitors. *Cancer Biol Ther* **18**, 379-391 (2017).

57. Erdos, G. & Dosztanyi, Z. Analyzing Protein Disorder with IUPred2A. *Curr Protoc Bioinformatics* **70**, e99 (2020).

58. Meszaros, B., Erdos, G. & Dosztanyi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* **46**, W329-W337 (2018).

59. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

60. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679-682 (2022).

61. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444 (2022).

62. Brautigam, C.A., Zhao, H., Vargas, C., Keller, S. & Schuck, P. Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat Protoc* **11**, 882-94 (2016).

63. Bohm, G., Muhr, R. & Jaenicke, R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng* **5**, 191-5 (1992).

64. Panjkovich, A. & Svergun, D.I. CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics* **34**, 1944-1946 (2018).

65. Petoukhov, M.V. et al. New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* **45**, 342-350 (2012).

66. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* **50**, 1212-1225 (2017).

67. Semenyuk, A.V. & Svergun, D.I. GNOM - a program package for small-angle scattering data processing. *Journal of Applied Crystallography* **24**, 537-540 (1991).

68. Svergun, D., Barberato, C. & Koch, M.H.J. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography* **28**, 768-773 (1995).

69. Franke, D. & Svergun, D.I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* **42**, 342-346 (2009).

70. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).

71. Evans, P.R. & Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-14 (2013).

72. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).

73. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

74. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).

75. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

76. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).

77. Murshudov, G.N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-67 (2011).

78. Vranken, W.F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-96 (2005).

79. Hyberts, S.G., Takeuchi, K. & Wagner, G. Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *J Am Chem Soc* **132**, 2145-7 (2010).

80. Orekhov, V.Y., Ibraghimov, I. & Billeter, M. Optimizing resolution in multidimensional NMR by three-way decomposition. *J Biomol NMR* **27**, 165-73 (2003).

81. Ulrich, E.L. et al. BioMagResBank. *Nucleic Acids Res* **36**, D402-8 (2008).

82. Keller, R.L. CANTINA verlag. *The computer aided resonance assignment tutorial* (2004).

83. Lee, W., Tonelli, M. & Markley, J.L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-7 (2015).

84. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

85. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).

86. Croll, T.I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530 (2018).

87. Chojnowski, G. Sequence-assignment validation in cryo-EM models with checkMySequence. *Acta Crystallogr D Struct Biol* **78**, 806-816 (2022).

88. Davis, I.W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-83 (2007).

89. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97 (2007).

90. el-Deiry, W.S. et al. WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**, 817-25 (1993).

91. Cerami, E. et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-4 (2012).

92. Besler, K. et al. Expression Patterns and Corepressor Function of Retinoic Acid-induced 2 in Prostate Cancer. *Clin Chem* **68**, 973-983 (2022).

Figures and legends

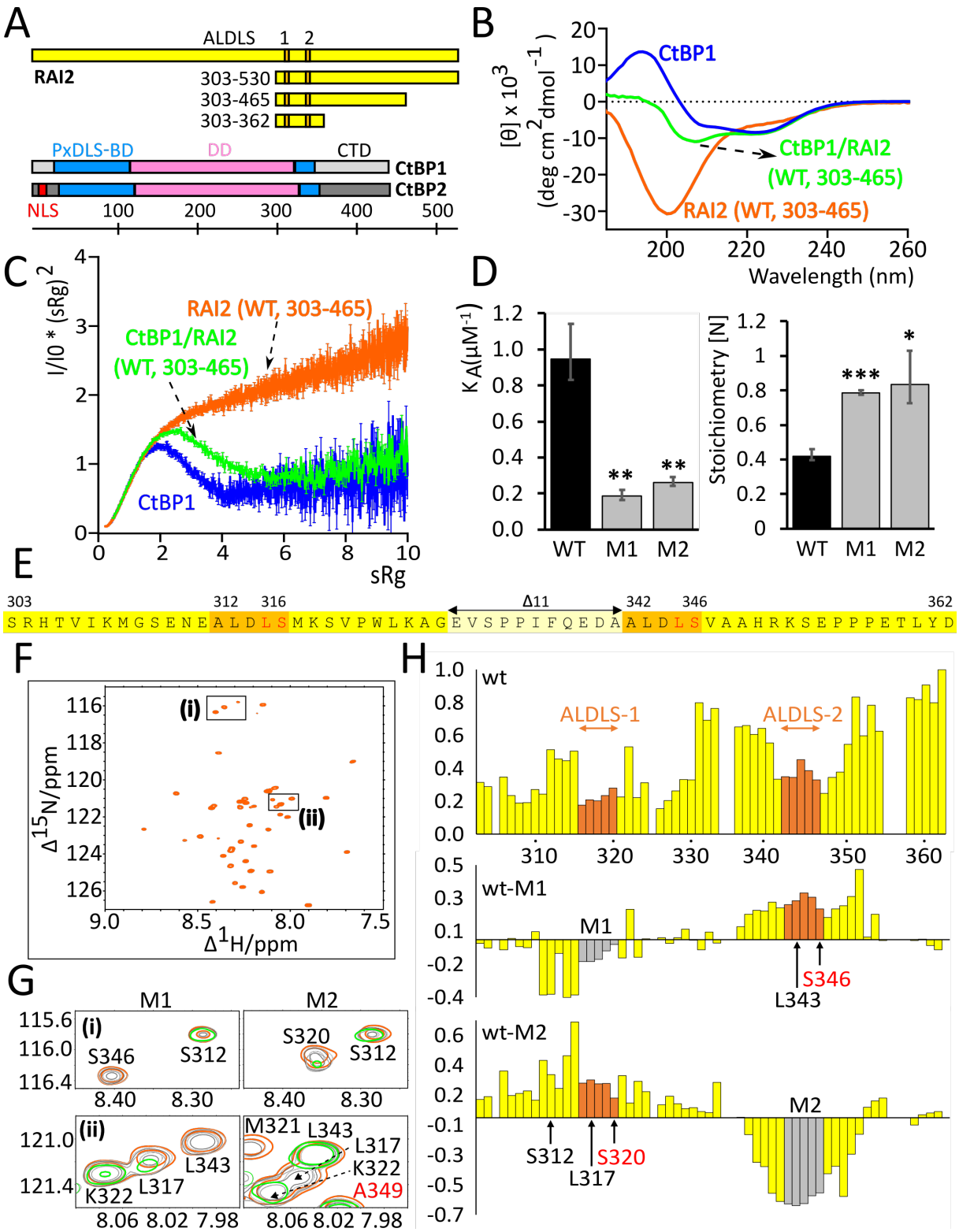

Figure 1: Tandem ALDLS motif-containing RAI2 binding to CtBP is autonomous. A. Protein constructs with sequence segments highlighted in proportion to their length. Colors: RAI2,

yellow: ALDLS motifs 1 and 2, orange; CtBP1 and CtBP2: split PxDLS binding domain (BD), blue; dehydrogenase domain (DD), pink; C-terminal domain (CTD), grey; nuclear localization sequence (NLS), red. The CTD grey shades of CtBP1 and CtBP2 indicate CTD diversity. **B.** CD spectra of CtBP1(WT, FL) (blue), RAI2(WT, 303-465) (orange) and CtBP1(WT, FL)/RAI2(WT, 303-465) complex (green). **C.** SAXS Kratky plots of the protein constructs used in panel B. **D**. Association constants $\mathbf{K_A}$ **(**equivalent to $1/K_D$, left panel, *cf*. **Extended Data Table 2**) and stoichiometry values (right panel) obtained from ITC upon titrations of RAI2(303-362) WT, M1 and M2 to CtBP1**.** *: $p < 0.05$, **: $p < 0.002$, ***: $p < 0.001$. **E.** Sequence of RAI2(WT, 303-362, yellow) comprising two-fold repeated ALDLS motifs (orange). Residues mutated in RAI2 M1, M2 and M1+M2 variants are in red. Linker residues removed in RAI2(Δ11, 303-362) are in light yellow and are indicated by a double arrow. **F.** [^{1}H,^{15}N]-HSQC spectra of RAI2(WT, 303-362) in the absence of CtBP1 (for further details, see **Extended Data Fig. 5**). **G.** Effects of CtBP1 titrations highlighted on selected RAI2 residues (boxed and labeled in panel F) on M1 (left) and M2 (right) variants. Colors of increasing amounts of titrated CtBP1 (WT, 28-353) on to RAI2(303-362) variants: orange (1:0, absence of RAI2), dark grey (1:0.4), light grey (1:0.7), and green (1:1). **H.** Histogram of chemical shift perturbation data upon titration of CtBP1(28-353) on to RAI2(WT, 303-362), indicating intensity (I) ratios for each RAI2 residue calculated by $I_{complex}/I_{free}$ (upper panel). Low intensities ratios indicate residue RAI2 residues perturbation in the presence of CtBP. Middle and bottom panels show intensity ratio differences of CtBP titrations with RAI2(M1) and RAI2(M2), using RAI2(WT) as reference. Mutated ALDLS sites are in grey.

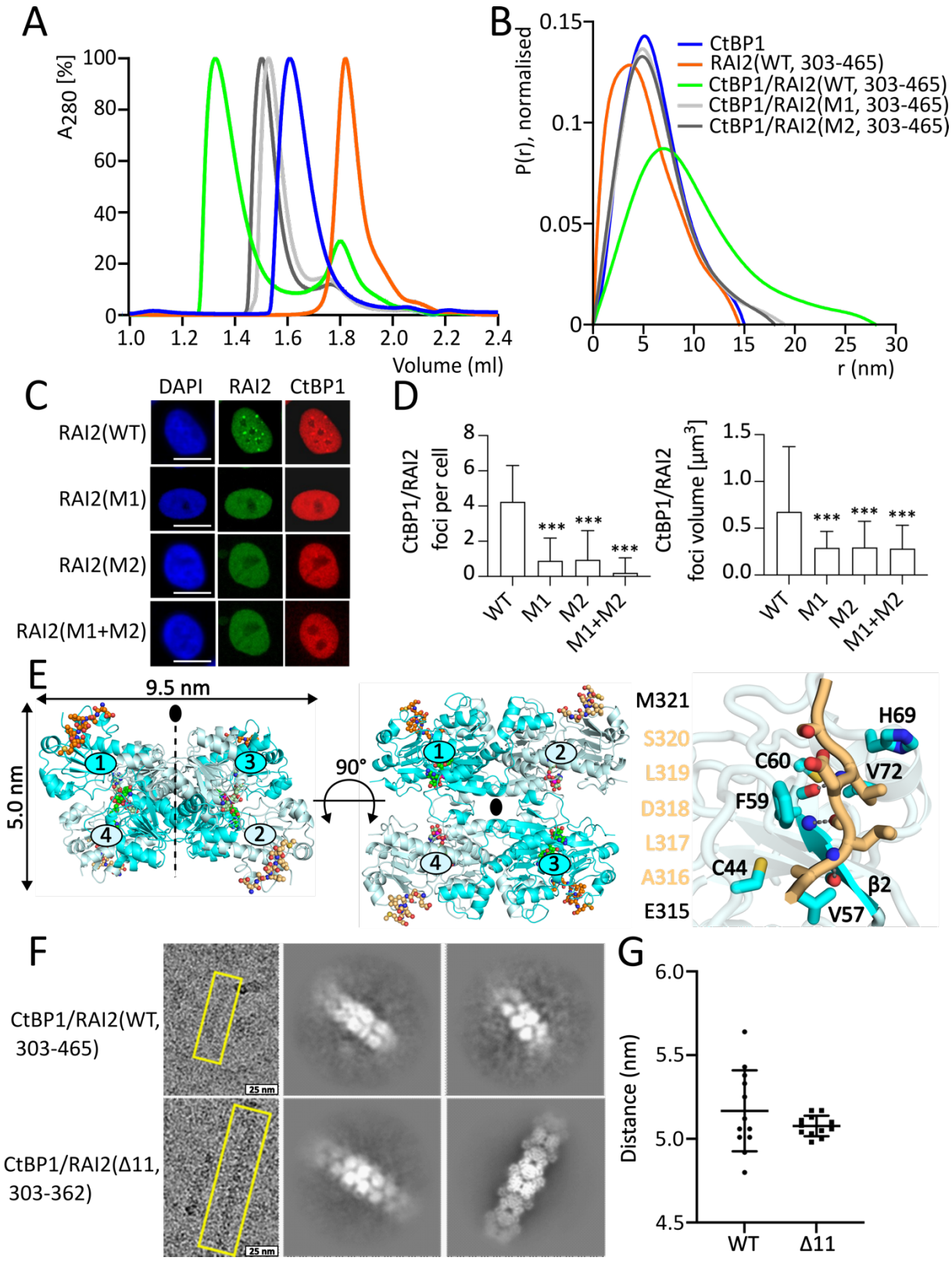

Figure 2. Tandem ALDLS motif-containing RAI2 induces CtBP polymerization. A. Normalized SEC profiles of CtBP1(WT, blue), RAI2(WT, 303-465, orange) or complexes of CtBP1(WT) with

RAI2(WT, green), RAI2(M1, light grey), RAI2(M2, dark grey). **B**. SAXS distance distribution P(r) profiles of CtBP1 in complex with RAI2(303-465) variants. **C**. Immunofluorescence staining of CtBP1 (red) and HA tagged-RAI2 variants (green) in KPL-1 cell nuclei. Nuclei are stained by DAPI (blue). Scale bars: 10 µm. **D**. Number of CtBP1/RAI2 foci per cell (left) and foci volumes (right). *: $p < 0.05$, ***: $p < 0.001$, using RAI2(WT) as reference. **E**. Crystal structure of tetrameric $[(\text{CtBP2(WT, 31-364)/NAD}^+)_2]_2$/RAI2(M2, 303-465)$_4$ complex in side view (left) and top view (middle). CtBP2 molecules are in cartoon presentation, colored cyan and pale-cyan, and numbered. The visible segments of RAI2 (315-321, orange) and NAD^+/NADH ligands (green) are in ball-and-stick presentation. Oxygen, nitrogen and phosphorus atoms are in red, blue and magenta, respectively. Right panel, zoom of CtBP/RAI2 binding site, highlighting specific CtBP-RAI2 interactions within 4 Å distance. The visible RAI2 sequence around the interacting ALDLS motif is indicated on the left. Hydrogen bonds are shown by dashed lines. CtBP strand β2 is highlighted and labeled (*cf*. **Extended Data Fig. 11)**. **F.** Cryo-EM micrographs (left panel) and selected 2D classes of CtBP1(WT, FL)/RAI2(WT, 303-465) and CtBP1(WT, FL)/RAI2(Δ11, 303-362) complexes. Yellow boxes (left) show a representative fiber for respective complexes. **G.** CtBP1 tetramer layer distances, based on the cryo-EM 2D classes shown (panel F), showing diminished distance variability for RAI2(Δ11) compared to RAI2(WT) used for filament formation with CtBP1.

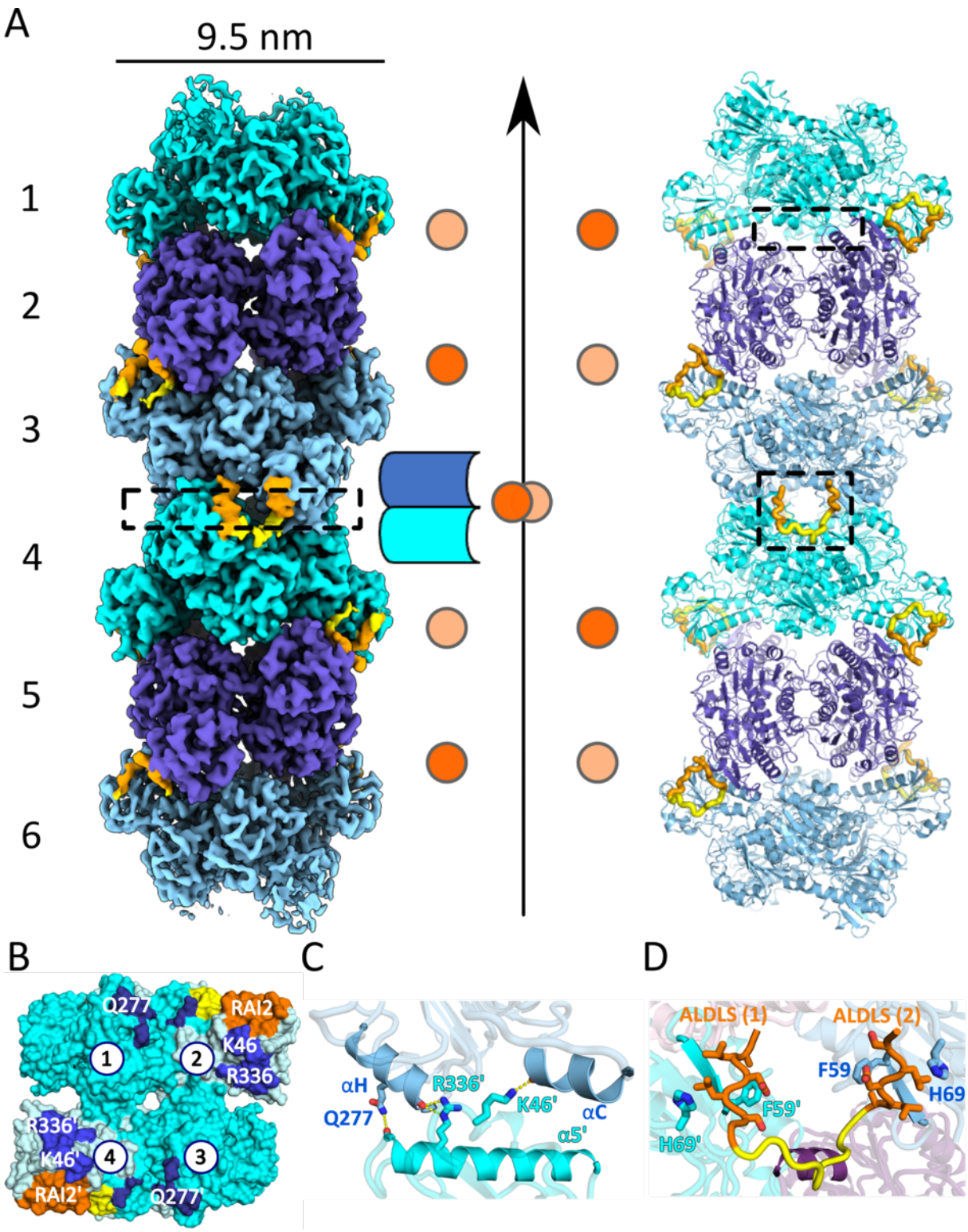

Figure 3: Molecular basis of RAI2-induced CtBP polymerization. A. Cryo-EM density (left) and cartoon representation of the CtBP1(WT, FL)/RAI2(Δ11, 303-362) filament limited to six CtBP1 tetramer layers (right). CtBP1 layers are in cyan, violet, and light blue. Neighboring CtBP1 layers are rotated by a vertical 3-fold filament axis and are connected by a pair of RAI2

peptides (yellow) with a two-fold repeated ALDS motif (orange) in opposite positions to the central filament axis. In the middle panel, the positions of the CtBP1-connnecting RAI2 peptides are schematically indicated for each CtBP1(n)/CtBP1(n+1) interface, highlighting the three-fold symmetry along the filament axis imposed on RAI2-mediated CtBP1 polymerization. Book-like opening of an CtBP1(n)/CtBP1(n+1) interface (boxed) is indicated schematically (*cf*. panel B) **B**. Top view of the RAI2-mediated CtBP1(n)/CtBP1(n+1) interface (boxed in panel A, left), where the surface of one CtBP tetramer is on display (cyan, pale-cyan). Areas interacting with the next tetrameric CtBP layer are in complementary colors (blue, dark blue). The four CtBP subunits are numbered as in **Fig. 2E**. Residues involved in specific CtBP1/CtBP1 interface interactions are indicated (*cf*. panel C). RAI2-interacting surface segments are shown in orange (ALDLS motif) and yellow (other residue segments). **C.** Side view of CtBP1(n)/CtBP1(n+1) interface (boxed in panel A, left), highlighting secondary structural elements and residues contributing to specific interface interactions **(*cf*. Extended Data Fig. 11)**. **D.** Side view of the visible RAI2 segment interacting with two neighboring CtBP layers (boxed in panel A, right). The two RAI2(ALDLS, orange)/CtBP interactions are identical to those observed in non-polymerized CtBP/RAI2(M2) **(*cf*. Fig. 2E)**. The connecting RAI2 ALDLS/ALDLS linker (yellow) is in close contact with a third CtBP subunit (violet).

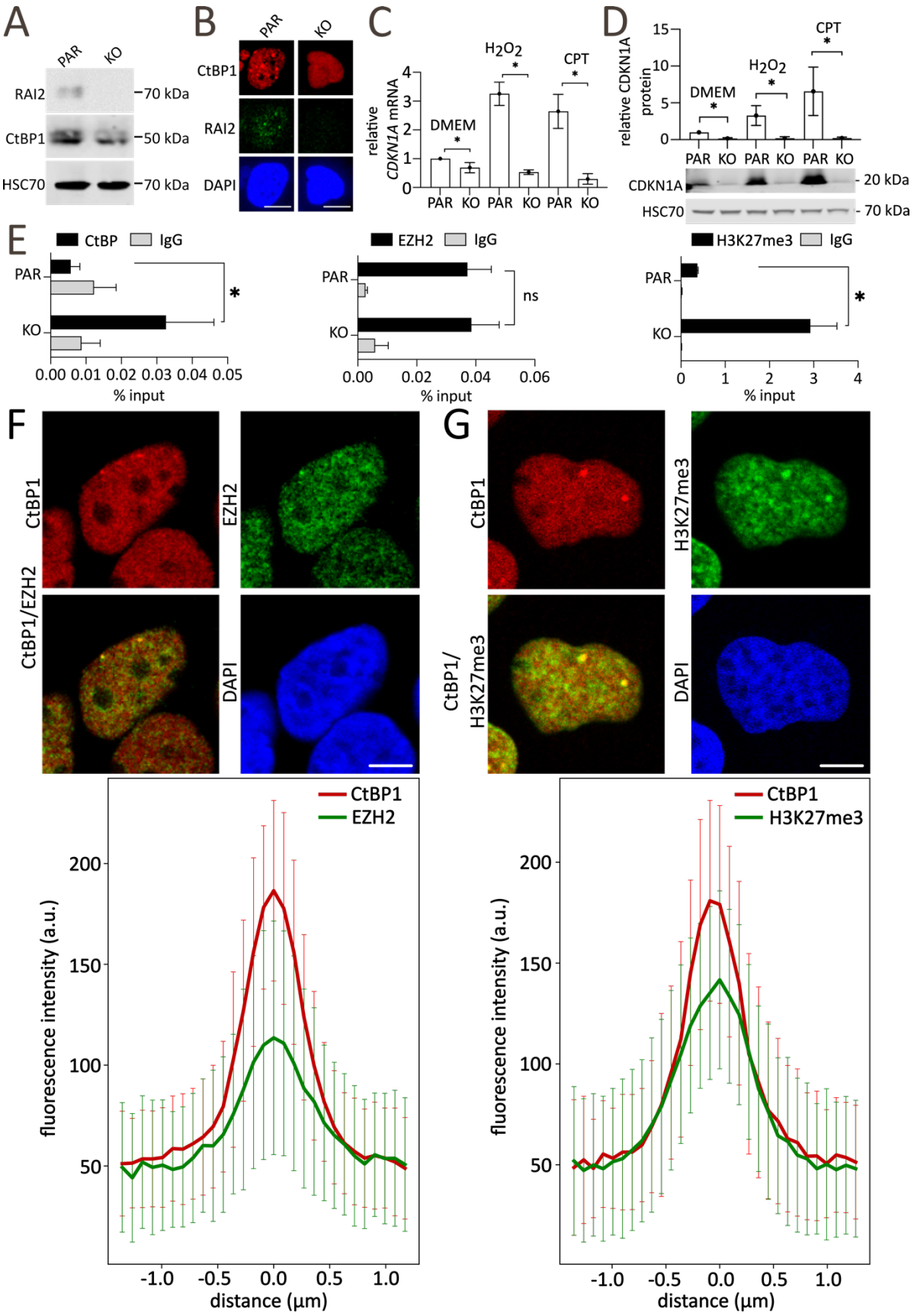

Figure 4. RAI2-induced CtBP polymerization relieves the *CDKN1A* gene repression by downregulating EZH2 activity. A. Expression levels of RAI2, CtBP1 and HSC70 (loading

control) in PAR and RAI2 KO VCaP PC cells. **B**. Immunofluorescence staining of CtBP1 and RAI2 in PAR and RAI2 KO VCaP cell nuclei; DAPI staining was used for localization of nuclei. Scale bars: 10 µm. **C.** *CDKN1A* gene expression levels in PAR and RAI2 KO VCaP cells, untreated (DMEM), H_2O_2 and CPT for 48h. **D.** Lower panel, CDKN1A protein expression levels in PAR and RAI2 KO VCaP cells, untreated (DMEM), H_2O_2 and CPT for 48h. Upper panel, densitometry-based quantification of lower panel. **E**. ChIP-qPCR analysis of CtBPs (left), EZH2 (middle) and H3K27me3 (right) binding to the proximal *CDKN1A* promotor in PAR and RAI2 KO VCaP cells. The respective negative control of IgG binding has also been shown. P-values were calculated using Student's *t*-test. *: $p < 0.05$, "ns" indicates a non-significant value. **F. and G.** Immunofluorescence staining of VCaP PAR cell nuclei, **(F)** stained for CtBP1 (red), EZH2 (green) and DNA/DAPI (blue) and **(G)** for CtBP1 (red), H3K27me3 (green) and DNA/DAPI (blue). The bottom panels depict fluorescence intensity profiles (arbitrary units) of overlay images of foci with EZH2 and H3K27me3 acquired by line scanning of 100 foci images each over 3 µm distance. Colors are as in the images used for analysis. Scale bars: 10 µm.

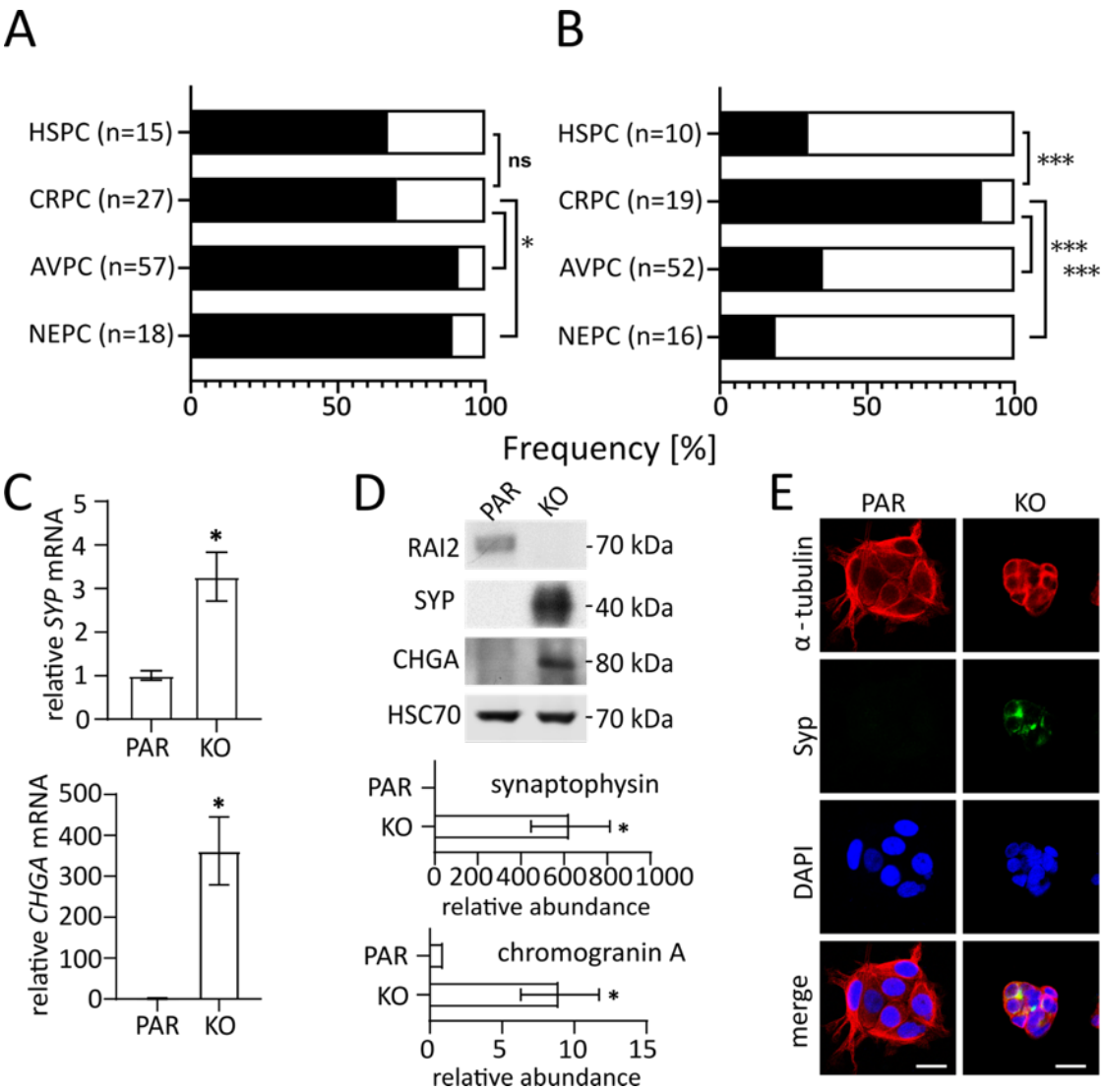

Figure 5: Impact of loss of RAI2 in PC progression. A. Frequency (black bars) of CTC detection in blood samples from HSPC, CRPC, AVPC and NEPC patients. **B**. Frequency (black bars) of *RAI2* mRNA detection in CTC-positive blood samples. **C.** *SYP* and *CHGA* gene expression levels in PAR and KO VCaP cells. **D.** Upper panel, Protein expression levels of RAI2, synaptophysin (SYP), chromogranin A (CHGA) and HSC70 (loading control) in PAR and RAI2 KO VCaP PC whole cell extracts. Lower panel, densitometry-based quantification of upper panel. **E.** Immunofluorescence staining of PAR and KO VCaP cells stained for α-tubulin (red), synaptophysin (green), DAPI (blue) and merge of α-tubulin-DNA (Red-blue). Scale bars: 25 µm. P-values in the experiments were calculated using Student's *t*-test. *: $p < 0.05$, ***: $p < 0.001$, "ns" indicates a non-significant value.

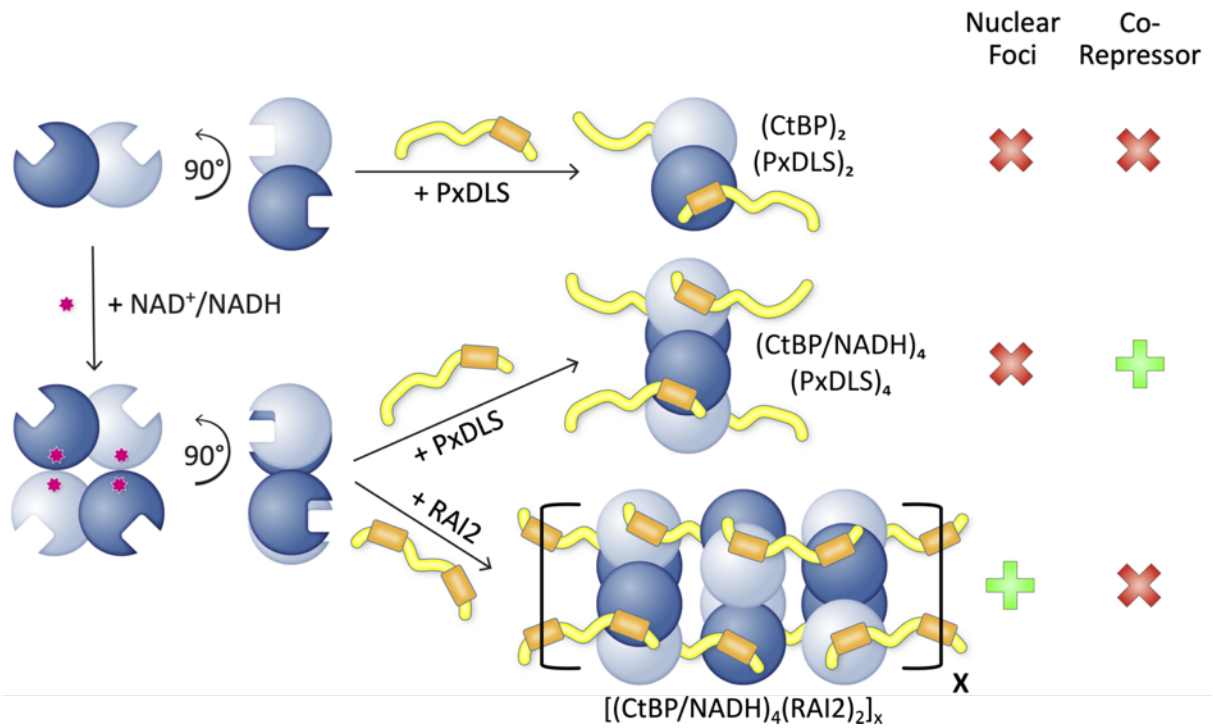

Figure 6: **Model of CtBP corepressor activity regulation by RAI2-induced polymerization.** While CtBP dimerization is NAD^+/NADH-independent, CtBP tetramerization is mediated by redox sensitive NAD^+/NADH-binding (indicated by asterisk). Binding of PxDLS SLiM-containing proteins does not require a specific CtBP assembly state and leads to CtBP/PxDLS-target complexes with multiple 1:1 stoichiometry. Tandem ALDLS motif-containing RAI2 induces CtBP polymerization into CtBP/RAI2 filaments with overall $[4:2]_x$ stoichiometry, which leads to nuclear CtBP/RAI2 foci and inactivates CtBP as a master PRC2 corepressor. Colors are adopted from **Fig. 3**.