


Mechanistic insights into PCBP1-driven unfolding of selected i-motif DNA at G₁/S checkpoint

Received: 30 April 2025

Accepted: 16 January 2026

Published online: 02 February 2026

 Check for updatesPallabi Sengupta¹, Natacha Gillet², Ikenna Obi¹ & Nasim Sabouri^{1,3} 

I-motifs are non-canonical, four-stranded DNA structures in cytosine-rich genomic regions, yet their protein-mediated regulation remains underexplored. Here, we identify PCBP1 (Poly(rC)-binding protein 1) as a selective i-motif-binding protein that unfolds specific i-motifs depending on their protonation and hairpin-forming propensities. Systematic truncation reveals that individual K-homology (KH) domains of PCBP1 cannot selectively bind or unfold i-motifs, but their coordinated actions restore wild-type PCBP1 functions. Using biochemical, biophysical, and molecular dynamics studies, we demonstrate that KH1+2 domains remodel i-motifs, recruiting KH3 to facilitate unfolding and efficient DNA replication. Chromatin and cell-based investigations reveal that PCBP1-knockdown increases i-motif formation at specific genomic loci, coinciding with G₁/S arrest and elevated γ H2AX, indicative of genomic instability. During G₁/S transition, PCBP1 occupancy peaks at these i-motif loci, ensuring i-motif resolution in early S phase. These findings establish PCBP1 as a critical regulator of i-motif dynamics, directly linking its unfolding activity to G₁/S transition and genome stability.

Protein-DNA interactions regulate essential processes like transcription, replication, and DNA repair. These interactions often induce conformational changes in DNA¹⁻³, leading to the emergence or resolution of various DNA structures, like four-stranded G-quadruplexes (G4s) or intercalated-motifs (i-motifs). Their spatiotemporal dynamics and processing by specialized proteins maintain genomic integrity across diverse organisms^{4,5}.

I-motifs are unique non-canonical DNA structures found within cytosine (C)-rich sequences, in the complementary strands of guanine (G)-rich G4-forming regions^{6,7}. They contain two parallel-stranded duplexes intercalating in an anti-parallel manner via hemi-protonated C:CH⁺ base pairs (Fig. 1A), predominantly under acidic conditions^{8,9}. Their biological relevance is controversial due to their requirement of non-physiological acidic pH. However, research has advanced the cellular study of i-motifs by developing iMab antibody^{10,11}, which specifically detects i-motifs in cellular nuclei. Further, growing evidence of i-motif formation at physiological pH¹²⁻¹⁴ within live cells and their

distinct distribution in cancer versus normal cells^{15,16} renewed interest in their biological roles.

A significant proportion of i-motif-forming sequences (iMfs) cluster within oncogene promoters¹⁵, where they may modulate transcription by influencing G4 folding on the complementary strands¹⁷. Moreover, i-motifs form transiently during DNA replication, posing topological obstacles that, if unresolved, may cause deletion and genomic instability¹⁶⁻¹⁸. In a study by Richter's group using Cut-and-Tag sequencing¹⁵, i-motifs are reported to play significant roles in cell cycle regulation. Similarly, Christ's group evidenced i-motif's peak abundance in the nuclei during G₁/S transition, followed by their resolution during replication initiation in early S phase¹⁰. This temporal pattern suggests a correlation between i-motif dynamics and cell cycle, implicating specialized proteins capable of stabilizing and/or resolving i-motifs to ensure the precision and fidelity of cellular activities.

Despite the growing interest in i-motif research, the identification of proteins that bind and regulate i-motifs remains limited due to the

¹Department of Medical Biochemistry and Biophysics, Umeå University, Umeå, Sweden. ²CNRS, ENS de Lyon, Laboratoire de Chimie, UMR 5182, 46 Allée d'Italie, F-69342 Lyon, France. ³Science for Life Laboratory, Umeå University, Umeå, Sweden. ✉ e-mail: nasim.sabouri@umu.se

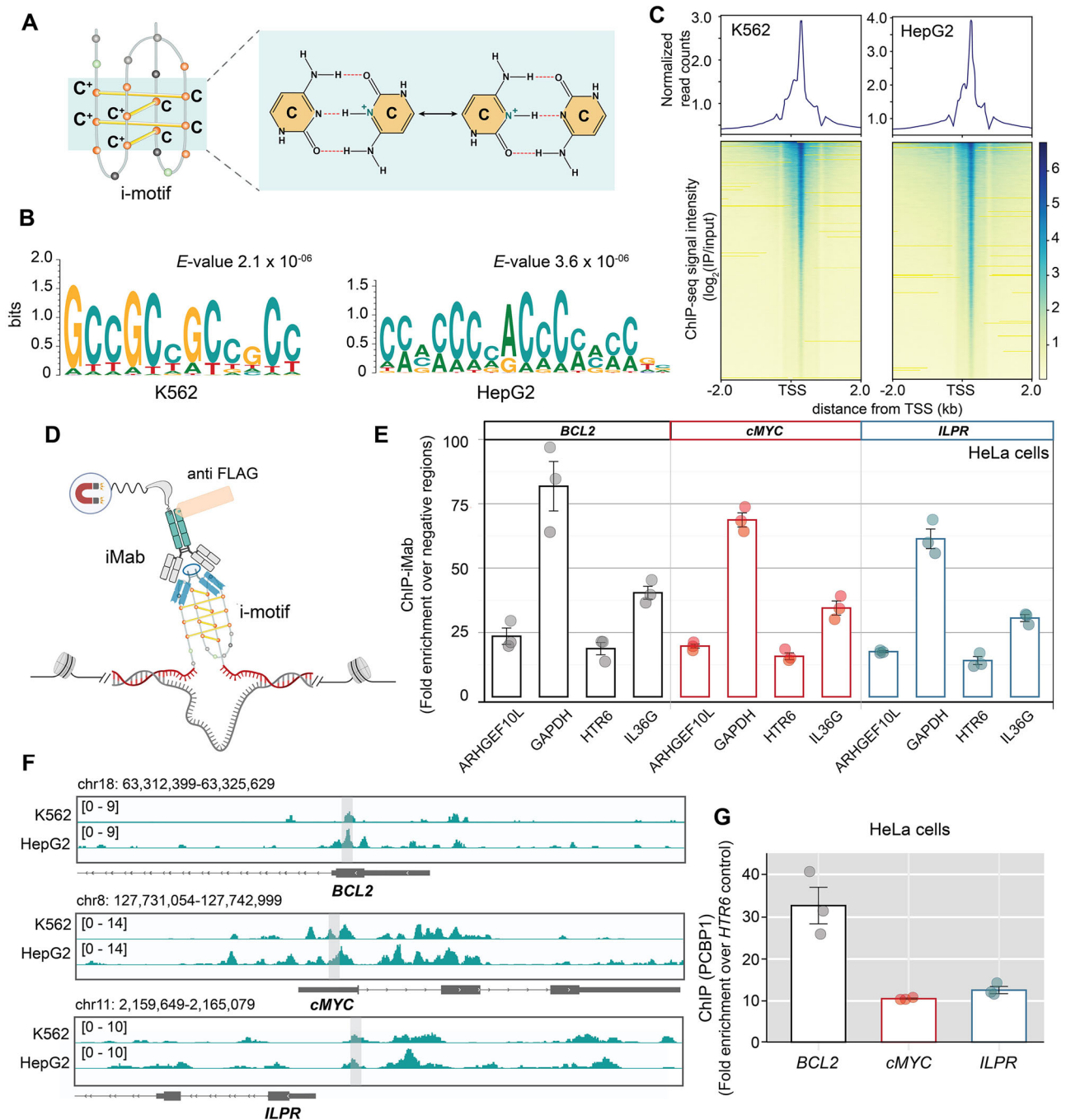


Fig. 1 | PCBPI occupancy is enriched at C-rich genomic sequences that potentially form i-motif structures. **A** Schematic illustration of i-motif DNA structure, highlighting intercalation of C:CH⁺ base pairs between two Cs, where one C is in its protonated form (CH⁺). **B** MEME (Multiple Expectation Maximizations for Motif Elicitation)-based motif-enrichment analysis showing significant enrichment of C-rich sequences in PCBPI-bound ChIP-seq peaks from human K562 (GSE174944) and HepG2 (GSE106035) cells; *E*-value indicates statistical significance. **C** Heatmap profile centered around TSS (transcription start site) for PCBPI-ChIP-seq peaks. Read counts were extracted within a region spanning ± 2 kb around TSS. The gradient blue-to-yellow color indicates high-to-low counts in the corresponding region. Plots at the top of the heatmaps show the normalized read counts at genomic regions centered at TSS. **D** Schematic diagram of iMab-ChIP, illustrating iMab binding to i-motif regions in crosslinked chromatin, captured by anti-FLAG antibody on magnetic beads. Created in BioRender. Sabouri, N. (2026) <https://BioRender.com/84wqzqv>. **E** Quantification of iMab-ChIP efficiency at i-motif

harboring promoters of *cMYC*, *BCL2* oncogenes and insulin gene *ILPR* versus several non-i-motif negative control sites (*ARHGEF10L*, *GAPDH*, *HTR6*, *IL36G*) in HeLa cells. The y-axis shows fold enrichment over these negative regions. **F** IGV (Integrative Genomic Viewer) tracks visualizing ChIP-seq data of PCBPI-bound regions at i-motif-containing promoters in K562 and HepG2 cells. Each track shows the signal intensity range, with a scale noted in brackets for each genomic region, indicating the level of PCBPI occupancy across those genomic regions for different cell types. Highlighted light gray regions overlap with corresponding i-motif-forming sequences under investigation. **G** Quantification of PCBPI-ChIP at three i-motif-containing promoters (*cMYC*, *BCL2*, *ILPR*) compared to the *HTR6* control site in HeLa cells to determine fold enrichment over the *HTR6* site. ChIP-qPCR bar plots in (E, G) represent the means \pm SD of three technical replicates ($n = 3$) and error bars indicate standard deviation (SD). Raw C_t values from input, IP, and mock (IgG) samples obtained from three independent biological replicates, each measured with three technical replicates, are provided in Supplementary Data 1 and 3.

lack of robust i-motif-detecting tools. In addition, the pH sensitivity of proteins makes interaction studies challenging, as reduced protein stability at lower pH can impede accurate assessment of i-motif binding. Recent proteomic studies uncovered several candidate proteins potentially involved in i-motif binding¹⁹. Among them, hnRNP (Heterogeneous Nuclear Ribonucleoprotein) or PCBP (Poly-(rC)-binding Protein)-family proteins showed the most promising *in vitro* interactions with C-rich sequences^{20–25}. However, their precise roles in i-motif binding and regulation are underexplored. Immunofluorescence studies demonstrated PCBP1's (hnRNP-E1) co-localization with i-motif foci^{26,27}. HnRNP-K has been reported to unfold *cMYC*-i-motif *in vitro*²¹. While these studies provided indirect evidence of PCBP's potential role in i-motif binding, they did not offer a mechanistic understanding of how PCBP regulates i-motifs at specific genomic loci. This is particularly important in the evolving field of i-motif research, because several independent studies observed i-motif's transient nature at specific loci, wherein the same genomic regions may not consistently fold into i-motifs across various cell types or under different intracellular conditions^{15,16,24}. This may cause oversight of specific folding events at particular loci or the dismissal of sporadic observations, necessitating a systematic investigation to characterize PCBP's role in i-motif regulation at specific loci and its cellular implications.

Here, we investigated the multifunctional protein, PCBP1 (Poly(rC)-binding protein 1), which belongs to the hnRNP family^{28,29} and is known to bind C-rich RNA/DNA sequences^{30,31}. PCBP1 contains three K-homology (KH) domains and plays key roles in oxidative stress response³², genomic integrity maintenance^{33,34}, and cell cycle regulation^{31,35–37}. While aberrant PCBP1 expression is linked to cell cycle arrest in various cancers^{31,37}, its molecular mechanisms remain unclear. Our study illustrates the mechanism of i-motif regulation by different domains of PCBP1 and establishes its implications to regulate G₁/S transition during the cell cycle. Using chromatin immunoprecipitation (ChIP) and cell-based studies, we demonstrated that PCBP1 is enriched at specific i-motif foci in the genome in a cell cycle-dependent manner, allowing these i-motifs to transiently accumulate during G₁/S transition and unfold in early S phase to facilitate replication. PCBP1 depletion caused G₁/S arrest and increased replication stress, reinforcing that PCBP1-mediated i-motif unfolding is essential for proper G₁/S transition. We also illustrated the molecular basis of specific interactions between PCBP1 and i-motifs by integrating spectroscopic studies, molecular dynamics (MD) simulations, and biochemical assays, which collectively revealed that PCBP1 selectively unfolds specific i-motifs with distinct kinetics, depending on their protonation states and hairpin-forming tendencies. We demonstrated a hierarchical mechanism for i-motif resolution, where the cooperative action of KHI + 2 domains remodels i-motifs, priming them for KH3-driven unfolding. Our findings lay the foundation for future studies exploring how PCBP1 orchestrates i-motif dynamics to ensure genomic stability and fidelity of G₁/S transition.

Results

PCBP1 is enriched at C-rich, potential i-motif-forming genomic sites

PCBP1 is well known for its affinity toward C-rich DNA sequences^{31,38}. However, its specific role in i-motif formation and regulation remains unclear. To investigate PCBP1's genomic binding preferences, we performed MEME-based motif-enrichment analysis³⁹ on published PCBP1-ChIP-seq datasets (Fig. 1B), which identified significantly enriched C-rich motifs with interspaced 3C/2C-repeats, resembling consensus iMfs (Fig. 1A, B). A global analysis of ChIP-seq signals further revealed significant enrichment of PCBP1-peaks around ± 2 kb of TSS (transcription start site) (Fig. 1C) overlapping with several promoters, including *cMYC*, *BCL2*, *ILPR*, *PDGFa*, *VEGFa*, and *HIF1a* (Fig. 1F, Supplementary Fig. 1A) that harbor C-rich sequences capable of forming i-motifs *in vitro*^{40–43}. To assess if these promoter sequences form

i-motifs in cells, we performed ChIP-quantitative PCR (ChIP-qPCR) with iMab antibody. While iMab is widely regarded for its selective i-motif binding within cells¹¹, recent studies argued over its potential interactions with non-i-motif-forming-C-rich sequences⁴⁴. To address this, we compared iMab occupancy in the target regions to several control sites (Fig. 1D) with moderate C-richness (~50%) (Supplementary Table 2), but no evidence of i-motif formation based on prior G4-ChIP-seq⁴⁵ and iMab-Cut-and-Tag-seq datasets¹⁵. We detected no significant immunoprecipitation (IP)-iMab signal at *HIF1a* promoter, while positive IP-iMab signals for *VEGFa* and *PDGFa*, albeit only against two of the four C-rich control regions (Supplementary Fig. 1B). For *cMYC*, *BCL2*, and *ILPR* promoters, we observed ~10-fold higher iMab occupancy compared to all control sites (Fig. 1E), suggesting their strong potential to form i-motifs in HeLa cells. Subsequently, we validated PCBP1's interactions with these sequences using ChIP-qPCR with PCBP1-specific antibody, which demonstrated robust PCBP1 occupancy at these sites (Fig. 1G). Collectively, our findings indicate that PCBP1 preferentially binds to C-rich genomic sequences that potentially form i-motifs.

PCBP1 selectively binds i-motifs over their unfolded forms *in vitro*

Since PCBP1 preferentially interacts with i-motif-forming genomic regions, we hypothesized that PCBP1 would bind iMfs *in vitro*. As both proteins and i-motifs' stability are pH-dependent, we first investigated their pH tolerance. At pH 6.4, 7.0, or 8.0, purified PCBP1 showed the characteristic double-dip α -helix structure in circular dichroism (CD) spectra⁴⁶ (Supplementary Fig. 2A, B). At pH 6.0 and 5.5, the negative peak around 210–220 nm became less pronounced, suggesting partial structural changes (Supplementary Fig. 2B). This was supported by thermal-shift assays, which revealed about 5 °C decrease in PCBP1's melting temperature (T_m) at pH 6.0 or 5.5, compared to pH 6.4 and higher (Fig. 2A, Supplementary Table 3), suggesting that PCBP1 maintains optimal functions and structural integrity at \geq pH 6.4. Next, we obtained CD spectra of *cMYC*, *BCL2*, *ILPR*, hTeloC (C-rich telomeric strand)-iMfs that yielded robust iMab-positive signals to evaluate pH-dependent stability and transition pH (pH_T) (where 50% of i-motifs remains folded at a particular pH) (Fig. 2B, Supplementary Fig. 3). All sequences demonstrated intramolecular i-motif spectra with a signature CD profile of positive and negative maxima around 287 and 260 nm, respectively⁴⁰, and pH_T values above pH 6.4 (Fig. 2B, Supplementary Figs. 3 and 4). These findings allowed us to investigate PCBP1 binding to iMfs at pH 6.4 and 8.0, ensuring PCBP1's stability. At pH 6.4, a significant proportion of i-motifs remained folded, whereas at pH 8.0, they were unfolded, allowing comparison of PCBP1's affinity to both structural states of the same sequence.

The electrophoretic mobility shift assays (EMSAs) at pH 6.4 and 8.0 revealed that PCBP1 binds to both i-motifs and their unfolded states; nevertheless, complete band-shifts occurred at lower PCBP1 concentrations at pH 6.4, indicating PCBP1's stronger preference for i-motifs (Fig. 2C, D). Notably, EMSA gels revealed distinct migration patterns of PCBP1-iMfs complexes between pH 6.4 and 8.0. At pH 8.0, only a single-shifted band was detected. However, at pH 6.4, two slow-migrating shifted complexes appeared for all iMfs, most prominently for *BCL2*, that increased in intensity with rising PCBP1 concentrations. These multiple bands at pH 6.4 likely arise from i-motif's structural heterogeneity induced by PCBP1 interactions or transient PCBP1 multimer formation. To clarify this, we performed ITC (Isothermal titration calorimetry) (Supplementary Fig. 5), which fitted into one-site binding models at both pH levels, thereby excluding higher-order binding stoichiometries.

To quantify binding affinities, we performed microscale thermophoresis (MST). Because subtle pH changes between pH 6.4 and 8.0 may influence both i-motif equilibrium and protein charge distribution, we also measured affinities at pH 7.0 as an intermediate condition (Fig. 2E, Supplementary Fig. 6A–E). To further evaluate PCBP1's selectivity towards i-motifs, we tested G4-forming or non-GC-rich

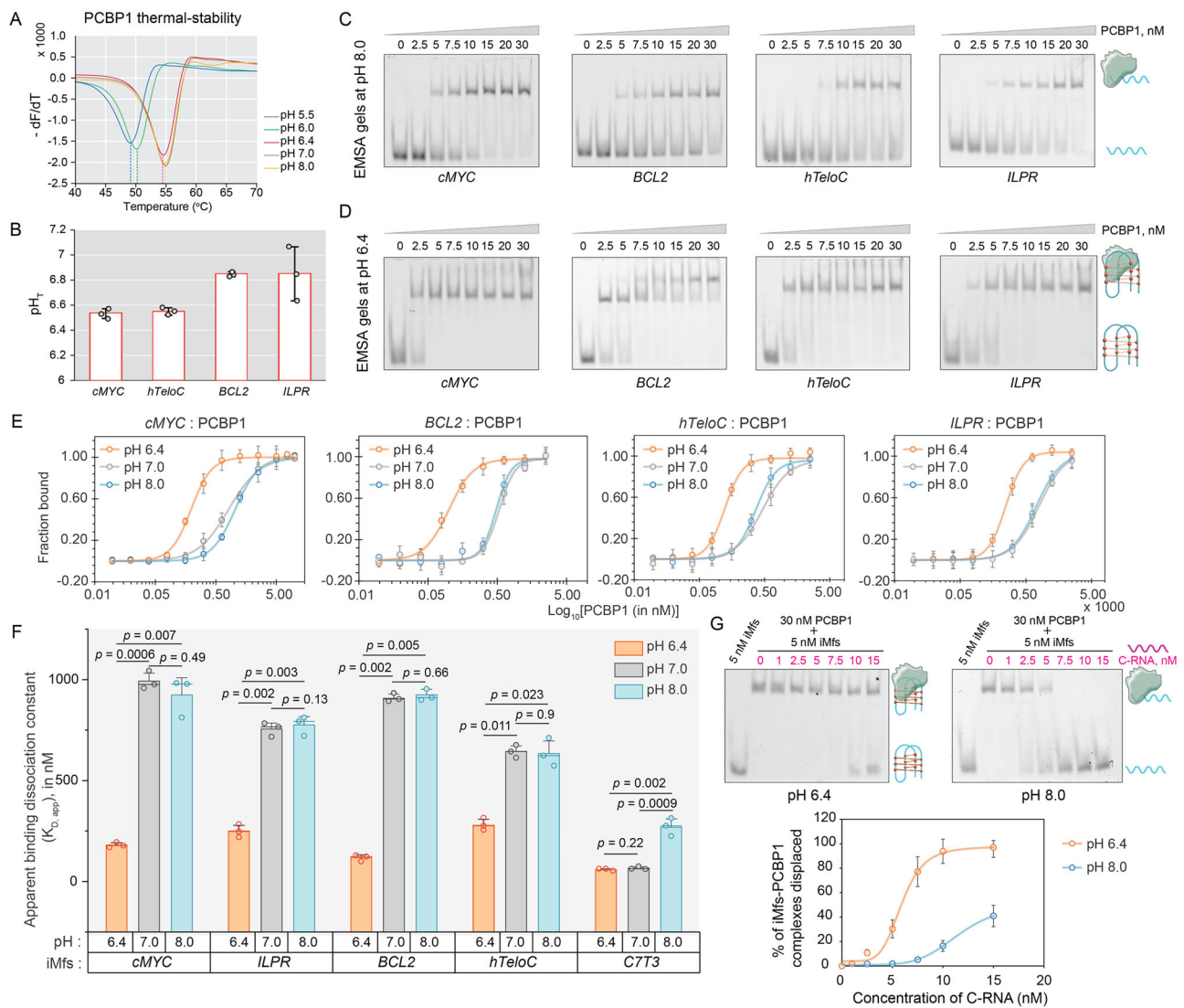


Fig. 2 | PCBPI exhibits higher affinity for i-motif DNA compared to its unfolded states in vitro. **A** Thermal-shift assay of PCBPI at different pH levels. The y-axis shows $-\frac{dF}{dT}$ representing the negative derivative of Sypro-Orange fluorescence signal with respect to temperature; melting temperature (T_m) defined by negative maxima at different pH values. **B** pH_T values indicate the transition pH, at which 50% of i-motifs remain folded at 25°C, derived from CD spectroscopy for *cMYC*, *BCL2*, *ILPR*, and *hTeloC*-i-motifs. **C** EMSA gels at pH 8.0 and **D** at pH 6.4, showing 5'-Cy5-labeled iMfs with band-shifts upon incubation with increasing concentrations of recombinantly purified PCBPI. **E** Sigmoidal curves from MST binding experiments between

PCBP1 and iMfs at pH 8.0, 7.0, and 6.4. **F** $K_{D,app}$ estimated from MST binding curves for iMfs at pH 8.0, 7.0, and 6.4. Statistical significance and p value calculated using two-tailed Student's t -test, where p values < 0.05 are considered statistically significant. **G** Competitive EMSA showing displacement of iMfs-PCBP1 complexes with increasing concentration of C-RNA performed at pH 6.4 and 8.0. Native gels (top) and densitometric analyses of gels (bottom). Data in **(B, E, F, G)** represent the mean \pm SD from three independent biological replicates ($n = 3$). Error bars indicate standard deviation. Uncropped and unprocessed gel images corresponding to EMSA experiments shown in **(C, D, G)** are provided in the Source data.

sequences, which showed significantly weaker binding than i-motifs (Supplementary Fig. 7, Supplementary Table 4). Consistent with EMSA, PCBPI showed >2-fold higher affinity for i-motifs at pH 6.4 than their unfolded counterparts at pH 7.0 or 8.0, with apparent dissociation constants ($K_{D,app}$) ranging between 150 and 300 nM (Fig. 2F, Supplementary Table 4). To verify whether PCBPI's selectivity depends on structural folding rather than sequence composition, we included a synthetic and well-characterized C7T3-i-motif^{47,48} that remains stable at pH 7.0 due to its high C-density. For promoter and *hTeloC* iMfs, which unfold at pH 7.0 and higher, PCBPI binding showed no significant difference between pH 7.0 and 8.0. However, C7T3, which forms a stable i-motif at pH 6.4 and 7.0 but unfolds at pH 8.0, displayed stronger $K_{D,app}$ values at pH 6.4 and 7.0 than at pH 8.0 (Fig. 2F, Supplementary Fig. 8), reinforcing PCBPI's binding preference for i-motifs over unfolded forms. Notably, C7T3's unfolded form at pH 8.0 still

bound to PCBPI with sub-micromolar binding affinity, which is comparable to other i-motifs at pH 6.4, suggesting that both folded and unfolded states may contribute to PCBPI's biological interactions, influenced by i-motif stability and C-density of the sequence. To probe the thermodynamic basis of this preference, we performed ITC, revealing a free energy difference ($\Delta\Delta G$) of 0.5–0.6 kcal mol⁻¹ between pH 6.4 and pH 8.0 (Supplementary Fig. 9). This corresponded to ~two-fold higher affinity at pH 6.4 (i-motif folded state), consistent with MST findings.

We further assessed binding selectivity by competitive EMSA using a canonical C-rich RNA (C-RNA) substrate (Supplementary Table 1) of PCBPI³¹. At pH 8.0, equimolar RNA concentrations displaced ~50% of the pre-formed PCBPI–iMfs complex with complete displacement at a two-fold RNA excess (Fig. 2G). In contrast, at pH 6.4, even a three-fold RNA excess displaced only ~40%, indicating that

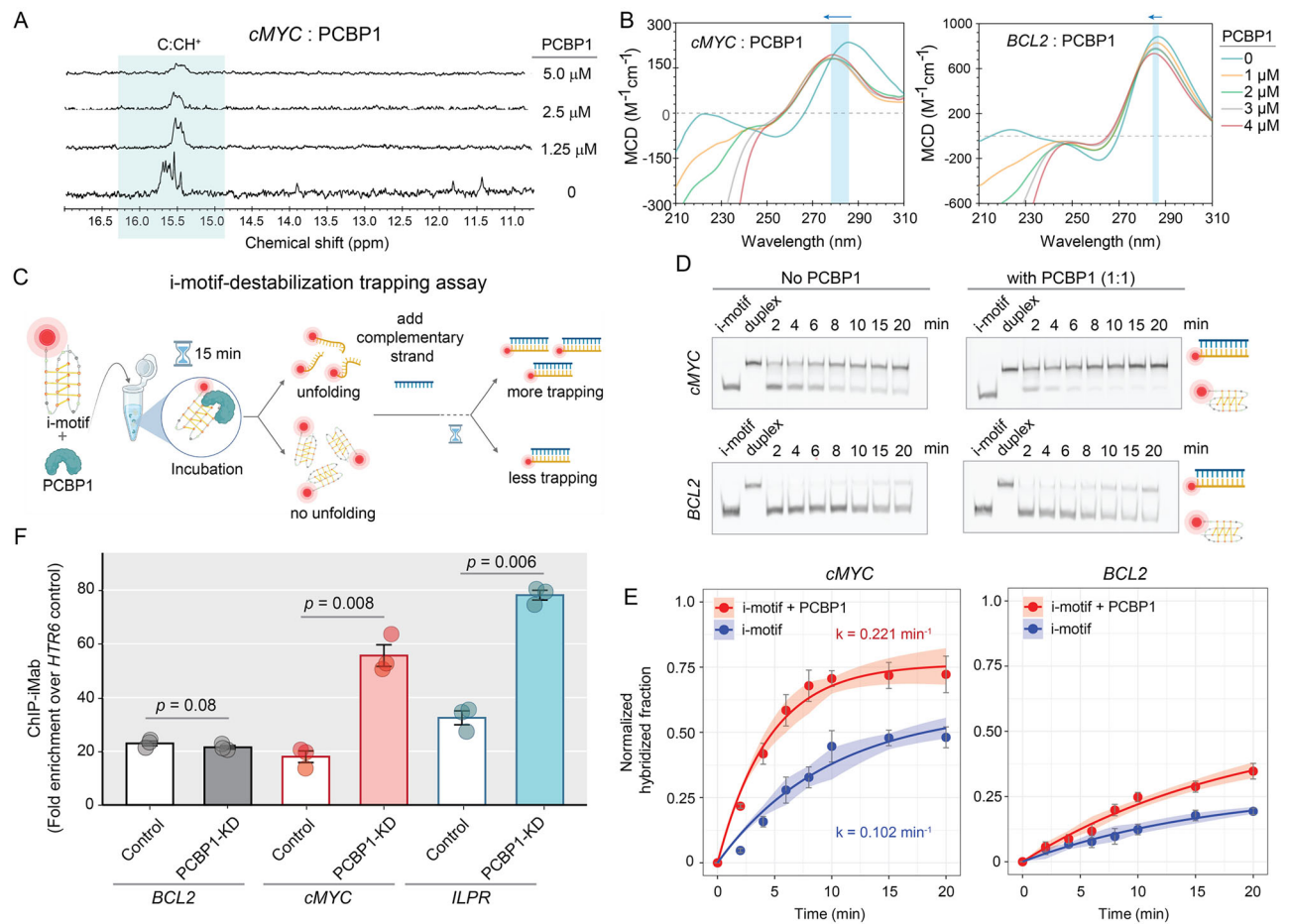


Fig. 3 | PCBPI unfolds *cMYC*-i-motifs with faster kinetics than *BCL2*-i-motifs in vitro. **A** ^1H NMR spectra. Broadening of imino proton resonances (15–16 ppm defining C:CH $^+$ bonds) of *cMYC*-i-motif during PCBPI titrations. **B** CD spectra of *cMYC* and *BCL2*-i-motifs upon PCBPI titrations, with highlighted regions showing blue-shift (with arrows) and hypochromic effects. **C** Schematic diagram of i-motif-destabilization trapping assay. PCBPI and i-motifs are incubated for 15 min, followed by the addition of complementary strands or traps; reactions were stopped at various time points until 20 min. PCBPI-driven i-motif unfolding led to increased hybridization rates (k) between i-motifs and their corresponding complementary strands. Created in BioRender. Sabouri, N. (2026) <https://BioRender.com/oaozsm>. **D** Native PAGE visualizing aliquots of hybridization reactions at different time intervals upto 20 min, alongside negative control (i-motifs without complementary strands) and positive controls (hybridized duplex for 24 h) in the absence and presence of PCBPI in 1:1 molar ratio. Uncropped and unprocessed gel images

corresponding to native PAGE experiments are provided in the Source data. Created in BioRender. Sabouri, N. (2026) <https://BioRender.com/oaozsm>. **E** Densitometric analysis of i-motif-trapping with and without PCBPI. Normalized intensities of hybridized fractions calculated from three independent biological replicates ($n = 3$) at different time points, fitted to single-exponential function equations. Error bands/bars denote standard deviation. **F** Fold enrichment of iMab-ChIP at i-motif-harboring promoters (*cMYC*, *BCL2*, *ILPR*) versus non-i-motif control site (*HTR6*); error bars from standard deviation of three technical replicates ($n = 3$). Statistical differences and p values between control and PCBPI-KD determined by one-way ANOVA followed by Tukey–Kramer post hoc test, where p values < 0.05 , are considered statistically significant. Raw C_t values from input, IP, and mock (IgG) samples obtained from three independent biological replicates, each measured with three technical replicates, are provided in Supplementary Data 2.

unfolded i-motifs form weaker complexes with PCBPI that are readily outcompeted by C-rich RNA, whereas folded i-motifs provide a more stable and preferred binding platform for PCBPI. To address the possibility that PCBPI binding arises from non-specific electrostatic interactions with protonated Cs, we evaluated a previously studied control C-rich sequence (mutC) 44 that does not fold into an i-motif under our acidic experimental conditions despite its C-richness (Supplementary Fig. 10A, Supplementary Table 1). EMSA at both pH 6.4 and pH 8.0 showed no detectable complex between mutC and PCBPI (Supplementary Fig. 10B). Together, these data demonstrate that PCBPI binding is structure-specific and not driven by general electrostatic interactions, and it directly interacts with i-motifs in vitro with higher affinities than their unfolded C-rich forms.

PCBPI unfolds different i-motifs with varied kinetics

Having established PCBPI's preferential recognition for i-motifs, we sought to examine how its binding influences i-motif structures. To

directly probe the conformational effects of PCBPI binding on i-motifs, we performed ^1H -NMR titrations of PCBPI into i-motifs. We observed significant broadening and up-field shifts in C:CH $^+$ resonances around 15–16 ppm, suggesting altered i-motif dynamics upon PCBPI binding (Fig. 3A, Supplementary Fig. 11A–C). However, since spectrum broadening results from both strong binding and protein-induced conformational alterations, we acquired CD spectra of *cMYC*, hTeloC, and *ILPR*-i-motifs, which exhibited hypochromism and distinct blue-shifts upon PCBPI titrations, suggesting PCBPI-induced structural destabilization of these i-motifs (Fig. 3B, Supplementary Fig. 12A, B). In contrast, *BCL2*-i-motif displayed hypochromism but lacked the pronounced blue-shift (Fig. 3B) in CD, suggesting a variable impact of PCBPI binding on *BCL2*-i-motif. To quantify the unfolding kinetics, we designed an i-motif-destabilizing-trapping assay (Fig. 3C), where we added i-motifs' complementary strands as traps to hybridize with unfolded i-motifs, preventing their refolding. We hypothesized that if PCBPI accelerates the unfolding of specific i-motifs, it would

accumulate more hybridized products within a shorter incubation period and faster hybridization rates (k) (Fig. 3C). In the absence of PCBPI, unfolding was generally slow, but *BCL2*-i-motif was strikingly slower compared to other iMfs. PCBPI addition significantly accelerated their unfolding, with varying k across four i-motifs (Fig. 3D, E). Consistent with CD, *cMYC* and hTeloC-i-motifs unfolded most rapidly, *ILPR*-i-motif showed an intermediate rate, while *BCL2*-i-motif remained comparatively resistant to PCBPI-induced destabilization, showing only a modest increase in trap formation (Fig. 3D, E, Supplementary Fig. 13A, B). These results indicate that *BCL2* adopts an intrinsically kinetically persistent fold that impedes spontaneous unfolding, and that this intrinsic stability limits the ability of PCBPI to promote its unfolding. Next, we examined the binding preferences within cells and performed iMab-ChIP-qPCR following siRNA-mediated knockdown of endogenous PCBPI (PCBPI-KD) within HeLa cells (Supplementary Fig. 14). Aligning with *in vitro* observations, PCBPI-KD enhanced iMab occupancy by 3-fold at *cMYC* and *ILPR* promoters, suggesting more stable i-motif formation at these regions upon PCBPI-KD (Fig. 3F, Supplementary Fig. 15). However, no significant change in iMab occupancy was observed at the *BCL2* promoter, reinforcing that PCBPI differentially regulates i-motif stability in a sequence-dependent manner. Together, these results reveal that while PCBPI preferentially recognizes i-motif structures, its binding exerts variable effects on their conformational stability and unfolding kinetics, likely reflecting sequence-intrinsic structural features within individual iMfs.

PCBPI-mediated i-motif-unfolding depends on their protonation and hairpin-forming potential

Since PCBPI resolves different i-motifs at varying rates, we investigated why certain i-motifs resist PCBPI-mediated unfolding. We hypothesized that i-motifs' intrinsic stability could be a contributing factor. To test this, we examined C7T3-i-motif, which has a pH_T of 7.1⁴⁹, and is more stable at pH 6.4 than other i-motifs tested. CD titrations revealed stronger blue-shifts and hypochromism in C7T3 (Supplementary Fig. 16), compared to *BCL2*, though less pronounced than *cMYC* or hTeloC, and similar to *ILPR*, suggesting efficient but variable PCBPI-driven unfolding. C7T3's greater protonation state, due to its high C-richness at pH 6.4, may contribute to its distinct unfolding. However, *BCL2*, despite having similar pH_T as *ILPR*, showed negligible unfolding, indicating that factors beyond protonation, such as structural features, may influence its resistance to PCBPI-mediated unfolding.

Since i-motifs often co-exist with hairpin structures, their unfolding may facilitate hairpin formation^{22,42,50}. To explore this, we examined *BCL2*-i-motif using ¹H NMR, which revealed proton resonances between 12–14 ppm, indicating Watson-Crick base-pairing in hairpin structures (Supplementary Fig. 11A), consistent with earlier studies⁴². To determine if PCBPI-mediated i-motif-unfolding is influenced by hairpin formation, we performed high-resolution primer extension assays⁵¹ with *cMYC* and *BCL2*-i-motifs, and non-i-motif controls, with or without PCBPI (Fig. 4A–D, F, Supplementary Figs. 17 and 18). In this assay, DNA polymerase stalls at replication obstacles (e.g., i-motif or hairpin), producing distinct bands on denaturing gels. To enhance i-motif stability under the conditions dictated by the replication assay buffer, we conducted these assays at a more acidic pH (6.0 instead of 6.4), where PCBPI undergoes only a mild structural rearrangement compared to pH 6.4 (Supplementary Fig. S19). No pausing signal was detected in the non-i-motif control template (Supplementary Fig. 17), while in *cMYC*, we observed a single stalling/pausing signal one-nucleotide preceding i-motif, suggesting that *cMYC*-i-motif acts as a stable replication barrier. PCBPI-binding significantly enhanced replication through *cMYC*, yielding $\sim 1.5\times$ more full-length products, suggesting that efficient i-motif-unfolding by PCBPI facilitates replication (Fig. 4B–D). Conversely, *BCL2* exhibited three stalling/pausing signals at C2, G6, and G10 (Fig. 4B, F–H), with fewer full-length products compared to *cMYC*. Although PCBPI slightly

increased the full-length product in *BCL2*, it was less efficient at unfolding *BCL2*-i-motif (Fig. 4F). While the pausing preceding C2 attributes to i-motif formation, we investigated whether other pausing signals were due to hairpin formations that inhibit PCBPI-driven unfolding. We designed two *BCL2* mutants: G10TG12T and $G_{all} \rightarrow T$ (all Gs mutated to Ts) that disrupt hairpin formation. ¹H NMR confirmed no Watson-Crick-associated proton resonances at 12–14 ppm in these mutants (Fig. 4E). Primer extension assays showed reduced G10-associated stalling in both mutants and $\sim 1.5\times$ more full-length products with PCBPI, indicating that G10 contributes to hairpin formation and acts as a structural barrier to PCBPI-driven unfolding. Nevertheless, stalling at G6, before the second C-tract, was unaffected and may represent another i-motif because *BCL2*-i-motifs form dynamic and heterogeneous conformers involving different C-tract combinations, similar to polymorphic G4s (Fig. 4F–H)^{24,42,52}. These findings were validated by CD, indicating stronger blue-shifts and hypochromism in *BCL2* mutants upon PCBPI titrations than wild-type sequence (Figs. 3B and 4I). Therefore, PCBPI-driven unfolding depends on i-motifs' protonation states, while sequences with co-existing hairpins significantly hinder the unfolding (Fig. 4J).

PCBPI depletes i-motif formation at specific loci during the cell cycle

Functional enrichment analysis of a recent proteomic study linked candidate i-motif-binding proteins to cell cycle regulation, which complemented previous observations of cell cycle-specific spatio-temporal changes in global i-motif foci^{8,10,19}. PCBPI is known from earlier studies to control cell cycle^{31,35–37,53,54}, and is now shown by us to regulate i-motif dynamics within cells. To explore this connection, we examined the impact of PCBPI-KD on cell cycle progression. Consistent with previous reports³⁷, PCBPI-KD caused G₁/S arrest (Fig. 5A, B). This was substantiated by dot-blot assays, which showed minimal BrdU (5-bromo-2'-deoxyuridine) incorporation even 4 h after the synchronized G₁/S cells were released upon PCBPI-KD, indicating a failure to initiate replication (Fig. 5C). In parallel, PCBPI-KD cells showed elevated γ H2AX levels (Fig. 5D), suggesting DNA damage accumulation, which potentially contributed to G₁/S arrest as a protective checkpoint response, allowing time for DNA repair before S phase entry. To determine whether PCBPI's roles in cell cycle control and i-motif regulation are linked, we synchronized cells using a double-thymidine block at G₁/S or S phase (Supplementary Fig. 20A, B) and analyzed PCBPI occupancy at three oncogenic i-motif loci by PCBPI-ChIP-qPCR. PCBPI occupancy was significantly higher at *cMYC* and *ILPR* in G₁/S, and decreased in S phase compared to asynchronous cells, while its occupancy at *BCL2* remained unchanged (Fig. 5E, F). To examine whether i-motif formation at these loci followed a similar pattern, we performed iMab-ChIP-qPCR in synchronized cells. Except for *BCL2*, iMab occupancy was higher at *cMYC* and *ILPR* during G₁/S, and decreased in early S phase (Fig. 5G). Since i-motif formation influences G4 dynamics in the complementary strands⁵⁵, we also examined BG4 occupancy to detect G4 formation⁵⁶, revealing pronounced BG4 occupancy at all three promoters compared to non-G4 control sites during both G₁/S and S phases. However, at *cMYC*, BG4 occupancy was significantly increased during S phase, compared to G₁/S transition (Supplementary Fig. 21A), aligning with earlier reports on the cell cycle-dependent dynamics of G4 structures^{56,57}. In PCBPI-KD cells, BG4 occupancy at *cMYC* was significantly reduced, while *BCL2* and *ILPR* remained unchanged (Supplementary Fig. 21B). This likely reflects the reduced proportion of S phase cells upon PCBPI-KD (Fig. 5A, B), rather than a direct regulatory role of PCBPI on G4s, considering its weak *in vitro* binding to G4s (Supplementary Fig. 7). In contrast, the temporal dynamics of i-motif formation at *cMYC* and *ILPR* loci during the G₁/S and S phases closely mirrored PCBPI occupancy, suggesting a coordinated regulation. To further assess the functional consequences of PCBPI-mediated resolution of specific i-motifs, we

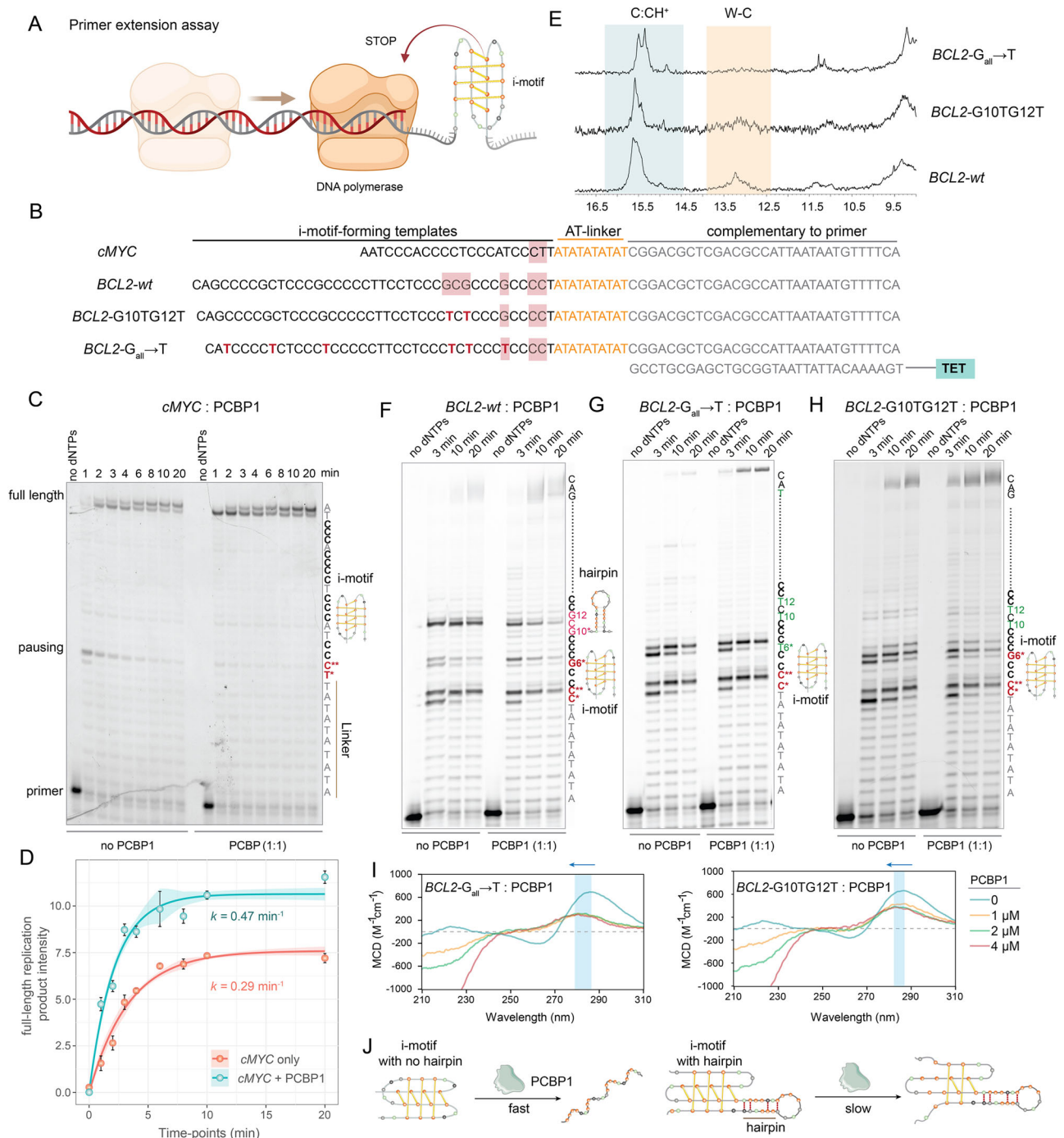


Fig. 4 | I-motifs co-existing with hairpin structures resist PCBP1-mediated unfolding. **A** Schematic of primer extension assay on i-motif-forming templates, where i-motif formation sterically blocks or slows DNA polymerase progression. Created in BioRender. Sabouri, N. (2026) <https://BioRender.com/vwd8yby>. **B** Primer extension templates containing i-motif-forming sequences, AT-linker, and TET (tetrachlorofluorescein)-labeled primer complementary sequences. G-to-T mutations are marked in red. Pausing signals are highlighted in pink boxes. **C** Primer extension assay with *cMYC*-i-motif-forming template at various time points until 20 min. AT-linker preceding *cMYC*-i-motif and sequence provided with pausing signals (red, asterisk), full-length products, and primer. **D** Densitometric analyses of full-length replication product formation over time in the presence of PCBP1; results expressed as means \pm SD from three biological replicates ($n = 3$), fitted to single-exponential function equations to calculate the rate of reaction (k). Error bars/bands denote standard deviation. **E** Imino proton NMR resonances (15–16 ppm) and Watson-Crick (W-C) pairing (12–14 ppm) of wild-type and mutated

BCL2-i-motifs. **F** Primer extension assay with wild-type *BCL2*-i-motif-forming template (*BCL2*-wt) at different time points. AT-linker preceding *BCL2*-wt-i-motif, the template sequence with three pausing signals highlighted with red asterisks; full-length product and primer are shown with potential involvement of G10 and G12 in hairpin formation that blocks PCBP1. Primer extension assays involving **F** *BCL2*-wt, **G** *BCL2*-G_{all} \rightarrow T, or **H** *BCL2*-G10TG12T i-motif-forming templates at different time points. Gels shown in (**F–H**) are representative of three independent biological replicates. Mutations at specific Gs (marked in green) to prevent hairpin formation; pause signals for i-motif formation marked in red and asterisks. **I** CD titrations with increasing PCBP1 concentrations on *BCL2*-G_{all} \rightarrow T and *BCL2*-G10TG12T; blue-shifts and hypochromic effects indicated by blue bars and arrows. **J** Schematic illustration showing hairpin-forming potential of i-motif significantly slows down PCBP1’s ability to unfold i-motifs. Uncropped and unprocessed gel images corresponding to primer extension assays shown in (**C, F, G, H**) are provided in the Source data.

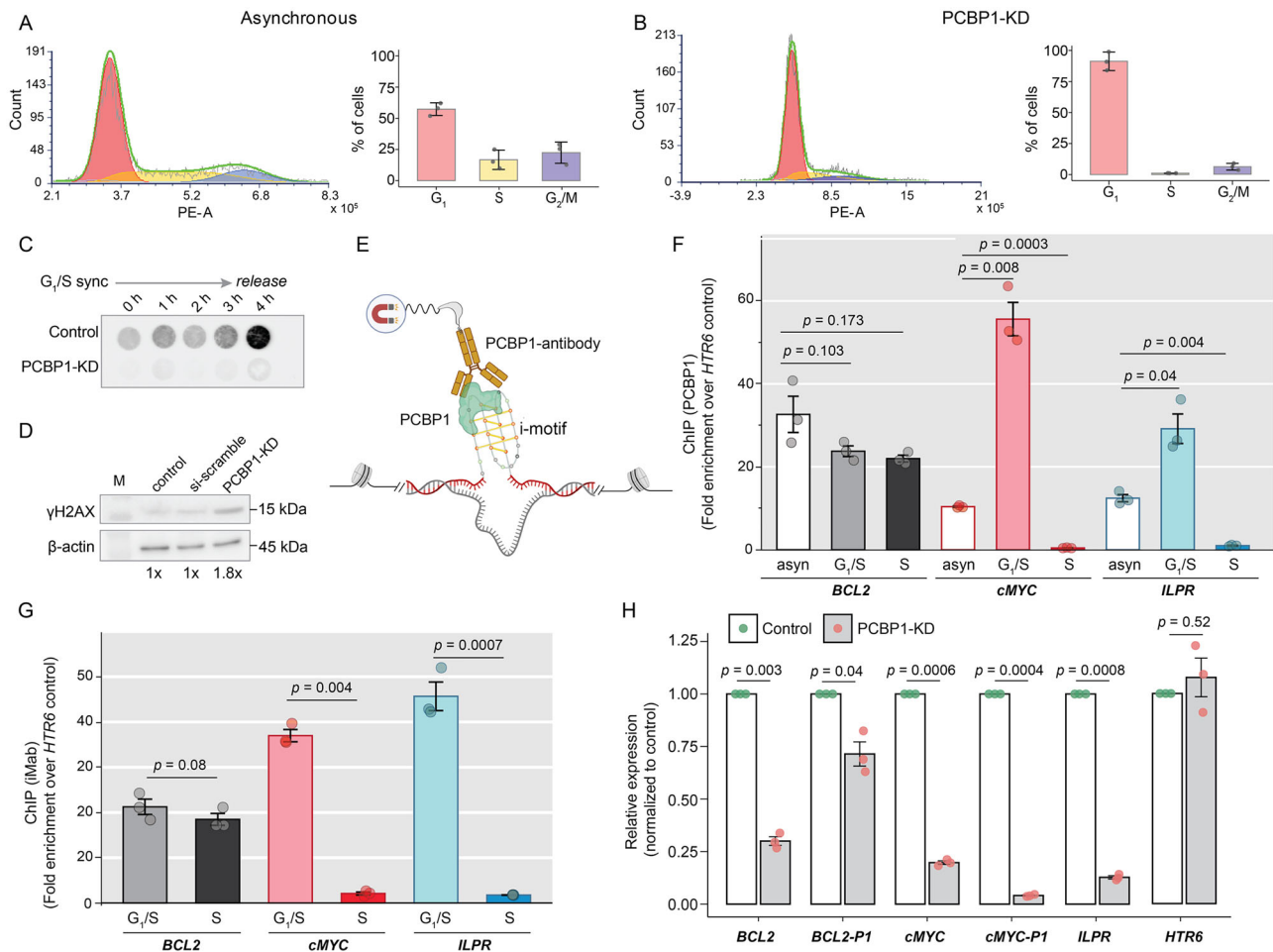


Fig. 5 | PCBPI regulates i-motif folding in HeLa cells in a cell cycle-dependent manner. Flow cytometry-based analysis of DNA content and cell cycle distribution in **A** asynchronous and **B** PCBPI-KD-treated HeLa cells; bar plots showing cell distribution in G_1 , S, and G_2/M phases. PE-A Phycoerythrin–Area. Results expressed as means \pm SD from three biological replicates ($n = 3$). **C** Anti-BrdU dot-blot assay on G_1/S -synchronized genomic DNA, followed by synchrony release at different time points up to 4 h in control and PCBPI-KD-treated HeLa cells. **D** Western blot of γ H2AX upon 48 h of PCBPI-KD. Si-scrambled siRNA as negative control; β -actin as housekeeping protein. Fold-change increase in γ H2AX below the blot. Blots shown are representative of three independent biological replicates. Uncropped and unprocessed images corresponding to dot-blot and western blot experiments shown in (**C**, **D**) are provided in the Source data. **E** Schematic illustration of PCBPI-ChIP using PCBPI-specific antibody. Created in BioRender. Sabouri, N. (2026) <https://BioRender.com/uvavqla>. **F** Fold enrichment of PCBPI-ChIP at i-motif

containing promoters compared to *HTR6* control site in asynchronous (asyn), G_1/S - and S-synchronized HeLa cells. **G** Fold enrichment of iMab-ChIP at i-motif-harboring promoters (*cMYC*, *BCL2*, *ILPR*) over non-i-motif control site (*HTR6*) in asynchronous, G_1/S - and S-synchronized HeLa cells. ChIP bar plots and error bars in (**F**, **G**) represented as means \pm SD of three technical replicates ($n = 3$). Statistical differences and p values determined by one-way ANOVA followed by Tukey–Kramer post hoc test, where p values < 0.05 are considered statistically significant. Raw C_i values from input, IP, and mock (IgG) samples obtained from three independent biological replicates, each measured with three technical replicates, are provided in Supplementary Data 3 and 4. **H** Relative expression ($2^{-\Delta\Delta C_t}$) in i-motif-containing and non-i-motif-containing gene expression upon PCBPI-KD using RT-PCR. Data and error bars represent the mean \pm SD from three independent biological replicates ($n = 3$). P values calculated using two-tailed Student's t tests, where p values < 0.05 are considered statistically significant.

examined their transcript levels upon PCBPI-KD. Because promoter activity at *cMYC* and *BCL2* is influenced by G_4/i -motif structural transitions at their P1 promoters^{58,59}, we specifically quantified P1-derived transcripts by RT-qPCR. We observed reduced transcript levels in all i-motif-containing genes upon PCBPI-KD, despite PCBPI's selectivity toward i-motif unfolding (Fig. 5H). This likely reflects a combination of direct effects, stemming from impaired i-motif resolution at these promoters, and indirect consequences of PCBPI-KD, given PCBPI's multifunctional nature and the involvement of additional regulatory proteins at these loci (e.g., Nucleolin, hnRNP-K, hnRNP-LL)^{21,22,60}. Collectively, our findings indicate that PCBPI facilitates unfolding of specific i-motifs in a cell cycle-dependent manner and that disruption of this regulation not only alters the temporal dynamics of i-motif formation but also impacts transcriptional output from these loci, likely through a combination of direct structural effects and broader PCBPI-dependent regulatory pathways.

Role of PCBPI-KH domains in i-motif-binding and unfolding in vitro

Although we show that PCBPI unfolds i-motifs with varied kinetics, its molecular mechanism remains unclear. To address this, we truncated PCBPI into KH domains—critical for DNA binding and examined their contributions in i-motif selectivity and unfolding^{36,61,62}. Previous research on hnRNP-K (a PCBPI-family protein) demonstrated that individual KH domains unfold *cMYC*-i-motif, albeit less efficiently than the full-length protein²¹. Considering the distinct nucleic acid binding behaviors and amino acid compositions of PCBP proteins (Supplementary Figs. 29 and 30A, B), we purified PCBPI's KH1+2 and KH3 domains (Fig. 6A, Supplementary Fig. 22A–C). Unfortunately, due to precipitation, likely caused by a long unstructured region between KH2 and KH3, KH2+KH3 purification was unsuccessful.

We used thermal-shift assays (Supplementary Fig. 23A, B, Supplementary Table 5) to assess the pH tolerance of KH mutants,

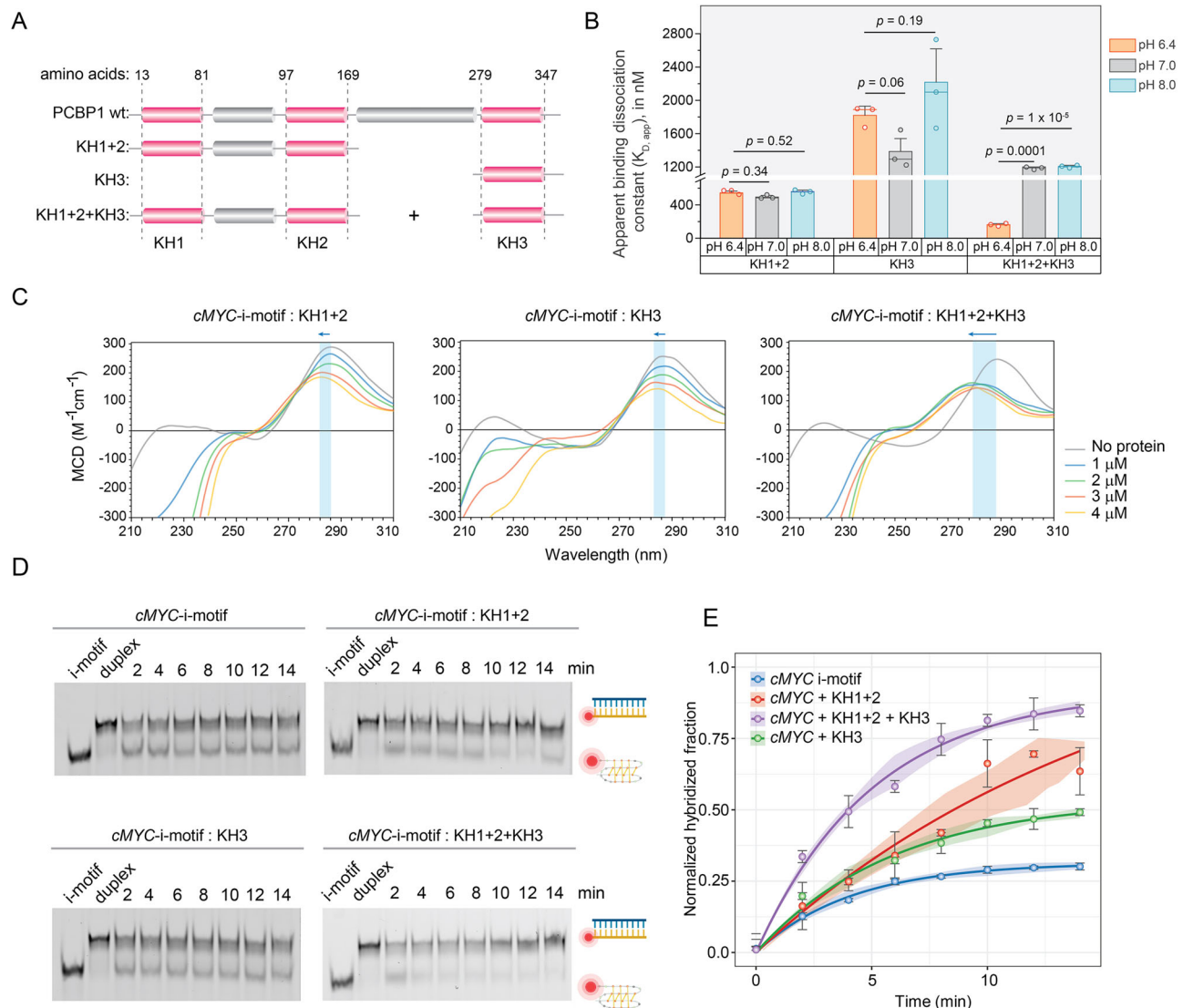


Fig. 6 | Role of KH domains in i-motif binding and unfolding in vitro.

A Schematic representation of purified wild-type PCBP1 and KH mutants (KH1 + 2 and KH3). KH domains are marked in pink, and the unstructured interconnecting domains are marked in gray. **B** $K_{D,app}$ from MST binding curves between KH mutants and i-motifs at pH 8.0, 7.0, and 6.4; data represent the mean \pm SD from three independent biological replicates ($n = 3$). Error bars indicate standard deviation. Statistical significance and p values were calculated using a two-tailed Student's t -test, where a p value < 0.05 indicates statistically significant results. **C** CD spectra of *cMYC*-i-motifs upon titrations of KH1 + 2, KH3, and a combination of KH1 + 2 and KH3, with highlighted regions showing blue-shift (with arrows) and

hypochromism. **D** Native PAGE visualizing hybridization reaction aliquots of *cMYC*-i-motifs in the presence of different combinations of KH mutants at different time intervals up to 14 min, alongside negative control (i-motifs without complementary strands) and positive controls (hybridized duplex). Uncropped and unprocessed gel images corresponding to native PAGE experiments are provided in the Source data. Created in BioRender. Sabouri, N. (2026) <https://BioRender.com/oazozsm>. **E** Densitometric analysis of i-motif-trapping. Normalized intensities of hybridized fractions calculated from three independent biological replicates ($n = 3$) at different time points, fitted to single-exponential function equations. Error bars indicate standard deviation.

confirming the feasibility of conducting MST assays between KH mutants and *cMYC*-i-motif at pH 6.4, 7.0, and 8.0 (Supplementary Fig. 24). KH1 + 2 domain bound both i-motifs and their unfolded forms showing lack of structural selectivity, while KH3 showed minimal binding and no preference to either of them (Fig. 6B). KH3 failed to unfold *cMYC*-i-motif, as evident from the negligible blue-shifts in the i-motif CD spectra upon KH3 titrations (Fig. 6C) and the lack of hybridized product accumulation in the i-motif-destabilization trapping assays, similar to the no protein control (Fig. 6D, E). KH1 + 2 also caused a negligible blue-shift in i-motif CD spectra; however, we detected more accumulation of hybridized products starting at 10 min in i-motif-destabilization trapping assays compared to control (Fig. 6C–E), suggesting that KH1 + 2 cannot unfold i-motif as efficiently as PCBP1, but upon interaction, destabilizes i-motif, causing delayed

unfolding. Strikingly, combining KH1 + 2 and KH3 restored full-length PCBP1 function, demonstrating enhanced affinity and selectivity for i-motifs in MST assays (Fig. 6B). This reconstitution also induced significant blue-shifts and hypochromism in CD spectra (Fig. 6C) and increased the accumulation of hybridized products, mirroring the unfolding efficiency like full-length PCBP1 (Fig. 6D, E). These findings highlight that all three KH domains orchestrate selective recognition and efficient unfolding of i-motifs.

Molecular basis of i-motif unfolding by coordination of KH domains

We proposed two models to elucidate the molecular coordination of KH domains in i-motif regulation. The first model suggested that KH1 + 2 and KH3 form a complex to selectively bind and unfold

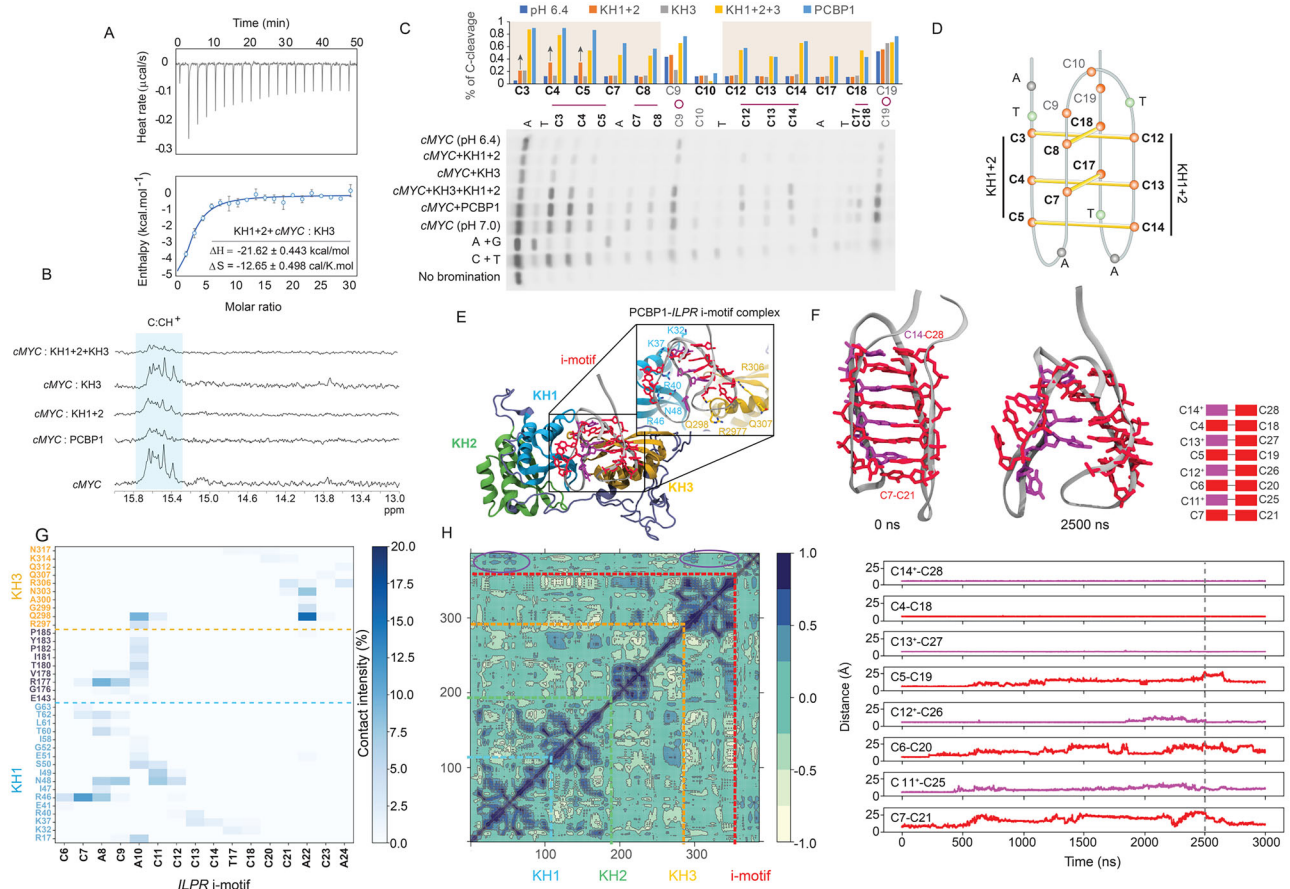


Fig. 7 | Molecular mechanism of i-motif resolution by coordination of KH domains. **A** ITC interaction between a 1:1 mix of KHI + 2 and *cMYC*-i-motif in the ITC cell, and KH3 in the syringe at pH 6.4. Top: enthalpic heat released versus time at 25 °C during titrations. Bottom: thermogram of integrated peak intensities against molar ratio. ΔH (enthalpy) and ΔS (entropy) are provided. Data and error bars calculated from mean \pm SD from three independent biological replicates ($n = 3$). **B** ^1H NMR spectra showing imino proton resonances (15–16 ppm) of *cMYC*-i-motif during PCBPI and KH mutants' titrations. **C** Bromine footprinting gel of *cMYC*-i-motif at pH 6.4 and 7.0 in 1:1-bound complexes of PCBPI, KHI + 2, KH3, and KHI + 2 + KH3 at pH 6.4. A + G, C + T ladders, and unreacted *cMYC*-i-motif, i-motif sequence are shown; Cs within C:CH⁺ and i-motif-core-connecting loops marked with line and round shapes, respectively. Bar plot shows the % of C-cleavage by piperidine upon bromination in the i-motif. An uncropped and unprocessed gel image corresponding to bromine footprinting experiments is provided in the Source data. **D** Schematic of *cMYC*-i-motif at pH 6.4, based on bromine footprinting; C:CH⁺ base pairs in first C-tract involving C3–C5 and C12–C14 are destabilized

upon KHI + 2 interactions. **E** In silico modeled complex of PCBPI-bound partially hemi-protonated *ILPR*-i-motif with the following modifications: G-to-A at G2, G16, and G30. C-bases in red and violet are unprotonated and protonated, respectively. KHI-3 domains in PCBPI are in green, sky blue, and yellow, respectively. Potential amino acids in the interactions within the KHI and KH3 domains are labeled. **F** During simulation, starting *ILPR*-i-motif structure at 0 ns and i-motif at 2500 ns of simulation are shown. Inter-residual distances between Cs in *ILPR*-i-motif base-pairing throughout simulation (b)–l. **G** Binding mode analysis of PCBPI with *ILPR*-i-motif throughout 2500 ns simulations with individual contribution of residues within KHI, KH3, and the unstructured region between KH2 and KH3. The darker the blue color of contact intensity (%), the more the frequency of intermolecular interactions. **H** Dynamical Cross-correlation Matrix between the atomic displacements considering C α for protein residues and phosphorus for i-motif residues during simulation (b)–l. Marked violet circles highlight correlated motions between protein domains and part of the i-motif.

i-motifs. The second model proposed that KHI + 2 initiates remodeling, facilitating KH3 recruitment for subsequent unfolding. To test the first model, we performed ITC to assess binding thermodynamics. No significant interaction was detected between KHI + 2 and KH3 in isolation (Supplementary Fig. 25). PCBPI–i-motif interaction was highly exothermic, driven by enthalpic contributions (Supplementary Fig. 26A). KHI + 2 alone also bound i-motif, displaying a negative ΔH (enthalpy) exceeding ΔS (entropy). However, full-length PCBPI exhibited stronger ΔH , suggesting additional interactions, potentially involving KH3 (Supplementary Fig. 26A, C). KH3 alone did not bind i-motif, indicating that a prior remodeling by KHI + 2 was required (Supplementary Fig. 26B). When KHI + 2 was pre-incubated with *cMYC*-i-motif before KH3 titration, ΔH closely matched that of full-length PCBPI, supporting a sequential remodeling and recruitment mechanism, ruling out the first model (Fig. 7A). To validate further, we performed ^1H NMR titrations of KH mutants upon *cMYC*-i-motif,

individually, in combination, and with full-length PCBPI. Consistent with ITC, KHI + 2 induced significant proton resonance shifts and broadening, indicative of structural changes in *cMYC*-i-motif upon binding. Contrarily, KH3 induced minimal spectral changes, reinforcing its lack of direct binding. However, a combination of KHI + 2 and KH3 generated pronounced spectral broadening comparable to full-length PCBPI (Fig. 7B), suggesting cooperative i-motif disruption by KH domains. Bromine footprinting further mapped destabilized C:CH⁺ pairs during protein interactions. KHI + 2 increased cleavage at C3–C5 and C12–C14 within the first and third C-tracts, while KH3 alone had no significant effect. Their combination enhanced cleavage across all Cs, mirroring the unfolding pattern observed with full-length PCBPI and the unfolded form at pH 7.0 (Fig. 7C). These suggest that KHI + 2 destabilizes specific C:CH⁺ pairs in the first and third C-tracts (Fig. 7D), facilitating KH3 recruitment, which ultimately unfolds i-motif.

To illustrate the molecular basis of PCBP1–i-motif interactions, we modeled five PCBP1–i-motif complexes using *ILPR*-i-motif crystal structure⁴⁰ as a template (Fig. 7E, Supplementary Fig. 27A) because of its demonstrated ability to bind PCBP1 with moderate unfolding (Fig. 2D, E). However, we reverted three G-to-A mutations (G2, G16, G30) to mimic wild-type *ILPR*-i-motif and performed microsecond-scale MD simulations. To assess the impact of i-motif protonation on PCBP1-mediated unfolding, we modeled each complex under three conditions: (a) fully unprotonated Cs, (b) four protonated C:CH⁺ pairs (C11–C14), and (c) eight protonated C:CH⁺ pairs (C11–C14 and C18–C21). Condition (b) resembled our in vitro experiments at pH 6.4, less acidic and protonated than the published i-motif crystal structure (–pH 5.5)⁴⁰. The simulations revealed that PCBP1 interacts with i-motif in distinct binding modes depending on its protonation. Across all simulations, the DNA-protein complexes maintained relatively stable conformations, as evidenced by the root mean square deviation (RMSD) profiles of PCBP1 and i-motif (Supplementary Fig. 27B, C). C:CH⁺ pairs exhibited greater stability and lower flexibility than C:C pairs, as reflected in the stable RMSD values (–3 Å) observed for (c) (Supplementary Fig. 27D). In condition (a), spontaneous i-motif destabilization was observed in structures 2 and 3 (Supplementary Fig. 28A–D, Supplementary Movies 1 and 2). In (b), the i-motif and its C:CH⁺ pairs remained largely intact, having minor rearrangements. However, in structure 1, an unfolding event was detected at C7–C21 pair propagating in a zipper-like manner (Fig. 7F), as characterized by progressive inter-base distance increases, with C:C pairs unfolding earlier than C:CH⁺ pairs (C6–C20 at –250 ns, C11–C25 at –400 ns, C5–C19 at –500 ns, and C12–C26 at –1800 ns) (Fig. 7F, Supplementary Movie 3). In contrast, no unfolding was observed in condition (c), reinforcing that higher protonation stabilizes i-motif, delaying PCBP1-mediated unfolding (Supplementary Movie 4). We further examined KH-domain contributions to PCBP1–*ILPR*-i-motif interactions. I-motif primarily contacted KH1 and KH3 via A8–C9–A10 and A22–C23–A24 loops, respectively (Fig. 7G), which acted as initiation points for unfolding. These loop dynamics correlated with key residues from KH1 and KH3 (Fig. 7H). We mapped frequently observed amino acids to KH1, KH3, and the unstructured region between KH2 and KH3, implicating these residues as being involved in i-motif unfolding. Consistent with in vitro findings, these simulations highlight the cooperative role of KH1 and KH3 in i-motif remodeling and unfolding, while underscoring the influence of protonation in regulating i-motif unfolding.

Discussion

Intracellular regulation of i-motifs is controversial due to their transient nature and sensitivity to non-physiological pH, which complicates efforts to define their protein-mediated interactions and dynamics. Although PCBP-family proteins showed promising in vitro results^{21,22,24,26,63}, their ability to bind i-motifs has been questioned, likely because they have restricted access to hydrogen-bonding within the i-motif core^{20,24,62}. To address this, we systematically examined PCBP1 interactions across diverse i-motifs with varying stability and C-content across different pH, which revealed PCBP1's selective binding to i-motifs, with a distinct preference for highly C-rich sequences. Longer C-tracts strengthened PCBP1's affinity for unfolded forms, compared to folded i-motifs with shorter C-tracts. We also demonstrate that PCBP1 retains appreciable affinity for unstructured C-rich DNA, indicating a degree of binding plasticity that may be relevant under physiological conditions where i-motif stability fluctuates. Nonetheless, quantitative and competitive analyses demonstrate a consistent energetic and competitive advantage for folded i-motifs, suggesting that i-motif stability and C-density together modulate PCBP1 binding. Our findings expand previous studies that focused on individual iMfs^{20,22,23}, by demonstrating how sequence-context, structure, and protonation collectively influence PCBP1's selectivity. To place our findings in a broader biological framework, we compared

PCBP1 with 19 hnRNP proteins, recently identified as putative i-motif interactors in a proteomics screen¹⁹. Many hnRNPs share roles in mRNA binding, recognition of poly-CT or poly-AG elements, underscoring the functional diversity within this family (Supplementary Fig. 29). Our findings reveal that PCBP1 unfolds different i-motifs with varied kinetics, consistent with previous reports of its co-localization with i-motif foci in cells²⁶—an unlikely observation if PCBP1 indiscriminately resolved all i-motifs. This challenges the view that a single protein universally unfolds all i-motifs, and instead supports a selective regulatory mechanism.

In comparison with other hnRNP proteins, such as hnRNP-K or hnRNP-LL, which both preferentially interact with C-rich sequences, PCBP1 shows a stronger preference for folded i-motifs. Specifically, hnRNP-K can both stabilize⁶⁴ and unfold i-motifs, suggesting its sequence-specific functions, whereas hnRNP-LL unfolds i-motif structures²² although its mechanistic determinants remain poorly defined. By contrast, PCBP1 recognizes residues in the lateral loops of folded i-motifs, promoting their destabilization and unfolding (Fig. 7E, F). PCBP1 remains stable and active across a broad pH range, including a mildly acidic environment that favors i-motif formation, supporting its robust adaptability. Further, hnRNP-K and -LL have mainly been studied in relation to *cMYC* and *BCL2* i-motifs, respectively; therefore, their broader sequence preferences are so far limited.

We also demonstrated that PCBP1-mediated i-motif unfolding is modulated by i-motif protonation status and hairpin-forming propensity (Fig. 8), an aspect not well studied for hnRNP-K or -LL. These differences likely arise from distinct domain architectures. Unlike hnRNP-LL, which utilizes RRM (RNA recognition motif) domains to interact with C-rich sequences⁶⁵, PCBP1 and hnRNP-K rely exclusively on KH domains for nucleic acid binding^{66,67} (Supplementary Fig. 30A). Structural alignment showed that both proteins share conserved KH folds, yet differ in amino acid composition and domain organization (Supplementary Fig. 30A, B). Specifically, the KH2–KH3 linker is highly flexible, enabling variable KH3 orientations in PCBP1 and hnRNP-K. This is also supported by MD simulations showing KH3's dynamic positioning. Thus, despite conservation within individual KH folds, differences in their domain flexibility, sequence, and interdomain arrangement may underlie distinct mechanisms of i-motif interaction and unfolding by PCBP1 and hnRNP-K (Supplementary Fig. 30C, D). Unlike hnRNP-K, whose individual KH domains independently bind and unfold *cMYC* i-motif²¹, PCBP1 requires coordinated action of all three KH domains: KH1 + 2 initiate docking and local rearrangement, allowing KH3 to complete unfolding. Our MD simulations complemented the in vitro results by supporting the dependence of PCBP1-driven unfolding on i-motif protonation. Analysis of interaction frequencies and persistence generated a residue-level interaction heatmap, identifying candidate amino acid residues for future mutagenesis to define their roles in i-motif selectivity and regulation.

A significant challenge in i-motif research is the discrepancy between in vitro and in-cell observations regarding i-motif formation²⁴. Sequencing across different cell lines showed considerable variation in genomic i-motif distribution, possibly due to experimental differences and potential non-specific interactions of iMab⁴⁴. To address this, we established an iMab-CHIP workflow using rigorously selected i-motif-forming and moderately C-rich non-i-motif-forming genomic regions, inspired by prior BG4 and iMab-based ChIP-seq and CUT-and-Tag studies^{15,45}. Using this approach, we investigated i-motif dynamics in specific loci following PCBP1-KD in HeLa cells. Only a subset of loci—those that behaved consistently with in vitro properties exhibited enhanced i-motif levels, despite robust PCBP1 occupancy across all regions. However, PCBP1-KD downregulated gene expression of all examined i-motif-containing genes, suggesting that their transcriptional outputs reflect a combination of structural effects and broader PCBP1-dependent regulatory pathways.

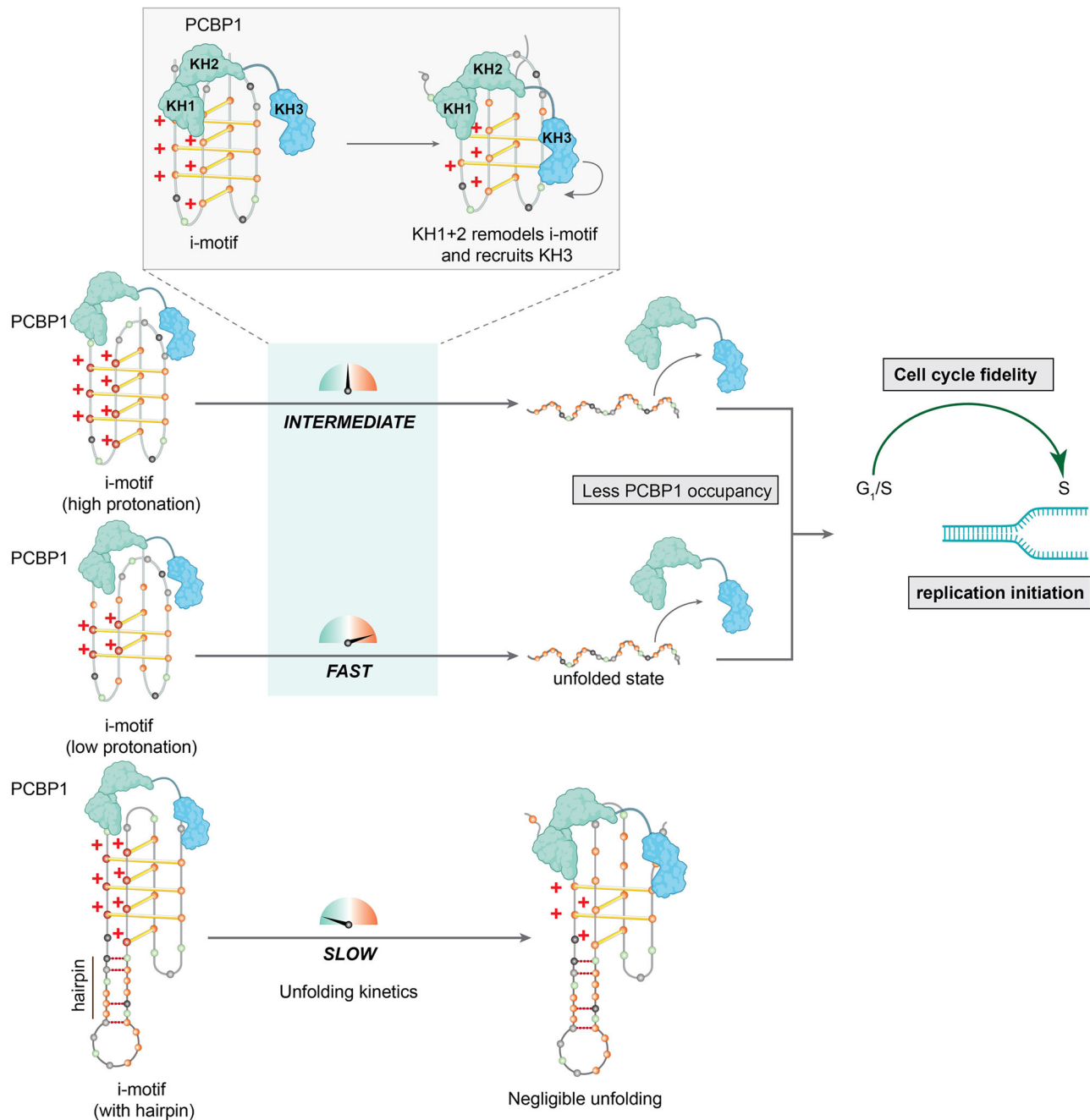


Fig. 8 | Schematic model of PCBP1-mediated unfolding of selective i-motifs during the cell cycle. PCBP1 binds i-motif DNA with higher affinity than its corresponding single-stranded forms. Upon binding, PCBP1 promotes i-motif unfolding in a manner that depends on both the protonation state of i-motifs and their propensity to form hairpin conformations. I-motifs that exist in equilibrium with hairpin structures remain accessible to PCBP1; however, stable hairpin formation markedly suppresses PCBP1-mediated unfolding. Protonation further modulates the kinetics of this process, with higher protonation states slowing unfolding, whereas lower protonation accelerates the unfolding. Although PCBP1 can associate with single-stranded forms, these interactions are weaker, and i-motif unfolding

consequently reduces PCBP1 occupancy at these genomic sites, thereby permitting replication initiation. PCBP1 contains three KH domains—KH1, KH2, and KH3—that act through a hierarchical two-step mechanism. Initial engagement of the i-motif by KH1–KH2 domains induces local structural destabilization, which subsequently recruits the KH3 domain to drive complete unfolding. This PCBP1-mediated spatiotemporal dynamics of i-motifs is proposed to support genomic integrity and cell-cycle progression by facilitating replication initiation and enabling the G₁/S transition into S phase. Created in BioRender. Sabouri, N. (2026) <https://BioRender.com/b3vfrkt>.

Variability in cellular i-motif detection may also stem from asynchronous cell populations, where cell cycle heterogeneity introduces differences in chromatin state and transcriptional activity. Since i-motif formation is cell cycle-regulated, this heterogeneity may obscure locus-specific dynamics. By synchronizing cells at G₁/S and S phases, we captured precise temporal folding patterns in i-motifs while preserving cell cycle phase-dependent changes. PCBP1-regulated i-

motifs peaked transiently at G₁/S and resolved in early S phase, coinciding with increased PCBP1 occupancy at the same loci at G₁/S. These findings highlight the physiological relevance of PCBP1 interactions with both folded and unfolded i-motifs. Because PCBP1 binds folded i-motifs with high affinity, and the PCBP1–i-motif complex remains stable even in the presence of C-rich RNA, transient increases in i-motif abundance at G₁/S likely bias PCBP1 toward binding folded structures.

On the contrary, PCBPI–ssDNA complexes are more susceptible to competition with alternative substrates. As i-motifs are generally rare under physiological conditions²⁴, this narrow temporal window at G₁/S provides an opportunity for PCBPI to preferentially associate with i-motifs over ssDNA. Thus, PCBPI acts as a critical regulator of specific i-motifs contributing to the fidelity of G₁/S transition (Fig. 8). PCBPI-KD caused G₁/S arrest and enhanced γ H2AX levels, thereby preventing DNA replication. In vitro, PCBPI facilitated replication on i-motif-forming templates, which otherwise hinder polymerase progression. These observations align with previous reports showing reduced replication efficiency upon failure to resolve i-motif structures and increased γ H2AX accumulation upon PCBPI-KD²⁶. So far, there is no direct evidence that other hnRNP proteins, such as hnRNP-K regulates i-motif dynamics during G₁/S, although it is reported to influence cell cycle progression through distinct mechanisms, such as stabilizing *CDK6* transcripts⁶⁸, or transactivating cell cycle arrest genes (*I4-3-3 α* , *GADD45*, and *CDKN1A/p21*)⁶⁹.

We simultaneously investigated G4 dynamics upon PCBPI-KD using BG4-ChIP at the same loci because i-motif formation influences G4 stability on complementary strands and PCBPI has been implicated in recruiting G4 helicase-DHX9 to G4-forming regions²⁷. We observed significant changes only at *cMYC*, while *BCL2* and *ILPR* remained unaffected. G4 structures at all loci appeared more robust than their i-motif counterparts, aligning with in-cell NMR findings that i-motifs are highly transient in cells²⁴. Under synchronized conditions, only *cMYC* showed expected G4 fluctuations—reduced at G₁/S and increased in S phase—whereas *BCL2* and *ILPR* did not change. While our iMab-ChIP partially aligned with expected trends for i-motif dynamics upon PCBPI-KD at *cMYC* and *ILPR*, our BG4-ChIP results revealed reduced G4 levels at *cMYC*, contrary to recent reports suggesting PCBPI-KD enhances G4 abundance by impairing DHX9 recruitment²⁷. These results suggest that PCBPI's influence on G4 dynamics may be indirect, potentially mediated by its impacts on i-motifs or broader role in cell cycle regulation. Thus, our findings challenge simple models proposing universal reciprocal regulation between i-motifs and G4s or generalizable dynamics of these structures across loci during the cell cycle. In fact, they highlight the need for caution when interpreting cellular i-motif or G4 dynamics, which may vary by locus, cell type, and physiological context. While our cellular experiments were performed in HeLa cells, which are widely used in mechanistic studies, future validation in other mammalian cell types will be important to determine the extent of conserved versus context-specific regulation. This raises an open question: while G4s and i-motifs may influence each other at shared sites, do they always function in opposition, or can they co-exist or form independently under different conditions?

In conclusion, our study highlights the importance of investigating specific i-motif loci and their spatiotemporal dynamics in relation to protein interactions. PCBPI-mediated i-motif resolution appears critical during G₁/S transition. These insights provide a foundation for further research into PCBPI's roles in maintaining cell cycle fidelity through i-motif unfolding.

Methods

Electrophoretic mobility shift assay (EMSA)

PCBPI binding to folded and unfolded i-motif structures was assessed by EMSA. 5 nM 5'-Cy5-labeled i-motif-forming sequences (iMfs), (Eurofins Genomics Europe; Supplementary Table S1) were incubated with increasing concentrations of purified recombinant PCBPI in a total volume of 50 μ L. Reactions were performed in assay buffers containing either 20 mM Bis-Tris (pH 6.4) (Zellbio B-020-250) or 20 mM Tris-HCl (pH 8.0) (Sigma T1503), supplemented with 100 mM NaCl and 0.25 mg/mL bovine serum albumin (BSA) (Merck A9647). Samples were incubated at 25 °C for 30 min, after which 10 μ L of 80% glycerol was added, and 20 μ L of each reaction mixture was resolved

on 8% native PAGE (Polyacrylamide gel electrophoresis) in 50 mM Bis-Tris (pH 6.4) or 50 mM Tris-HCl (pH 8.0), supplemented with 100 mM NaCl, matching the respective assay buffer conditions and pH. Electrophoresis was conducted at 70 V at 6 °C, using a 1 \times running buffer of identical composition. Gels were imaged using an Amersham Typhoon scanner (Cy5 laser; 300 V; 25 μ m resolution) and processed in ImageJ. All experiments were performed in three biological replicates.

Thermal-shift assay

pH tolerance of purified PCBPI and KH mutants (KH1 + 2 and KH3) was assessed using a Sypro-orange-based protein thermal-shift dye kit (Applied biosystems; 4461146). Protein samples (10 μ M; 20 μ L) were prepared in 20 mM MES buffer (pH 5.5, 6.0, or 6.4) or 20 mM Tris-HCl buffer (pH 7.0 or 8.0), each supplemented with 100 mM NaCl, and loaded into flat-bottom 96-well plates (Bio-Rad). Thermal unfolding was monitored using a CFX96 real-time PCR system (Bio-Rad) by heating from 25 to 95 °C at 0.2 °C increments while recording SYPRO Orange fluorescence. The melting temperatures (T_m) (Supplementary Tables S3 and S5) were determined from the maximum of the negative first derivative of fluorescence with respect to temperature ($-dF/dT$), corresponding to the midpoint of the unfolding transition. Proteins prepared in storage buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl) were used as reference controls. All measurements were performed in three biological replicates.

Microscale thermophoresis (MST)

We performed MST to study the interaction between PCBPI/KH mutants and 5'-Cy5-labeled iMfs and other oligonucleotides (Supplementary Table S1) and calculate the apparent binding affinities ($K_{D,app}$). 40 nM 5'-Cy5-labeled oligos were incubated in 20 mM MES (pH 6.4) or phosphate buffer (pH 7.0) or Tris-Cl (pH 8.0) with 100 mM NaCl and varying concentrations of PCBPI or KH mutants (35 nM to 2.5 μ M) in 20 μ L reactions at 25 °C for 15 min. Then, samples were centrifuged at 11,000 \times g for 10 min at 25 °C, and loaded into Monolith NT Capillaries (NanoTemper; SKU:MO-K022). Fluorescence intensities and MST were measured in Monolith NT.LabelFree (NanoTemper) instrument at LED excitation 30%. Data were normalized with MO. Control (v1.6.1) and fitted using 4PL equations that yield the $K_{D,app}$, where C is the inflection point or $K_{D,app}$ in nM.

$$y(x) = A + \frac{B - A}{1 + \left(\frac{x}{C}\right)^D} \quad (1)$$

Results were expressed as mean \pm SD from three biological replicates.

Circular dichroism (CD)

We conducted CD experiments to observe i-motif-spectral changes upon PCBPI/KH mutants' interactions, spectral analyses of PCBPI and KH mutants, and pH-dependent changes in PCBPI using a Jasco J1700-CD spectrophotometer. All CD scans were performed with 4 μ M iMfs or proteins within a 320–210/200 nm range and at 100 nm/min scanning speed. Data points were obtained at 0.5 intervals and averaged over three accumulations. We used 1 mm-pathlength quartz cuvettes (Hellma) carrying 200 μ L samples. Digital integration time and bandwidth were 2 s and 1 nm, respectively. Oligos were first annealed into 20 mM MES (pH 6.4) and 100 mM NaCl.

To monitor i-motif-spectral changes upon interactions, PCBPI or KH mutants were titrated into iMfs in increasing concentrations (1–4 μ M) and incubated for 30 min at 25 °C.

To measure the transition pH (pH_T) of i-motifs, we acquired i-motif spectra (4 μ M) within a pH range from 5.4 to 7.2. CD values were corrected for respective buffer contributions and converted to molar ellipticities (MCD) using the following Formula (2), where θ is CD ellipticity in millidegrees, c is DNA concentration (mol/L), and l is

pathlength (cm):

$$\Delta\varepsilon(\text{M}^{-1}\text{cm}^{-1}) = \frac{\theta}{32980 \times c \times l} \quad (2)$$

pH_T values were calculated by plotting fraction of folded i-motifs (calculated from MCDs at 287 nm (positive maxima)) vs pH, using the equation below:

$$f(\text{folded}) = \frac{\text{MCD}_i - \text{MCD}_{\text{pH}7.2}}{\text{MCD}_{\text{pH}5.4} - \text{MCD}_{\text{pH}7.2}} \quad (3)$$

We fitted these data points using the Slogistic1 Eq. (1) in OriginPro 2020, where $x_c = \text{pH}_T$. Results were shown as means \pm SD of three biological replicates.

Concentration-dependent CD melting assay

To assess whether the four i-motifs undergo intermolecular folding at 4 μM concentration under our experimental conditions, we conducted concentration-dependent CD melting assays across a concentration range of 2–24 μM . CD melting experiments were conducted under identical instrumental parameters as described above. Spectra were recorded between 20 and 95 $^\circ\text{C}$ at 5 $^\circ\text{C}$ intervals with a heating rate of 1 $^\circ\text{C}/\text{min}$. To ensure thermal equilibration, each temperature point was maintained for 180 s prior to acquisition. CD signals at the positive ellipticity maximum (-287 nm) were normalized by calculating the fraction of folded using the following relation:

$$\text{Fraction of folded} = \frac{(\text{ellipticity}_i - \text{ellipticity}_{95^\circ\text{C}})}{(\text{ellipticity}_{20^\circ\text{C}} - \text{ellipticity}_{95^\circ\text{C}})} \quad (4)$$

Then melting curves were generated by plotting the fraction of folded values versus temperature range; and fitting the data points using 4PL equations that yield the melting temperatures (T_m), C is the inflection point or T_m in $^\circ\text{C}$

$$y(x) = A + \frac{B - A}{1 + \left(\frac{x}{C}\right)^D} \quad (5)$$

I-motif-destabilization trapping assay

To estimate the i-motif unfolding kinetics by PCBPI and KH mutants, we optimized an i-motif-destabilization trapping assay. In this assay, 5'-Cy5-labeled iMfs substrates were prepared at 5 nM in 20 mM MES (pH 6.4), 50 mM NaCl, 2 mM MgCl₂, and 0.25 mg/mL BSA, followed by incubation for 30 min at 25 $^\circ\text{C}$ in the presence of 5 nM PCBPI or KH mutants. Reactions were performed by adding 10 nM i-motif-trap-complementary oligos, specific for each iMfs, at 25 $^\circ\text{C}$ upto 20 min. At indicated time points, 10 μL of reaction mix was quenched with 1:1 stop buffer (40% glycerol, 60 mM EDTA, 0.6% SDS, 0.5 μM unlabeled iMfs) and analyzed by 20% native PAGE at 110 V. Gels were imaged using an Amersham Typhoon scanner with a Cy5 laser at 300 V with a resolution of 25 μm per pixel and quantified with ImageJ. Negative controls lacked i-motif-trap-complementary oligos, while in positive controls, iMfs were incubated with i-motif-trap-complementary oligos for 24 h at room temperature to allow complete hybridization. The fraction of hybridization between unfolded i-motif and its complementary trap oligos was estimated and normalized by dividing the intensity of trapped or hybridized products by the sum of trapped and unhybridized band intensity. Data points plotted against time points were fitted using a single-exponential function Eq. (4), to calculate the rate of i-motif-unfolding (k), where $F_{\text{max}}(t)$ denotes the fractional maximal value (or normalized signal) at time t . Experiments were repeated three

times.

$$F_{\text{max}}(t) = 1 - e^{-kt} \quad (6)$$

Bromine footprinting

To map the destabilized C:CH⁺ pairs in *cMYC*-i-motif upon PCBPI or KH mutant binding, we performed a bromine footprinting assay. 10 μM of 3'-Cy5-labeled *cMYC*-iMfs were annealed in 20 mM MES (pH 6.4) or 20 mM phosphate buffer (pH 7.0) with 50 mM NaCl by heating at 95 $^\circ\text{C}$ for 5 min, followed by slow-cooling overnight. 5 μM iMfs pre-incubated with or without 5 μM PCBPI/KH mutants for 30 min at 25 $^\circ\text{C}$ were treated with molecular bromine formed in situ by mixing equimolar KBr (Sigma Aldrich; 221864) and KHSO₅ (Sigma Aldrich; 228036) (20 μM) for 20 min at room temperature. Reactions were stopped by 80 μL of 0.3 M sodium acetate (pH 7.0) and 25 $\mu\text{g}/\text{mL}$ calf thymus DNA on ice. Unreacted bromine was removed by two successive ethanol precipitations. The DNA pellet was resuspended in 100 μL of 100 mM piperidine (Sigma; 104094), heated at 90 $^\circ\text{C}$ for 20 min to induce bromination-specific strand cleavage, then washed thrice and dried using speed-vac. Samples were dissolved in 95% formamide (VWR; A2156) and 20 mM EDTA, heated at 95 $^\circ\text{C}$ for 5 min, and snap-cooled. Purine- and pyrimidine-specific reactions were performed using 4% formic acid (Sigma Aldrich; 5438040100) and hydrazine (Sigma Aldrich; 309400) to generate A+G and C+T sequencing markers, respectively, following Maxam-Gilbert reactions⁷⁰. Samples were analyzed on 20% denaturing PAGE at 60 W, imaged using Amersham Typhoon scanner with Cy5 laser (500 V, 25 μm pixel resolution), and quantified with ImageJ. Each assay was performed in three biological replicates.

Isothermal titration calorimetry (ITC)

Thermodynamic binding profiles between iMfs and proteins (PCBPI and KH mutants) were determined using ITC with a MicroCal Auto iTC200 at 25 $^\circ\text{C}$ at both pH 8.0 and pH 6.4 using similar buffer conditions. For the binding studies involving KH1+2 and KH3, 5 μM of KH1+2 was prepared in 20 mM MES (pH 6.4) and 150 mM NaCl and placed in the calorimetric cell. The syringe contained 350 μM KH3 in the same buffer. To investigate the interactions between KH mutants (KH1+2 and KH3) and i-motif, 5 μM KH1+2 was pre-incubated with an equimolar concentration of *cMYC*-i-motifs for 30 min before being placed in the cell, while the syringe was filled with 350 μM KH3. For binding studies between proteins and iMfs, 5 μM of each protein was loaded into the cell, and 300 μM of iMfs was placed in the syringe. Control experiments were conducted simultaneously by injecting the same concentrations of substrates into a buffer without ligand to account for the heat of dilution. Oligonucleotides were injected 20 times at 150-s intervals into the calorimeter cell to achieve binding saturation. Data analysis was performed using the Malvern Microcal PEAQ-ITC software, employing a "one-site" binding model to obtain the best-fit values for the number of binding sites (N), changes in enthalpy (ΔH), entropy (ΔS), and free energy (ΔG) of binding reactions.

Purification of recombinant PCBPI and KH mutants

cDNAs of PCBPI, KH1+2, and KH3 were cloned into the pET-His1a vector for overexpression of recombinant PCBPI in *Escherichia coli* BL21(DE3) cells. PCBPI was induced at 18 $^\circ\text{C}$ before harvesting by centrifugation. The cell pellet was resuspended in lysis buffer (50 mM NaP 8.0, 500 mM NaCl, 10% glycerol, 0.2% Triton X100, 10 mM imidazole, 5 mM β -mercaptoethanol (β -me), DNase) followed by sonication. The sample was centrifuged (30 min, 20,000 $\times g$, 4 $^\circ\text{C}$), and the supernatant was incubated (1.5 h, 4 $^\circ\text{C}$) with ThermoS NiNTA previously equilibrated in lysis buffer. The mixture was then poured into a gravity flow column and the resin was washed with wash buffer (wash buffer 1: 50 mM NaP pH 8.0, 500 mM NaCl, 10% glycerol, 0.2% Triton

X100, 10 mM Imidazole, 5 mM β -me; wash buffer 2: 50 mM NaP pH 8.0, 1 M NaCl, 10% glycerol, 0.2% Triton X100, 10 mM Imidazole, 5 mM β -me). The elution was performed stepwise with two elution buffers (Elution buffer 1: 50 mM NaP, pH 8.0, 500 mM NaCl, 20 mM imidazole, 5 mM β -me; elution buffer 2: 50 mM NaP, pH 8.0, 300 mM NaCl, 300 mM imidazole, 5 mM β -me). The eluates were analyzed by SDS-PAGE, pooled together according to their purity, and desalted by dialysis in 20 mM NaP, pH 8.0, 150 mM NaCl, 5 mM β -me. His tag was cleaved by TEV protease, and final dialysis was done in 20 mM Tris-Cl 8.0, 150 mM NaCl, 20% Glycerol, 1 mM DTT (Dithiothreitol).

High-resolution-primer extension assay

To examine the effect of i-motif/hairpin formation to slow down in vitro replication, we performed primer extension assays. 5'-TET (Tetrachloro fluorescein)-labeled primer (1 μ M) was annealed to i-motif-forming templates (1.5 μ M) in 75 mM NaCl and 6 mM MgCl₂ by heating at 95 °C for 5 min followed by slow-cooling overnight. 40 nM of annealed samples were incubated in 50 μ l reactions containing 20 mM MES (pH 6.0), 6 mM MgCl₂, 0.2 mg/ml BSA, and 0.05 μ U/L Klenow fragment (Thermo Scientific; EPO051) with/without 40 nM PCBPI or KH mutants for 20 min at 25 °C. Primer extension was started with 0.2 mM dNTPs and continued for specific time points. Reacted samples were collected at designated time intervals, quenched with stop solution (95% formamide, 20 mM EDTA), denatured at 95 °C for 5 min, and snap-cooled. 5 μ l was loaded on 12% denaturing PAGE (8 M urea (VWR; 28877.292), 25% formamide, 1 \times TBE), and run at 60 W in 1 \times TBE running buffer. The gel was visualized using Amersham Typhoon (GE Healthcare) with Cy3 laser (500 V, 25 μ m pixel resolution) and quantified using ImageJ. Each experiment was performed in three biological replicates.

¹H 1D NMR

NMR experiments were conducted using a Bruker DRX 850 MHz NMR spectrometer, equipped with a 0.7 mm ultra-fast MAS probe, to understand the impact of PCBPI/KH mutants binding on i-motif's structural dynamics. 100 μ M of iMfs were prepared and annealed in 20 mM MES buffer (pH 6.4) containing 100 mM NaCl in a solvent mixture of 90% water and 10% D₂O. The 1D ¹H NMR experiments were performed in 3 mm NMR tubes (Bruker; Z172598) with an active sample volume of 200 μ l. Spectral referencing was performed using an internal standard, TSP [3-(trimethylsilyl)-2,2',3,3'-tetra-deuteriopropionic acid] (Sigma; 450510), set at 0.0 ppm. In the 1D proton spectra, imino proton resonances corresponding to C:CH⁺ bonds in the i-motifs were observed at 15–16 ppm, while resonances indicative of hairpins involving Watson-Crick base pairs were detected at 12–14 ppm. These measurements were obtained using the standard Bruker pulse program “zgesgp” with a spectral width of 20 ppm, 256 scans (ns), acquisition time 2 s, and a calibrated pulse length (pl) of 9.48 μ s. NMR titrations were performed by incrementally adding aliquots of purified proteins (PCBPI and KH mutants) to the 100 μ M iMfs. Samples were thoroughly mixed and allowed to reach thermal equilibrium. Proton spectra were recorded at each titration point, following an incubation period of 15 min at 298 K. Data acquisition and processing were carried out using Topspin 4.2.0 (Bruker).

Docking and MD simulations

MD simulations were performed using the Amber20 package on the structure obtained from AlphaFold⁷¹ and RoseTTAFold⁷² or RoseTTAFoldNA⁷³. Amberff14SB⁷⁴ and parmbsc1⁷⁵ force fields were used to model protein and DNA i-motif, respectively, with the CUFIX correction for ionic interaction⁷⁶ and Amberff14IDPSFF⁷⁷ force field for the intrinsically disordered part of PCBPI between domains 2 and 3 (residues 170–268). The protonation state of the amino acids of the PCBPI protein was determined on the basis of PropKa⁷⁸ calculations. All the systems were solvated in a cubic box of TIP3P water with a

solvent buffer of at least 12 Å and about 0.10 M of NaCl salt, taking into account the neutralization of the box. Periodic conditions were applied in combination with the Particle Mesh Ewald method for electrostatics and a cutoff of 10 Å for intermolecular interactions. All the systems were first minimized for 10,000 steps (5000 steepest descent, 5000 conjugated gradient) and then heated from 0 to 300 K during 30 ps in NVT ensemble with a time step of 1 fs. Equilibration (100 ns) and production (1–3 μ s) were run in NPT ensemble at 300 K and 1 bar using Langevin thermostat. The SHAKE algorithm was applied to constrain hydrogen-heavy atom bonds and maintain a time step of 2 fs. Trajectories were analyzed using CPPTRAJ⁷⁹. First, MD simulations were performed on 6 conformations of PCBPI alone, one from AlphaFold prediction, 5 from RoseTTAFold prediction. We selected the representative structure of the most important cluster from a cluster analysis based on the RMSD of the protein residues and a hierarchical agglomerative approach for each trajectory of 100 ns. These structures were combined with the 8AYG⁴⁰ structure for i-motif DNA with the ILPR sequence using RoseTTAFoldNA docking. Five docked conformations (Fig. S21A) were selected based on the localization of the DNA towards the different domains, especially KH1 and KH3, and the DNA-RNA binding part of these domains. Three simulation setups were created with only neutral cytosines in the i-motif: one neutral (set a), one with one positively charged cytosine for half of the cytosine pair (set b), and one with one positively charged cytosine per i-motif cytosine pair (set c). The production runs last 1 μ s for sets a structures 1, 4, and 5; 2 μ s for set a structures 2 and 3, and for all structures of set c; 3 μ s for all structures of set b. To compare the structural features of the KH folds between hnRNP-K and PCBPI, we aligned them using VMD software (<https://www.ks.uiuc.edu/Research/vmd/>)⁸⁰.

Bioinformatic analyses

De novo motif discovery was performed on the bed file extracted from the ChIP-seq datasets using the MEME-ChIP web interface. We used the plotheatmap tool in Galaxy to visualize ChIP-enrichment values around 2000 bp around the TSS (transcription start site). To compare functional overlap and uniqueness between PCBPI and other hnRNP family proteins, we used previously published i-motif interactome proteomics data¹⁹ that identifies candidate i-motif-binding proteins. From this dataset, all proteins annotated to the hnRNP family were extracted for gene ontology (GO) enrichment of Biological Process (BP) terms using the PANTHER Classification System, applying Fisher's exact test with FDR correction. Enriched GO terms were visualized using a Sankey diagram generated in RStudio 2024.09.0 + 375 and R 4.4.0.

Cell culture and PCBPI-KD

HeLa cell line was a generous gift from Dr. Sjoerd Wanrooij's group at Umeå University. The cell line was originally procured from ATCC and propagated in academic laboratories. Cells were cultured at the required density in Dulbecco's modified Eagle's medium (DMEM) (Gibco; 10565018) in the presence of 10% FBS (Fetal bovine serum) (Gibco; A5256701) at 37 °C and 5% CO₂ in the incubator. PCBPI knockdown (KD) was performed using small interfering RNA (siRNA) (PCBPI Human siRNA Oligo Duplex (Locus ID 5093, Origene; SR303372)) transfection with Lipofectamine 3000 (Thermo Fisher Scientific; L3000008) according to the manufacturer's protocol. Briefly, cells were seeded in a 6-well plate at a density of 10⁶ cells per well and allowed to adhere overnight. The following day, 5 ng of siRNA targeting PCBPI (siPCBPI) or 5 ng of non-targeting scrambled siRNA (siScramble, negative control) was diluted separately in Opti-MEM reduced serum medium (Thermo Fisher Scientific; 31985070). The diluted siRNA was then combined with the diluted Lipofectamine 3000 reagent and incubated for 10–15 min at room temperature. The siRNA-lipid complexes were added to the cells and incubated at 37 °C

with 5% CO₂ for 48 h, and the KD was confirmed by western blot analyses with anti-PCBP1 antibody.

qRT-PCR

To inspect the transcript levels of *cMYC*, *BCL2*, and *ILPR* genes upon PCBP1-KD, we performed quantitative real-time PCR. HeLa cells grown at a density of 1×10^6 cells per well were treated with PCBP1-KD. Total RNA is isolated from both untreated and KD cells using Qiagen RNeasy Mini Kit (Qiagen; 74104) as per the manufacturer's instructions. cDNA was prepared using UltraScript 2.0 cDNA synthesis kit (PCR Biosystems; PB30.31-02) by incubating a 20 μ L reaction containing 2 μ g RNA, cDNA synthesis mix, and Ultrascrip 2.0 reverse transcriptase at 50 °C for 30 min, followed by a denaturation at 95 °C. qRT-PCR reactions contain 1x SyGreen mix (PCR Biosystems), 0.5 μ M forward and reverse primers, and cDNA. The PCR program was 95 °C for 3 min (1 cycle) followed by a three-step reaction of 95 °C for 10 s (denaturation), 57–60 °C for 25 s (annealing), and 72 °C for 20 s (elongation) (35 cycles) and performed in CFX Real Time System (C1000 Thermal cycler (Bio-Rad)). The housekeeping gene, *GAPDH*, is used as an internal control to normalize the variability in target mRNA expression levels, and *HTR6* transcript levels are shown as a negative control. qRT-PCR primers (Supplementary Table S7) are designed using Primer-BLAST, NCBI, and analyzed in OligoAnalysier 3.1-IDT. All experiments are performed in three biological replicates.

Western blot

To confirm PCBP1-KD and estimate DNA double-strand breaks upon PCBP1-KD, we performed western blot analyses of PCBP1 and γ H2AX (H2AX histones phosphorylated at Ser139), respectively, in HeLa cells upon 48 h of PCBP1-KD. Cells were seeded at a density of 7×10^5 cells per 60 mm dish. At 48 h post-transfection, total protein was extracted from the cells using freshly prepared RIPA lysis buffer (50 mM Tris (pH 8.0), 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) supplemented with Phenylmethylsulfonyl fluoride and protease inhibitor (complete mini EDTA-free tablets, Roche; 5892791001). 30 μ g protein lysates were resolved on mini protean TGX precast SDS-PAGE gels (Bio-Rad) and transferred onto nitrocellulose membranes (Amersham™ Protran® Premium Western blotting membranes). Membranes were blocked with 5% non-fat milk in 1x Tris-buffered saline (TBS) with Tween 20 (Amresco; SKU QBIC20726) for 1 h at 4 °C. Then, membranes were briefly washed with 1x TBST and incubated overnight at 4 °C with rabbit monoclonal anti-PCBP1 primary antibody (dilution 1:1000, Abcam; clone number: EPR11049(B), catalog number: ab168377, IgG isotype; Lot number: 1050664-5), or rabbit phosphor-histone H2AX (Ser139) antibody (dilution 1:1000, Cell signaling technologies; Catalog number 2577S; clonality polyclonal; Isotype: IgG; lot number 14) and mouse monoclonal anti- β -Actin antibody (dilution 1:5000, Abcam; catalog number: ab8224; isotype: IgG1; Lot number: 1051636-14). Then, membranes were incubated with HRP-conjugated Goat anti-mouse/anti-rabbit polyclonal IgG (H+L) (dilution 1:5000, Thermo Fisher Scientific; catalog number: 31460) for 2 h at room temperature. Target proteins were visualized on membranes using Western Supersignal West Pico chemiluminescence substrate (Thermo Fisher), and images were captured using ChemiDoc MP system with Image Lab™ software. After incubation with each antibody, membranes were washed for 15 min at room temperature three times with 1x TBST. β -Actin was used as a loading control in the same blots. All experiments are performed in three biological replicates.

iMab, PCBP1, and BG4 Chromatin immunoprecipitation (ChIP)
 2×10^6 HeLa cells were crosslinked in 1% formaldehyde (Thermo Fisher; 28908) for 10 min at room temperature, followed by quenching with 125 mM glycine. After quenching, cells were washed with ice-cold 1x phosphate buffered saline (PBS) and collected by centrifugation at $500 \times g$ for 5 min at 4 °C. Cells were lysed on ice in ChIP lysis buffer

(50 mM HEPES, pH 7.4; 140 mM NaCl; 1 mM EDTA; 1% Triton X100; 0.1% Na-deoxycholate) for 10 min at 4 °C. Chromatin was isolated and sheared to 200–600 bp using the Covaris E220 system. After RNase A treatment (10 μ g/mL) (Thermo Fisher scientific; EN0531) at 37 °C for 30 min, chromatin was incubated overnight (4 °C) with 2 μ g iMab antibody (Absolute Antibody) or 500 ng BG4 antibody (in-house prepared)⁸¹ or rabbit monoclonal anti-PCBP1 primary antibody (1:500, Abcam) in ChIP lysis buffer containing 1% non-fat milk in 200 μ L, shaking at $25 \times g$. For iMab and BG4, DYKDDDDK Tag (D6WSB) Rabbit mAb anti-FLAG antibody (Cell Signaling, catalog number: 14793S, Clonality: Monoclonal; IgG isotype; Lot number: 7) (8 μ g) bound to 10 μ L dynabeads protein-G (Invitrogen; 10003D) was prepped by shaking at $130 \times g$ for 2–4 h at 4 °C, then washed and incubated with 80 μ L chromatin-iMab complex at $33 \times g$ for 4 h at 4 °C. Beads were washed four times with 100 μ L ChIP lysis buffer and twice with 100 μ L wash buffer at 4 °C, then immunoprecipitants were eluted with TE buffer (10 mM Tris-Cl, pH 8.0; 1 mM EDTA) containing 0.5 mg/mL Proteinase K by incubation at 37 °C for 1 h and 65 °C for 2 h. Elutes were purified using the ChIP-DNA Clean and Concentrate kit (ZYMO Research; D5201) following the manufacturer's protocol. qPCR was conducted using primers (Supplementary Table S6) targeting promoters with iMfs or non-i-motif-forming control regions. ChIP enrichment was quantified using the percentage of input method, where input DNA corresponds to a fraction of total chromatin collected before immunoprecipitation. Rabbit anti-FLAG IgG was used as the mock control for iMab and BG4, and anti-rabbit IgG was used for anti-PCBP1 ChIP. For each sample, the percentage of input was first calculated for both target loci and negative control regions. Fold enrichment of target loci was then determined by dividing the percentage of input of the target by the corresponding negative control. This two-step calculation allowed us to account for both IP efficiency and background signal, ensuring that the observed enrichment was specific to i-motif-forming regions. The raw C_t values of input, immunoprecipitated (IP) samples, and mock (IgG) in ChIP experiments are provided in detail in Supplementary Data 1–6 alongside the two-step calculations.

Flow cytometry-based cell cycle analyses

HeLa cells were grown to a density of 1×10^6 cells per well. Were subjected to a double-thymidine block or PCBP1-KD, followed by trypsinization and subsequently washed with cold 1x PBS. The cells were then fixed with chilled 70% ethanol and stored at -20 °C overnight. After fixation, the cells were centrifuged at $850 \times g$ for 15 min at 4 °C. The resulting pellets were resuspended in 500 μ L of FxCycle PI/RNase staining solution (Invitrogen; F10797) and incubated in the dark at 4 °C with gentle rocking for 4 h. DNA content was analyzed using BD CSampler Plus Software version:1.0.34.1. Gating of single cells, histograms of DNA content (Propidium iodide (PI) fluorescence intensity) were analyzed in FCS Express. The percentage of cells in each phase of the cell cycle (G_1 , S, and G_2/M) was determined by fitting the data into Cox multivariate analyses.

Cell synchronization using the double-thymidine block

HeLa cells were synchronized at the G_1/S boundary using a double-thymidine block. Thymidine (2 mM final concentration) (Sigma Aldrich; 89270) was added to the cells grown in 30–40% confluency into DMEM supplemented with 10% FBS overnight, and incubated for 18 h. Cells were then washed thrice with 1x PBS and released into fresh, thymidine-free medium for 9 h. Subsequently, a second thymidine treatment (2 mM final concentration) was applied for another 15–18 h to achieve synchronization at the G_1/S transition. After the second block, cells were either harvested at the G_1/S boundary after 2 h or released into fresh medium for 4 h to obtain synchronized populations in S phase. Synchronization efficiency was confirmed by flow cytometry analysis of DNA content as described above.

Anti-BrdU dot-bot assay

HeLa cells were synchronized at G₁/S using the aforementioned double-thymidine block and released into fresh medium at four time points at an interval of 1 h upto 4 h. 1.3 µg genomic DNA isolated from cells at each time-point using DNeasy Blood & Tissue Kit (Qiagen; 69504), and were heated at 95 °C and immediately cooled on ice. 0.3 M final concentration of NaOH was added, and DNA was loaded onto a Hybond-N⁺ membrane (GE Healthcare) using a Bio-Dot Microfiltration Apparatus (Bio-Rad). The membrane was blocked overnight with 1% non-fat milk in 1× PBS and incubated with Rat anti-BrdU BUI/75 (ICRI) antibody (Abcam; catalog number: ab6326, Clonality: monoclonal, subtype: IgG2a) (1:1000 dilution) diluted in 1% non-fat milk. After washing the membrane for 15 min twice with 1× PBS, it was incubated with goat anti-rat IgG antibody peroxidase (Merck; catalog number: DC01L-200UG) (1:5000 dilution) diluted in 1% non-fat milk. Dots were developed using Western Supersignal West pico chemiluminescence substrate (Thermo Fisher; 34580), and images were captured using ChemiDoc MP system with Image Lab™ software. Dot-blot assays were repeated in three biological replicates.

Statistical analyses

All statistical analyses were performed using OriginPro 2020. One-way ANOVA followed by Tukey–Kramer post hoc test was used for ChIP enrichment analyses to compare multiple groups. In qRT-PCR studies and comparison of $K_{D,app}$ between PCBP1/KH mutants and multiple i-motifs, statistical significance was assessed using a two-tailed Student's *t*-test. Data from biochemical and cellular studies are presented as mean ± standard deviation (SD) from three biological replicates. A *p* value < 0.05 was considered statistically significant.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available within the paper and its Supplementary Information. For ChIP-qPCR experiments, raw *C_t* values for Input, mock (IgG), and IP of all three biological replicates (having three technical replicates each) and related calculations are provided separately in Supplementary Datasets 1–6. MD simulation trajectories are available at Zenodo (<https://doi.org/10.5281/zenodo.17737326>)⁸². Source data are provided with this paper.

References

- Yu, B., Pettitt, B. M. & Iwahara, J. Dynamics of ionic interactions at protein–nucleic acid interfaces. *Acc. Chem. Res.* **53**, 1802–1810 (2020).
- Tan, C., Terakawa, T. & Takada, S. Dynamic coupling among protein binding, sliding, and DNA bending revealed by molecular dynamics. *J. Am. Chem. Soc.* **138**, 8512–8522 (2016).
- Alexandrov, B. S. et al. DNA breathing dynamics distinguish binding from nonbinding consensus sites for transcription factor YY1 in cells. *Nucleic Acids Res.* **40**, 10116–10123 (2012).
- Niu, K. et al. Bmi1F and i-motif structure are involved in transcriptional regulation of BmPOUM2 in *Bombyx mori*. *Nucleic Acids Res.* **46**, 1710–1723 (2018).
- Marsico, G. et al. Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res.* **47**, 3862–3874 (2019).
- Mergny, J.-L., Lacroix, L., Han, X., Leroy, J.-L. & Helene, C. Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA motif. *J. Am. Chem. Soc.* **117**, 8887–8898 (1995).
- Gehring, K., Leroy, J.-L. & Guéron, M. A tetrameric DNA structure with protonated cytosine–cytosine base pairs. *Nature* **363**, 561–565 (1993).
- Abou Assi, H., Garavís, M., González, C. & Damha, M. J. i-Motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.* **46**, 8038–8056 (2018).
- Tao, S., Run, Y., Monchaud, D. & Zhang, W. i-Motif DNA: identification, formation, and cellular functions. *Trends Genet.* <https://doi.org/10.1016/j.tig.2024.05.011> (2024).
- Zeraati, M. et al. I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* **10**, 631–637 (2018).
- Ruggiero, E. et al. The iMab antibody selectively binds to intramolecular and intermolecular i-motif structures. *Nucleic Acids Res.* **53**, gkae1305 (2025).
- Ma, X. et al. Genome-wide characterization of i-motifs and their potential roles in the stability and evolution of transposable elements in rice. *Nucleic Acids Res.* **50**, 3226–3238 (2022).
- Serrano-Chacón, I., Mir, B., Escaja, N. & González, C. Structure of i-motif/duplex junctions at neutral pH. *J. Am. Chem. Soc.* **143**, 12919–12923 (2021).
- Dzatzko, S. et al. Evaluation of the stability of DNA i-motifs in the nuclei of living mammalian cells. *Angew. Chem. Int. Ed.* **57**, 2165–2169 (2018).
- Zanin, I. et al. Genome-wide mapping of i-motifs reveals their association with transcription regulation in live human cells. *Nucleic Acids Res.* **51**, 8309–8321 (2023).
- Peña Martínez, C. D. et al. Human genomic DNA is widely interspersed with i-motif structures. *EMBO J.* **43**, 4786–4804 (2024).
- El-Khoury, R. et al. Telomeric i-motifs and C-strands inhibit parallel G-quadruplex extension by telomerase. *Nucleic Acids Res.* **51**, 10395–10410 (2023).
- Martella, M. et al. i-Motif formation and spontaneous deletions in human cells. *Nucleic Acids Res.* **50**, 3445–3455 (2022).
- Ban, Y. et al. Profiling of i-motif-binding proteins reveals functional roles of nucleolin in regulation of high-order DNA structures. *Nucleic Acids Res.* **52**, 13530–13543 (2024).
- Yazdani, K. et al. Decoding complexity in biomolecular recognition of DNA i-motifs with microarrays. *Nucleic Acids Res.* **51**, 12020–12030 (2023).
- Wu, W.-Q., Zhang, X., Bai, D., Shan, S.-W. & Guo, L.-J. Mechanistic insights into poly(C)-binding protein hnRNP K resolving i-motif DNA secondary structures. *J. Biol. Chem.* **298**, 102670 (2022).
- Kang, H.-J., Kendrick, S., Hecht, S. M. & Hurley, L. H. The transcriptional complex between the *BCL2* i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.* **136**, 4172–4185 (2014).
- Miglietta, G., Cogoi, S., Pedersen, E. B. & Xodo, L. E. GC-elements controlling HRAS transcription form i-motif structures unfolded by heterogeneous ribonucleoprotein particle A1. *Sci. Rep.* **5**, 18097 (2015).
- Víšková, P. et al. In-cell NMR suggests that DNA i-motif levels are strongly depleted in living human cells. *Nat. Commun.* **15**, 1992 (2024).
- Gao, B. et al. Remodeling the conformational dynamics of I-motif DNA by helicases in ATP-independent mode at acidic environment. *iScience* **25**, 103575 (2022).
- Mohanty, B. K. et al. Heterogeneous nuclear ribonucleoprotein E1 binds polycytosine DNA and monitors genome integrity. *Life Sci. Alliance* **4**, e202000995 (2021).
- Karam, J. A. Q. et al. The RNA-binding protein PCBP1 modulates transcription by recruiting the G-quadruplex-specific helicase DHX9. *J. Biol. Chem.* **300**, 107830 (2024).
- Schmoldt, A., Benthe, H. F. & Haberland, G. Digitoxin metabolism by rat liver microsomes. *Biochem. Pharmacol.* **24**, 1639–1641 (1975).
- Piñol-Roma, S. & Dreyfuss, G. hnRNP proteins: localization and transport between the nucleus and the cytoplasm. *Trends Cell Biol.* **3**, 151–155 (1993).

30. Ji, X. et al. α CP binding to a cytosine-rich subset of polypyrimidine tracts drives a novel pathway of cassette exon splicing in the mammalian transcriptome. *Nucleic Acids Res.* **44**, 2283–2297 (2016).
31. Ji, X., Humenik, J., Yang, D. & Liebhaber, S. A. PolyC-binding proteins enhance expression of the CDK2 cell cycle regulatory protein via alternative splicing. *Nucleic Acids Res.* **46**, 2030–2044 (2018).
32. Ishii, T. et al. PCBP1 and PCBP2 both bind heavily oxidized RNA but cause opposing outcomes, suppressing or increasing apoptosis under oxidative conditions. *J. Biol. Chem.* **295**, 12247–12261 (2020).
33. Hahm, J. Y., Park, J., Jang, E.-S. & Chi, S. W. 8-Oxoguanine: from oxidative damage to epigenetic and epitranscriptional modification. *Exp. Mol. Med.* **54**, 1626–1642 (2022).
34. Link, L. A., Howley, B. V., Hussey, G. S. & Howe, P. H. PCBP1/HNRNP E1 protects chromosomal integrity by translational regulation of CDC27. *Mol. Cancer Res.* **14**, 634–646 (2016).
35. Zhang, T. et al. PCBP-1 regulates alternative splicing of the CD44 gene and inhibits invasion in human hepatoma cell line HepG2 cells. *Mol. Cancer* **9**, 72 (2010).
36. Song, Y. et al. Molecular and functional characterization of porcine poly C binding protein 1 (PCBP1). *BMC Vet. Res.* **20**, 25 (2024).
37. Shi, H. et al. PCBP1 depletion promotes tumorigenesis through attenuation of p27Kip1 mRNA stability and translation. *J. Exp. Clin. Cancer Res.* **37**, 187 (2018).
38. Patel, S. J. et al. The iron chaperone and nucleic acid-binding activities of poly(rC)-binding protein 1 are separable and independently essential. *Proc. Natl. Acad. Sci. USA* **118**, e2104666118 (2021).
39. Ma, W., Noble, W. S. & Bailey, T. L. Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat. Protoc.* **9**, 1428–1450 (2014).
40. Guneri, D. et al. Structural insights into i-motif DNA structures in sequences from the insulin-linked polymorphic region. *Nat. Commun.* **15**, 7119 (2024).
41. Kendrick, S., Akiyama, Y., Hecht, S. M. & Hurley, L. H. The i-motif in the *bcl-2* P1 promoter forms an unexpectedly stable structure with a unique 8:5:7 loop folding pattern. *J. Am. Chem. Soc.* **131**, 17667–17676 (2009).
42. Kendrick, S. et al. The dynamic character of the *BCL2* promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.* **136**, 4161–4171 (2014).
43. Dai, J., Hatzakis, E., Hurley, L. H. & Yang, D. I-motif structures formed in the human c-MYC promoter are highly dynamic—insights into sequence redundancy and I-motif stability. *PLoS ONE* **5**, e11647 (2010).
44. Boissieras, J. et al. iMab antibody binds single-stranded cytosine-rich sequences and unfolds DNA i-motifs. *Nucleic Acids Res.* **52**, 8052–8062 (2024).
45. Hänsel-Hertsch, R., Spiegel, J., Marsico, G., Tannahill, D. & Balasubramanian, S. Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* **13**, 551–564 (2018).
46. Wei, Y., Thyparambil, A. A. & Latour, R. A. Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230nm. *Biochim. Biophys. Acta Proteins Proteom.* **1844**, 2331–2337 (2014).
47. Sengupta, P. et al. Site-specific amino acid substitution in dodecameric peptides determines the stability and unfolding of c-MYC quadruplex promoting apoptosis in cancer cells. *Nucleic Acids Res.* **46**, 9932–9950 (2018).
48. Školáková, P. et al. Systematic investigation of sequence requirements for DNA i-motif formation. *Nucleic Acids Res.* **47**, 2177–2189 (2019).
49. Sengupta, P., Jamroskovic, J. & Sabouri, N. A beginner's handbook to identify and characterize i-motif DNA. in *Methods in Enzymology* (eds Raney, K. D., Eoff, R. L., Byrd, A. K. & Kendrick, S.) Vol. 695, 45–70 (Elsevier, 2024).
50. Singh, A. & Kukreti, S. Homoduplex to i-motif structural switch exhibited by a cytosine rich strand of the *MYH7* heavy chain β gene promoter at physiological pH. *RSC Adv.* **8**, 34202–34214 (2018).
51. Jamroskovic, J., Deiana, M. & Sabouri, N. Probing the folding pathways of four-stranded intercalated cytosine-rich motifs at single base-pair resolution. *Biochimie* **199**, 81–91 (2022).
52. Grün, J. T. et al. Unraveling the kinetics of spare-tire DNA G-quadruplex folding. *J. Am. Chem. Soc.* **143**, 6185–6193 (2021).
53. Carvalho, L. et al. Conserved role for PCBP1 in altered RNA splicing in the hippocampus after chronic alcohol exposure. *Mol. Psychiatry* **28**, 4215–4224 (2023).
54. Huang, S. et al. PCBP1 regulates the transcription and alternative splicing of metastasis-related genes and pathways in hepatocellular carcinoma. *Sci. Rep.* **11**, 23356 (2021).
55. Sengupta, P., Bose, D. & Chatterjee, S. The molecular Tête-à-Tête between G-quadruplexes and the i-motif in the human genome. *ChemBioChem* **22**, 1517–1537 (2021).
56. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013).
57. Di Antonio, M. et al. Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat. Chem.* **12**, 832–837 (2020).
58. Esain-Garcia, I. et al. G-quadruplex DNA structure is a positive regulator of *MYC* transcription. *Proc. Natl. Acad. Sci. USA* **121**, e2320240121 (2024).
59. Sengupta, P., Chattopadhyay, S. & Chatterjee, S. G-Quadruplex surveillance in *BCL-2* gene: a promising therapeutic intervention in cancer treatment. *Drug Discov. Today* **22**, 1165–1186 (2017).
60. Chen, L. et al. Structural basis for nucleolin recognition of *MYC* promoter G-quadruplex. *Science* **388**, eadr1752 (2025).
61. Ishii, T., Hayakawa, H., Igawa, T., Sekiguchi, T. & Sekiguchi, M. Specific binding of PCBP1 to heavily oxidized RNA to induce cell death. *Proc. Natl. Acad. Sci. USA* **115**, 6715–6720 (2018).
62. Braddock, D. T. Molecular basis of sequence-specific single-stranded DNA recognition by KH domains: solution structure of a complex between hnRNP K KH3 and single-stranded DNA. *EMBO J.* **21**, 3476–3485 (2002).
63. Sutherland, C., Cui, Y., Mao, H. & Hurley, L. H. A Mechanosensor mechanism controls the G-quadruplex/i-motif molecular switch in the *MYC* promoter NHE III₁. *J. Am. Chem. Soc.* **138**, 14138–14151 (2016).
64. Ruggiero, E. et al. A dynamic i-motif with a duplex stem-loop in the long terminal repeat promoter of the HIV-1 proviral genome modulates viral transcription. *Nucleic Acids Res.* **47**, 11057–11068 (2019).
65. Roy, B. et al. Interaction of individual structural domains of hnRNP LL with the *BCL2* promoter i-Motif DNA. *J. Am. Chem. Soc.* **138**, 10950–10962 (2016).
66. Du, Z. et al. Crystal structure of the first KH domain of human poly(C)-binding protein-2 in complex with a C-rich strand of human telomeric DNA at 1.7 Å. *J. Biol. Chem.* **280**, 38823–38830 (2005).
67. Yoga, Y. M. K., Traore, D. A. K., Wilce, J. A. & Wilce, M. C. J. Mutation and crystallization of the first KH domain of human polycytosine-binding protein 1 (PCBP1) in complex with DNA. *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **67**, 1257–1261 (2011).
68. Kawasaki, Y. et al. MYU, a target lncRNA for Wnt/c-Myc signaling, mediates induction of CDK6 to promote cell cycle progression. *Cell Rep.* **16**, 2554–2564 (2016).
69. Enge, M. et al. MDM2-dependent downregulation of p21 and hnRNP K provides a switch between apoptosis and growth arrest induced by pharmacologically activated p53. *Cancer Cell* **15**, 171–183 (2009).

70. Maxam, A. M. & Gilbert, W. Sequencing end-labeled DNA with base-specific chemical cleavages. in *Methods in Enzymology* (eds Grossman, L. & Moldave, K.) Vol. 65, 499–560 (Elsevier, 1980).
 71. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
 72. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
 73. Baek, M. et al. Accurate prediction of protein-nucleic acid complexes using RoseTTAFoldNA. *Nat. Methods* **21**, 117–121 (2024).
 74. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
 75. Ivani, I. et al. Parmbsc1: a refined force field for DNA simulations. *Nat. Methods* **13**, 55–58 (2016).
 76. Yoo, J. & Aksimentiev, A. New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Phys. Chem. Chem. Phys.* **20**, 8432–8449 (2018).
 77. Song, D., Luo, R. & Chen, H.-F. The IDP-specific force field *ff14IDPSFF* improves the conformer sampling of intrinsically disordered proteins. *J. Chem. Inf. Model.* **57**, 1166–1178 (2017).
 78. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
 79. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
 80. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
 81. Deiana, M. et al. A new G-quadruplex-specific photosensitizer inducing genome instability in cancer cells by triggering oxidative DNA damage and impeding replication fork progression. *Nucleic Acids Res.* **51**, 6264–6285 (2023).
 82. Gillet, N., Sengupta, P., Obi, I. & Sabouri, N. Molecular dynamics simulations of PCBP1 protein and DNA i-motif. Zenodo, <https://doi.org/10.5281/ZENODO.17737326> (2025).
- 2019-AURA-P5B and AXELERA Pôle de compétitivité) and GENCI—IDRIS (project A0150800609) for computational resources.

Author contributions

P.S. and N.S. conceived the project and designed the experiments. P.S. and I.O. performed the experimental studies. N.G. performed the MD simulations. All authors contributed to the data analyses. P.S. and N.S. wrote the manuscript with input from all authors. N.S. supervised the work. All authors read and approved the manuscript.

Funding

Open access funding provided by Umea University.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-68822-5>.

Correspondence and requests for materials should be addressed to Nasim Sabouri.

Peer review information *Nature Communications* thanks Dennis Gomez and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

Acknowledgements

The authors acknowledge Biochemical Characterization Umeå (BMCU) for ITC and CD instruments, Swedish NMR Centre, Protein Production Sweden (funded by the Swedish Research Council) for purification of PCBP1, KH mutants, and BG4 at Umeå University, and Biorender for graphics. The authors thank Wenner-Gren Foundations (UPD2020-0097) and Swedish Cancer Society (24 0907 PT 01 H) for P.S.'s work. N.S. lab received support from Swedish Cancer Society (22 2380 Pj 01 H), Swedish Research Council (VR-MH 2021-02468), and Knut and Alice Wallenberg foundations (KAW 2021.0173). The authors acknowledge the PSMN mesocenter in Lyon (CPER/SYSPROD 2015–2022 project No.