Targeting RNA:protein interactions with an integrative approach leads to the identification of potent YBX1 inhibitors

Krystel El Hage, Nicolas Babault, Olek Maciejak, Bénédicte Desforges, Pierrick Craveur, Emilie Steiner, Juan Carlos Rengifo-Gonzalez, Hélène Henrie, Marie-Jeanne Clément, Vandana Joshi, et al.

Targeting RNA:Protein Interactions with an Integrative Approach Leads to the Identification of Potent YBX1 Inhibitors

Krystel El Hage[1*], Nicolas Babault[2†], Olek Maciejak[1†], Bénédicte Desforges[1†], Pierrick Craveur[2], Emilie Steiner[1], Juan Carlos Rengifo-Gonzalez[1], Hélène Henrie[1], Marie-Jeanne Clément[1], Vandana Joshi[1], Ahmed Bouhss[1], Liya Wang[1], Cyril Bauvais[2], David Pastré[1*]

***For correspondence:**
krystel.elhage@univ-evry.fr (KEH);
david.pastre@univ-evry.fr (DP)

[†]These authors contributed equally to this work

[1]SABNP, Univ Evry, INSERM U1204, Université Paris-Saclay, 91025 Evry, France;
[2]SYNSIGHT, 4 rue Pierre Fontaine, 91058 Evry, France

Abstract RNA-protein interactions (RPIs) are promising targets for developing new molecules of therapeutic interest. Nevertheless, challenges arise from the lack of methods and feedback between computational and experimental techniques during the drug discovery process. Here, we tackle these challenges by developing a drug screening approach that integrates chemical, structural and cellular data from both advanced computational techniques and a method to score RPIs in cells for the development of small RPI inhibitors; and we demonstrate its robustness by targeting Y-box binding protein 1 (YB-1), a messenger RNA-binding protein involved in cancer progression and resistance to chemotherapy. This approach led to the identification of 22 hits validated by molecular dynamics (MD) simulations and nuclear magnetic resonance (NMR) spectroscopy of which 11 were found to significantly interfere with the binding of messenger RNA (mRNA) to YB-1 in cells. One of our leads is an FDA-approved poly(ADP-ribose) polymerase 1 (PARP-1) inhibitor. This work shows the potential of our integrative approach and paves the way for the rational development of RPI inhibitors.

Introduction

Targeting RNA:protein interactions (RPIs) critically involved in pathological mechanisms is a promising strategy to find novel classes of drug candidates that remains largely unexploited (***Einstein et al., 2021***). RPIs in cells are highly diverse encompassing interactions with messenger RNA (mRNA) (***Baltz et al., 2012***), ribosomal RNA (rRNA) (***Simsek et al., 2017***), and non-coding RNA (ncRNA) (***Lu et al., 2019***), which are critical to fine tune the spatiotemporal gene expression. As revealed by genomic approaches (***Van Nostrand et al., 2020***; ***Castello et al., 2012***), the human genome contains more than 1000 transcripts encoding RNA-binding proteins (RBPs), thus providing a large variety of interactions with coding or non-coding RNAs. However, while the diversity of RNA:Protein interfaces may allow the development of RPIs inhibitory molecules (***Wu, 2020***), only scarce studies have already been undertaken and were restricted to few complexes such as LIN28/let-7 (***Roos et al., 2016***; ***Wang et al., 2018***), MUSASHI (MSI)/RNA (***Minuesa et al., 2019***) and heterogeneous nuclear ribonucleoprotein A18 (hnRNP A18)/RNA (***Solano-Gonzalez et al., 2021***).

Several challenges arise from the drug discovery process such as finding a druggable pocket in

RNA-binding interfaces (***Minuesa et al., 2019***), the quality of the computational models, the strategies used in the *in silico* screening, and the lack of experimental feedback and validation of computationally predicted inhibitors essential to orient the rational drug design procedure toward the most relevant molecules. Besides the above listed issues, new experimental assays must be developed to screen molecules targeting RPIs which ideally would work in a cellular context and be amenable to high content screening (HCS) (***Usaj et al., 2016***; ***Julio and Backus, 2021***). Indeed, to find potent inhibitors of RNA:protein interfaces, previous approaches used *in vitro* assays such as fluorescence polarization assay complemented by pull-down experiments with cell lysates or RNA enzyme-linked immunosorbent assay (ELISA) to test the effectiveness or selectivity of few hits (***Roos et al., 2016***; ***Minuesa et al., 2019***). While *in vitro* approaches are important to define putative hits and lead to the validation of effective compounds, deciphering whether the selected molecules are effective in a cellular context generally relies on indirect measurements using techniques such as cellular engagement thermal shift assay (CETSA) or functional assays where the putative consequences of disrupting RPIs on cellular function bear a considerable uncertainty. Indeed, multiple functions are associated to RBPs, which renders the interpretation of the results of functional assays tricky. In addition, toxicity and off-target effects are putative biases which are always difficult to get rid of, notably when using small molecules with a K_d in the low micromolar range, which is generally the case for RPI inhibitors. To fill the gap between *in vitro* and functional assays, cellular approaches initially used to detect protein:protein interactions (PPIs) such as fluorescence resonance energy transfer (FRET) or proximity ligation assay (PLA) have been adapted to detect RPIs (***Jung et al., 2013***; ***Camborde et al., 2017***) in cells but several technical issues have hampered their application such as the requirement of an adapter to RNA in FRET and PLA, the proximity of the donor and acceptor proteins in FRET, and the use of antibodies in PLA.

The aim of this paper is to tackle these challenges by introducing an experimental assay amenable to HCS to score RPIs in cells and a drug screening approach that integrates chemical, structural and cellular data from both advanced computational and experimental techniques for the development of small molecules that target RPIs. As an application model we chose to target Y-Box binding protein 1 (YB-1) of the YBX1 gene. As other abundant nucleic acid binding proteins, YB-1 participates in many DNA/RNA-dependent processes such as mRNA translation, splicing, transcription, long ncRNA (lncRNA) functions, and DNA repair (***Lyabin et al., 2014***). However, YB-1 is mostly a core component of untranslated messenger ribonucleoprotein particles (mRNPs) in the cytoplasm (***Singh et al., 2015***) which, according to crosslinking immunoprecipitation coupled to sequencing (CLIP) analysis (***Wu et al., 2015***), preferentially binds coding sequences and 3'-UTRs across most transcripts with a weak specificity. Since YB-1 binds to and regulates the activation of dormant mRNAs (***Budkina et al., 2021***) which are particularly enriched in gene controlling transcription (***Roos et al., 2016***), YB-1 is possibly involved in cellular decisions; and consistently, YB-1 was recently identified as one of the few key genes that control gene expression plasticity in rats subjected to caloric restriction (***Ma et al., 2020***). Interestingly, YBX1 is also one of the genes whose gene-protein expression is the most correlated in cancers vs. normal tissues (***Kosti et al., 2016***), and YBX1 was identified among the few genes in a clustered regularly interspaced short palindromic repeats (CRISPR) screen showing the highest sensitivities with broad proteome co-expression in cancer cell lines ((***Nusinow et al., 2020***), Figure S4 of this reference), pointing toward a possible role for YBX1 in cancer. The involvement of YB-1 in the progression and resistance to stress and chemotherapy (***Kang et al., 2013***; ***Yang et al., 2010***; ***El-Naggar et al., 2019***), notably after its translocation in the nucleus in certain cancers (***Bargou et al., 1997***), has also been documented. Together, these data make YB-1 a relevant target for cancer treatment (***Lasham et al., 2013***) and a subject of ongoing research to identify YB-1 inhibitors (***Khan et al., 2014***; ***Tailor et al., 2021***). Moreover, YB-1 is one of the host proteins implicated in viral replication of human immunodeficiency virus (HIV) (***Jung et al., 2018***; ***Poudyal et al., 2019***) and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (***Schmidt et al., 2021***) and hence targeting it along with targeting specific viral proteins can help reduce viral replication to a higher extent than just targeting the viral proteins. Our choice in targeting YB-1 was

also guided by the availability of structural data on RNA:YB-1 complexes to probe *in vitro* whether small molecules can interact with the cold-shock domain (CSD) of YB-1 (***Kretov et al., 2019***; ***Yang et al., 2019***).

We started this work by addressing the drug screening challenge and developing an integrative approach that uses in synergy advanced computational and experimental techniques in a concerted manner (as illustrated in Figure 1). Based on our discovery of a druggable pocket by molecular dynamics simulations (MD) located on the outside surface of the CSD β-barrel (which is also part of the RNA binding interface (***Yang et al., 2019***)), we implemented a large-scale computational approach that balances accuracy and computational cost to virtually screen potent compounds from small molecule libraries containing more than 7 million molecules. Next we addressed the abovementioned lack of methods able to score RPIs in a cellular context. To this end, we adapted the microtubule bench (MT bench) assay to score protein interactions with endogenous mRNAs in cells and implemented a robust HCS-based detection scheme. The MT bench was first introduced in 2015 to probe PPIs in cells with conventional fluorescence microscope by using microtubules as intracellular nanoplatforms (***Boca et al., 2015***; ***Rengifo-Gonzalez et al., 2021***).

The results presented here, show that the physics-based *in silico* approach allowed the identification of 22 potential hits that we subsequently tested *in vitro* by nuclear magnetic resonance (NMR) spectroscopy and in cells using the adapted MT bench assay by scoring the interaction of YB-1 with mRNA in the cytoplasm. Of these 22 potential YB-1 inhibitors, 15 compounds were found to bind YB-1 *in vitro* and 11 of them were found to efficiently interfere with the interaction of YB-1 with mRNA in cells at low micromolar concentrations; and with a notable specificity when compared with two other RBPs, Human antigen R (HuR) and fused in sarcoma (FUS). The potency of the selected compounds was further demonstrated by in depth MD and NMR analyses. The results also validate that the MT bench allows to robustly and automatically score RBP-specific interactions with endogenous mRNAs by using high-resolution HCS imagers.

Interestingly, compound P1, an FDA-approved poly(ADP-ribose) polymerase 1 (PARP-1) inhibitor (***Zandarashvili et al., 2020***), was found to interact with YB-1 with higher selectivity compared to the other hits. Whether P1 interferes with YB-1 cellular functions in cells therefore merits further investigations. Together, these results demonstrate the validity of our integrative approach and the efficacy of the MT bench assay that critically complements computational and structural approaches to identify compounds targeting RPIs in cells.

Results

A druggable pocket found in YB-1 CSD, a conserved RNA-binding domain

The first challenge was to find a druggable pocket in the structured cold-shock domain of YB-1 located at the RNA-binding interface. We started by taking into consideration small molecules that were reported to target YB-1 in the literature. The only molecule for which a structural validation was available, though only *in silico*, is the flavonoid, Fisetin (***Khan et al., 2014***). In this paper, using refined docking, Fisetin was found to inhibit YB-1 activation by Akt-mediated phosphorylation at S102 with a binding pocket located inside the β-barrel structure of YB-1 CSD (51-129 aa). Having in hand the longest YB-1 fragment (1-180 aa) amenable to NMR spectroscopy (***Kretov et al., 2019***), we then analyzed the interaction between Fisetin and YB-1 fragment *in vitro*. Significant chemical shift perturbations (CSPs) were indeed observed but not within the previously predicted pocket (***Khan et al., 2014***). The observed CSPs implicated residues located in a hydrophobic pocket on the outside surface of the β-barrel; these are W65, V84, F85, V86, G116, K118, G119 and A120 (pocket residues shown on the top right of Figure 1).

Quercetin, a Fisetin analog with an additional hydroxyl group capable of forming new H-bond interactions with YB-1, was also tested. Since it showed higher CSPs with the same pocket, compared to Fisetin, we decided to subsequently name it the “Quercetin-pocket” (average CSP of 0.032 for Quercetin (F1) compared to 0.028 for Fisetin (F4)). To delineate the characteristics of the Quercetin-

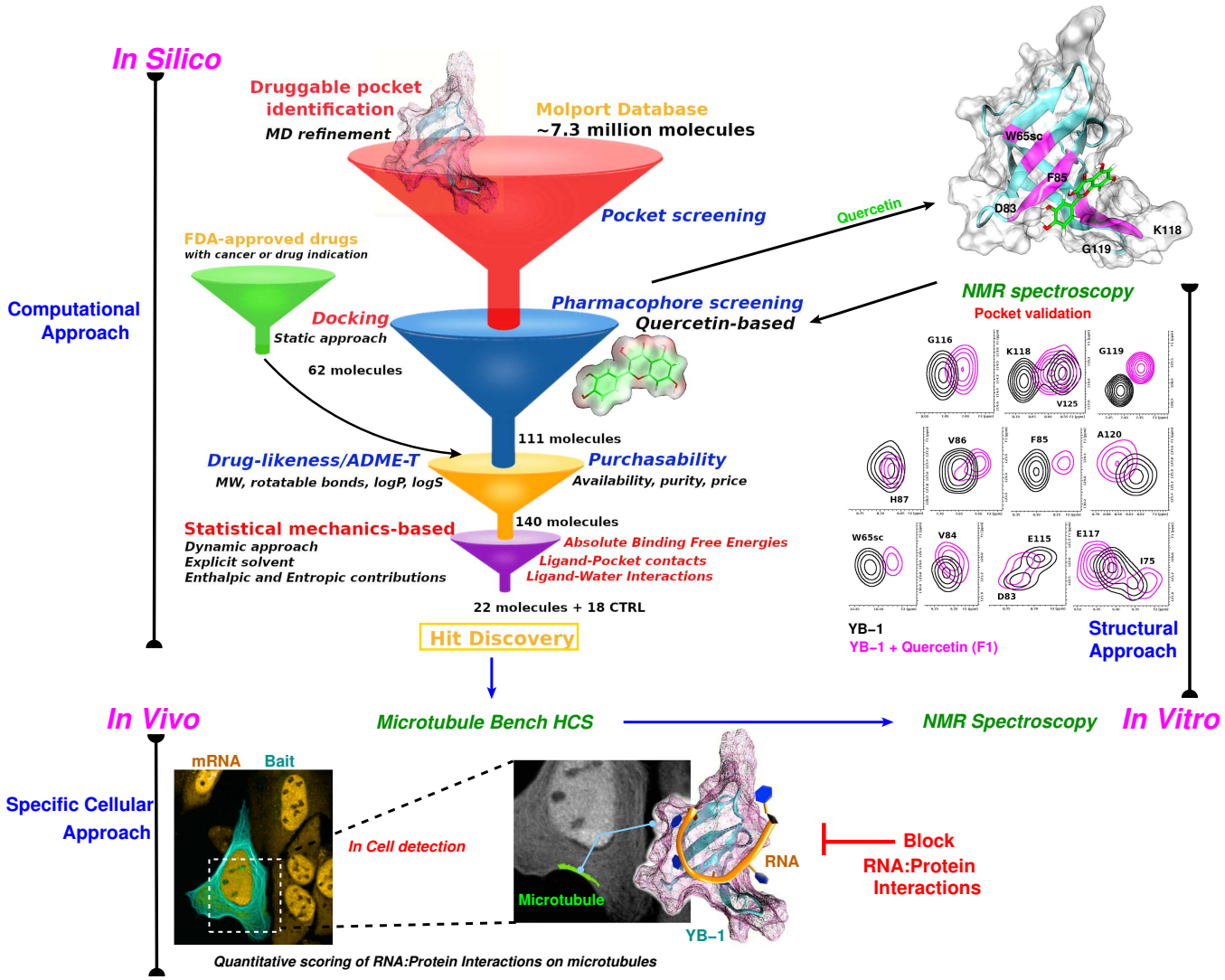


Figure 1. Schematic representation of an integrative approach for screening RPI inhibitors. This approach combines information from three data sources: computational (*in silico*, top left), cellular (*in vivo*, bottom), and structural (*in vitro*, top right). Blue arrows indicate the data flow. *In silico*: Starting by a large-scale computational approach that uses Docking (static approach), Molecular Dynamics and Free Energy Simulations (dynamic approach), using a computational model to virtually screen large libraries of small molecules (here, Molport and FDA-approved drugs) with the prior knowledge of a validated pocket and where several filters are used to reduce the selection to the most pertinent ligands that are then proposed as hits to be tested experimentally. Filters are represented as funnels. *In vivo*: In cell validation of the efficiency of the proposed hits in blocking RPIs with the MT Bench assay. This technology can quantify RPIs at the single cell level by using microtubule filaments as intracellular nanoplatforms (lower left inset, the bait, here YB-1, is shown in cyan and mRNA in orange). Lower right inset: enlarged view on mRNAs (orange) brought on microtubules using YB-1 as bait (cartoon representation: YB-1 in dark cyan with a violet surface mesh is complexed with RNA (orange ribbon)). *In vitro*: Experimental validation of binding the target pocket using solution NMR spectroscopy. A *zoom in* on pocket residue signals in a 2D ^{1}H-^{15}N-SOFAST-HMQC of YB-1 alone (black) and in the presence of Quercetin F1 (magenta). The top right 3D structure shows the binding of Quercetin (green stick) to YB-1 (cartoon representation in cyan combined with a transparent surface). Residues showing chemical shifts upon F1 binding are colored in magenta and depict what we identified as the Quercetin-pocket.

Figure 1–Figure supplement 1. Conformational study of YB-1 in its unbound/free form using MD simulations.

Figure 1–Figure supplement 2. Structural and energetic study of YB-1:RNA (C5) complex using MD simulations.

pocket, we used extensive MD simulations of YB-1 CSD either in its unbound or RNA-bound form (Figure 1-Figure supplement 1 and Figure 1-Figure supplement 2, respectively; detailed MD analysis can be found in Appendix 1). Results show that the Quercetin-pocket in its unbound form presents an open and a closed state. This pocket is located at the third β-hairpin and is monitored by K118 and F85 side chains. The opening mechanism is controlled by an electrostatic cation-π interaction formed between the cationic side chain of K118 (NH^{3+}) and the π-electron ring system of F85 (Figure 1-Figure supplement 1). The sampled structures of both open and closed states of CSD were also captured by NMR in the published 3D solution structure of ***Kloks et al.*** (***2002***) which is consistent with our findings. MD and NMR analysis of YB-1 in complex with 5-nt long poly(C) RNA (C5) show that some of the CSD key residues implicated in RNA binding are located in the Quercetin-pocket; these residues include W65, Y72, F74, F85, H87, K118, and E121 (Figure 1-Figure supplement 2). These residues are evolutionary conserved as shown by the ConSurf (***Chorin et al., 2020***; ***Goldenberg et al., 2009***) analysis reported in Appendix 1-section III and illustrated in Appendix 1-Figure 1. Together, MD and NMR analysis evidence the validity of the Quercetin-pocket as a potential target for the development of small molecules interfering with RNA:YB-1 interactions.

Prediction of potent inhibitors of mRNA:YB-1 interactions using a large scale computational approach

Having identified a druggable pocket at the RNA:YB-1(CSD) interface, we next sought to target it pharmacologically. Therefore, based on these atomistic and structural data, we implemented a large-scale computational strategy to propose putative inhibitors of RNA:YB-1 interactions. This approach is illustrated in Figure 1 and detailed in the Computational Methods section and in Appendix 2.

We started by using a pharmacophore approach to virtually screen a database composed of 208 million pharmacophores representing the conformers of around 7.3 million distinct commercially available molecules from MolPort: i) a “pocket-based” pharmacophore screening built from the prediction of a pseudo-ligand in the binding site of the MD refined structure of the open-state pocket and (ii) a distinct “ligand-based” pharmacophore built on the 3D structure of Quercetin (F1) with YB-1. The 3D structure of the YB-1:F1 complex was obtained by docking followed by refinement MD simulations and the binding site was confirmed by NMR spectroscopy (Figure 1). 249 and 407 distinct molecules were selected from the "pocket-based" and the "ligand-based" screening, respectively. Next these molecules were reduced to a final selection by predicting ADME-T (absorption, distribution, metabolism, excretion, and toxicity) endpoints and using computed molecular docking in the Quercetin-pocket (details in Appendix 2-section I). At the end, 111 molecules were retained from this static virtual screen after visual inspection and rational selection of structurally promising candidates.

In a second step, we applied physico-chemical filters to keep only molecules belonging to a drug-like chemical space (molecular weight, number of rotational bonds, number of proton donors and acceptors, lipophilicity and solubility). Purchasability filters were also applied based on availability, purity and price in order to facilitate and optimize the conditions for the *in vitro* and *in vivo* assays. From the 7.3 million MolPort molecules, 78 molecules were finally retained. In parallel, we executed an automated blind docking of 4700 FDA-approved drugs (Drugs-lib library (***Lagarde et al., 2018***)) using the MTiOpenScreen web server (***Labbé et al., 2015***) which lead to the selection of 62 molecules that may target the Quercetin-pocket and may be suitable for a repositioning strategy (details in Appendix 2-sections II and III).

In the last step, the pre-selected molecules using the above static approach, 140 in total (62 FDA-approved and 78 molPort molecules), were subject to a statistical mechanics-based filter that relies on MD and free energy simulations (dynamic approach). First, the docked poses were chosen after visual inspection of the docking results (***Fischer et al., 2021***). Second, short 10 ns MD simulations were run, in the presence of explicit water molecules, in order to refine the poses and check the stability of the ligands in the targeted pocket. Only ligands that stayed in the pocket during the

short MD were retained for the next step (87 out of the 140), where a weighted score (S) based on two observables that describes the ability of the ligand to bind and reside in the pocket was derived (this is detailed in the computational methods section). Ligands with a positive S were considered as hits, and ligands with $S < 0$ were only considered as "possible" if S becomes positive when we take into account the statistical error. From the 87 molecules tested, only 26 potential hits were retained (of which 6 "possible"). Finally, absolute free energy simulations (ABFE) were used to compute the protein-ligand binding free energies (ΔG_{bind}) and rank the ligands in terms of affinity (in kcal.mol^{-1}). ABFE simulations were done using the all-atom point charge CHARMM force field (***MacKerell et al., 1998***) and BAR (***Bennett, 1976***) for ΔG estimation. Here potential hits were selected for having a ΔG value > 5.50 kcal/mol. However, the 6 "possible" potential hits evaluated using S were considered as hits if they have a $\Delta G > 6.5$ kcal/mol (this is the case of F3: low S (6.15) and high ΔG_{bind} (-10.82 kcal/mol); C11 and C12 represent a similar case). The selection of the hits at the end took into account both evaluation methods (S and ΔG_{bind}) and their corresponding selection criteria. For example: A3 that was not considered a hit by S, was considered a "possible" potential hit due to its high ΔG.

Based on these criteria, 22 potential inhibitors were selected to be tested *in vitro* and in cells where their efficiency to inhibit mRNA:YB-1 interactions can be measured. To this list, 18 molecules, predicted inefficient, were also added as negative controls (CTRL) in order to have a total number of 40 molecules which is convenient for the experimental assays. However, these 18 molecules were rationally selected from the 87 molecules that stayed in the pocket and for which we have calculated and applied the statistical mechanics-based filter described above and computed their ΔG_{bind}. The selection criteria was based on their structural similarity to F1 (hit validated by NMR spectroscopy) in order to generate an initial QSAR that will help us rationally optimize these molecules later. As for the FDA-approved drugs, we chose all PARP inhibitors, in order to compare with P1; the other 2 non-PARP inhibitors (D2 and D3) were chosen for their scaffold. Figure 2 and Figure 2-source data 1 show the classification and chemical structures of these selected 40 molecules along with their resulting scores and free energy values.

In summary, this computational approach allowed us to identify 22 potential hits from $\sim$ 7 million molecule candidates.

Robust HCS Scoring of endogenous mRPIs in cells with the MT bench assay

In order to score the interaction between mRNAs and YB-1 in cells with an HCS imager, we adapted a method that we recently developed, the MT bench (***Boca et al., 2015***). Briefly, an RBP is brought to the microtubules (MTs) after its fusion to a microtubule binding domain (MBD) so it can be used as a bait for a prey (here, mRNA). In our constructs, an RBD was fused via its C-terminus to a GFP-tag itself fused to the MBD (MBD-GFP-RBP). As MBD, we used the longest isoform of MAPT gene (2N4R-tau), which allows the binding of microtubules in a non-cooperative manner (***Butner and Kirschner, 1991***) and enables the bait protein, for example YB-1, to protrude outward the MT surface several nm away from the microtubule surface, which increases the bait accessibility to ligands (***Boca et al., 2015***); the RBP brought on MTs subsequently interacts with mRNAs in the cytoplasm which results in an enrichment of endogenous mRNAs along the MT network in cells (Figure 3a). To measure the enrichment of poly(A)-mRNA on microtubules, we used *in situ* hybridization with a cy3-labeled poly(dT) probe in fixed U2OS cells (***Lubeck et al., 2014***) which have a well-extended MT network. Importantly, an HCS imager equipped with a water immersed lens (40x, NA=1.1) operating in confocal mode was necessary to reach a sufficiently high lateral resolution and thus clearly distinguish the microtubule network in fluorescence microscopy images (Figure 3b and Figure 3-Figure Supplement1). To detect the presence of baits on MTs, an automatic detection scheme has been implemented using specific criteria such as a low width-to-length ratio of the detected GFP-rich spots (<0.22) keeping only MT-shaped spots (Figure 3b). Details on image acquisitions and statistical analysis are provided in Appendix 3.

Results indicate an accurate detection of MBD-GFP-YB-1-decorated MTs in U2OS cells. In the

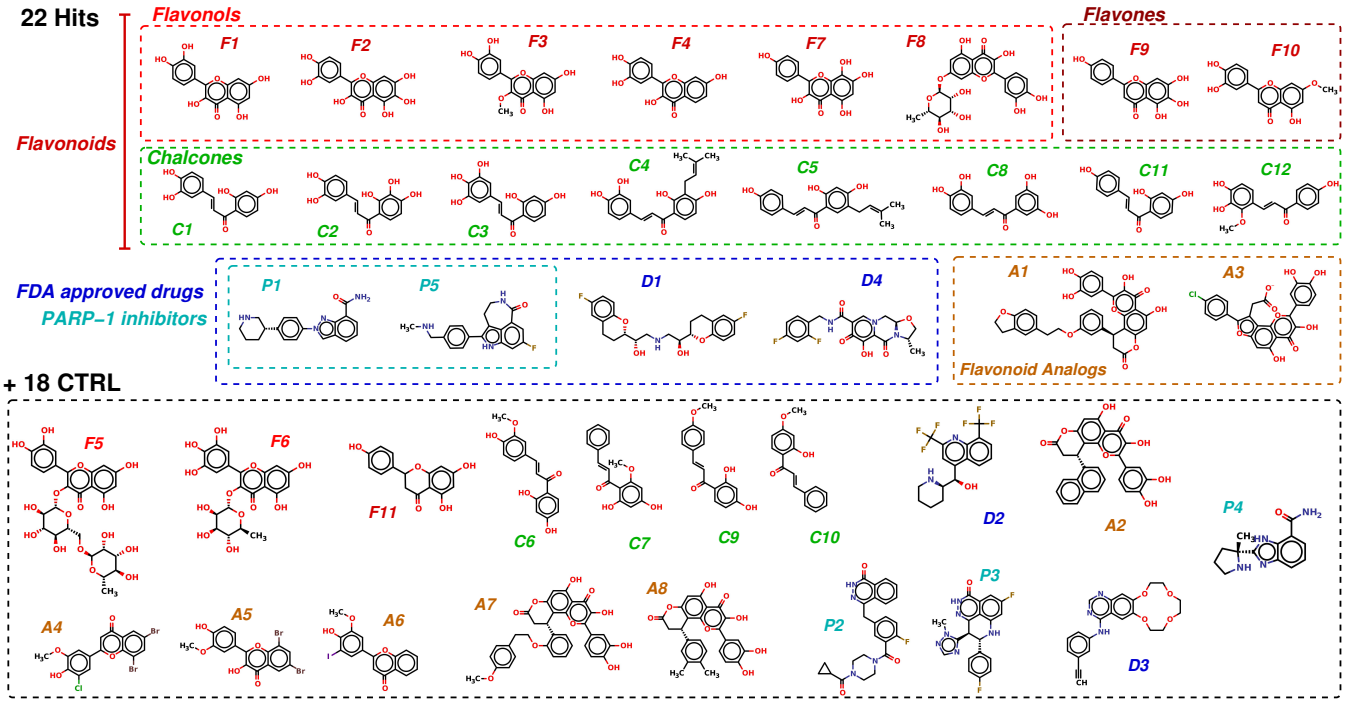


Figure 2. Chemical structures of the selected 40 molecules separated into 22 potential hits and 18 CTRL. Colored dashed boxes assemble hits by subclass and the black dashed box regroups the 18 CTRL. Labels and colored boxes are color coded as function of the family classification: Flavonoids (red) are divided into Flavonols (light red), Flavones (dark red) and Chalcones (green); Flavonoid Analogs in orange, FDA-approved drugs (blue) of which PARP-1 inhibitors (cyan).

Figure 2–source data 1. Classification of the 40 molecules selected using the computational approach. The scoring function $\mathcal{S}$ and the free energy of binding ΔG_{bind} from ABFE simulations used to identify potential hits are reported, along with the average pocket <CSP> from NMR and the average <mRNA enrichment slope> from MT bench assay. The structures of the molecules are illustrated in Figure 2. Confirmed hits *in vitro* and/or *in vivo* are emphasized using a bold font and a color code: red for hits confirmed *in vitro* and *in vivo*, black for *in vitro* only, green for *in vivo* only and blue for a negative control that was found potent *in vivo* but not *in vitro*. Errors on the computed ΔG_{bind} range from 0.34 to 0.97 kcal.mol^{-1}.

selected spots, the mean bait intensity and enrichment in mRNA (ratio of the mean intensity of cy3 in the spots to that in the cytoplasm) were measured (Figure 3c). In contrast to MBD-GFP spots, the enrichment of mRNA in MBD-GFP-YB-1 spots located on MTs increased linearly with GFP spot fluorescence. This result demonstrates the positive correlation between the number of YB-1 brought on MTs and the relative enrichment of mRNAs on the same MTs. Interestingly, the slope thus depends directly on the binding affinity of the bait for mRNAs. We therefore considered the slope as a mRNA affinity score for RBPs brought on MTs. We next estimated the sensitivity of this scoring method by measuring the slopes of 48 positive (MBD-GFP-YB-1) and 48 negative (MBD-GFP) controls from a 96-well plate (Figure 3c; data from all wells are given in Figure 3- Figure Supplement 2). The measured SSMD value (strictly standardized mean difference) for this assay is 8.1, which is the difference of the mean values of the positive and negative controls divided by the standard deviation. A SSMD value of 8.1 corresponds to an efficient assay whatever the estimated strength of the positive controls (***Bray and Carpenter, 2017***). The SSMD value also indicates the sensitivity of the MT bench assay. Here, only molecules that decrease the slope by more than 1/8 of the positive control can be detected. Additional negative control experiments were also conducted using, as baits, 3 different DNA-binding proteins that should not bring mRNAs onto microtubules in the MT bench assay. These proteins are DNA topoisomerase 1 (TOP1), Apurinic/apyrimidinic endonuclease 1 (APE1), and DNA ligase 1 (LIG1). The results represented in Figure 3-Figure supplement 3 confirm that DNA-binding proteins indeed fail to bring mRNA onto the microtubules. In summary, the automatic image analysis that we implemented for the MT bench assay can reliably detect and score the interaction of YB-1 with mRNAs in the cytoplasm with HCS capacity.

MT bench assays measure RBP-specific interactions with mRNAs in cells

Although mRNAs can be detected on microtubules in a 96-well plate setting with an HCS imager, it is critical to estimate whether fusion proteins that confine RBPs to microtubules do not lead to artificial interactions with non-specific transcripts. To this end, we designed an experiment to estimate the enrichments of mRNAs on microtubules in cells expressing MBP-GFP-RBP (mRNA brought on the microtubule with the bait protein). Briefly, cell lysates were incubated with purified MTs reconstituted *in vitro* from sheep brains (Figure 3-Figure Supplement 4a). Therefore, mRNAs were brought onto MTs owing to the presence of MBP-GFP-RBP in cell extracts and subsequently detected from MT pellets by RT-PCR after centrifugation. As a control to probe the influence of RBP confinement on microtubules, we also measured mRNA enrichments by classical RNA immunoprecipitation using magnetic beads (RIP36) with anti-GFP antibody in HEK cells expressing GFP-RBP without the MBD domain. Finally, to analyze whether mRNA enrichment profiles are RBP-specific, classical RIP and microtubule co-sedimentation experiments were performed for YB-1 but also for two additional RBPs, FUS and HuR. RT-PCR analysis were performed over 13 mRNAs including actin and GADPH as abundant mRNA controls (Figure 3d and Figure 3-Figure Supplement 4b and c). Due to their high transfection efficiency, HEK cells were used to perform these experiments. Together the results show a similar profile of mRNA enrichment when the same RBP is expressed in cells, regardless of whether classical mRNA IP or microtubule co-sedimentation was used to purify mRNAs (Figure 3d). In contrast, as expected since each RBP binds differentially to mRNAs, enrichment profiles are much more different when different RBPs were used as baits, regardless of the method used for mRNA purification (MT co-sedimentation or mRNA IP). Therefore, we could reasonably assume that the specific binding of RBPs to mRNAs is at least partly preserved for YB-1, FUS and HuR. However, MBD fusion and the vicinity of MTs can interfere with the binding of RBPs to certain mRNAs. For example, we do observe anti-correlations in the enrichment score for some mRNAs such as CALR mRNA (Figure 3d). In addition, MT co-sedimentation or mRNA IP requires cell lysis, the mRNA enrichment profile that we measured may therefore not totally reflect what is occurring in cells and in the vicinity of microtubules. Finally, the MT bench assay is obviously more adapted to detect interactions of RBPs with cytoplasmic RNAs than nuclear RNAs.

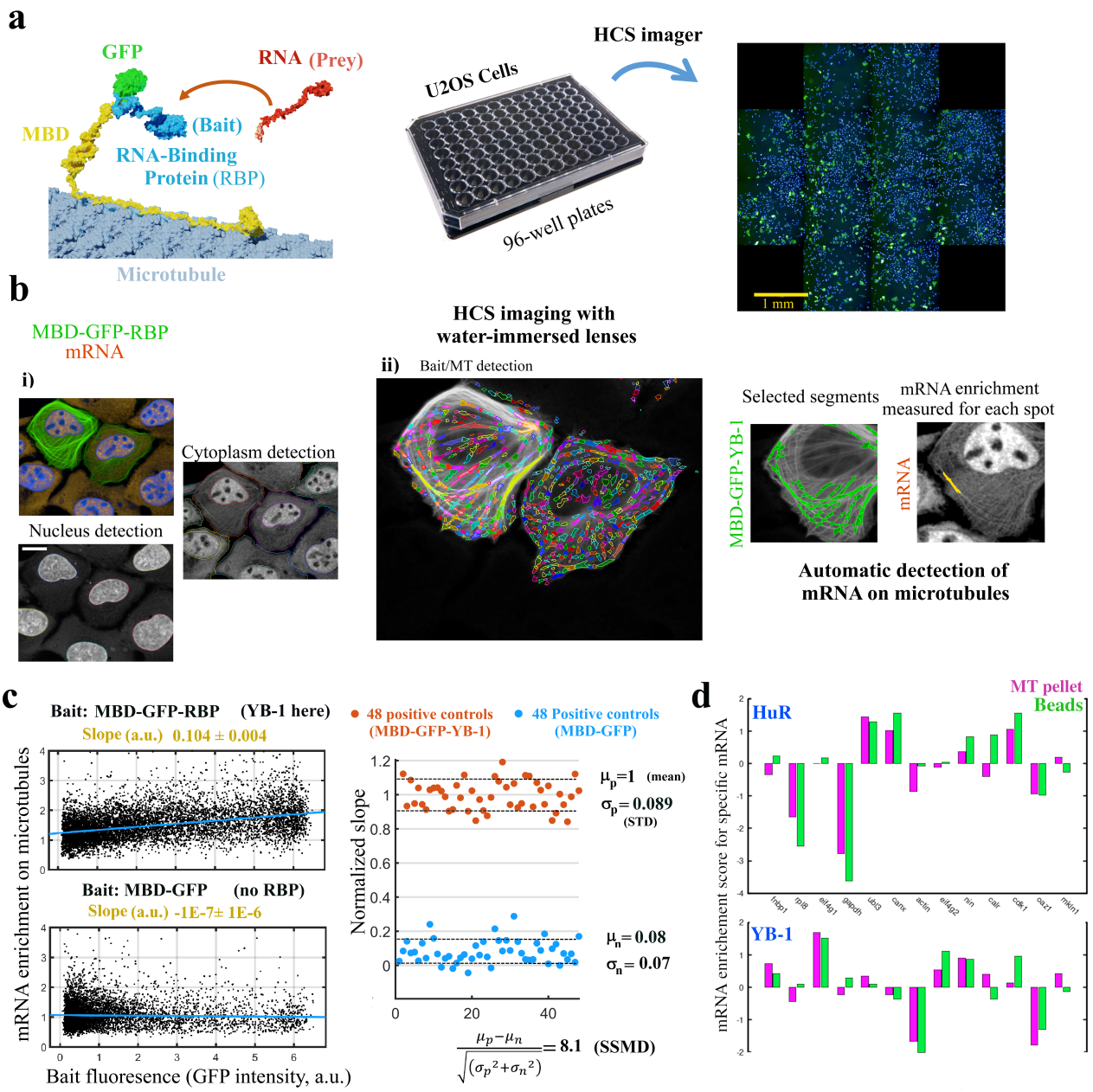


Figure 3. MT bench assay scores mRPIs in cells in a 96-well plate. (a) Left panel: Schematic view of the MT bench technology. A GFP-labeled RBP fused to MBD (Microtubule-Binding domain, yellow) was brought to microtubules in U2OS cells to attract endogenous mRNAs (in red) on the microtubule network (grey). Middle panel: Image of a 96-well plates seeded with U2OS cells. Right panel: Image of a single well processed by HCS imager showing the expression of MDP-GFP-YB-1 in U2OS cells (green). (b) U2OS cells expressing MBD-GFP-YB-1 (bait in green, GFP). mRNAs in red (in situ hybridization, poly(dT) probe). Nuclei in blue (DAPI). The images were obtained with an HCS imager (40x, water immersed objective operating in confocal mode). i) DAPI and the red channel (mRNA) were used to detect automatically the nuclei and cytoplasm, respectively. ii) Using HARMONY "find spots" procedure, elongated spots along the microtubules were detected using the green channel (the bait, RBP). Spots were selected owing to their width-to-length ratio (<0.22) and their enrichment in GFP (YB-1). Scale bar: 20 μm. (c) Left panels: The enrichment of mRNAs in single selected spots (spot/cytoplasm intensity ratio, red channel) and spot bait intensity on microtubules (green channel) show a linear relationship when YB-1 was used as bait. The slope of the regression line reflects the affinity of an RBP for mRNAs. A large number of cells can be analyzed by HCS (>500 cells *per* well with in average 10-50 spots *per cell*). Slopes from linear regression were measured for each well with a 95% confidence interval. Right panel: SSMD value estimated by measuring the normalized slopes in 48 negative controls (MBD-GFP used as bait) and 48 positive controls (YB-1 was used as bait). The SSMD value is 8.1 for a 96-well plate. Spot data from all wells are shown in Figure 3-Figure Supplement 2a. (d) Bar diagram representing the enrichments of 13 different mRNAs measured by RT-PCR after two different purification procedures, co-sedimentation (MT pellet) and immunoprecipitation (Beads), and for 2 different RBPs, YB-1 and HuR; the purification procedures are illustrated in Figure 3-Figure Supplement 4, data and correlation analysis are provided in Appendix 5-Table 5 for 3 RBPs (YB-1, HuR, and FUS). *(continued)*

Figure 3–Figure supplement 1. Image analysis process to quantify mRPIs in cells.

Figure 3–Figure supplement 2. Quality assessment of the MT bench cell assay.

Figure 3–Figure supplement 3. DNA-binding proteins do not bring mRNAs on MTs when used as baits.

Figure 3–Figure supplement 4. The mRNAs brought onto microtubules by MBD-GFP-RBP are RBP-specific.

Figure 3–source data 1. Slope of mRNA enrichement on MTs versus bait fluorescence in selected spots for Figure 3-Figure supplement 3b.

Identification of potent mRNA:YB-1 interaction inhibitors in cells

With the additional 18 CTRL, the 22 ligands that fulfill all the above-mentioned *in silico* criteria were screened by using the MT bench assay. Compound concentration and level of purity were confirmed by NMR spectroscopy. These 40 molecules were then scored in two 96-well plates containing U2OS cells with 4 replicates *per* molecule in cells expressing MBD-GFP-YB-1. Cells were treated with 10 μM of the indicated molecules for 4 h before fixation and analysis with the HCS imager (Figure 4a). Results show a significant decrease in the slope of the mRNA enrichment on microtubules versus bait expression level for 11 of the tested molecules, all of them were already considered as putative hits *in silico*, except C6. These 11 significant hits include 2 flavonols (F2 and F3), 7 chalcones (C1, C2, C3, C6, C8, C11, and C12), a flavonoid analog (A3) and one FDA-approved drug (P1) known as Niraparib, which is a PARP-1 inhibitor notably prescribed for advanced ovarian cancers. Moreover, 17 among the 18 CTRL did not lead to a significant decrease of the slope as expected. The remaining CTRL (compound C6) could be considered a false positive since it does not interact with YB-1 *in vitro*. To ensure that the decrease in the slope was specific to YB-1, we performed the same experiment using two other RBPs, HuR and FUS (Figure 4b). HuR and FUS bind to mRNA via RNA Recognition Motif (RRM), an RNA-binding domain of a different structure (four-stranded antiparallel β-sheet, stacked on two α-helices) that does not harbor a Quercetin-pocket. The presence of HuR and FUS on MTs after their fusion to MBD and their interactions with specific mRNAs onto MTs was confirmed beforehand (Figure 4c). Ten molecules were tested, of which 5 hits (F2, C3, C8, A3, P1) and 5 negative controls (P2, F11, F5, D4, C12), and none of them did significantly affect mRNA:HuR or mRNA:FUS interaction scores (Figure 4b). Hence, the 5 selected hits specifically target mRNA interactions with YB-1.

In vitro validation of targeting the Quercetin-pocket

An *in vitro* validation of the binding of the above selected compounds to the Quercetin-pocket was also conducted using NMR spectroscopy. Here, ligand binding was detected via changes of protein resonances in 2D ^{1}H-^{15}N spectra upon ligand addition using a 1:4 protein:ligand ratio. However, from the 40 molecules selected by the *in silico* approach and tested with the MT bench assay, only 25 (of which 8 CTRLs) were amenable to solution NMR studies (15 of them presented solubility issues, notably the *in cellulo* hit A3). Analysis of the chemical shift data show significant CSPs for the pocket residues for 15 of the molecules being tested including all the 11 putative hits identified with the MT bench assay in cells (except for A3). Average pocket CSPs are reported in Figure 2-source data 1 along with MT bench scores. F1, F4, F6, F8 and F9 did not significantly decrease mRNA:YB-1 interactions in cells even though the amplitude of the CSPs *in vitro* indicate a significant binding to the pocket. Parameters related to the cellular context in which the MT bench assays were performed such as half-life of compounds, potential off-target interactions, membrane permeability and/or selectivity toward the YB-1 targeted pocket most likely account for the discrepancy between cellular and *in vitro* data.

Data mining of ligand-induced CSPs reveals P1 selectivity

To analyze compound selectivity towards the Quercetin-pocket and make a parallel between *in vitro* and cellular results, we implemented an in-depth structural analysis based on the NMR chemical shift perturbations using data mining techniques. The aim is to (i) examine how these ligands target the Quercetin-pocket differently and (ii) identify key residues relevant to differential ligand selectivity. To this end, a principal component analyses (PCA) was performed on a 15 by 20 2D

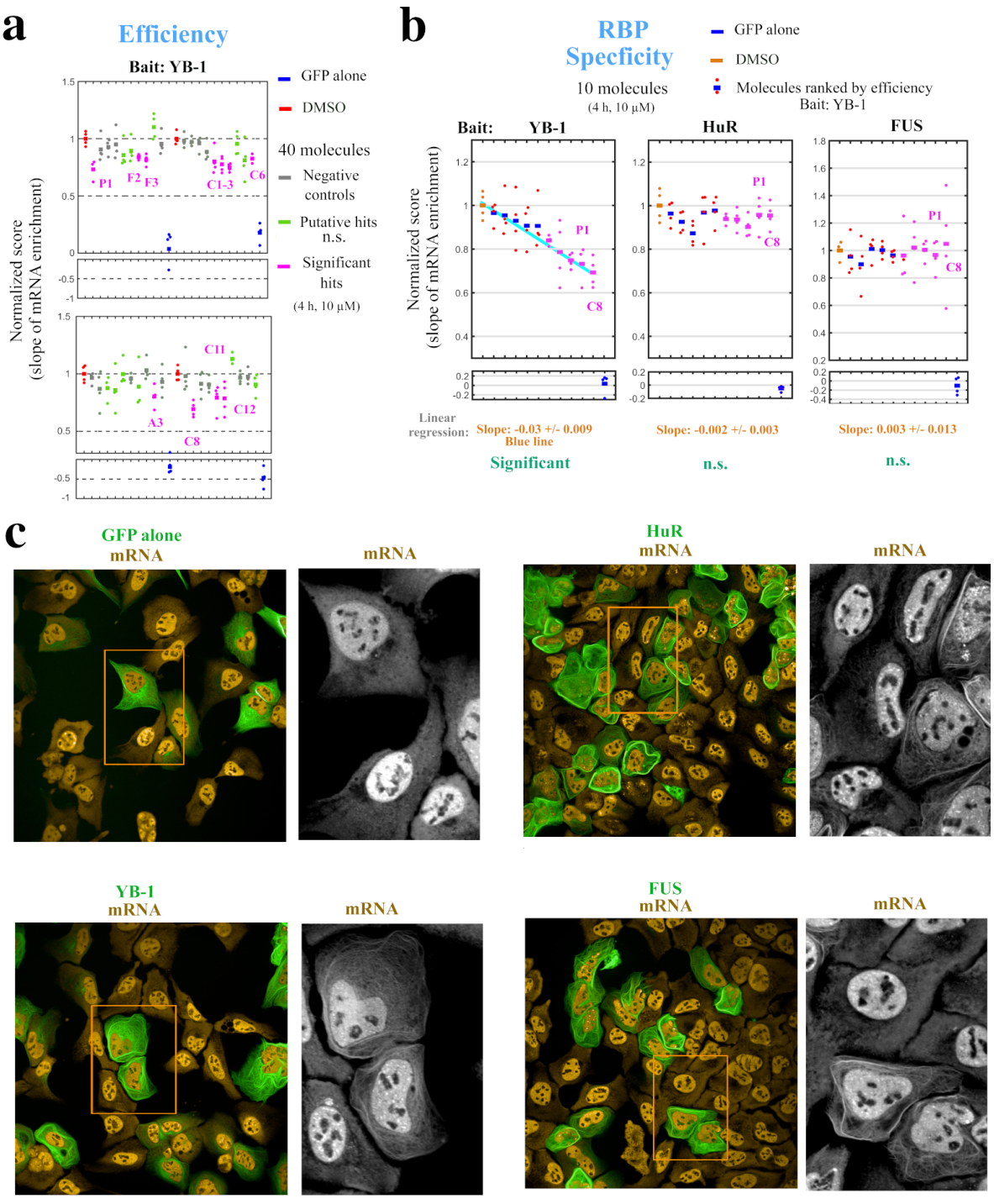


Figure 4. Hits identified with the MT bench assay target selectively mRNA:YB-1 interactions. (a) Normalized mRNA enrichment slope measured with MT bench in U2OS cells expressing YB-1 as bait after 4 h exposure to the selected 40 molecules at 10 μM. The slope of the mRNA enrichment on MT versus bait expression was measured in quadruplet in two 96-well plates (20 molecules *per* plate). Each plot represents a plate; negative controls (grey), non-significant (n.s.) putative hits (green), significant hits (magenta), DMSO control (red), GFP control (blue). Compounds were selected as significant hits (magenta) when p <0.05 according to a paired t-test relative to DMSO controls. (b) The specificity of the molecules to YB-1 in (a) was tested against two other RBP baits, HuR and FUS. Left panel: 5 negative controls (blue, red) and 5 significant hits (magenta), from (a) were selected and ordered on the x-axis according to their efficiency to affect mRNA:YB-1 interactions according to (a). Their interaction score is shown on the y-axis. The blue line represents the decreasing slope with 95% confidence intervals. Middle and right panels: 5 negative controls and 5 significant hits for YB-1, from (a), tested against HuR and FUS, respectively. The same ordering of compounds in x-axis from the left panel was used and a non-significant slope was measured (n.s.). No significant hit was detected for both RBPs. P1 and C8 are labeled in all 3 panels. (c) Images representing the expression and localization of 3 different RBPs used as bait (MBP-GFP-RBP) and a negative control (MDB-GFP) in U2OS cells. The 3 baits used here are HuR (top right panel), YB-1 (lower left panel), and FUS (lower right panel). All the RBP baits tested were successfully detected on MTs (green) and efficiently brought mRNAs onto MTs (orange).

Figure 4–source data 1. Normalized slope values for Figure 4b.

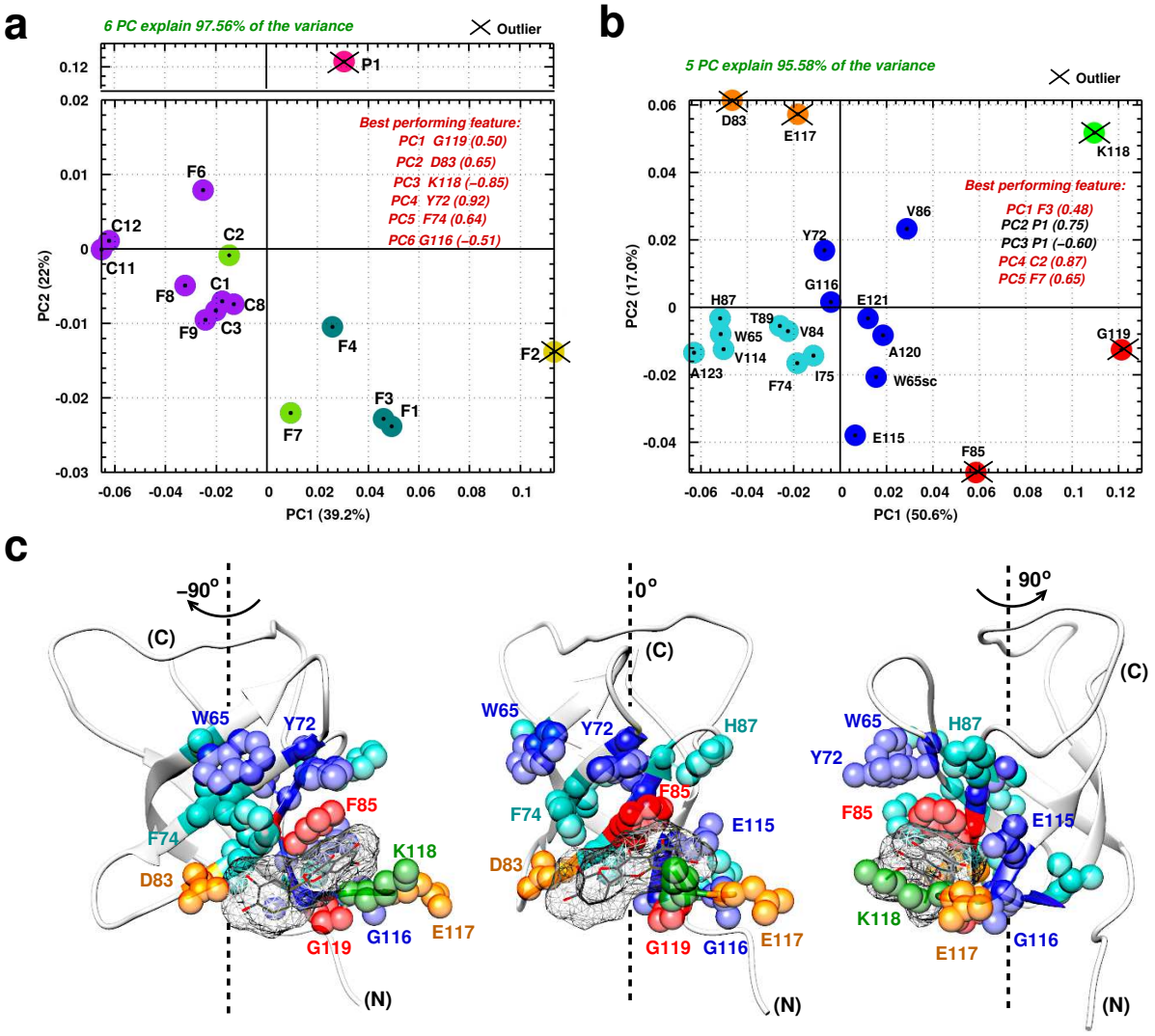


Figure 5. Principal component analysis of ligand-induced CSPs from ^{1}H-^{15}N NMR spectra of ^{15}N-labeled YB-1. (a, b) A 2D scatter plot of PCA results obtained on matrix A, where A= [15 ligands x 20 residues], and its transpose A^T [20 residues x 15 ligands]. The scatter plot shows PCA projected vectors for different ligands (a) and residues (b) on the first and second PC dimension (PC1 and PC2). Different colors indicate 5 clusters found for the 6 (a) and 5 (b) PCs by cluster analysis. The best performing features by PC with absolute highest loading are also listed. Outliers based on SPE and Hoteling's T^2 tests are also indicated. (c) Projection of PCA results on the 3D structure of a YB-1:Ligand complex. The protein is represented in light grey cartoon; the ligand (here F3, taken as an example) in dark grey sticks and a mesh surface. Residues showing significant loadings from PCA are represented in spheres and colored based on the colors of the clusters formed in (b). (N) and (C) indicate the N-and C-terminal, respectively.

Figure 5–Figure supplement 1. Correlation-matrix-based hierarchical clustering of NMR ligand-induced CSPs.

matrix (denoted A) and its transpose A^T built from the average CSPs ($\Delta\delta_{avg}$) of 15 ligands and 20 YB-1 residues; analysis details are described in Appendix 4 section I and results are illustrated in Figure 5. Here, linear dimensionality reduction using singular value decomposition of the data to project it to a lower dimensional space aims to reveal hidden simplified structures. As a result, the accumulative contribution ratio of the first 6 principal components (PCs, linear combinations of the CSPs of YB-1(CSD) residues) for matrix A was 0.97, meaning that these 6 dominant dimensions are likely to describe most contributions to the signal. YB-1 residues that correspond to the best performing features by PC are K118, G119, G116, Y72, F74, and D83 (Figure 5a). On the ligand side, two outliers were detected, P1 (magenta) and F2 (yellow), meaning that each of these 2 ligands target the Quercetin-pocket differently compared to the rest of the ligands. PC results also reveals 5 clusters where each gathers group of similarly-acting ligands (all clusters are color coded in Figure 5). Consistent with their identification as outliers, P1 and F2 each belong to a cluster of their own.

A similar analysis can be performed with the matrix transpose A^T. In this case, 5 PCs explain 95% of the variance with an accumulative ratio of 0.95 (Figure 5b). The ligands that correspond to the best performing features by PC are P1, F3, C2, and F7. In addition, clustering of the PC results shows 5 clusters of similarly-affected residues, and where the 5 PC outliers detected are grouped into 3 sets of YB-1 residues: K118 (green), [F85;G119] (red) and [D83;E117] (orange).

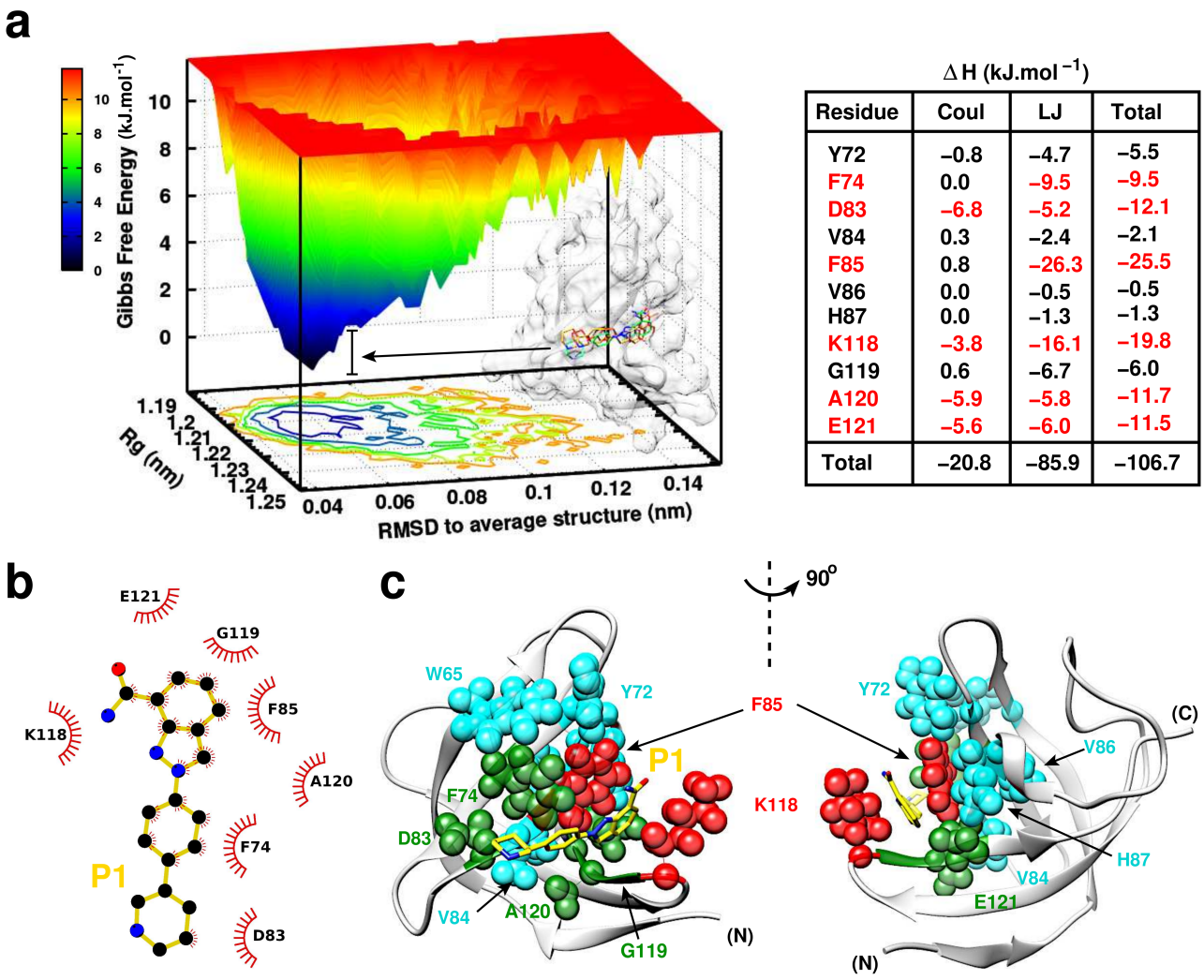
In order to explain the observed residue-related results, a direct comparison with the binding modes obtained by MD was essential. Figure 5c shows a color-coded projection of A^T PCA results on a 3D structure of a YB-1:Ligand complex, where F3 is taken as an example. The 5 residue outliers (forming the first 3 clusters) are residues involved in the direct binding of YB-1 to the ligands (green, red, orange). The two remaining clusters with higher populations (blue and cyan) divide the binding site residues into three sets: (i) direct neighbors of residues making direct interactions with the ligands (such as V86 and G116 that are direct neighbors of outliers F85 and K118), (ii) residues located in the vicinity of the pocket that interact with residues that bind the ligand (such as Y72 and W65sc that are related to F85 via $\pi - \pi$ stacking) and (iii) residues located further away (such as T89, I75, and V114); and where in these cases the observed CSP is due to indirect binding or structural rearrangement.

Cross validation of PCA results between matrix A and its transpose yields striking observations. For instance, P1 is distinguished as the only ligand to appear as an outlier in A (magenta) and as the best performing ligand, with highest loadings, in 2 of the 5 dominant PCs in A^T. Looking from the residue side of things, K118 (green), G119 (red), and D83 (orange) are highly distinguished as best performing residues in A and outliers in A^T. Hence, according to the PCA analysis, the higher selective binding of P1 to the Quercetin-pocket, compared to the other tested ligands, may be due to the interaction of P1 with central (F85, K118, G119) and peripheral (D83, E117) residues.

FDA-approved P1 binds YB-1(CSD) with a K_d of $6\,\mu M$ *in silico* and *in vitro*

Since the FDA-approved P1 was found to have the highest selectivity toward the Quercetin-pocket compared to flavonoid- and chalcone-like molecules, we then further scrutinized the interaction of P1 with YB-1. First, in order to assess whether P1 presents one or multiple binding modes, we conducted 2D ^{1}H-^{15}N NMR titrations. The superposition of the titration spectra produced a straight line (and not curved plots) which is indicative of a single binding mode (Figure 6-Figure Supplement 1a). The multiple binding in general produces curved plots, because the secondary interactions will almost always have different effects on the chemical shifts than the primary interaction (***Craven et al., 1996***; ***Williamson, 2013***). Second, Saturation Transfer Difference (STD) NMR (***Mayer and Meyer, 2001***) was carried out to investigate the binding of P1 to YB-1 (Figure 6-Figure Supplement 1b). The obtained epitope mapping of P1 illustrates which chemical moieties of the ligand are key for molecular recognition in the binding site and allowed us to unequivocally orient P1 in the Quercetin-pocket. The resulting epitope mapping confirms the 3D structure of P1 bound to YB-1 obtained by MD simulations.

Next, the binding mode was assessed by an extended 200 ns MD simulation. The free energy



ΔH (kJ.mol^{-1})

Residue	Coul	LJ	Total
Y72	–0.8	–4.7	–5.5
F74	0.0	–9.5	–9.5
D83	–6.8	–5.2	–12.1
V84	0.3	–2.4	–2.1
F85	0.8	–26.3	–25.5
V86	0.0	–0.5	–0.5
H87	0.0	–1.3	–1.3
K118	–3.8	–16.1	–19.8
G119	0.6	–6.7	–6.0
A120	–5.9	–5.8	–11.7
E121	–5.6	–6.0	–11.5
Total	–20.8	–85.9	–106.7

Figure 6. P1:YB-1 binding mode assessed by MD simulations. (a) Left panel: Free Energy Landscape (FEL) of P1 in complex with YB-1 computed from 200 ns MD simulations. FEL is represented using two structural reaction coordinates: the radius of gyration of the system and the root mean square deviation (RMSD) with respect to the average structure. The zero energy is at 0 kJ.mol^{-1} and corresponds to the most stable conformational states. The free energy scale highlights energy differences (0-12 kJ.mol^{-1}) relative to the global minimum. Radius of gyration and RMSD values are reported in nm. The 3D representation shows "valleys" of low-free energy corresponding to the metastable conformational states of the system, and "hills" account for the energetic barriers connecting these states. The free energy surface is also projected as a 2D "contour plot" on *x*- and *y*-axis. The inset shows an overlay of several conformational states sampled from two low energy wells (indicated by black arrows); the protein is shown as transparent light grey surface. Right panel: Interaction energy contribution (ΔH) of the residues implicated in the binding, along with its Coulomb (Coul) and Lennard-Jones (LJ) contributions, averaged over 200 ns of MD simulation with variant of fluctuations being ± 1.6 kJ.mol^{-1}. The most contributing residues are marked in red. (b) 2D interaction diagram between P1 (gold) and YB-1 residues. (c) 3D representation of the zero-energy complex. The protein is represented in light grey cartoon, P1 in gold sticks. Residues implicated in the binding and showing significant CSP in NMR 2D ^{1}H-^{15}N-SOFAST-HMQC spectrum (Figure 7-Figure supplement 1, top left panel) are represented in spheres: residues with high interaction energy and/or high CSPs are in red, intermediate (green), lower (cyan). (N) and (C) indicate the N- and C-terminal, respectively.

Figure 6–Figure supplement 1. NMR investigation of P1 binding to YB-1.

Figure 6–Figure supplement 2. Mapping the effect of F85A mutation on P1 binding to YB-1 by MD and NMR

Figure 6–Figure supplement 3. NMR investigation of P1 binding to LIN28(CSD) and HuR(RRM2).

landscape of YB-1:P1 complex sampled from MD (Figure 6a) shows a local energy minima over a large free energy space (deep basins, dark blue) indicating that the protein structure has become a minimum energy structure during the simulation period. The overlay of several structures sampled and extracted from the basins shows stability and the same binding mode within fluctuations. The interaction energies (ΔH) averaged over the simulation between P1 and the pocket residues along with its electrostatic (Coul) and Lennard-Jones (LJ) contributions show that $\sim$ 80% of the binding is due to van der Waals (vdW) contacts (-85.9 kJ.mol^{-1} for LJ vs -20.8 kJ.mol^{-1} for Coul). The highest contributing residues to the binding are F85 with -25.5 kJ.mol^{-1} and K118 with -19.8 kJ.mol^{-1}, where P1 is retained in the Quercetin-pocket by a hydrophobic sandwich (Figure 6b and c). A strong $\pi - \pi$ stacking between the indazole ring and F85 from one side and strong vdW interactions with K118 from the opposite side. Moreover, P1 is also retained/pinned by E121 located at the right side of the pocket due to electrostatic and vdW interactions. The middle benzene ring engages in a perpendicular $\pi - \pi$ stacking with F74 (-9.5 kJ.mol^{-1}). In addition, the piperidine cycle of P1 is pinned by D83 located at the left side of the pocket.

Isothermal titration calorimetry (ITC) measurements were also conducted in order to determine the binding affinity of P1 to YB-1 (Appendix 5 Table 1). Results indicate a binding free energy (ΔG_{bind}) of -7.14 kcal.mol^{-1}; which translates into a K_d of $\sim 6\,\mu M$. These results are in line with the calculated ΔG_{bind} from ABFE simulations of -7.24 kcal.mol^{-1} (Appendix 5 Table 2). In addition, both results (ITC and ABFE) show that the driving force for P1's association with its target is enthalpic, meaning an enthalpy-driven association. The observed low enthalpy and entropy values from simulations compared to ITC results are due to an underestimation of the $\pi - \pi$ stacking interactions involving aromatic and non-aromatic groups computed with a point charge force field. However, this did not affect the resulting calculated free binding energy which reproduces the experimental ITC value, within statistical errors.

Since, the key interaction in P1's binding to YB-1 is the strong $\pi - \pi$ stacking with F85, we next decided to change this residue into alanine. MD and NMR results show that YB-1 mutant F85A no longer interacts with P1 (Figure 6-Figure Supplement 2a and b, respectively). The structural investigation of MD results reveal that the binding pocket collapses due to F85A mutation since F85 maintained the aromatic side chains of residues H87, Y72, F74 and W65 along with K118 in place (Figure 6-Figure Supplement 2c). Finally, to ascertain the specificity of YB-1:P1 interactions *in vitro*, we tested whether P1 interacts with two other RBPs LIN28A(CSD) and HuR(RRM2) using 2D NMR. YB-1(CSD) is different in residue composition compared to the LIN28 family (***Moss and Tang, 2003***) (LIN28A and LIN28B), two other CSD proteins with a high degree of sequence homology with YB-1 in humans ($\sim$ 40% residue identities). LIN28(CSD) is structurally similar to YB-1(CSD) with few residue mismatches located in the Quercetin-pocket (mainly Y72/F, G116/S, E117/A, and A120/L) that generate a different structural rearrangement of the side chains. HuR(RRM2) is an RNA-binding domain of a different structure (four-stranded antiparallel β-sheet, stacked on two α-helices) with no Quercetin-pocket. NMR results show that some residues in the LIN28(CSD) pocket were experiencing CSPs but to a significantly lesser extent than YB-1(CSD). In addition, we noticed that CSPs in LIN28(CSD) residues are located outside of the Quercetin-pocket, demonstrating a weak and nonspecific binding to LIN28(CSD) (Figure 6-Figure Supplement 3a). Regarding P1 binding to HuR(RMM2), no relevant interaction was detected (Figure 6-Figure Supplement 3b).

To sum up, P1 is found to bind YB-1(CSD) via vdW interactions (mostly) with a high affinity (K_d of $\sim 6\,\mu M$, measured *in vitro* (ITC) and *in silico* (ABFE)) and with a certain specificity when compared to the two other RBPs used in this study, LIN28(CSD) and HuR(RRM2). Here, MD simulations provided a resolved atomistic picture of the binding mode and revealed the inhibition mechanism. Furthermore, MD and NMR analysis of the F85A YB-1 mutant in complex with P1 emphasize on the leading role of F85 in targeting the Quercetin-pocket.

P1 interferes with RNA:YB-1 interactions *in vitro* and in cells

To put to the test whether P1 can significantly interfere with the binding of YB-1 to mRNAs *in vitro*, the CSPs of YB-1 residues located in the Quercetin-pocket in the presence of 5-nt long poly(C) RNA with or without P1 was analyzed. To calculate the CSPs induced by P1 in the presence of RNA in solution, the YB-1:RNA spectrum was used as a reference (see inset in Figure 7a). Results show additional CSPs associated with P1, apart from those due to RNA:YB-1 interactions, which indicates the presence of both YB-1:RNA and YB-1:P1 complexes in solution. In order to analyze and assess the competitive behavior of P1, we thus considered in detail the chemical shift variations from different ^{1}H-^{15}N-SOFAST-HMQC NMR spectra of ^{15}N-labeled YB-1 in the absence and/or presence of P1 and/or 5-nt long poly(C) RNA. For this, the ligand's ability to compete with RNA on YB-1 binding was evaluated using the scalar product of pair displacement vectors, here denoted $SP(residue)$ (Figure 7a; a detailed description of the analysis is provided in Appendix 4 section II). These vectors correspond to the chemical displacement induced after adding the ligand ($\vec{u}$), RNA ($\vec{v}$) and both RNA+ligand ($\vec{w}$) to YB-1. Hence, YB-1 residues that display chemical shifts moving in opposite directions in the presence of P1 and RNA compared to RNA alone will have a negative SP such as G119, and residues displaying chemical shifts moving in the same direction will have a positive SP such as V84 (see insets in Figure 7a). These observations can be translated into a "competition" or "additive" regime for $SP < 0$ and $SP > 0$, respectively. Among the residues showing competition, several are directly involved in the interaction with RNA such as W65sc, V84, F85, and E82. Their observed CSPs have negative SP values, which is what is expected in a competition for binding. Similar competing behavior was also observed for C8, another hit used here as a positive control (Figure 7-Figure supplement 1 (middle panels) and 2), while P2 (Olaparib, another PARP inhibitor used here as a negative control) showed no effect (Figure 7-Figure supplement 1, lower panels). Although *in vitro* results show that P1 and C8 can interfere with the binding of CSD to RNA, the Quercetin-pocket represents only a part of the RNA:CSD interface which involves at least 4 consecutive nucleotides. However, in agreement with the results of the MT bench assay, many RBPs have to compete with each other to gain access to mRNAs in cells. A compound that interferes slightly with the RNA:YB-1 interface may thus dramatically shift the balance toward RBP competitors and lead to an apparent decrease in the affinity of YB-1 for mRNA in cells which cannot be observed *in vitro*. To further test whether P1 significantly decrease the affinity of YB-1 for mRNA in cells owing to its binding to YB-1(CSD) and not to its PARP inhibition activity, we planned a series of experiments. First, the dose responses of P1, C8, and P2 were analyzed (Figure 7b). The mRNA enrichment slope was measured in quadruplicate in 96-well plates. Results show that P1 and C8 but not P2 displayed a classical dose response with a critical concentration of about $10\,\mu M$, consistent with a low μM range affinity for YB-1 as calculated by free energy simulations and/or by ITC ($K_d \sim 6\,\mu M$ for P1 by ABFE simulations and ITC; and $K_d \sim 4\,\mu M$ for C8 by ABFE simulations). In addition P3, P4, P5, three others PARP inhibitors did not affect mRNA:YB-1 interactions in our first screen at $10\,\mu M$ (Figure 4b) and no significant CSPs in YB-1 residues were detected in the presence of P2, P3 or P4 *in vitro* (Figure 2-source data 1). Only P5 was found to bind to the Quercetin-pocket of YB-1 but with a significantly lower affinity than P1, which was also confirmed by ABFE simulations (estimated $K_d \sim 60\,\mu M$; but it can range between 30 and $90\,\mu M$ considering the 0.5 kcal.mol^{-1} error on the computed value). Therefore any potential effect resulting from the inhibition of PARP by P1 on mRNA:YB-1 interactions in cells can be ruled out. Altogether, the results obtained *in silico* by MD simulations, *in vitro* by NMR and, in cells with the MT bench assay point toward the ability of P1 to compete with mRNA for binding YB-1 at μM concentrations.

May P1 affect YB-1 cellular functions related to mRNAs?

Finding functional cellular assays that would reveal a phenotype specific to a general RBP (such as YB-1) is not an easy task, and it is even more difficult with YB-1 since it binds non-specifically to most mRNAs (***Singh et al., 2015***) as shown from CLIP analysis (***Wu et al., 2015***). In addition, YB-1 is an abundant protein in cancer cells. In HeLa cells, the cellular model chosen here, YB-1 abundance

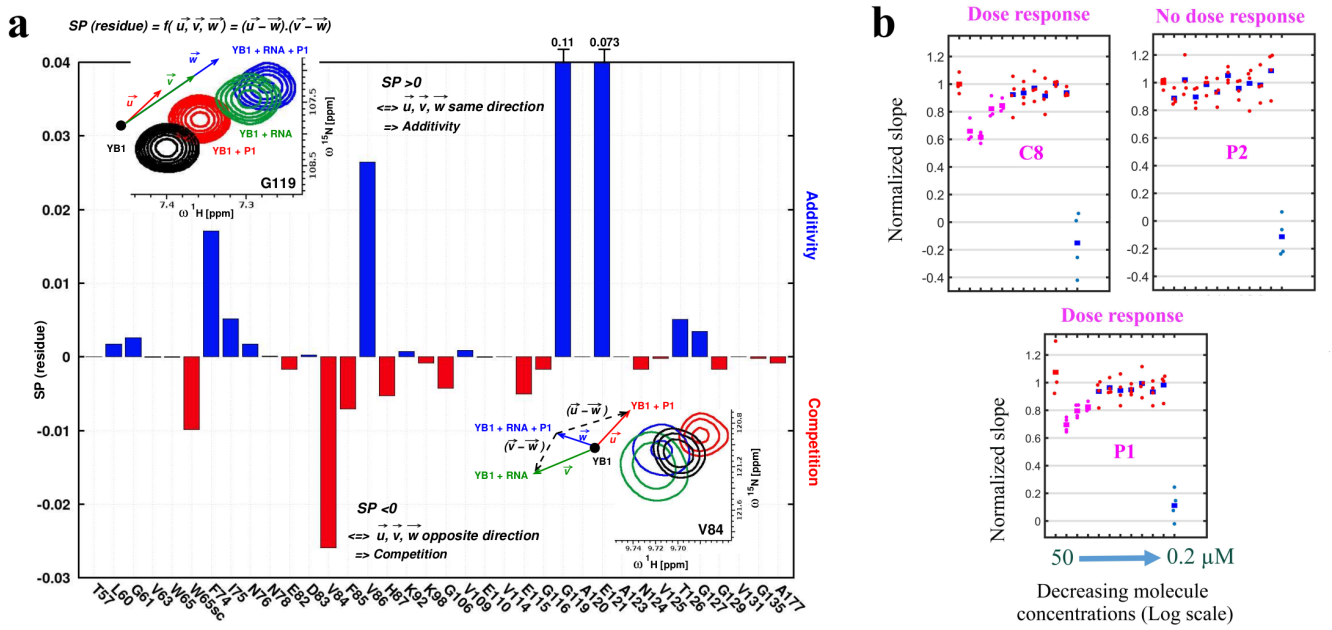


Figure 7. P1 interferes with RNA:YB-1 interactions *in vitro* and in cells. (a) Histogram showing the competitive behavior of P1 using $SP(residue)$ calculated based on CSPs extracted from ^{1}H-^{15}N-SOFAST-HMQC spectra of ^{15}N-labeled YB-1 in the absence and/or presence of P1 and 5-nt long poly(C) RNA. The insets show a zoom-in on residue V84 (exhibiting competition) and G119 (exhibiting additivity) from overlaid NMR spectra of YB-1 alone (black) and in the presence of: P1 (red), RNA (green), and both P1+RNA (blue); an illustration of the associated displacement vectors $\vec{w}$, $\vec{v}$ and $\vec{u}$ (same color code as NMR), relative to YB-1 alone (large black dot), and the pair vectors ($\vec{v}$ - $\vec{w}$) and ($\vec{u}$ - $\vec{w}$) (black dashed arrows) on which the scalar product was calculated are also indicated. (b) Dose response plots of P1, C8, and P2 in U2OS cells expressing YB-1 as bait (MT bench) following 4 h exposure at decreasing concentrations from 50 to 0.2 μM (quadruplicate in 96-well plate).

Figure 7–Figure supplement 1. Mapping Ligand/RNA competition on binding YB-1 using NMR.

Figure 7–Figure supplement 2. Investigating C8 binding to YB-1 in the presence of RNA by MD and NMR.

Figure 7–source data 1. Normalized slope values for Figure 7b.

is ranked 248 among all proteins with 1.7 million copies *per* cell (***Nagaraj et al., 2011***).

In order to reveal a phenotype related to the interaction between YB-1 and mRNA, we decided to expose HeLa cells to elevated P1 concentrations (increase from 20 to $100\,\mu M$) during 2 h. Below $50\,\mu M$, no change in YB-1 and mRNA distribution could be noticed in the cytoplasm of HeLa cells (Figure 8-Figure supplement 1a and b). Above $200\,\mu M$, cells underwent massive death. However, at $100\,\mu M$ of P1, even if nonspecific activity cannot be avoided so close to the toxic concentration threshold, we detected the presence of YB-1-rich granules in the cytoplasm with two different anti-YB-1 antibodies (Figure 8a and Figure 8-Figure supplement 1c). YB-1 granules appeared only in few cells treated with P1, which may reflect a cell phase dependency, but repeatedly in many different and independent experiments. While YB-1 granules can be considered nonspecific stress granules related to cellular stress, they were distinct from stress granules (SGs) triggered by Arsenite (***Khong et al., 2017***; ***Bounedjah et al., 2014***), a potent and widely used inducer of SGs in cells (large cytoplasmic SG in Figure 8a, upper right panel). Consistent with a decreased affinity of YB-1 for mRNAs at high P1 concentrations, mRNAs were poorly recruited in YB-1 granules compared to SGs formed in the presence of Arsenite. This is illustrated in the scatter plots of Figure 8a for SG (Arsenite, orange scatter plot) and YB-1 granules (P1, blue scatter plot). None of the other PARP-1 inhibitors led to the formation of YB-1-rich granules (Figure 8a, Figure 8-Figure supplement 1a and b). In addition, YB-3, which shares an identical CSD with YB-1, is also significantly recruited in these granules. As a control, HuR which is not a target of P1 is recruited to a lesser extent than YB-3 (Figure 8-Figure supplement 1d).

We then considered whether P1 affects YB-1 function related to mRNA translation. As YB-1 binds to most non polysomal mRNAs (***Singh et al., 2015***), YB-1 may regulate the overall translation rates in cells by controlling the switch from a polysomal state (active) to a non polysomal state (dormant). When mRNAs are blocked in a non polysomal state, cellular translation rates should decrease. Accordingly, in a recent work, we showed that YB-1 unwinds non polysomal mRNAs in a way that facilitates the translation from dormant to active state. In agreement with another report in myeloma cells (***Bommert et al., 2013***), we also showed that decreasing the expression of YB-1 reduces mRNA translation in HeLa cells (***Budkina et al., 2021***). Hence, we tested whether P1 may interfere with mRNA translation. In a previous report, but after long PARP inhibitor treatment (72 h), a decrease in translation level was measured by puromycin incorporation because of the activation of PARP-1 by small nucleolar RNAs (snoRNAs) in the nucleolus (***Kim et al., 2019***). To limit this bias, we chose to measure mRNA translation after short P1 treatment (2 h) and compared the results obtained with P1 with two other PARP-1 inhibitors that do not target the Quercetin-pocket, P2 and P3. The incorporation of puromycin to nascent peptide chains during translation is significantly reduced in cells treated with P1 but not P2 and P3 at concentrations as low as $2.5\,\mu M$ (Figure 8-Figure supplement 2). We then decided to directly probe whether the inhibition of mRNA translation detected with P1 was YB-1-dependent. For this HeLa cells were pre-treated with two different siRNA to decrease endogenous YB-1 levels and with siNEG as a negative control. Cells were then exposed to indicated molecules ($10\,\mu M$, 2 h) and briefly exposed with puromycin before fixation to estimate global mRNA transition at the single cell level. In cells treated with two different siRNAs targeting YB-1, we observed that P1 did not significantly impair mRNA translation whereas a significant decrease in mRNA translation was measured with the negative control siRNA, (siNEG)-treated cells (Figure 8b, Figure 8-Figure supplement 3). In contrast, P2 had no measurable impact on mRNA translation in both siRNA- and siNEG-treated cells.

Since YB-1 expression is associated to elevated cancer cell proliferation (***Evdokimova et al., 2009***; ***Alkrekshi et al., 2021***) , we also probed whether P1 reduces the cell number in a YB-1 dependent manner. To this end, we measured the number of HeLa cells plated at low density in 12-well plates after having decreased, or not, YB-1 levels with siRNA. In addition, we used P2 and P3 as negative controls as they inhibit PARP-1 but do not target YB-1 like P1. Figure 8-Figure supplement 5 shows that all PARP-1 inhibitors decrease the cell number, albeit to a higher extent with P3. However, both P2 and P3 further decrease the number of cells in siRNA-treated cells compared to

siNEG-treated cells (with significant differences at 5 μM), which may be due to reduced resistance to stress when YB-1 expression is decreased. In contrast, P1 rather further decreases the number of cells in siNEG-treated cells when YB-1 levels are high (non-significant variations but opposite to those observed with P2 and P3). The separation distance between cells also decreases significantly in YB-1-poor cells (siRNA) treated with P1 compared to siNEG-treated cells, in contrast to P2 and P3 (Figure 8-Figure supplement 4). A short distance of separation between cells may be due to colony formation when cells were plated at low density and allowed to grow for 48 h. Therefore, while we may have expected a higher sensitivity of cells to P1 when the YB-1 expression is low, in contrast, P1 seems to further decrease cell number when YB-1 level is high, which may be due to a gain of toxic or cytostatic function, notably a decreased translation rate as observed in Figure 8b. However, further analyses need to be undertaken to document this point.

Therefore, the appearance of YB-1-rich granules and the inhibition of the YB-1-dependent mRNA translation in HeLa cells are consistent with P1 interfering with mRNA:YB-1 interactions. To which extent P1 may affect YB-1-related functions in cells remains to be investigated in details.

Discussion

In this study, we introduce an integrative approach that leads to the identification of several effective YB-1 inhibitors in the low micromolar range selected computationally and validated *in vitro* by NMR spectroscopy and in cells using the MT bench assay. Here, the MT bench was adapted to score small molecules targeting RBP interactions with endogenous mRNA in cells. The MT bench assays can notably fill the gap between *in vitro* and functional assays by probing whether the interaction of a selected RBP with mRNAs is affected in a cellular context but not that of other RBPs. Our results validate the reliability of the MT bench assay in detecting and scoring YB-1 interactions with mRNA in 96-well plates (SSMD >8, Figure 3c).

Using a rationally designed large-scale computational approach, 22 potentially effective compounds (along with 18 CTRL) targeting a druggable pocket located at the YB-1(CSD):RNA interface (the Quercetin-pocket) were selected to be tested (Figures 1 and 2). An *in vitro* structural validation using protein-based NMR data, which is necessary to ascertain their capability in targeting the Quercetin-pocket, was also conducted when possible. The MT bench assay revealed that 11 out of the 22 selected hit compounds significantly decrease the interaction of YB-1 with mRNA in cells. In contrast, when 5 hits were tested with two other RBPs (FUS and HuR), no decrease of RPIs was observed (Figure 4). Here, endogenous poly(A)-tailed mRNA was used as bait to detect mRNA:YB-1 interactions since YB-1 is a general mRNP factor in the cytoplasm. New developments may enable to target RBPs interacting with specific RNA (mRNA encoding a specific gene, ncRNA, etc.), which may be helpful for the challenging issue of developing molecules that would target an interaction between a specific RNA and a specific RBP.

All of the molecules selected in this study are multi-aromatic ring systems that are sandwiched in the Quercetin-pocket. Besides their common anchoring key residues, F85 and K118, other interactions with neighboring residues, from both sides of the pocket, are needed to stabilize the ligands in the binding pocket, as clearly indicated by the PCA analysis (Figure 5). The PCA analysis also identified key residues implicated in the high selectivity of P1 toward the Quercetin-pocket that can thus be used to rationally optimize our leads. As for the computational approach implemented in this study, the validity of its predictive potential was challenged *in vitro* and in cells. *In vitro*, 15 predicted hits out of 17 were confirmed to bind YB-1 in the targeted pocket by NMR (yielding 88% success rate); and in cells, 11 out of 22 were found to inhibit RNA:YB-1 interactions (50% success rate, while only C6 of the 18 negative controls emerged as a significant hit). Other factors may play a role in yielding negative results for predicted hits such as off-targets and cell permeability, which is precisely the point of using the MT bench assay. Here, we managed to balance computational accuracy and cost by using the point-charge force field CHARMM. However further efforts can be applied to optimize the computational approach by using advanced multipolar and polarizable force fields in order to improve the ranking and reduce the errors (***Gresh et al., 2015***).

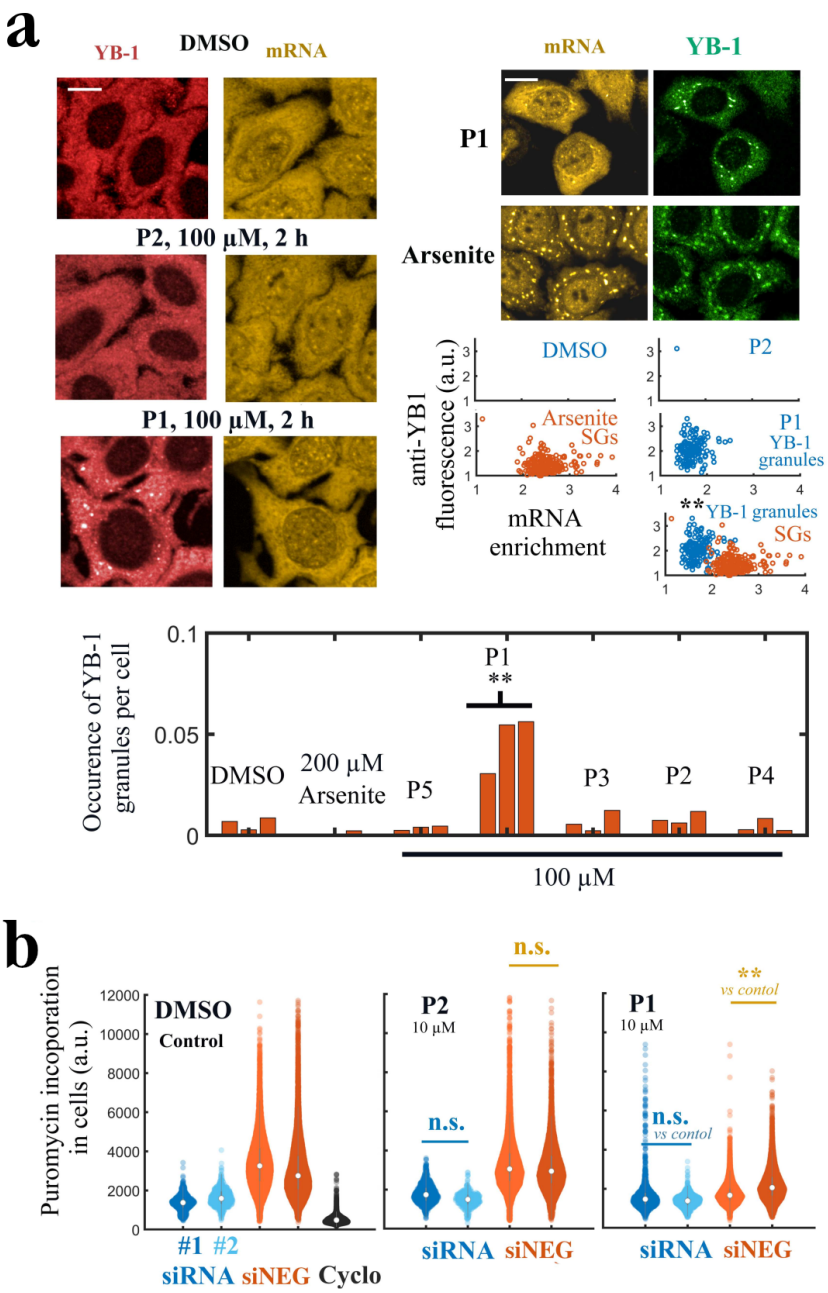


Figure 8. P1 alters cellular YB-1 functions in HeLa cells independently of PARP activity. (a) YB-1 distribution in cells exposed to P1. Upper left panels: HeLa cells treated with P1 and P2 at 100 μM for 2 h. Upper right panels: Images of cells exposed to P1 and Arsenite. The corresponding scatter plots indicate a difference in mRNA-enrichment of stress granules (Arsenite, orange scatter plot) and YB-1 granules (P1, blue scatter plot). Lower panel: Occurrence of YB-1-rich granules in cells exposed to indicated treatment for 2 h. **, p<0.01; t-test with two tails (triplicates). Scale bars: 20 μm. (b) HeLa cells pretreated with two different siRNAs to decrease endogenous YB-1 levels or siNEG (negative control), and then exposed to indicated molecules (10 μM, 2 h) followed by a brief exposure to puromycin before fixation to estimate global mRNA transition at the single cell level. **, p<0.01; t-test with two tails.

Figure 8–Figure supplement 1. Cytoplasmic YB-1-rich granules in HeLa cells at high P1 concentrations (≥ 20 μM).

Figure 8–Figure supplement 2. Monitoring puromycin incorporation in HeLa cells.

Figure 8–Figure supplement 3. Detection of puromycin incorporation in HeLa cells.

Figure 8–Figure supplement 4. P1 increases the separation distance between cells when comparing YB-1-rich with YB-1-poor HeLa cells.

Figure 8–Figure supplement 5. Compared to other PARP-1 inhibitors, P1 further decreases the number of cells *per* well in YB-1 rich cells.

Figure 8–source data 1. Anti-YB-1 fluorescence intensity values for Figure 8-Figure supplement 4b.

Figure 8–source data 2. Data points of the calculated distance between closest neighbors for Figure 8-Figure supplement 4c.

Figure 8–source data 3. Number of cells per well data points for Figure 8-Figure supplement 5.

Owing to the rapid feedback between atomistic, chemical, structural, and cellular data integrated here, our ligand screening strategy for RPI inhibitors may also be refined.

Apart from the FDA-approved inhibitors, these molecules belong to the flavonoid family with known anti-inflammatory or anti-tumor activity in humans (***Panche et al., 2016***), except for C3, C4, C8-C10, and the A series for which no known activity was reported. However, given the many biological processes on which these compounds interfere (***Panche et al., 2016***) and their numerous targets, their selectivity may be doubtful. Nevertheless, a rational optimization of these Chalcone- and Flavonol-like molecules guided by quantum chemical calculations and relative free energy simulations may increase their affinity and selectivity to YB-1.

The top lead, on which we focused the rest of our analysis, is P1, an FDA-approved PARP-1 inhibitor. Based on an exhaustive structural analysis, P1 displays the highest selectivity by targeting key residues from all sides of the pocket via mostly vdW interactions (Figures 5 and 6). Moreover, it represents a clear specificity to mRNA:YB-1(CSD) interactions when compared to LIN28(CSD) and HuR(RRM2) *in vitro*. Altogether, the results obtained *in silico* by MD simulations, *in vitro* by NMR, and in cells with the MT bench point toward an inhibition of mRNA:YB-1 interaction by P1 at low micromolar concentrations, which is consistent with the moderate affinity of P1 for the YB-1 Quercetin-pocket (K_d $\sim$6 μM). In addition, functional assays show a global decrease in YB-1-related mRNA translation, ~~cell proliferation,~~ and the appearance of YB-1-rich granules in Hela cells treated with P1. Given the positive regulation of mRNA translation and the negative regulation of stress granule assembly exerted by YB-1 in HeLa cells (***Budkina et al., 2021***), these results thus did not exclude the possibility that P1 may target YB-1 functions related to mRNA in HeLa cells.

As YB-1 is a secondary target of P1 and finding secondary targets of FDA-approved PARP-1 inhibitors has been a recent concern due to their different indications and multiple adverse effects, especially P1 (***LaFargue et al., 2019***; ***Knezevic et al., 2016***), we may consider whether impairing YB-1 function in cells may provide a rational explanation for the observed adverse effects of P1. For instance, the pronounced hematological adverse effects, particularly thrombocytopenia (***LaFargue et al., 2019***) may be explained by the role YB-1 in megakaryocyte versus erythroid differentiation (***Bhullar and Sollars, 2011***). However, more data are needed to explore the putative involvement of YB-1 in P1 adverse effects.

In summary, we have developed an integrative approach to specifically target RPIs in cells with small molecules. While the data are promising for RNA:YB-1 interactions and provide a first proof of concept, we would like to stress out that this is not yet sufficient to assert that this approach could be successful with all RBPs. Separate studies are needed to validate the MT bench for other RBPs. YB-1 is an "ideal" target because it has a single cold-shock domain and a druggable pocket, which may not be the case for other RBPs. In addition, many RNPs harbor several RNA-binding domains, which may reduce the sensitivity of our method when a specific domain is targeted by small molecules because the other domains would contribute to the binding to mRNAs. However, a single RNA-binding domain may be isolated and used as bait for the MT bench assay to overcome this obstacle. Developing molecules that would target a specific domain may be sufficient, to modulate the biological function exerted by the full length protein.

Methods and Materials

Computational Methods

System Setup and Molecular Dynamics Simulations

For this study, the following systems were considered for MD simulations: WT YB-1 protein (apo form), WT YB-1:RNA(C5) complex (holo form), YB-1:Ligand complexes and YB-1-F85A mutant in complex with P1 Niraparib.

The starting 3D coordinates of YB-1 CSD used the NMR solution structure PDB code 1H95 (***Kloks et al., 2002***) and the YB-1:RNA(C5) complex was constructed using as a template the crystal structure of YB-1 cold-shock domain in complex with UCAACU (PDB ID 5YTX (***Yang et al., 2019***), resolution

1.55 Å). The protein sequence is of 85 amino acid in length going from A45 to G129. The YB-1:Ligand complexes were generated, in a next step, using as a building block for docking screened ligands, an MD sampled open-state of YB-1. As for the mutant F85A bound to P1, it was generated from the lowest energy state MD refined WT YB-1 structure in complex with P1.

All MD simulations were carried out with GROMACS software package version 2018.2 (***Abraham et al., 2015***) using the additive force field CHARMM27 for proteins (***MacKerell et al., 1998***) and nucleic acids (***Hart et al., 2012***) with periodic boundary conditions. Ligands parameters were obtained using SwissParam (***Zoete et al., 2011***) which provides topology and parameters for small organic molecules compatible with the CHARMM all atoms force field. The protonation states of the residues were adjusted to pH 7.6 (pH used in our NMR experiments). The systems were centered and solvated in a triclinic box of TIP3P (***Jorgensen et al., 1983***) water model with 1.4 nm distance between the boundary of the box and the system in question. A [KCl] of $100\,\mathrm{mM}$ was used and counter-ions were added to neutralize the system. Each system was first energy minimized using 50000 steps of steepest descent, then heated to 298 K at constant volume for 500 ps and equilibrated in the NPT ensemble at p = 1 atm for 500 ps which was followed by 10 or 200 ns of NPT production run depending on the aim of the computational protocol. A 10 ns of MD production run was used for MD pose refinement of YB-1/ligand complexes in order to allow protein rearrangement upon ligand binding. These refined poses were then used for subsequent ΔG calculations, respectively. The long MD simulations were used to study the evolution as a function of time of YB-1 apo state, YB-1:RNA and YB-1:P1 complex. The Velocity Rescaling (***Bussi et al., 2007***) (with τ = 0.1 ps) and Parrinello-Rahman (***Parrinello and Rahman, 1981***) methods were used for temperature and pressure control, respectively. The equations of motion were propagated with the leap-frog (***Van Gunsteren and Berendsen, 1988***) algorithm and the time step was Δt = 2 fs. The particle mesh Ewald (PME) (***Darden et al., 1993***; ***Essmann et al., 1995***) method was used for electrostatic interactions, with grid spacing of 1.6 Å, a relative tolerance of 10^{-5}, an interpolation order of 4 for long-range electrostatics, and a cutoff of 14 Å together with a 12 Å switching threshold for LJ interactions. All bonds with hydrogen atoms were constrained with LINCS (***Hess, 2008***).

Virtual Screening

The virtual screening part of the *in silico* approach is divided into 3 parts: (i) the pharmacophore-based screening; (ii) the automated blind docking of FDA-approved drugs; and (iii) the physico-chemical and purchasability filters applied. The procedure is detailed in Appendix 2.

Statistical Mechanics-Based Filter

After a visual inspection of the docking results (***Fischer et al., 2021***) and carefully selecting the docked poses of the ligands, these docked complexes were reassessed using a dynamic view. Compared to the static docking approach, here the complex is simulated in presence of an explicit solvent, where real enthalpic and entropic contributions are taken into account. Short MD simulations of 10 ns are sufficient to refine the docked pose and to check the stability of the ligand in the binding site. Ligands that left the binding site during this 10 ns were discarded, and only ligands that stayed in the binding site were considered for the following.

Knowing that the ligand can be destabilized by the water molecules interacting with its unbound side (on the water accessible surface), we defined a first observable that allows us to estimate if the ligand will reside in the pocket for a longer time. This can be translated into the ability of the water network interacting with the bound ligand to extract it from the binding site. In order to estimate this, we calculated the difference between ligand-protein (ΔH_{LP}) and ligand-water (ΔH_{LW}) interaction energies along the MD simulation. This enthalpic observable ($\Delta\Delta H_{(LPvsLW)} = \Delta H_{LP} - \Delta H_{LW}$) was then averaged over the 10 ns of MD, in order to score the ligand by its preference to reside in the pocket $\Delta\Delta H_{(LPvsLW)} < 0$ or in the solvent $\Delta\Delta H_{(LPvsLW)} > 0$. Standard deviations ($\sigma$) fluctuated between 2 and 7 kcal.mol^{-1} depending on the size and the complexity of the ligand.

Another observable denoted c was also defined based on the ability of the ligand to make more

than one interatomic contact *per* interacting atom. This was expressed as the ratio between (i) the total number of unique atomic contacts made between the ligand and the pocket residues (ii) and the number of ligand atoms in contact. Ligand-Pocket interatomic contacts were calculated from the MD refined structure. Interatomic contacts are defined based on their vdW radius, and englobe polar and nonpolar interactions.

With these two observables we were able to define a simple weighted scoring function in order to determine the best-performing ligands:

$$S = \sum_{i=1}^{2} w_i(Obs_i) \tag{1}$$

with $w_c = 4$ and $w_{\Delta\Delta H} = -1$ introduced to differently weight the two observables. This scoring function is rather more qualitative than quantitative and allowed us to sort our ligands into potential hits or not. Ligands with $S > 0$ were considered as potential hits. However, ligands with $S < 0$ were also considered as potential hits as long σ for $\Delta\Delta H_{(LPvsLW)}$ was higher than the actual average value. This filter proved to be computationally efficient (short 10 ns MD) for screening a large number of ligands.

Absolute Binding Free Energy Simulations

Using a non-physical thermodynamic cycle, the absolute binding free energy of YB-1:ligand complexes (ΔG_{bind}) was calculated as sum of free energy change of formation of protein-ligand complex formation ($\Delta G_{complex}$) and the free energy of desolvating the ligand (ΔG_{solv}); to which an analytical correction term for adding restrains on the decoupled ligand was added ligand (ΔG_r).

The free energy difference between two end states was estimated using the Bennett Acceptance Ratio (BAR) (***Bennett, 1976***). Here, the ratio of weighted average of Hamiltonian difference of two given states is calculated using multiple intermediate states defined by the coupling parameter λ to monitor the alchemical transformation. Hence, the Hamiltonians for the states were determined by combined Hamiltonians for the end states A and B. The linear relationship $H_\lambda = H_A + \lambda(H_B - H_A)$; $0 \leq \lambda \leq 1$ leads to a Hamiltonian representing states A and B, respectively. The initial and final states are defined as A ($\lambda = 0$) where the ligand is absent and B ($\lambda = 1$) where the ligand is fully grown.

In these equilibrium simulations, the system is coupled/decoupled by applying a scaling parameter λ to the nonbonded interactions, which switches between the initial ($\lambda = 0$, state A) and final state ($\lambda = 1$, state B). The interval $0 < \lambda < 1$ was divided into 40 equally spaced windows. First, the LJ interactions with soft-core potentials (***Beutler et al., 1994***) are fully grown, followed by the electrostatics in the presence of the full vdW interactions, thereby avoiding the need for soft-core electrostatic potentials. For each of these steps, the system was re-equilibrated for 500 ps followed by 2 ns of dynamics in the NPT ensemble during which information was accumulated. For the solvation free energies, the system was-re-equilibrated in the NPT ensemble for 200 ps and information was accumulated for 1 ns.

Auxiliary restraints were used to prevent the ligand from leaving the binding site when the native ligand-receptor interactions were turned off alchemically. These restraints restrict both the position and the orientation of the ligands and are defined relative to the receptor. This free energy cost can be evaluated analytically using Equation 2 (***Boresch et al., 2003***):

$$\Delta G_r^{VBA,0} = -RT \ln \left[\frac{8\pi^2 V^0}{(r_{\alpha A,0}^2 \sin\theta_{A,0} \sin\theta_{B,0}} \frac{(K_r K_{\theta A} K_{\theta B} K_{\Phi A} K_{\Phi B} K_{\Phi C})^{1/2})}{(2\pi RT)^3} \right] \tag{2}$$

,where R refers to ideal gas constant, T is temperature in Kelvin, V^0 is standard system volume for 1 molar concentration, r_0 is reference distance for restraints, θ_A and θ_B are reference angles for restraints, K_x refers to strength constant of distance (r_0), two angles (θ_A, θ_B), and three dihedrals (Φ_A, Φ_B, Φ_C). The ligands were restrained by means of one distance and force constant of

1000 kcal.mol^{-1}.nm^{-2}, two angles, and three dihedral harmonic potentials with force constant of 10 kcal.mol^{-1}.rad^{-2}.

Protein expression and purification

The recombinant His6-tagged YB-1 fragment (^{1}Met-^{180}Gly) from the human full-length YB-1 was first cloned into the pET22b expression vector at NdeI/XhoI sites. BL21 (DE3) competent E. coli cells were transformed with the constructed plasmid pET22b-YB1_1-180 and grown at $37\,^{\circ}\mathrm{C}$ in 1L 2YT-ampicillin medium (non-labeled proteins) or in minimal medium M9 supplemented with 15NH4Cl (labeled proteins). The protein expression was induced by IPTG $1\,\mathrm{mM}$ added at OD_{600nm} = 0.7. The culture was grown at $37\,^{\circ}\mathrm{C}$ for 4 h and cells were harvested and washed with 20 mL of cold $20\,\mathrm{mM}$ Tris-HCl buffer, pH 7.6, containing $100\,\mathrm{mM}$ KCl. The cell pellet was resuspended in 10 mL of buffer A ($20\,\mathrm{mM}$ Tris-HCl, pH 7.6, 2 M KCl, 0.5 mM DTT, 1 mM PMSF, 10 mM Imidazole, and EDTA-free protease inhibitor Cocktail (Roche)) and cells were disrupted by sonication on ice (Bioblock Vibracell sonicator, model 72412). The cell lysate was centrifuged at $4\,^{\circ}\mathrm{C}$ for 30 min at 150,000 × g in a TL100 Beckman centrifuge.

The YB-1 (1-180) protein fragment was purified following the manufacturer's recommendations (Qiagen). Briefly, the supernatant was incubated for 2 h at $4\,^{\circ}\mathrm{C}$ with Ni^{2+}-NTA-agarose (Qiagen) (20 mg of proteins/ml of resin) pre-equilibrated in buffer A. The resin was then washed extensively with buffer A containing 10 mM imidazole and by reducing progressively the KCl concentrations (from 2 M till 0.5 M). The elution of the protein was performed by adding 250 mM imidazole in buffer A and fractions were pooled and diluted 25x with 20 mM Tris-HCl, pH 7.6, 0.5 mM DTT, 1 mM PMSF in order to incubate them with protease-free RNase A (Thermo Scientific) for 90 min at room temperature. The protein pool was finally re-purified on the same conditions as described above, dialyzed against 20 mM Tris-HCl, pH 7.6, 0.5 M KCl and stored at $-80\,^{\circ}\mathrm{C}$.

Site-directed mutagenesis of the human YB-1 coding gene was carried out directly on the pET22b-YB-1_1-180 expression plasmid by using the "Quikchange II XL site-directed mutagenesis kit" from Stratagene and appropriate oligonucleotides (Eurofins Genomics). The introduced mutation (F85A) was validated by DNA sequencing (Eurofins Genomics). Overexpression and purification of YB-1 (1-180 aa) mutant F85A were performed by following the same protocol detailed above.

NMR spectroscopy

All NMR experiments, protein- or ligand-based were performed at 600 MHz on a Bruker AVIII HD spectrometer equipped with a triple-resonance cryoprobe. All samples were prepared in a final volume of $200\,\mu\mathrm{L}$ using 3 mm diameter tubes. NMR data were processed with Topspin 4.0 (Bruker). Assignment of ^{1}H and ^{15}N chemical shifts of YB-1(1-180 aa) was retrieved from our previous study (***Kretov et al., 2019***) and from the results obtained by ***Zhang et al.*** (***2020***).

Characterization of the purchased compounds, solubility and stability assessment

All compounds were purchased from Molport, except for F3 and F6 that were purchased from SigmaAldrich, C2, C5, C6, C7, C8 from CarboSynth and C3 from Ambinter, and had purity > 97% (compound IDs and supplier codes are provided in Appendix 5-Table 3; their chemical structures are displayed in Figure 2). Marvin was used for drawing, displaying and characterizing chemical structures, Marvin version 19.16.0, ChemAxon (https://www.chemaxon.com). Purity and solubility were verified by acquiring ^{1}H NMR spectra for each compound dissolved in 100% DMSO-d6. Next compound solubility was checked in aqueous buffer (50 mM phosphate buffer at pH 6.8 and 298 K), to ensure a 1 mM final concentration, by measuring peak integrals from ^{1}H-NMR spectra compared to an internal reference. All occurring peak variations due to instability or solubility issues were monitored over time within a 48 h time period by acquiring 1D ^{1}H NMR spectra at regular intervals (***Sreeramulu et al., 2020***). Instability issues due to fast degradation were mostly observed for some of the flavonoids.

From the 40 molecules purchased, 15 represented solubility issues and thus were not amenable for NMR studies with YB-1. These molecules are: F10, C4, C5, C7, C9, C10, A1 to A8 and D3. Hydrophobic buffers such as MOPS can be used instead of phosphate buffer to solubilize these ligands, however this exceeds the scope of this paper. A prior testing of how the YB-1(1-180) fragment will behave in a different buffer environment should be evaluated beforehand.

Chemical Shift Perturbation Analysis and Titration

Free, ligand- and RNA-bound protein samples were prepared in NMR buffer (20 mM Tris, pH 7.6, containing 100 mM KCl, 10% D2O) supplemented with SUPERase·in RNase Inhibitors (Eurogentec) for RNA samples. All the protein-ligand samples were prepared in a 1:4 protein:ligand ratio. Typically, the final protein and ligand concentrations were $50\,\mu M$ and $200\,\mu M$ respectively, and the final volume was $200\,\mu L$. However, due to low solubility a ratio of 1:1.8 and 1:1.25 was used for F7 and F9, respectively. For samples containing RNA, the protein:RNA ratio was 1:1.2 and the protein:RNA:ligand ratio for competition experiments was 1:1.2:4. A DMSO-d6 percentage of 2 % was maintained in all experiments. For P1 titration essay, a $50\,\mu M$ protein solution was incubated with increasing ligand concentrations from 10 to $1000\,\mu M$, where a constant percentage of DMSO-d6 (2%) was maintained. The number of titration points was 16. Ligand and RNA binding to YB-1 were investigated using 2D ^{1}H-^{15}N SOFAST-HMQC (***Schanda and Brutscher, 2005***) at 298 K. The number of dummy scans and scans was respectively set to 16 and 256. Data were acquired with 2048 points along the direct dimension and with 128 t_1 increments with a relaxation delay of 0.2 s. Spectral widths were set to 12.5 ppm (centered at 4.7 ppm) in the ^{1}H direction and 30 ppm (centered at 118 ppm) in the ^{15}N dimension. Shaped pulse length and power were set by considering an amide ^{1}H bandwidth of 4.5 ppm and a chemical shift offset of 8.5 ppm.

Ligand binding was followed by CSP analysis for which the weighted average chemical shift values were calculated and normalized according to Equation 3:

$$\Delta\delta_{avg} = \sqrt{0.5[\Delta\delta_H^2 + (0.14\Delta\delta_N)^2]} \tag{3}$$

where $\Delta\delta_{avg}$ is the average CSP at a given ratio, $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical shift changes in the ^{1}H and ^{15}N dimension, respectively (***Williamson, 2013***). The NMR data analysis and interpretation approaches implemented and adapted in the purpose of this article are presented in full detail in Appendix 4. These include three data mining techniques: principal component analysis (PCA), correlation-matrix-based hierarchical clustering and scalar similarity measure. Even though the results of the correlation-matrix-based hierarchical clustering performed here are in line with the PCA, it was less sensitive and informative for ligand selectivity (Figure 5-Figure supplement 1 and Appendix 4 Section I).

STD Experiment and P1 Ligand Mapping

P1 ligand resonances were assigned through 2D ^{1}H COSY, ^{1}H-^{13}C-HSQC and ^{1}H-^{13}C-HMBC spectra acquired on a 2 mM P1 solution in 50 mM phosphate buffer (pH 6.8) and 100 mM KCl at 298 K.

STD experiments were acquired on a YB-1:P1 sample prepared in a 50-fold excess of ligand ($500\,\mu M$) at 283 K using a pseudo-2D Bruker pulse scheme (stddiffesgp.3) with excitation sculpting (***Hwang and Shaka, 1995***) for water suppression and a spinlock to suppress protein signals. The number of dummy scans and scans was respectively 32 and 1024. On-resonance irradiation was applied on one of the protein methyl resonance arising at -0.53 ppm, where no signal coming from the ligand is observed. The off-resonance carrier was set to 40 ppm, where no protein signals are visible. Selective pre-saturation of the protein was achieved by a cascade of 50 ms Gaussian-shaped pulses (***Bauer et al., 1984***) corresponding to a total saturation time of 2 s.

Bruker AU program "stdsplit" was used to process data. Integrals corresponding to the reference spectrum off-resonance spectrum (I_{off}) and to the difference spectrum between (off- and on-resonance (I_{diff}) were extracted and used to calculate the fractional STD (A_{STD}) and the STD

amplification factor STD_{AF} (***Mayer and Meyer, 2001***) using Equations 4 and 5, respectively:

$$A_{STD} = \frac{I_{diff}}{I_{off}} = \frac{I_{off} - I_{on}}{I_{off}} \tag{4}$$

and

$$STD_{AF} = A_{STD} \frac{[L]_T}{[E]_T} \tag{5}$$

where $[L]_T$ and $[E]_T$ are the total ligand and protein concentrations, respectively. Relative STD percentages were derived by normalizing all STD integrals against the highest one obtained (assigned to a value of 100 %).

Isothermal Titration Calorimetry measurements of YB-1/P1 binding

ITC experiments were carried out at $25\,^{\circ}\mathrm{C}$ with a MicroCal PEAQ-ITC isothermal titration calorimeter (Malvern Instruments). The protein sample was dialyzed against the ITC buffer (20 mM Tris-HCl, pH 7.6 containing 100 mM KCl and 2% DMSO). The protein concentration in the microcalorimeter cell (0.2 mL) was fixed at $14\,\mu\mathrm{M}$. 26 injections of 1.5 μL of P1 at $200\,\mu\mathrm{M}$ (resuspended in ITC buffer) were carried out at 90-s intervals, with stirring at 650 rpm and a reference power set at 11 μcal.s^{-1}. In experiments with 5 nt-long poly(C) DNA (DNA(C5)), titration was carried out in the same buffer (without DMSO), 18 injections of 2.0 μL of C5 at $200\,\mu\mathrm{M}$ and a reference power set at 5 μcal.s^{-1}. Data were analyzed using the Microcal PEAQ-ITC Analysis Software and fitted using a one set-of-site binding model. All titrations were performed in triplicate.

MT bench

Cellular plate preparation for imagery

The MT bench assay was performed using bone osteosarcoma U2OS cells (ATCC HTB-96), a human cell line that was provided by O. Kepp (Gustave Roussy, Cell Biology Platform, Villejuif, France). U2OS cells were cultured at $37\,^{\circ}\mathrm{C}$ in a humidified atmosphere with 5% CO2 in Dulbecco's modified Eagle's medium (DMEM, Life Technologies) supplemented with 10% FBS (Fetal Bovine Serum, Life Technologies) and 1% penicillin/streptomycin. The confluence of cells was verified every 5 days and cell were confirmed mycoplasma-free.

Cells were seeded on black 96-well plates cell carrier ultra (PerkinElmer) at a density of 16.000 cells *per* well using the liquid handler BRAVO from Agilent equipped with a 96-LT (Large Tips) head. After 24 h incubation in a humidified incubator at $37\,^{\circ}\mathrm{C}$ with 5% CO2, cells were transfected with $0.4\,\mu\mathrm{g}$ of indicated MBD-GFP-RFP plasmid for the positive control condition or with $0.2\,\mu\mathrm{g}$ of MBD-GFP plasmids by using 0.2 μL lipofectamine 2000 (Invitrogen) in optiMEM based on the optimization of the transfection conditions. The transfection complexes were prepared manually and transfection was done using the liquid handler BRAVO with specific transfection protocols depending on the type of plate that was prepared (Optimization, SSMD value, Hit Identification, IC50 determination).

Cells were treated in quadruplicate during 4 h at $37\,^{\circ}\mathrm{C}$ using 0.1% DMSO for the control wells at 10 μM of the different compounds for hit identification and with 10 concentrations ranging from $0.098\,\mu\mathrm{M}$ to $50\,\mu\mathrm{M}$ of the same compound for dose response assessment. The molecules were diluted in the culture medium, with a 2-fold serial dilution for the IC50 determination, and the treatment was made using the liquid handler BRAVO. A double fixation methanol/ParaFormAldehyde (PFA) was used to maintain the cellular protein cytoskeletal structure and allowing a good visualization of the microtubules. Cells were first fixed with ice-cold methanol 100% for 10 min at $-20\,^{\circ}\mathrm{C}$, washed with PBS and then further fixed with 4% PFA in PBS freshly prepared for 10 min at RT. After fixation, cells were incubated with oligo-dT-[Cy3], diluted in SSC 2X, 1 mg.ml^{-1} yeast tRNA, 0.005% BSA, 10% dextran sulfate, 25% formamide, for 2 h at $37\,^{\circ}\mathrm{C}$ for RNA visualization. Wash steps were performed using 4X and then 2X SSC buffer (0.88% sodium citrate, 1.75% NaCl, pH 7.0). Cell nuclei were stained with DAPI ($0.1\,\mu\mathrm{g.mL}^{-1}$) for 5 min at RT. All the washing steps were performed with the

Thermo Scientific Wellwash Versa Microplate Washer and the additions of the different solutions were done with the VIAFLO Electronic multichannel pipettes from Integra. Image acquisition was performed atomically with the Opera Phenix® Plus High Content Screening System. Image analysis was performed with the HARMONY v4.8 software. Details on image acquisition and statistical analysis are provided in Appendix 3.

RT-PCR analysis of RBP specificity

10^6 HEK cells were plated in 6-well plates and transfected with the indicated plasmids with Lipofectamine 2000^{TM} reagent (Invitrogen). 24 hours after transfection, cells were placed on ice for 30 min and lysed in 200 μL of lysis buffer (50 mM TrisHCl, pH 7.0, 50 mM NaCl, 1 mM EDTA, 0.05% sodium deoxycholate, 1% Triton X-100, 0.1% SDS, 1 mM PMSF, protease and RNAse inhibitors). Tubulin was purified from sheep brain as previously described (***Méphon-Gaspard et al., 2016***). Tubulin concentration was determined by spectrophotometry using an extinction coefficient of 1.2 $mg^{-1}.cm^2$ at 278 nm. Tubulin polymerization was initiated by placing the ice-cold cuvette (1 cm light path) at $37\,^{\circ}C$ in a PTI QuantaMaster 2000-4 thermostated spectrofluorimeter. The kinetics of microtubule assembly was then immediately monitored by $90\,^{\circ}C$ angle light scattering at 370 nm. Microtubules were then taxol-stabilized ($5\,\mu M$ taxol, $40\,\mu M$ tubulin).

Cell lysates were centrifuged at 20, 000 g for 1 h at $16\,^{\circ}C$ and the supernatant was collected. 10 μL of microtubule solution was added to 200 μL of cell supernatant, incubated for 15 min at $16\,^{\circ}C$ and centrifuged at 20, 000 g for 30 min at $16\,^{\circ}C$. The microtubule pellet was resuspended in 100 μL of lysis buffer and again centrifuged at 20, 000 g for 30 min. After discarding the supernatant, RNA was purified from the pellet with Tri-Reagent (Molecular Research Center, Inc)) RNA quality and quantity was assessed by UV-spectrometry (nanodrop). RT-PCR reactions were performed using impromII Reverse transcriptase and GoTaq® qPCR Master Mix on a 7500 Applied BiosytemsTM block. RNA quantification results obtained with the microtubule pellet were compared to those obtained from the whole cell lysate. The oligo probes used for the RT-PCR analyzed are listed in Appendix 5-Table 4.

For mRNA purification classical magnetic beads, HEK cells expressing indicated plasmids we lysed under conditions mentioned above. The purification assays were performed using Dynabeads® Protein G Kit (Invitrogen) with anti-GFP antibody (monoclonal antibody, Invitrogen A11120, clone 3E6, IgG2a) in the same buffer used to isolate mRNA in a microtubule pellet, except the incubation time (here overnight in a cold room). RT-PCR analysis was performed as described above (results are listed in Appendix 5-Table 5).

Functional assays in HeLa cells

Cell culture and transfection

HeLa cell lines (American Type Culture Collection, USA) were cultured at $37\,^{\circ}C$ in a humidified atmosphere with 5% CO2 and maintained in the high glucose formulation of DMEM (Life Technologies) supplemented with penicillin G 100 $U.ml^{-1}$, streptomycin 100 $\mu g.mL^{-1}$ and fetal bovine serum (FBS) 5% (10% for HeLa cells; Thermo-Fisher). The absence of mycoplasma was tested regularly to prevent any inference with the obtained results. The cell line identity was tested and authenticated (see Supplementary file 1).

The cells were grown in 24 or 96-well plates and transiently transfected with siRNA to decrease endogenous YB-1 levels with 2 different siRNAs (siRNA-1: [sense 5'-(CCACGCAAUUACCAGCAAA)dTdT-3', anti-sense 5'-(UUUGCUGGUAAUUGCGUGG)dTdT-3']; siRNA-2 which targets the 3'UTR of YB-1 mRNA was used for the addback experiments [sense 5'-(GAUUGGAGCUGAAGACCUA)dTdT-3', antisense 5'-(UAGGUCUUCAGCUCCAAUC)dTdT-3']. The negative siRNA (1027310, Qiagen), siNEG, was applied in the same concentration as the two siRNAs. The mix of $1\,\mu g$ siRNA or siNEG in 300 μL optiMEM with 0.8 μL lipofectamine was left for 20 min at room temperature and added to cells for 3 h, after that the solution was removed and the usual media was added to the well. Efficiency control was performed by immunofluorescence (Figure 8-Figure supplement 3a). We obtained

clusters of cells expressing endogenous YB-1 coexisting in the same sample with clusters of cells that displayed a significantly reduced expression of endogenous YB-1. Only the cells with a low YB-1 expression were retained for analysis (Figure 8-Figure supplement 3b).

Cellular translation assays

Hela cells treated with puromycin (10 μg.ml^{-1}) for 10 min prior to fixation after washing out puromycin were fixed with 4% PAF for 30 min at 37 °C and subjected to immunoblotting using puromycin antibody (Merck, MABE343). For the negative control, cells were treated with cycloheximide (100 μg.mL^{-1}) prior to the addition of puromycin. The anti-puromycin fluorescence in the cytoplasm was detected automatically using the Opera Phenix® Plus High Content Screening System (PerkinElmer). The cytoplasm was detected automatically using the HARMONY v4.8 software.

Detection of YB-1-rich granules

HeLa cells were subjected to indicated treatments for 2 h. Cells were then fixed with methanol for 20 min at −20 °C, followed with 4% paraformaldehyde for 30 min at 37 °C. Immunofluorescence was performed with anti-YB-1-1 (rabbit polyclonal, Bethyl Laboratories, Montgomery, USA), anti-YB-1-2 (Anti-YBX1 antibody produced in rabbit, HPA040304, Sigma-Aldrich), Anti-YB-3 (Anti-YBX3 antibody produced in rabbit, HPA034838, Sigma-Aldrich) and Anti-HuR (antibody produced in mouse (3A2), 390600, Thermo Fisher Scientific). mRNA was detected by in situ hybridization as above-mentioned.

Quantifications were performed with Opera Phenix® Plus High Content Screening System (PerkinElmer) in confocal mode. The HARMONY v4.8 software was used to detect and measure the number of cells having YB-1-rich granules, and the fluorescence intensity in the granules and in the cytoplasm for both and/or the number of SGs *per cell* (These values are directly accessible by selecting them in the "spot analysis" parameters). The mRNA enrichment in YB-1-rich granules was measured by dividing the mean mRNA intensity in granules with the mean mRNA intensity in the cytoplasm.

Cell Number assay

HeLa cells were treated with siRNA-1 or siNEG overnight to decrease YB-1 level in most cells in siRNA-treated cells (Figure 8-Figure supplement 4b). Then, Hela cells were plated at low density (10^5) in 12-well plates and treated-with indicated molecules for 48 h. After cell fixation, cells were stained with anti-tubulin and DAPI. The number of cell was measured by an automatic detection of cell nuclei (HARMONY v4.8 software) as well as the distance between nearest neighbors.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 895024 (MSCA-IF-2019 to K.EH.). We thank the Centre Régional Informatique et d'Applications Numériques de Normandie (CRIANN), Project 2022010. This work was supported in part by the Région Ile-de-France (SESAME grant n°15013102), and Genopole (SATURNE grant 2020, HCS imager). This work was also supported by the INSERM PRI grant "RaPiD". We gratefully acknowledge the Genopole Evry, the University of Evry, and INSERM for constant support of the laboratory. We thank Pr. Olga I. Lavrik and Dr. Maria V. Sukhanova for the careful reading of the manuscript and valuable discussions. This paper is dedicated to the memory of our co-worker Guillaume Lambert who participated in our early work to decipher YB-1:RNA Interactions at the structural level.

Additional Information

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 895024 (MSCA-IF-2019 to K.EH.). This work was supported in part by the Région Ile-de-France (SESAME grant n°15013102),

and Genopole (SATURNE grant 2020, HCS imager). This work was also supported by the INSERM PRI grant “RaPiD”.

Additional Files

Data Availability

All data are available within the Article, Supplementary Files and Appendices, or available from the corresponding authors on reasonable request. Source data for figures 2, 4d, 7b, Figure 3-Figure supplement 3, Figure 8a, Figure 8-Figure supplement 4b-c, Appendix 5 Table 1 and Appendix 5 Figure 1 are also provided with the paper.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015; 1-2:19–25. https://www.sciencedirect.com/science/article/pii/S2352711015000059, doi: https://doi.org/10.1016/j.softx.2015.06.001.

Alkrekshi A, Wang W, Rana PS, Markovic V, Sossey-Alaoui K. A comprehensive review of the functions of YB-1 in cancer stemness, metastasis and drug resistance. Cellular Signalling. 2021; 85:110073. https://www.sciencedirect.com/science/article/pii/S0898656821001625, doi: https://doi.org/10.1016/j.cellsig.2021.110073.

Baltz AG, Munschauer M, Schwanhäusser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Molecular cell. 2012; 46(5):674–690.

Bargou RC, Jürchott K, Wagener C, Bergmann S, Metzner S, Bommert K, Mapara MY, Winzer KJ, Dietel M, Dörken B. Nuclear localization and increased levels of transcription factor YB-1 in primary human breast cancers are associated with intrinsic MDR1 gene expression. Nature medicine. 1997; 3(4):447–450.

Bauer C, Freeman R, Frenkiel T, Keeler J, Shaka AJ. Gaussian pulses. Journal of Magnetic Resonance (1969). 1984; 58(3):442–457. https://www.sciencedirect.com/science/article/pii/0022236484901483, doi: https://doi.org/10.1016/0022-2364(84)90148-3.

Bennett CH. Efficient estimation of free energy differences from Monte Carlo data. Journal of Computational Physics. 1976; 22(2):245–268. https://www.sciencedirect.com/science/article/pii/0021999176900784, doi: https://doi.org/10.1016/0021-9991(76)90078-4.

Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. Chemical Physics Letters. 1994; 222(6):529–539. https://www.sciencedirect.com/science/article/pii/0009261494003971, doi: https://doi.org/10.1016/0009-2614(94)00397-1.

Bhullar J, Sollars VE. YBX1 expression and function in early hematopoiesis and leukemic cells. Immunogenetics. 2011; 63(6):337–50. https://www.ncbi.nlm.nih.gov/pubmed/21369783, doi: 10.1007/s00251-011-0517-9.

Boca M, Kretov DA, Desforges B, Mephon-Gaspard A, Curmi PA, Pastré D. Probing protein interactions in living mammalian cells on a microtubule bench. Scientific reports. 2015; 5(1):1–11.

Bommert KS, Effenberger M, Leich E, Kuspert M, Murphy D, Langer C, Moll R, Janz S, Mottok A, Weissbach S, Rosenwald A, Bargou R, Bommert K. The feed-forward loop between YB-1 and MYC is essential for multiple myeloma cell survival. Leukemia. 2013; 27(2):441–50. https://www.ncbi.nlm.nih.gov/pubmed/22772059, doi: 10.1038/leu.2012.185.

Boresch S, Tettinger F, Leitgeb M, Karplus M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. The Journal of Physical Chemistry B. 2003; 107(35):9535–9551. https://doi.org/10.1021/jp0217839, doi: 10.1021/jp0217839.

Bounedjah O, Desforges B, Wu TD, Pioche-Durieu C, Marco S, Hamon L, Curmi PA, Guerquin-Kern JL, Piétrement O, Pastré D. Free mRNA in excess upon polysome dissociation is a scaffold for protein multimerization to form stress granules. Nucleic acids research. 2014; 42(13):8678–8691.

Bray MA, Carpenter A. Advanced assay development guidelines for image-based high content screening and analysis. Assay Guidance Manual [Internet]. 2017; .

Budkina K, El Hage K, Clément MJ, Desforges B, Bouhss A, Joshi V, Maucuer A, Hamon L, Ovchinnikov L, Lyabin DN, Pastré D. YB-1 unwinds mRNA secondary structures in vitro and negatively regulates stress granule assembly in HeLa cells. Nucleic Acids Research. 2021; https://doi.org/10.1093/nar/gkab748, doi: 10.1093/nar/gkab748.

Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. The Journal of Chemical Physics. 2007; 126(1):014101. https://aip.scitation.org/doi/abs/10.1063/1.2408420, doi: 10.1063/1.2408420.

Butner K, Kirschner MW. Tau protein binds to microtubules through a flexible array of distributed weak sites. The Journal of cell biology. 1991; 115(3):717–730.

Camborde L, Jauneau A, Briere C, Deslandes L, Dumas B, Gaulin E. Detection of nucleic acid–protein interactions in plant leaves using fluorescence lifetime imaging microscopy. Nature protocols. 2017; 12(9):1933–1950.

Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell. 2012; 149(6):1393–1406.

Chorin B, Masrati A, Kessel G, Narunsky A, Sprinzak A, Lahav A, Shlomtzion H, Ben-Tal N. ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. Protein Science. 2020; 29(1):258–267. https://onlinelibrary.wiley.com/doi/abs/10.1002/pro.3779, doi: https://doi.org/10.1002/pro.3779.

Craven CJ, Whitehead B, Jones SKA, Thulin E, Blackburn GM, Waltho JP. Complexes Formed between Calmodulin and the Antagonists J-8 and TFP in Solution. Biochemistry. 1996; 35(32):10287–10299. https://doi.org/10.1021/bi9605043, doi: 10.1021/bi9605043.

Darden T, York D, Pedersen L. Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. The Journal of Chemical Physics. 1993; 98(12):10089–10092. https://aip.scitation.org/doi/abs/10.1063/1.464397, doi: 10.1063/1.464397.

Dong J, Wang NN, Yao ZJ, Zhang L, Cheng Y, Ouyang D, Lu AP, Cao DS. ADMETlab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. Journal of Cheminformatics. 2018; 10(1):29. https://doi.org/10.1186/s13321-018-0283-x, doi: 10.1186/s13321-018-0283-x.

Einstein JM, Perelis M, Chaim IA, Meena JK, Nussbacher JK, Tankka AT, Yee BA, Li H, Madrigal AA, Neill NJ, Shankar A, Tyagi S, Westbrook TF, Yeo GW. Inhibition of YTHDF2 triggers proteotoxic cell death in MYC-driven breast cancer. Molecular Cell. 2021; 81(15):3048–3064.e9. https://www.sciencedirect.com/science/article/pii/S1097276521004937, doi: https://doi.org/10.1016/j.molcel.2021.06.014.

El-Naggar AM, Veinotte CJ, Cheng H, Grunewald TG, Negri GL, Somasekharan SP, Corkery DP, Tirode F, Mathers J, Khan D. Translational activation of HIF1α by YB-1 promotes sarcoma metastasis. Cancer cell. 2015; 27(5):682–697.

El-Naggar AM, Somasekharan SP, Wang Y, Cheng H, Negri GL, Pan M, Wang XQ, Delaidelli A, Rafn B, Cran J. Class I HDAC inhibitors enhance YB-1 acetylation and oxidative stress to block sarcoma metastasis. EMBO reports. 2019; 20(12):e48375.

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. The Journal of Chemical Physics. 1995; 103(19):8577–8593. https://aip.scitation.org/doi/abs/10.1063/1.470117, doi: 10.1063/1.470117.

Evdokimova V, Tognon C, Ng T, Ruzanov P, Melnyk N, Fink D, Sorokin A, Ovchinnikov LP, Davicioni E, Triche TJ. Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition. Cancer cell. 2009; 15(5):402–415.

Fischer A, Smieško M, Sellner M, Lill MA. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. Journal of Medicinal Chemistry. 2021 03; 64(5):2489–2500. https://doi.org/10.1021/acs.jmedchem.0c02227, doi: 10.1021/acs.jmedchem.0c02227.

Goldenberg O, Erez E, Nimrod G, Ben-Tal N. The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. Nucleic Acids Res. 2009 Jan; 37(Database issue):D323–7. doi: 10.1093/nar/gkn822.

Gresh N, El Hage K, Goldwaser E, de Courcy B, Chaudret R, Perahia D, Narth C, Lagardère L, Lipparini F, Piquemal JP. 1. In: Rivail JL, Ruiz-Lopez M, Assfeld X, editors. Addressing the Issues of Non-isotropy and Non-additivity in the Development of Quantum Chemistry-Grounded Polarizable Molecular Mechanics Cham: Springer International Publishing; 2015. p. 1–49. https://doi.org/10.1007/978-3-319-21626-3_1, doi: 10.1007/978-3-319-21626-3_1.

Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, Mackerell J Alexander D. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. Journal of chemical theory and computation. 2012; 8(1):348–362. https://pubmed.ncbi.nlm.nih.gov/22368531https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3285246/, doi: 10.1021/ct200723y.

Hess B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. Journal of Chemical Theory and Computation. 2008; 4(1):116–122. https://doi.org/10.1021/ct700200b, doi: 10.1021/ct700200b.

Hwang TL, Shaka AJ. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. Journal of Magnetic Resonance, Series A. 1995; 112(2):275–279. https://www.sciencedirect.com/science/article/pii/S1064185885710479, doi: https://doi.org/10.1006/jmra.1995.1047.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics. 1983; 79(2):926–935. https://aip.scitation.org/doi/abs/10.1063/1.445869, doi: 10.1063/1.445869.

Julio AR, Backus KM. New approaches to target RNA binding proteins. Current Opinion in Chemical Biology. 2021; 62:13–23.

Jung J, Lifland AW, Zurla C, Alonas EJ, Santangelo PJ. Quantifying RNA–protein interactions in situ using modified-MTRIPs and proximity ligation. Nucleic acids research. 2013; 41(1):e12–e12.

Jung YM, Yu KL, Park SH, Lee SD, Kim MJ, You JC. Investigation of function and regulation of the YB-1 cellular factor in HIV replication. BMB Rep. 2018 Jun; 51(6):290–295. doi: 10.5483/bmbrep.2018.51.6.231.

Kang Y, Hu W, Ivan C, Dalton HJ, Miyake T, Pecot CV, Zand B, Liu T, Huang J, Jennings NB. Role of focal adhesion kinase in regulating YB–1–mediated paclitaxel resistance in ovarian cancer. Journal of the National Cancer Institute. 2013; 105(19):1485–1495.

Khan MI, Adhami VM, Lall RK, Sechi M, Joshi DC, Haidar OM, Syed DN, Siddiqui IA, Chiu SY, Mukhtar H. YB-1 expression promotes epithelial-to-mesenchymal transition in prostate cancer that is inhibited by a small molecule fisetin. Oncotarget. 2014; 5(9):2462.

Khong A, Matheny T, Jain S, Mitchell SF, Wheeler JR, Parker R. The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. Molecular cell. 2017; 68(4):808–820. e5.

Kim DS, Camacho CV, Nagari A, Malladi VS, Challa S, Kraus WL. Activation of PARP-1 by snoRNAs Controls Ribosome Biogenesis and Cell Growth via the RNA Helicase DDX21. Mol Cell. 2019; 75(6):1270–1285 e14. https://www.ncbi.nlm.nih.gov/pubmed/31351877, doi: 10.1016/j.molcel.2019.06.020.

Kloks CPAM, Spronk CAEM, Lasonder E, Hoffmann A, Vuister GW, Grzesiek S, Hilbers CW. The solution structure and DNA-binding properties of the cold-shock domain of the human Y-box protein YB-111Edited by P. E. Wright. Journal of Molecular Biology. 2002; 316(2):317–326. https://www.sciencedirect.com/science/article/pii/S0022283601953343, doi: https://doi.org/10.1006/jmbi.2001.5334.

Knezevic CE, Wright G, Rix LLR, Kim W, Kuenzi BM, Luo Y, Watters JM, Koomen JM, Haura EB, Monteiro AN. Proteome-wide profiling of clinical PARP inhibitors reveals compound-specific secondary targets. Cell chemical biology. 2016; 23(12):1490–1503.

Kosti I, Jain N, Aran D, Butte AJ, Sirota M. Cross-tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues. Scientific Reports. 2016; 6(1):24799. https://doi.org/10.1038/srep24799, doi: 10.1038/srep24799.

Kretov DA, Clément MJ, Lambert G, Durand D, Lyabin DN, Bollot G, Bauvais C, Samsonova A, Budkina K, Maroun RC. YB-1, an abundant core mRNA-binding protein, has the capacity to form an RNA nucleoprotein filament: a structural analysis. Nucleic acids research. 2019; 47(6):3127–3141.

Labbé CM, Rey J, Lagorce D, Vavruša M, Becot J, Sperandio O, Villoutreix BO, Tufféry P, Miteva MA. MTiOpenScreen: a web server for structure-based virtual screening. Nucleic Acids Research. 2015 04; 43(W1):W448–W454. https://doi.org/10.1093/nar/gkv306, doi: 10.1093/nar/gkv306.

LaFargue CJ, Dal Molin GZ, Sood AK, Coleman RL. Exploring and comparing adverse events between PARP inhibitors. The Lancet Oncology. 2019; 20(1):e15–e28.

Lagarde N, Rey J, Gyulkhandanyan A, Tufféry P, Miteva MA, Villoutreix BO. Online structure-based screening of purchasable approved drugs and natural compounds: retrospective examples of drug repositioning on cancer targets. Oncotarget. 2018 08; 9(64):32346–32361. https://pubmed.ncbi.nlm.nih.gov/30190791, doi: 10.18632/oncotarget.25966.

Lasham A, Print CG, Woolley AG, Dunn SE, Braithwaite AW. YB-1: oncoprotein, prognostic marker and therapeutic target? Biochemical Journal. 2013; 449(1):11–23.

Lu S, Zhang J, Lian X, Sun L, Meng K, Chen Y, Sun Z, Yin X, Li Y, Zhao J. A hidden human proteome encoded by 'non-coding'genes. Nucleic acids research. 2019; 47(15):8111–8125.

Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L. Single-cell in situ RNA profiling by sequential hybridization. Nature methods. 2014; 11(4):360–361.

Lyabin DN, Eliseeva IA, Ovchinnikov LP. YB-1 protein: functions and regulation. Wiley Interdisciplinary Reviews: RNA. 2014; 5(1):95–110.

Lyons SM, Achorn C, Kedersha NL, Anderson PJ, Ivanov P. YB-1 regulates tiRNA-induced Stress Granule formation but not translational repression. Nucleic acids research. 2016; 44(14):6949–6960.

Ma S, Sun S, Geng L, Song M, Wang W, Ye Y, Ji Q, Zou Z, Wang S, He X, Li W, Esteban CR, Long X, Guo G, Chan P, Zhou Q, Belmonte JCI, Zhang W, Qu J, Liu GH. Caloric Restriction Reprograms the Single-Cell Transcriptional Landscape of Rattus Norvegicus Aging. Cell. 2020; 180(5):984–1001.e22. doi: 10.1016/j.cell.2020.02.008.

MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. The Journal of Physical Chemistry B. 1998; 102(18):3586–3616. https://doi.org/10.1021/jp973084f, doi: 10.1021/jp973084f.

Mayer M, Meyer B. Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. Journal of the American Chemical Society. 2001; 123(25):6108–6117. https://doi.org/10.1021/ja0100120, doi: 10.1021/ja0100120.

Minuesa G, Albanese SK, Xie W, Kazansky Y, Worroll D, Chow A, Schurer A, Park SM, Rotsides CZ, Taggart J. Small-molecule targeting of MUSASHI RNA-binding activity in acute myeloid leukemia. Nature communications. 2019; 10(1):1–15.

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry. 2009 12; 30(16):2785–2791. https://pubmed.ncbi.nlm.nih.gov/19399780, doi: 10.1002/jcc.21256.

Moss EG, Tang L. Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. Developmental biology. 2003; 258(2):432–442.

Méphon-Gaspard A, Boca M, Pioche-Durieu C, Desforges B, Burgo A, Hamon L, Piétrement O, Pastré D. Role of tau in the spatial organization of axonal microtubules: keeping parallel microtubules evenly distributed despite macromolecular crowding. Cellular and Molecular Life Sciences. 2016; 73(19):3745–3760.

Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Pääbo S, Mann M. Deep proteome and transcriptome mapping of a human cancer cell line. Molecular systems biology. 2011; 7(1):548.

Nusinow DP, Szpyt J, Ghandi M, Rose CM, McDonald ER, Kalocsay M, Jané-Valbuena J, Gelfand E, Schweppe DK, Jedrychowski M, Golji J, Porter DA, Rejtar T, Wang YK, Kryukov GV, Stegmeier F, Erickson BK, Garraway LA, Sellers WR, Gygi SP. Quantitative Proteomics of the Cancer Cell Line Encyclopedia. Cell. 2020; 180(2):387–402.e16. https://www.sciencedirect.com/science/article/pii/S0092867419313856, doi: https://doi.org/10.1016/j.cell.2019.12.023.

O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. Journal of Cheminformatics. 2011; 3(1):33. https://doi.org/10.1186/1758-2946-3-33, doi: 10.1186/1758-2946-3-33.

Panche AN, Diwan AD, Chandra SR. Flavonoids: an overview. Journal of nutritional science. 2016; 5:e47–e47. https://pubmed.ncbi.nlm.nih.gov/28620474https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5465813/, doi: 10.1017/jns.2016.41.

Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied Physics. 1981; 52(12):7182–7190. https://aip.scitation.org/doi/abs/10.1063/1.328693, doi: 10.1063/1.328693.

Pearson K. LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1901 11; 2(11):559–572. https://doi.org/10.1080/14786440109462720, doi: 10.1080/14786440109462720.

Poudyal D, Yang J, Chen Q, Goswami S, Adelsberger JW, Das S, Herman A, Hornung RL, Andresson T, Imamichi T. IL-27 posttranslationally regulates Y-box binding protein-1 to inhibit HIV-1 replication in human CD4+ T cells. AIDS. 2019; 33(12).

Ravindranath PA, Sanner MF. AutoSite: an automated approach for pseudo-ligands prediction-from ligand-binding sites identification to predicting key ligand atoms. Bioinformatics (Oxford, England). 2016 10; 32(20):3142–3149. https://pubmed.ncbi.nlm.nih.gov/27354702, doi: 10.1093/bioinformatics/btw367.

Rengifo-Gonzalez JC, El Hage K, Clément MJ, Steiner E, Joshi V, Craveur P, Durand D, Pastré D, Bouhss A. The co-operative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation. Elife. 2021; 10:e67605.

Roos M, Pradère U, Ngondo RP, Behera A, Allegrini S, Civenni G, Zagalak JA, Marchand JR, Menzi M, Towbin H. A small-molecule inhibitor of Lin28. ACS chemical biology. 2016; 11(10):2773–2781.

Schanda P, Brutscher B. Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds. Journal of the American Chemical Society. 2005; 127(22):8014–8015. https://doi.org/10.1021/ja051306e, doi: 10.1021/ja051306e.

Schmidt N, Lareau CA, Keshishian H, Ganskih S, Schneider C, Hennig T, Melanson R, Werner S, Wei Y, Zimmer M, Ade J, Kirschner L, Zielinski S, Dölken L, Lander ES, Caliskan N, Fischer U, Vogel J, Carr SA, Bodem J, et al. The SARS-CoV-2 RNA–protein interactome in infected human cells. Nature Microbiology. 2021; 6(3):339–353. https://doi.org/10.1038/s41564-020-00846-z, doi: 10.1038/s41564-020-00846-z.

Selth LA, Gilbert C, Svejstrup JQ. RNA immunoprecipitation to determine RNA-protein associations in vivo. Cold Spring Harbor protocols. 2009; 2009(6):pdb. prot5234.

Simsek D, Tiu GC, Flynn RA, Byeon GW, Leppek K, Xu AF, Chang HY, Barna M. The mammalian ribo-interactome reveals ribosome functional diversity and heterogeneity. Cell. 2017; 169(6):1051–1065. e18.

Singh G, Pratt G, Yeo GW, Moore MJ. The clothes make the mRNA: past and present trends in mRNP fashion. Annual review of biochemistry. 2015; 84:325–354.

Solano-Gonzalez E, Coburn KM, Yu W, Wilson GM, Nurmemmedov E, Kesari S, Chang ET, MacKerell AD, Weber DJ, Carrier F. Small molecules inhibitors of the heterogeneous ribonuclear protein A18 (hnRNP A18): a regulator of protein translation and an immune checkpoint. Nucleic acids research. 2021; 49(3):1235–1246. https://pubmed.ncbi.nlm.nih.gov/33398344https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7897483/, doi: 10.1093/nar/gkaa1254.

Sreeramulu S, Richter C, Kuehn T, Azzaoui K, Blommers MJJ, Del Conte R, Fragai M, Trieloff N, Schmieder P, Nazaré M, Specker E, Ivanov V, Oschkinat H, Banci L, Schwalbe H. NMR quality control of fragment libraries for screening. Journal of Biomolecular NMR. 2020; 74(10):555–563. https://doi.org/10.1007/s10858-020-00327-9, doi: 10.1007/s10858-020-00327-9.

Tailor D, Resendez A, Garcia-Marques FJ, Pandrala M, Going CC, Bermudez A, Kumar V, Rafat M, Nambiar DK, Honkala A. Y box binding protein 1 inhibition as a targeted therapy for ovarian cancer. Cell Chemical Biology. 2021; 28(8):2462. doi: doi: 10.1016/j.chembiol.2021.02.014.

Taminau J, Thijs G, De Winter H. Pharao: Pharmacophore alignment and optimization. Journal of Molecular Graphics and Modelling. 2008; 27(2):161–169. https://www.sciencedirect.com/science/article/pii/S109332630800048X, doi: https://doi.org/10.1016/j.jmgm.2008.04.003.

Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry. 2010 01; 31(2):455–461. https://pubmed.ncbi.nlm.nih.gov/19499576, doi: 10.1002/jcc.21334.

Usaj MM, Styles EB, Verster AJ, Friesen H, Boone C, Andrews BJ. High-content screening for quantitative cell biology. Trends in cell biology. 2016; 26(8):598–611.

Van Gunsteren WF, Berendsen HJC. A Leap-frog Algorithm for Stochastic Dynamics. Molecular Simulation. 1988; 1(3):173–185. https://doi.org/10.1080/08927028808080941, doi: 10.1080/08927028808080941.

Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen JY, Cody NA, Dominguez D. A large-scale binding and functional map of human RNA-binding proteins. Nature. 2020; 583(7818):711–719.

Wang L, Rowe RG, Jaimes A, Yu C, Nam Y, Pearson DS, Zhang J, Xie X, Marion W, Heffron GJ. Small-molecule inhibitors disrupt let-7 oligouridylation and release the selective blockade of let-7 processing by LIN28. Cell reports. 2018; 23(10):3091–3101.

Williamson MP. Using chemical shift perturbation to characterise ligand binding. Progress in Nuclear Magnetic Resonance Spectroscopy. 2013; 73:1–16. https://www.sciencedirect.com/science/article/pii/S0079656513000265, doi: https://doi.org/10.1016/j.pnmrs.2013.02.001.

Wu P. Inhibition of RNA-binding proteins with small molecules. Nature Reviews Chemistry. 2020; 4(9):441–458.

Wu SL, Fu X, Huang J, Jia TT, Zong FY, Mu SR, Zhu H, Yan Y, Qiu S, Wu Q. Genome-wide analysis of YB-1-RNA interactions reveals a novel role of YB-1 in miRNA processing in glioblastoma multiforme. Nucleic acids research. 2015; 43(17):8516–8528.

Yang JY, Ha SA, Yang YS, Kim JW. p-Glycoprotein ABCB5 and YB-1 expression plays a role in increased heterogeneity of breast cancer cells: correlations with cell fusion and doxorubicin resistance. BMC Cancer. 2010; 10(1):388. https://doi.org/10.1186/1471-2407-10-388, doi: 10.1186/1471-2407-10-388.

Yang XJ, Zhu H, Mu SR, Wei WJ, Yuan X, Wang M, Liu Y, Hui J, Huang Y. Crystal structure of a Y-box binding protein 1 (YB-1)–RNA complex reveals key features and residues interacting with RNA. Journal of Biological Chemistry. 2019; 294(28):10998–11010.

Zandarashvili L, Langelier MF, Velagapudi UK, Hancock MA, Steffen JD, Billur R, Hannan ZM, Wicks AJ, Krastev DB, Pettitt SJ. Structural basis for allosteric PARP-1 retention on DNA breaks. Science. 2020; 368(6486).

Zeng Z, Yu J, Jiang Z, Zhao N. Oleanic Acid (OA) Targeting UNC5B Inhibits Proliferation and EMT of Ovarian Cancer Cell and Increases Chemotherapy Sensitivity of Niraparib". Journal of Oncology. 2022; 85:12. https://doi.org/10.1155/2022/5887671.

Zhang J, Fan JS, Li S, Yang Y, Sun P, Zhu Q, Wang J, Jiang B, Yang D, Liu M. Structural basis of DNA binding to human YB-1 cold shock domain regulated by phosphorylation. Nucleic acids research. 2020; 48(16):9361–9371.

Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: A fast force field generation tool for small organic molecules. Journal of Computational Chemistry. 2011; 32(11):2359–2368. https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21816, doi: https://doi.org/10.1002/jcc.21816.

Appendix 1

I. Conformational study of YB-1 CSD in its unbound/free form

In order to study the dynamic behavior of YB-1 CSD and delignate the Quercetin-pocket, MD simulations were ran for 200 ns using as starting coordinates the NMR solution structure PDB ID 1H95 (***Kloks et al., 2002***). First, the free energy landscape (FEL) issued from the MD simulation was analyzed (Figure 1-Figure supplement 1a). FEL is represented using two variables that reflect specific properties of the system and measure conformational variability: the radius of gyration and the root mean square deviation (RMSD) with respect to the average conformation; and the Gibbs free energy is estimated from the probability distribution of sampled populations. The zero energy is at 0 kJ.mol^{-1} and corresponds to the lowest energy conformational state (dark blue). A local energy minimum is observed over a large free energy space (deep basin, dark blue) indicating that these conformational ensembles are stable during the simulation period. The comparison of 2 structures extracted from the basin with a different radius of gyration show two different conformational states of a pocket (called here "the quercetin-pocket"): a closed state, where K118 is interacting with F85 (upper structure), and an open state where K118 is moved away from F85 into the solvent (bottom structure). This pocket is located at the third β-hairpin and somehow monitored by K118 and F85. To better understand the relation between these residues, we monitored the distance between the side chain N_ζ of K118 and C_γ of F85 (red curve) along the simulation (Figure 1-Figure supplement 1b, left panel). The probability distribution shows a first peak at ~ 3.7 Å which confirms a strong cation-π interaction formed between the cationic side chain of K118 (NH^{3+}) and the electronegative benzene ring system of F85 (the cutoff being 6 Å), which is concomitant with a closed state pocket. The second is at 6.2 Å, meaning that the lysine is far away and thus the pocket exhibit an open state. The distance between the C_α of both residues was also monitored to see if the side chain movement is driven by the backbone. The distance probability distribution (black curve) shows one peak at ~ 7 Å, meaning that the observed cation-π interaction is driven by K118 side chain movement and not by a backbone structural change of the U-turn. The higher probability of the open state (~ 75% of the time, compared to 25% for the closed state) is important to keep the pocket accessible for RNA binding. The RMSD of the pocket U-turn was also calculated for the C_α atoms of the protein (green), the β-sheet (black), and the U-turn (red) Figure 1-Figure supplement 1b, right panel). Results show a high stability of the β-sheet of the CSD (<1 Å), a higher variation of protein C_α (~ 2.7 Å) owing to the flexibility of N- and C-terminal parts and U-turns. The RMSD of the pocket U-turn C_α atoms do not show high changes (broad peak 1-1.5 Å). In summary, these results show that the Quercetin-pocket presents an open and a closed state due to K118 side chain movement and that the opening mechanism is controlled by an electrostatic cation-π interaction formed between the cationic side chain of K118 (NH^{3+}) and the electronegative π-ring system of F85.

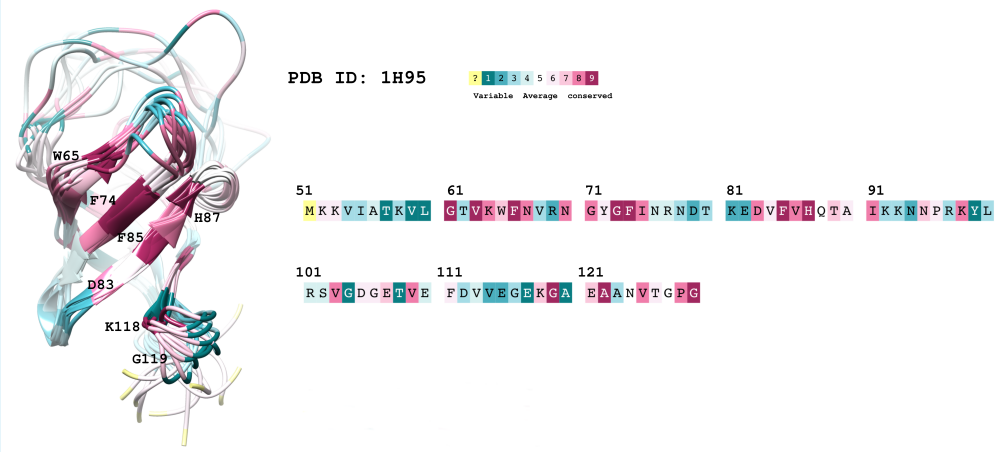
II. Structural and energetic study of YB-1 CSD bound to C5 RNA

The YB-1 CSD bound to RNA was also investigated using MD simulations. To make a link with our experiments, a 5-nt long poly(C) RNA (C5) was used. The system was build using as a template the crystal structure of YB-1 CSD in complex with UCAACU (PDB ID 5YTX (***Yang et al., 2019***)), and simulations were run for 200 ns. The FEL plot shows low energy basins with a ~ 2 kJ.mol^{-1} difference between their local minima (Figure 1-Figure supplement 2a). The structures extracted from the two observed wells show differences at the RNA extremities 5' and 3' which are highly flexible, and thus explains the two conformational ensembles. In order to identify the CSD residues implicated in the binding to C5 RNA, the interaction en-

ergy (ΔH) between individual residues and RNA was calculated and averaged along the MD simulation (Figure 1-Figure supplement 2b). Results show key residues highly implicated in the binding such as K64, W65, Y72, F74, F85, H87, K118, and E121. The energy decomposition into Coulomb (Coul) and Lennard-Jones (LJ) contributions show that RNA C5 binds the CSD via electrostatic and vdW interactions equally. Three types of binding are observed: (i) purely electrostatic (K64 and E121), (ii) purely vdW via $\pi - \pi$ stacking (Y72, F74, F85), and (iii) both electrostatic and vdW (W65 and K118). Figure 1-Figure supplement 2c, shows a 3D representation of the zero-energy complex on which a projection of significant CSPs obtained by NMR from 2D ^{1}H-^{15}N-SOFAST-HMQC spectrum of ^{15}N-labeld YB-1 in complex with C5 RNA is illustrated. The observed CSPs are in line with the binding mode from MD simulations, which in turn provides a resolved atomistic picture of the binding mechanism.

III. Evolutionary conservation of YB-1 CSD

The evolutionary conservation of YB-1 CSD was evaluated by the ConSurf-DB server (***Chorin et al., 2020***; ***Goldenberg et al., 2009***). Calculations were done using default parameters and results are illustrated in Appendix 1 Figure 1. The analysis show that: (a) the CSD in conserved in general; (b) the residues implicated in RNA-binding are the ones that are highly conserved (such as F74, F85, H87 and K118); (c) and the two residues that monitor the opening of the Quercetin-pocket (F85 and K118, who are also implicated in RNA-binding) are also highly conserved.



Appendix 1 Figure 1. The conservation coloring profile from the ConSurf-DB repository, mapped onto the protein. The conservation coloring scale is shown on the top right (conservation score varies from 1 to 9, where 1 corresponds to maximal variability and 9 to maximal conservation; "?" indicates insufficient data. Only highly conserved residues of interest are labeled on the 3D structure.

Appendix 2

Virtual Screening Details

I. Pharmacophore-Based Screening

All docking computations were based on the minimized structure of the open-state conformation identified from the MD simulation of the apo form of the CSD of YB1. This structure was used to build two distinct pharmacophores. A first pharmacophore using the prediction of a "pseudo ligand" in the binding site made by AutoSite (***Ravindranath and Sanner, 2016***) (see below), this approach is called "Pocket-based". And a second pharmacophore, called "ligand-based", built from the 3D structure of YB-1 in complex with Quercetin (F1). This structure was obtained by docking F1 in the binding pocket of YB-1 followed by MD refinement and the target pocket was confirmed by NMR spectroscopy. Then we used both pharmacophores to virtually screen an in-house database composed of 208 million pharmacophores, representing the conformers of around 7.3 million distinct commercially available molecules from MolPort (https://molport.com).

From the "pocket-based" screening, 249 distinct molecules were identified sharing at least 80% of their 3D pharmacophoric volume and sharing at least 5 pharmacophoric points. And the "Quercetin-based" (or Ligand-based) screening allowed us to identify 407 distinct molecules sharing at least 60% of their 3D pharmacophoric volume and sharing at least 7 pharmacophoric points. None of the identified molecules were found in both virtual screening. To reduce these molecules to a final selection, we predicted ADME-T endpoints and computed molecular docking in the binding sites of YB-1 for each of the identified molecules. This docking was used to avoid the selection of compounds that couldn't fit in the pocket, and to compute a first estimation of the ligand's affinity to the target.

Docking: Docking computations were performed using AutoDock 4.2.6 (***Morris et al., 2009***), with atom types grids generated using AutoGrid 4.2.6. Grid box was 74x66x82 points of size, centered on the binding site, with a spacing of 0.375 Å. Each docking computations performed 100 runs of the genetic algorithm, with a maximum number of generations and energy evaluations equal to 27000 and 2500000, respectively.

Pharmacophoric screening: The pharmacophore library was built using the commercially available "all-stock screening compounds" dataset from MolPort company. The corresponding 3D conformers were generated using Open Babel 2.3.2 (***O'Boyle et al., 2011***). Pharmacophore generation and pharmacophore 3D alignments were performed using Align-it software from Silicos-it (***Taminau et al., 2008***). The pharmacophore representation is composed of eight different types of pharmacophoric points: lipophilic region, hydrogen bond donor, hydrogen bond acceptor, positive charge center, negative charge center, hybrid type of aromatic and lipophilic, hybrid type of hydrogen bond donor and acceptor.

A total of 208 million pharmacophores were screened, representing around 7.3 million distinct molecules. The pharmacophore alignments are scored based on point types, and the overlap volumes. As a result, four metrics were computed:

1. PP: the number of pharmacophoric points that are in common between the reference representation and the database representation;
2. REF: the « percentage » of the volume of the pharmacophores generated from the reference molecule that is common and aligned to the pharmacophores from our database (ranging from 0 to 1);
3. DB: is the « percentage » of the volume of the pharmacophores generated from our database that is common and aligned to the pharmacophores from the reference molecule (ranging from 0 to 1);

4. Tanimoto: represents the similarity between the two pharmacophores (ranging from 0 to 1).

At the end, the top molecules based on Tanimoto metrics were selected for each screening, with a minimum threshold for PP.

ADME-T prediction: ADME-T predictions were performed using SAR/QSAR models from ADMETlab (***Dong et al., 2018***). In these models, the prediction is based on molecular descriptors computed from SMILEs.

II. Virtual Screening of FDA-Approved Drugs using MTiOpenScreen

For the drug repurposing part of our study, an automated blind docking of an FDA-approved drug library (Drugs-lib) (***Lagarde et al., 2018***) was considered using MTIiOpenScreen (***Labbé et al., 2015***), a web server that performs virtual screening using AutoDock Vina (***Trott and Olson, 2010***). The Drugs-lib library contains 7173 stereoisomers corresponding to 4574 single isomer drugs. A gradient-based conformational search approach is used and defines the search space by a grid box that was centered at the center of our protein and its dimensions were 20 Å in x, y and z. The grid resolution is internally assigned to 1 Å. A number of binding modes of 10 and an exhaustiveness of 8 were used. The scoring of the generated docking poses and ranking of the ligands is based on the Vina empirical scoring function approximating the binding affinity in $kcal.mol^{-1}$.

III. Physico-Chemical and Purchasability Filters

We applied physico-chemical filters to select molecules belonging to a preferred chemical space that has drug-like properties. This included compounds in the ranges: $250 < MW$ (Molecular Weight) < 650; $0 < tPSA$ (topological polar surface area) < 180; $-3 < logP < 6$; $0 <$ number of HBD (hydrogen bond donors) < 7; $0 <$ number of HBA (hydrogen bond acceptors) < 12; $0 <$ Rotatable Bonds < 10 ; $0 <$ Rigid Bonds < 30 ; Num Rings ≤ 6 ; Max Size Ring ≤ 18 ; $3 <$ Num Carbon Atoms < 35 ; $1 <$ Num HeteroAtoms < 15 ; $0.1 <$ Ratio H/C < 1.1 ; Num Charges ≤ 3 ; $-2 <$ Total Charge < 2. Compounds F5 and F6 were an exception for some of these criteria. We also made sure that the selected compounds were commercially available with a purity > 95%.

Appendix 3

MT bench: Image Acquisitions and Statistical Analysis

Images of the cellular fluorescent signals were acquired on the high content imaging system Opera Phenix Plus from Perkin Elmer on 40x water immersion objectives with a numerical aperture of 1.1, allowing us to obtain a good resolution in the confocal mode. 160 fields of views were taken for each well resulting in thousands of cells to be analyzed, by well, in a 96 well plate format in order to have the strongest statistical significance. The data were calculated and extracted with the HARMONY software version 5.0 using an analysis pipeline containing successive building blocks for image segmentation, selection of population of interest, and calculation of signal enrichment on the microtubules (MTs) (Figure 3-Figure supplement 1). The enrichment is calculated on identified spots using the GFP channel that corresponds to the signal of the bait protein forced to be localized at MT due to its fusion with a microtubule binding domain (MBD). Spots representing segments of MTs were selected based on their shape and on the intensity of the GFP channel signals (corresponding to the presence of the bait on MT). mRNA were detected with Cy3-labelled poly(dT) and could be brought on MT due to their potential interaction with the bait. The calculated GFP and Cy3 intensities in the spots and in the cytoplasm were extracted from the HARMONY software and treated subsequently in order to measure the slope of the mRNA enrichment on MTs (mean spot intensity divided by mean cytoplasm intensity versus mean bait spot intensity) (Figure 3c). The robustness of a screening assay is usually determined according to the value of a calculated SSMD (Strictly standardized mean difference). The SSMD measures the strength of the difference between two controls following Equation 6:

$$SSMD = \frac{(\mu_n - \mu_p)}{\sqrt{(\sigma_n^2 + \sigma_p^2)}} \tag{6}$$

Where μ_p and σ_p^2 are the mean and standard deviation values of the positive control and μ_n and σ_n^2 are those of the negative control. If the difference between the mean values is many times greater than the standard deviation, the assay is accurate. An assay with an SSMD value ≥ 7 is considered of excellent quality and with an extremely strong control.

Appendix 4

NMR Data Analysis

I. YB-1:Ligand Binding (complex formation)

Data analysis was performed on the 15 ligands and 20 residues exhibiting significant CSPs. The ligands and residues in question are shown in Figure 5. For this, two data mining approaches were used: principal component analysis (PCA) and a correlation-matrix-based hierarchical clustering. First, the chemical shift data from each spectrum were represented as a one-dimensional vector that contains $\Delta\delta_{avg}$, that corresponds to the normalized δN and δH values. Following this, 15 vectors from ligands and 20 residues were concatenated to build a two-dimensional matrix. Some row vectors lacking standardized chemical shift data due to disappearing NMR signals were replaced by a high CSP value of 0.1 in order to mark a different exchange regime. The matrix size was 300 [20 residues × 15 ligands]. In order to look at the changes affecting the residues as a function of the ligand and vice-versa, the above analysis approaches were performed on this matrix, denoted A, and on it transpose, denoted A^T [15 ligands × 20 residues].

A PCA standard singular value decomposition analysis was performed on matrix A and on the transpose A^T. PCA is a statistical method widely used in exploratory data analysis (***Pearson, 1901***). This non-parametric method compresses the dimension of a matrix by finding the directions that captures most of the variability in our data matrix and thus can reveal some simplified structures hidden in the dataset. For matrix A, results show that the first 6 PC dimensions represent 97% of the variance, meaning that these 6 PCs are likely to describe most contributions of the signal changes. And for the transpose A^T, results show that the first 5 PCs represent 95% of the variance. The results are illustrated in Figure 5. Outliers were detected using SPE and Hoteling's T^2 tests. These two tests are complementary to each other. A clustering analysis of the PC results was also conducted using k-means.

The matrices were also analyzed using a correlation-matrix-based hierarchical clustering, where Pearson's correlation coefficient was calculated based on the ligand-induced CSPs followed by an agglomerative hierarchical clustering to extract multiple correlation patterns. Pearson's correlation is used to measure similarity between different rows/columns. And the cluster analysis seeks to build a hierarchy of clusters, where each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. Features are thus grouped hierarchically according to their distances. Threshold was set at 0.7 and we were able to see five different clusters grouped in the main diagonal for matrix A and four for A^T. Thus, the corresponding data in Figure 5-Figure supplement 1 are represented as correlation heatmaps arranged using a dissimilarity matrix, which gives information on how far are two features, to improve the visual representation, and the relationship between features is illustrated in a dendrogram. For negative and positive correlations, the distance will be close to zero. If there is no correlation, the distance will be 0. The results obtained with the correlation-matrix-based hierarchical clustering are in line with PCA analysis. P1, C2 and F2 (black) are again found as outliers with different effect on the binding. Flavonols F1, F3 and F7 show an identical behavior. And the rest of the molecules represent a more diverse but similar behavior. When looking at A^T, residues W65, V84, F85, V86, G119, A120 and E121(green) form a highly correlated cluster. The first outliers with 100% similarity are D83 and E117 (red), followed by K118 and G116 (blue). F74, H87 and W65sc (grey) also manifest being neighbors. Even though we were able to identify the interacting pocket, and classify the ligands by binding mode, compared to the PCA analysis, this technique is less sensitive and less informative in extracting specific residue information related to ligand selectivity.

II. Ligand-RNA Competitive Binding to YB-1

The Ligand's ability to compete with RNA on YB1 binding was evaluated by comparing pair displacement vectors of affected residues using their scalar product (denoted $SP(residue)$). These vectors correspond to the chemical displacement induced after adding the ligand ($\vec{u}$), RNA ($\vec{v}$) and both RNA+ligand ($\vec{w}$) to YB-1 (see Figure 7a, example for P1). The aim is to compare the following pair ($\vec{u}$-$\vec{w}$; $\vec{v}$-$\vec{w}$) in order to see if the induced chemical shift displacement of YB-1 residues is closer to the YB-1 Ligand-bound state or to the YB-1 RNA-bound state.

And thus,

$$SP(residue) = f(\vec{u}, \vec{v}, \vec{w}) = (\vec{u} - \vec{w}).(\vec{v} - \vec{w}) \tag{7}$$

1. If $SP(residue) > 0$, this means that the angle formed by these two pair vectors is acute and that vectors $\vec{w}$, $\vec{v}$ and $\vec{u}$ move in the same direction, which can be translated in to the fact that the concerned residue show an additive effect between the ligand and the RNA.
2. If $SP(residue) < 0$, this means that the angle formed by these two pair vectors is obtuse and that vectors $\vec{w}$, $\vec{v}$ and $\vec{u}$ move in different directions and that vector $\vec{w}$ is closer to $\vec{u}$ rather than $\vec{v}$, which can be translated into the fact that the concerned residue shows a significant competition of the ligand on the binding site rather than an additivity.
3. If $SP(residue) = 0$, can mean three things:
   - Either that $(\vec{u} - \vec{w}) = 0$; $\implies \vec{u} = \vec{w}$; meaning that the concerned residue represents the exact displacement for YB-1+Ligand and for YB-1+RNA+Ligand. Here, we have a full competition where the Ligand's effect manifest at 100%. This is the case of residue A120.
   - Or $(\vec{v} - \vec{w}) = 0$; $\implies \vec{v} = \vec{w}$; meaning that the concerned residue represents the exact displacement for YB-1+RNA and for YB-1+RNA+Ligand. Here, the ligand does not compete with RNA. No residue representing this case was observed.
   - Or $(\vec{u} - \vec{w})$ is perpendicular to $(\vec{v} - \vec{w})$; meaning that the concerned residue reflects the ligand's competitive binding. This is the case of residue V114.

This is an unsupervised and systematic way to compare two displacements for each residue. We could on top of this reduction of dimensionality apply a clustering algorithm to identify different clusters. However, this is beyond the scope of this application since we have results for only two ligands (P1 and C8).

Appendix 5 Table 1. ITC measurements of YB1/P1 complex. Calorigrams of raw data are provided in Appendix 5 Figure 1.

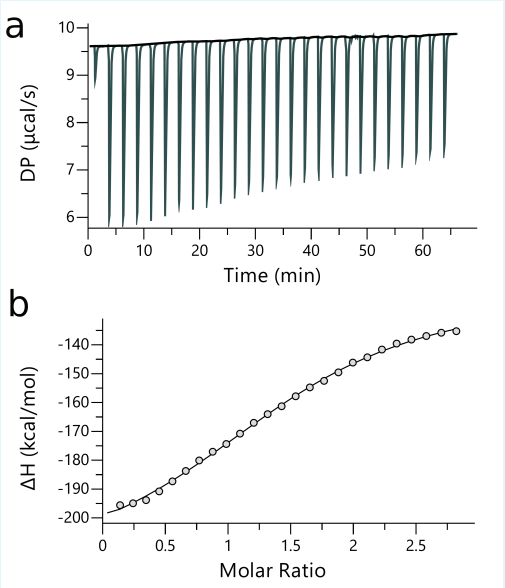
	[Ligand] (M)	[YB-1] (M)	N	K_d (M)	ΔH (kcal.mol^{-1})	ΔG (kcal.mol^{-1})	-TΔS (K.kcal.mol^{-1})	Red. χ^2
P1	200.10^{-6}	14.10^{-6}	1.54 $\pm 3.9 \times 10^{-2}$	$5.84 \times 10^{-6} \pm 674 \times 10^{-9}$	-98.1 ± 4.61	-7.14	91.00	0.47

Table 1–source data 1. ITC raw and fitted data obtained from P1 binding to YB-1. (see legends of Appendix 5 Figure 1 and Appendix 5 Table 1).

Appendix 5 Table 2. Thermodynamics of P1 binding to YB-1 obtained from ABFE simulations and ITC experiments. Comparison of P1/YB-1 binding free energies (ΔG) and their respective enthalpic (ΔH) and entropic (ΔS) contributions from ITC and computations. ΔS was calculated according to ΔG = ΔH - TΔS at 298 K; units are in kcal.mol^{-1}.

	ΔG_{bind} (kcal.mol^{-1})	ΔH_{bind} (kcal.mol^{-1})	$-T\Delta S_{bind}$ (K.kcal.mol^{-1})
ITC	-7.14 ± 0.47	-98.10 ± 4.61	91.00
ABFE	-7.24 ± 0.52	-28.62 ± 1.62	21.38

Appendix 5



Appendix 5 Figure 1. Isothermal Titration Calorimetry measurements of YB-1/P1 binding. (a) ITC raw data for titration of P1 into YB-1 in aqueous buffer solution. (b) Enthalpograms retrieved from (a). The solid line represents the fit to a single-site binding model.

Appendix 5 Table 3. Supplier list and compound IDs of the 40 tested molecules.

Compound	Traditional Name	Compound ID	Supplier
F1	Quercetin	MolPort-001-740-557	MolPort
F2	Quercetagetin	MolPort-006-147-776	MolPort
F3	3-O-methylquercetin	90081	SigmaAldrich
F4	Fisetin	MolPort-000-882-130	MolPort
F5	Rutin	MolPort-001-740-246	MolPort
F6	myricitrin	91255	SigmaAldrich
F7	Herbacetin	MolPort-019-998-217	MolPort
F8	Vincetoxicoside B	MolPort-035-758-036	MolPort
F9	Scutellarien	MolPort-003-724-680	MolPort
F10	Luteolin-7-methylether	MolPort-001-740-950	MolPort
F11	Naringenin	MolPort-000-861-091	MolPort
C1	Butein	MolPort-006-111-425	MolPort
C2	Okanin	FO66168	CarboSynth
C3	Robtein	Amb22172818	Ambinter
C4	Chorilifol B	MolPort-039-338-845	MolPort
C5	Bavachalcone	FB145210	CarboSynth
C6	homobutein	FM65711	CarboSynth
C7	Cardamonin	FC66017	CarboSynth
C8	3′,3,4,5′-tetrahydroxychalcone	FC66017	CarboSynth
C9	2′,4′-dihydroxy-4-methoxychalcone	MolPort-000-662-842	MolPort
C10	2′-hydroxy-4′-methoxychalcone	MolPort-000-779-850	MolPort
C11	2′,4,4′-trihydroxychalcone	MolPort-001-741-660	MolPort
C12	Lichochalcone B	MolPort-046-594-311	MolPort
A1	10-3-[2-(2,3-dihydro-1-benzofuran-5-yl)ethoxy]phenyl-2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-9H,10H-pyrano[2,3-h]chromene-4,8-dione	MolPort-035-700-332	MolPort
A2	2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-10-(naphthalen-1-yl)-9H,10H-pyrano[2,3-h]chromene-4,8-dione	MolPort-029-885-579	MolPort
A3	[8-(4-chlorophenyl)-2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-4-oxofuro[2,3-h]chromen-9-yl]acetic acid	MolPort-044-544-604	MolPort
A4	2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-10-2-[2-(4-methoxyphenyl)ethoxy]phenyl-9H,10H-pyrano[2,3-h]chromene-4,8-dione	MolPort-035-699-845	MolPort
A5	2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-10-isopropyl-9H,10H-pyrano[2,3-h]chromene-4,8-dione	MolPort-029-886-488	MolPort
A6	6,8-dibromo-2-(3-chloro-4-hydroxy-5-methoxyphenyl)chromen-4-one	MolPort-008-821-914	MolPort
A7	6,8-dibromo-3-hydroxy-2-(4-hydroxy-3-methoxyphenyl)chromen-4-one	MolPort-023-282-651	MolPort
A8	2-(4-hydroxy-3-iodo-5-methoxyphenyl)chromen-4-one	MolPort-002-521-806	MolPort
P1	Niraparib	MolPort-023-219-142	MolPort
P2	Olaparib	MolPort-009-679-395	MolPort
P3	Talazoparib	MolPort-028-600-028	MolPort
P4	Veliparib	MolPort-016-633-168	MolPort
P5	Rucaparib	MolPort-028-744-762	MolPort
D1	Nebivolol	MolPort-015-163-751	MolPort
D2	Mefloquine	MolPort-006-170-692	MolPort
D3	Icotinib	MolPort-039-139-676	MolPort
D4	Cabotegravir	MolPort-035-944-338	MolPort

Appendix 5 Table 4. Primers used for RT-PCR analysis. Sequences are from 5' to 3'.

	Forward	Backward
canx	GCAACCACTTCCCTTCCAT	TCCGCCTCTCTCTTTACTGC
calr	TGTCAAAGATGGTGCCAGAC	ACAACCCCGAGTATTCTCCC
oaz1	TACAGCAGTGGAGGGAGACC	GGATAAACCCAGCGCCAC
rpl8	AGATGGGTTTGTCAATTCGG	CAAGAAGACCCGTGTGAAGC
eif4g1	CCCAACTGTAGAAGGCATCC	CTCCAGGCCCTTGTAGTGAC
fnbp1	GCATGAAGTTATCTCCGAGAACA	CGGCCATCGTGAAAGTTTGAT
nin	GGAGGAACTCACCGACCTTTG	CGTCCGTAACGCTTCCCAC
cdkal1	GGGACTGAGTATCATTGGGGT	CCAAGCCGCCTTCCATTATC
mkln1	AGCCACGATGGAGTCAAATCA	TGGCACTAGGACCATTCTCTTT
eif4g2	AATCGCACTCTCCACTTTGG	GCTGCTGAGTTCTCGGTGA
ubl3	TGACAATTGGCCAATGGACTG	GCCACCAAATGCATCACTGT
actin	CATGTACGTTGCTATCCAGGC	CTCCTTAATGTCACGCACGAT
gapdh	CCTCCTGCACCACCAACTGCTTA	GTGATGGCATGGACTGTGGTCAT

Appendix 5 Table 5. RT-PCR analysis of 13 mRNAs isolated via MT bench pull-down or via magnetic beads RIP and for 3 different RBPs (YB-1, HuR, and FUS).

	MT bench			IP (beads)		
mRNA \ Bait	YB-1	HuR	FUS	YB-1	HuR	FUS
fnbp1	0.733	-0.339	1.032	0.429	0.227	0.641
rpl8	-0.436	-1.647	-1.976	0.098	-2.546	-1.232
eif4g1	1.697	0.009	0.024	1.527	0.176	1.258
gapdh	-0.228	-2.772	-2.834	0.294	-3.616	-1.609
ubl3	0.356	1.439	2.073	0.094	1.297	1.658
canx	-0.227	1.013	2.286	-0.367	1.555	0.770
actin	-1.661	-0.855	-3.051	-2.007	-0.074	-1.716
eif4g2	0.546	-0.118	1.458	1.117	0.037	0.324
nin	0.906	0.371	2.806	0.859	0.824	2.259
calr	0.403	-0.405	-1.801	-0.357	0.893	-1.163
cdk1	0.137	1.050	0.569	0.954	1.558	0.896
oaz1	-1.783	-0.940	-3.758	-1.310	-0.976	-2.189
mkln1	0.418	0.196	1.974	-0.127	-0.256	0.103

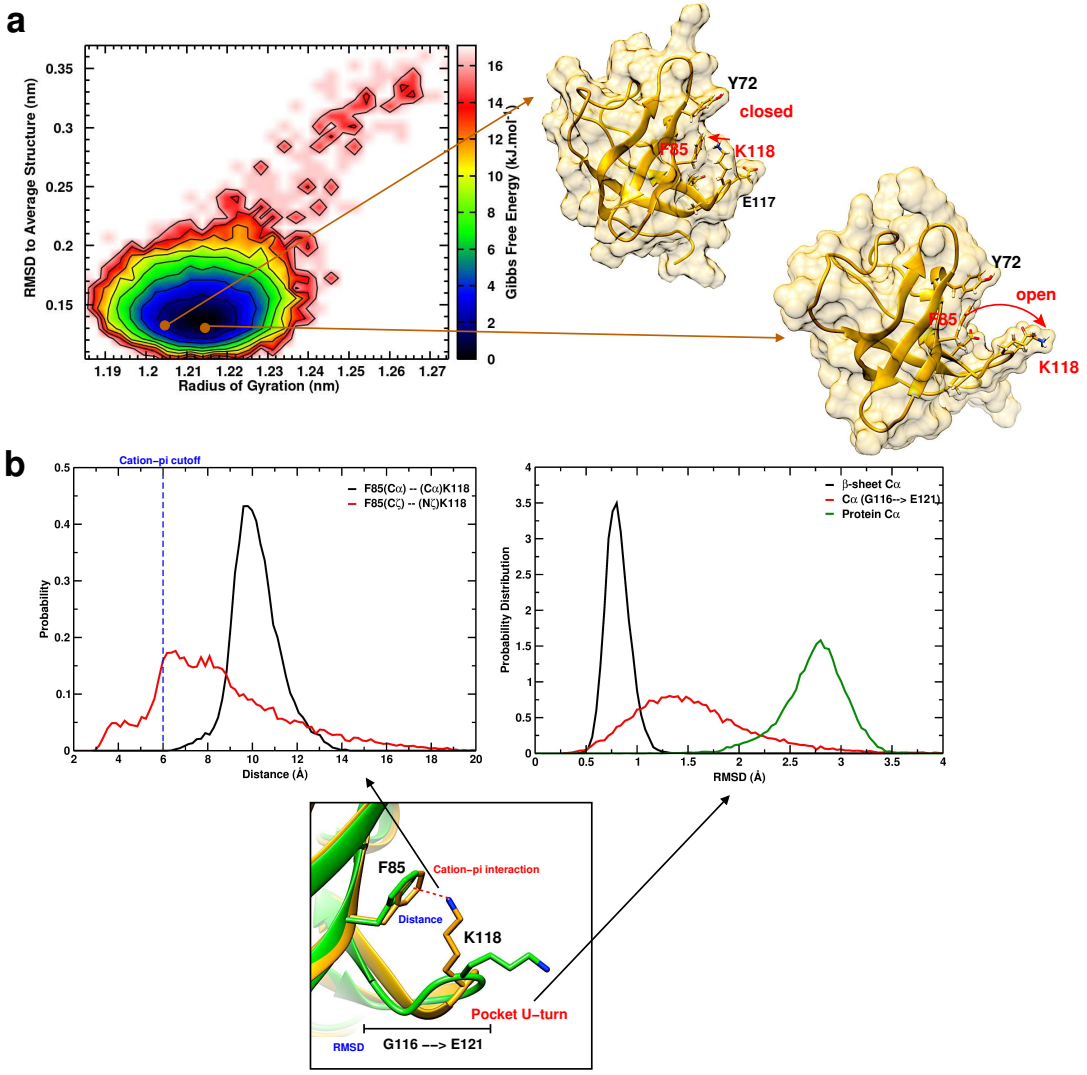
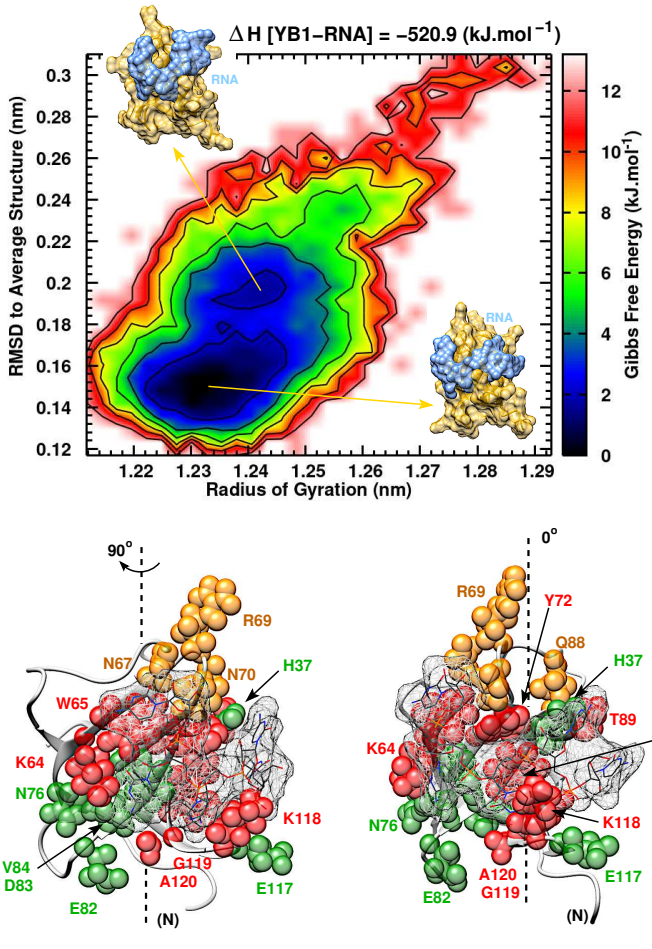


Figure 1–Figure supplement 1. Conformational study of YB-1 in its unbound/free form using MD simulations. (a) Free energy landscape (FEL) of unbound YB-1 computed from 200 ns of MD. FEL is represented using two structural reaction coordinates: the radius of gyration of the system and the RMSD with respect to the average structure. The zero energy is at 0 kJ.mol^{-1} and corresponds to the lowest energy conformational state. Two different conformational states of the Quercetin-pocket are sampled: a closed state, where F85 and K118 side chains are interacting (upper structure), and an open state where K118 is away from F85 (bottom structure). The red arrows indicate the moving direction of K118 side chain relative to F85. (b) Probability distribution of the cation-π interaction distance (left panel) and of the pocket RMSD (right panel) extracted from the 200 ns MD simulation. Left panel: The distance was monitored between Cα of F85 and K118 (black plot) and between F85(Cγ) and K118(Nζ) (red plot). The blue dashed line indicates the distance cutoff of a cation-π interaction. Right panel: The RMSD was calculated on the Cα atoms of the protein (green), the β-sheet (black) and the U-turn (red).



ΔH (kJ.mol⁻¹)

Residue	Coul	LJ	Total
K64	–38.4	–3.3	–41.7
W65	–28.0	–38.4	–66.4
N67	–19.1	–6.6	–25.7
R69	–6.1	–4.6	–10.7
N70	–6.9	–18.0	–24.8
Y72	–6.4	–33.4	–39.8
F74	–3.3	–31.8	–35.1
D83	–22.6	–3.1	–25.6
F85	–3.0	–35.9	–38.9
H87	–12.0	–23.8	–35.8
Q88	–4.8	–2.2	–7.0
E117	0.9	–4.0	–3.0
K118	–62.0	–26.1	–88.1
G119	–7.6	–5.8	–13.4
A120	–1.6	–4.9	–6.4
E121	–35.8	–4.5	–40.3
Total	–256.9	–251.7	–508.6

Figure 1–Figure supplement 2. Structural and energetic study of YB-1:RNA (C5) complex using MD simulations. (a) 2D Free energy landscape built from 200 ns MD simulation. The insets show 2 structures extracted from the two observed wells, where the conformational surface of the complex is represented (Protein in tan; RNA in light blue). The YB-1:RNA interaction energy (ΔH[YB1-RNA]) is also indicated. (b) Interaction energy contribution (ΔH) of the residues implicated in the binding, along with its Coulomb (Coul) and Lennard-Jones (LJ) contributions, averaged over 200 ns of MD simulation with variant of fluctuations being ±1.8 kJ.mol^{-1}. The most contributing residues are marked in red. (c) 3D representation of the zero-energy complex. The protein is represented in light grey cartoon, RNA in sticks and highlighted with a mesh surface. Residues implicated in the binding and showing significant CSPs in 2D ^{1}H-^{15}N-SOFAST-HMQC NMR spectrum are represented in spheres: residues with high interaction energy and/or high CSPs are in red, intermediate (green), lower (orange). (N) and (C) indicate the N-and C-terminal, respectively.

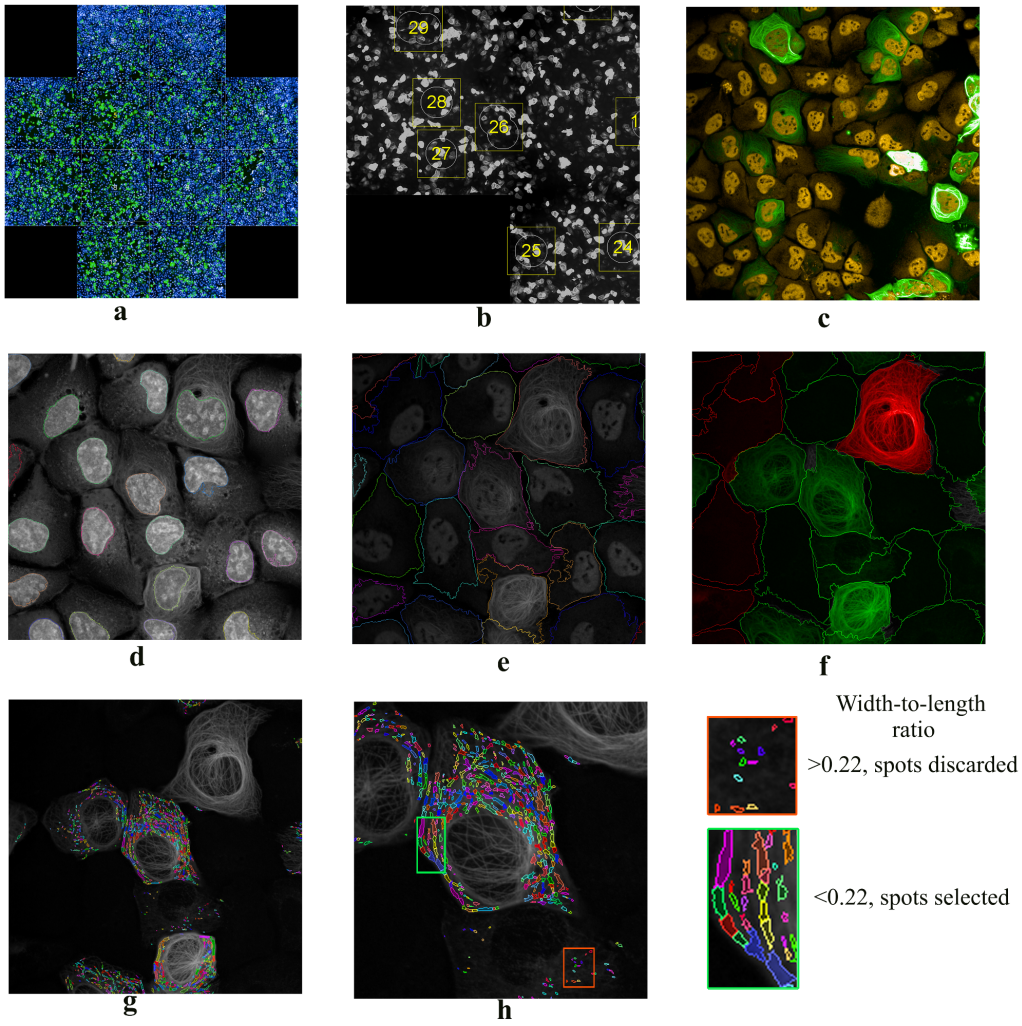
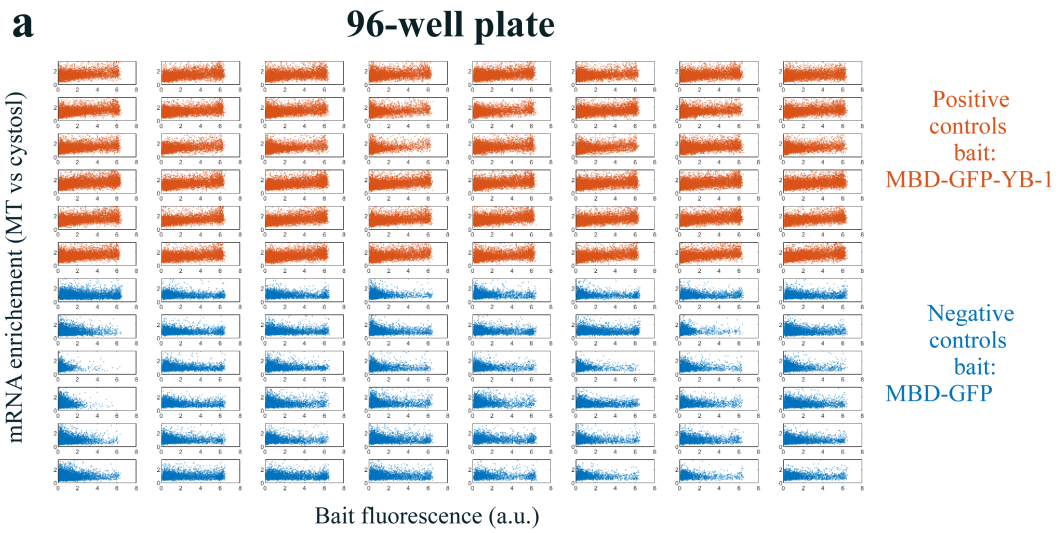


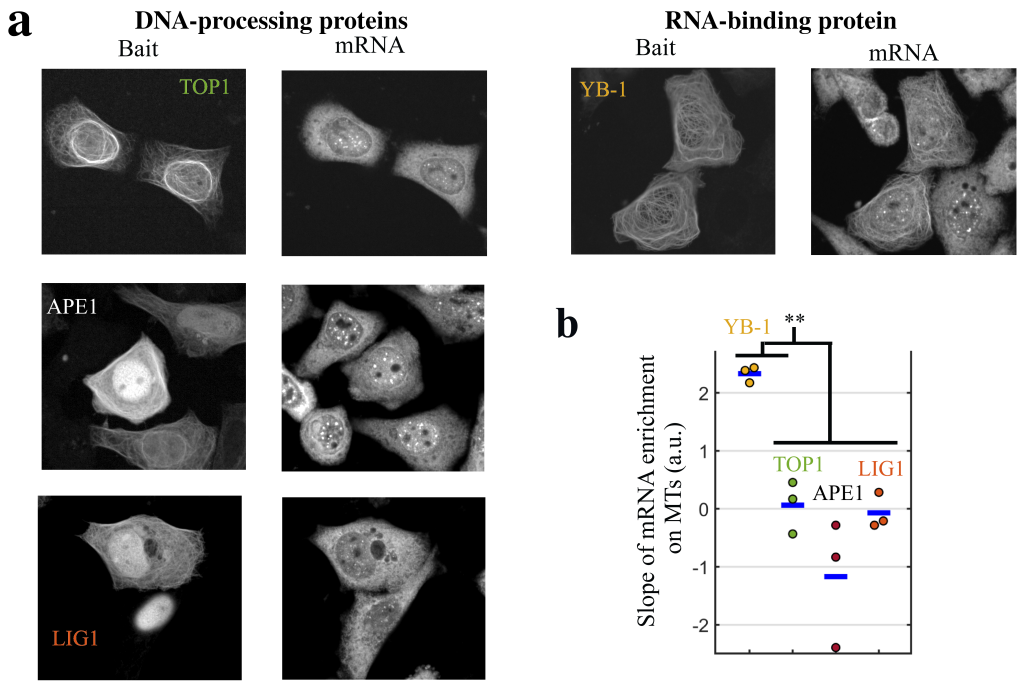
Figure 3–Figure supplement 1. Image analysis process to quantify mRPIs in cells. (a) Image of a single well in a 96-well plate showing the expression of MDP-GFP-YB-1 (geen) in U2OS cells at low magnification (juxtaposed images obtained with a 10x objective). (b) Automatic selection of smaller areas displaying several MBP-GFP-expressing cells with the HARMONY software. (c) High resolution image of one of the selected areas (green, MBP-GFP-RBP; red, mRNA (poly(dT) probe); blue, DAPI). Images obtained with a 10x water-immersed objective. (d) Automatic detection of nuclei using HARMONY and the DAPI signal. (e) For each nucleus, the cytoplasm was delineated using the mRNA signal (green). (f) Cells expressing MDP-GFP-YB-1 at low level (low signal to noise ratio) or at a too high level (saturation) were discarded (cell shown in red here). (g) The find spot analysis procedure of the HARMONY software was used for an automatic detection of a microtubule segment using the MBD-GFP-RBP signal in green. (h) To detect microtubule segments and discard putative aberrant structures (aggregates, noise, cell debris, etc.) in each spot, the width-to-length ratio was measured. Only spots with a ratio < 0.22 were selected to measure the relative enrichment of mRNA versus bait expression.



b slope values

1.21	1.12	0.93	1.09	0.95	1.04	0.94	0.91
1.03	1.04	1.04	1.02	0.99	0.90	0.91	1.05
0.98	1.12	0.90	1.02	0.85	0.97	0.91	0.88
0.98	1.11	1.08	1.19	1.04	0.94	1.07	0.92
0.95	1.11	1.11	0.93	1.07	1.03	0.94	1.05
0.85	0.89	1.00	0.94	0.84	0.99	1.12	1.02
0.02	0.08	0.14	0.07	0.08	0.03	0.13	0.04
0.24	0.16	0.07	-0.01	0.04	0.05	-0.02	0.01
0.04	0.06	-0.04	0.04	0.12	0.06	0.14	0.05
-0.02	0.21	0.15	0.08	0.15	0.09	0.29	0.14
0.07	0.04	0.03	0.03	0.13	0.09	0.18	0.01
0.10	0.12	0.00	0.07	0.03	0.07	0.02	0.17

Figure 3–Figure supplement 2. Quality assessment of the MT bench cell assay. (a) Correlation plots of mRNA enrichment versus GFP fluorescence in all 96 wells of a 96-well plate setting. The 48 positive (orange) and negative (blue) controls correspond to MBD-GFP-YB-1 and MBD-GFP, respectively. (b) Slope values *per* well. The slope value is measured with 95% confidence bounds and found an accuracy of about ± 5%. The data are homogenous for most wells.



1497

Figure 3–Figure supplement 3. mRNA- but not DNA-binding proteins bring mRNAs on MTs when used as baits. (a) Representative cell images after indicated treatments when DNA-binding or RNA-binding proteins were used as baits. Left: Bait's fluorescence (GFP). Right: endogenous mRNA (*in situ* hybridization with cy3-labelled poly(T) probes). (b) Slope of the mRNA enrichment versus bait fluorescence in selected spots (see Figure 3c for details). We note that YB-1, but not DNA-binding proteins (APE1, LIG1 and TOP1), brings mRNAs on MTs. **, $p < 0.01$; t-test with two tails. Each dot represents a different well (n=3 samples).

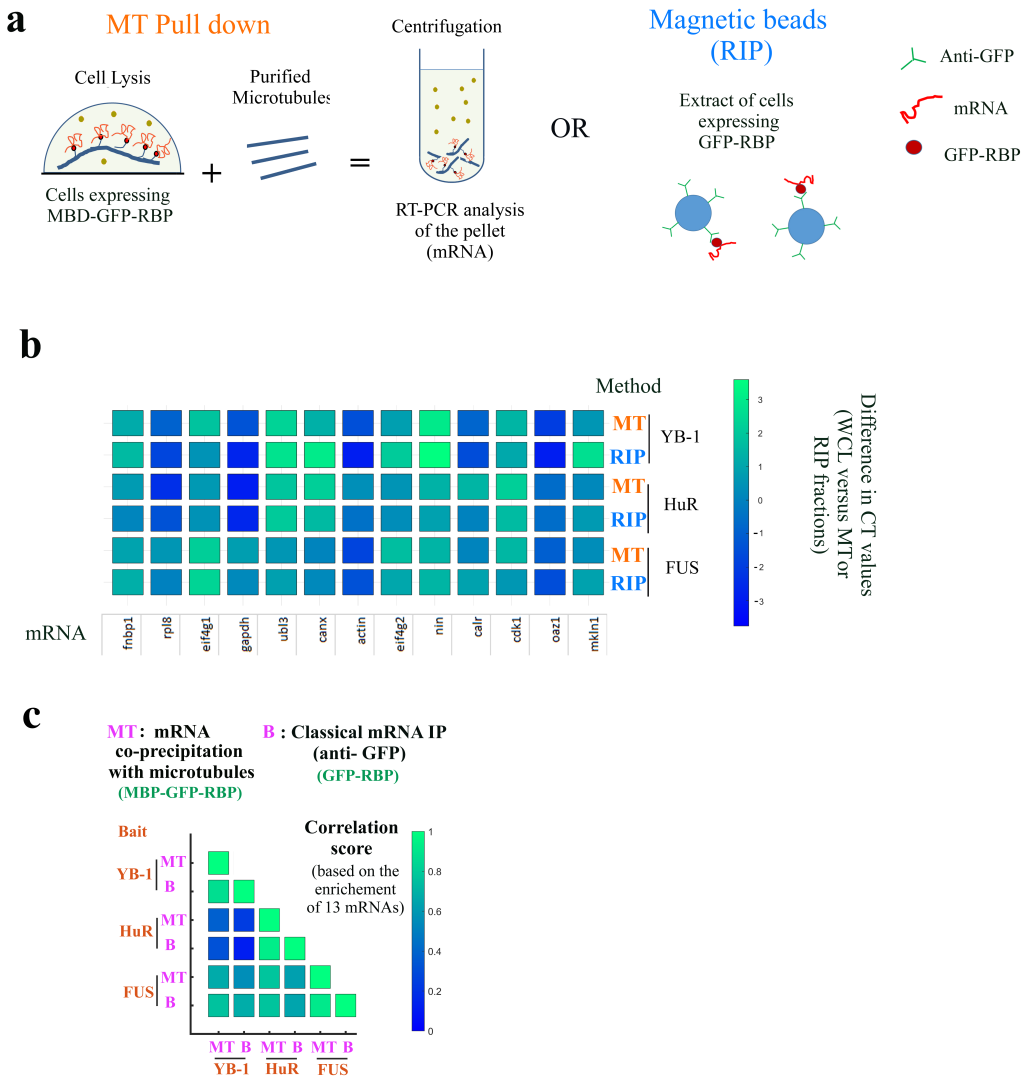


Figure 3–Figure supplement 4. The mRNAs brought onto microtubules by MBD-GFP-RBP are RBP-specific. (a) Schematic representation of the two mRNAs purification procedure. mRNAs were either purified from extracts of HEK293T cells expressing MBB-GFP-RBP by addition of sheep-brain microtubules and centrifugation or by magnetic beads with anti-GFP antibodies from extracts of HEK293T cells expressing MBB-GFP-RBP. (b) RT-PCR analysis of 13 mRNAs isolated in the microtubule pellet (MT) or via anti-GFP-coated magnetic beads (RIP) for indicated baits (YB-1, HuR and FUS). The differences in CT (cycle threshold) values are reported in Appendix 5 Table 5. MT: MT pull down. RIP: RNA-binding protein Immunoprecipitation. Scale bar: Differences in CT values recorded between the WCL (Whole cell lysate) and the RIP or MT fraction.(c) Correlation analysis of mRNA enrichments from (b).

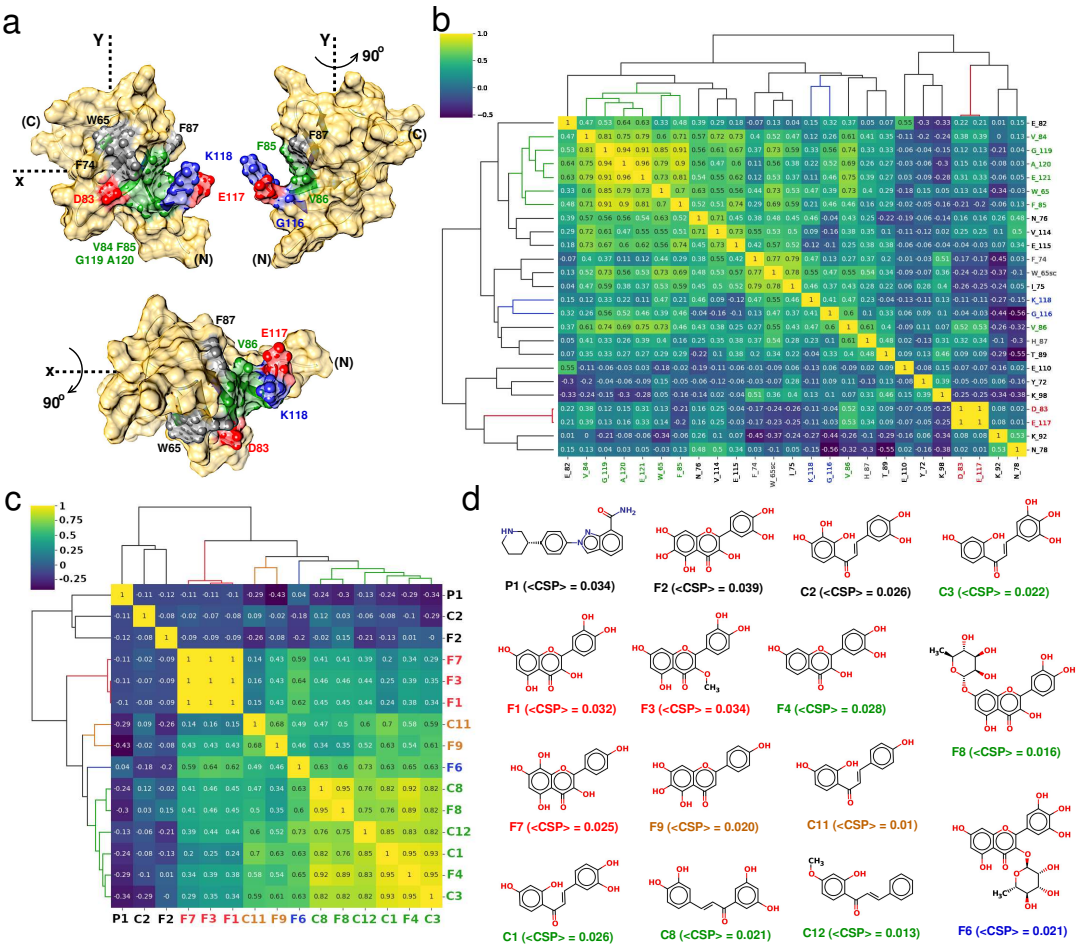


Figure 5–Figure supplement 1. Correlation-matrix-based hierarchical clustering of NMR ligand-induced CSPs. (a) 3D surface representation of YB-1 structure; emphasis on the Quercetin-pocket. (N) and (C) indicate the N-and C-terminal, respectively. The color coding of the pocket residues is that of the clusters formed in (b). (b) and (c) Correlation heatmaps of matrices A and A^T, respectively, arranged according to hierarchical clustering. The dendrograms illustrate the relationship between features. Labels are color-coded by clusters. (d) Chemical structures of the 15 ligands that bind the quercetin-pocket along with the estimated average <CSP>. Colored labels are those of the clusters identified in (c).

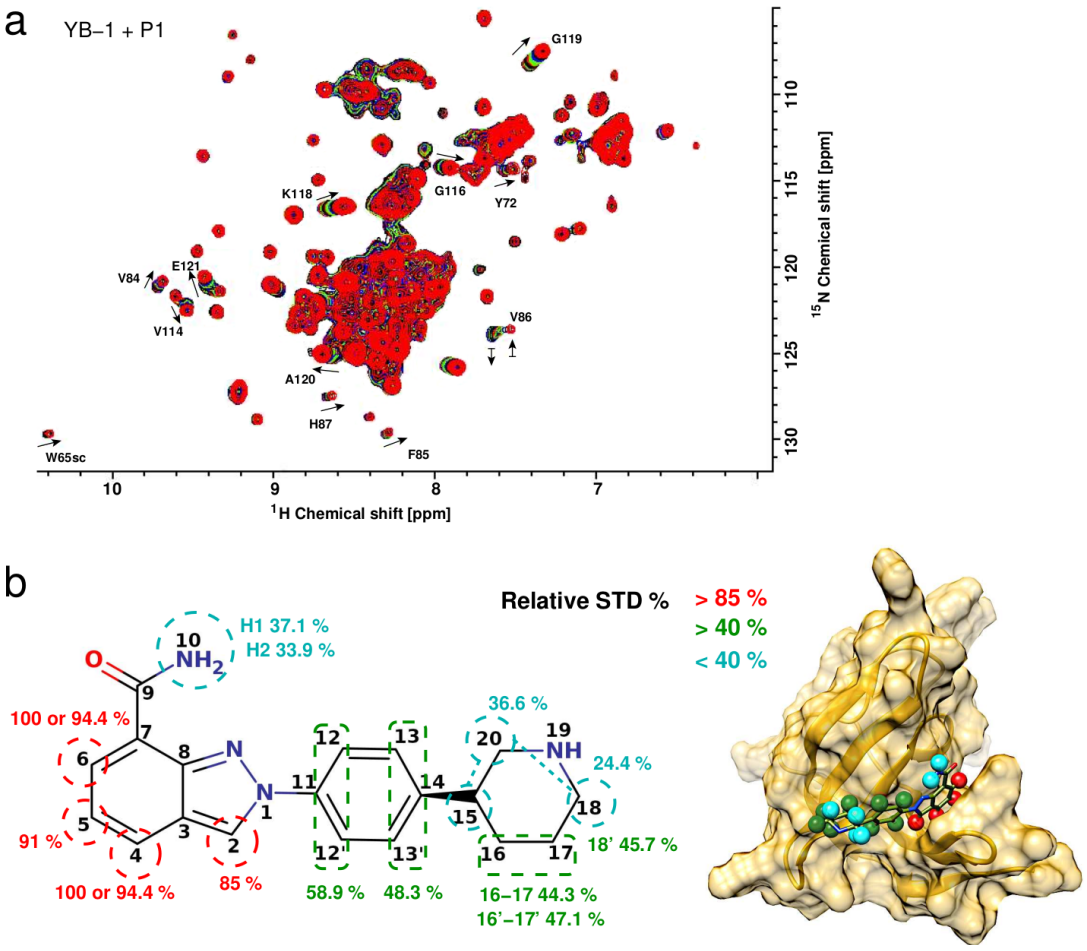


Figure 6–Figure supplement 1. NMR investigation of P1 binding to YB-1. (a) Overlays of 2D ^{1}H-^{15}N-SOFAST-HMQC spectrum monitoring ^{15}N-labelled YB-1 (50 μM) being titrated with P1. Increasing [P1] corresponds to color changes from black (free protein) to red (1:20 ratio). Black arrows indicate the shift direction of the peaks during the titration process for residues in fast exchange: peaks move smoothly from free to bound. Only specific pocket residues are indicated. For residue V86, up and down arrows illustrate the slow exchange regime: the free peak decreases in intensity (down arrow) as the bound peak (up arrow) increases. (b) STD-NMR analysis of P1. The epitope mapping is highlighted on the chemical structure of P1 (left) and on the 3D representation of YB-1:P1 complex (right). Relative STD percentages are conveyed by color code: red dashes indicate the most intense signals over 85% (including the most intense STD signal 100%), green dashes over 40%, and cyan under 40% relative to the most intense STD signal. YB-1 is represented as gold cartoon and surface. P1 is depicted in olive sticks and H atoms with STD signals in spheres.

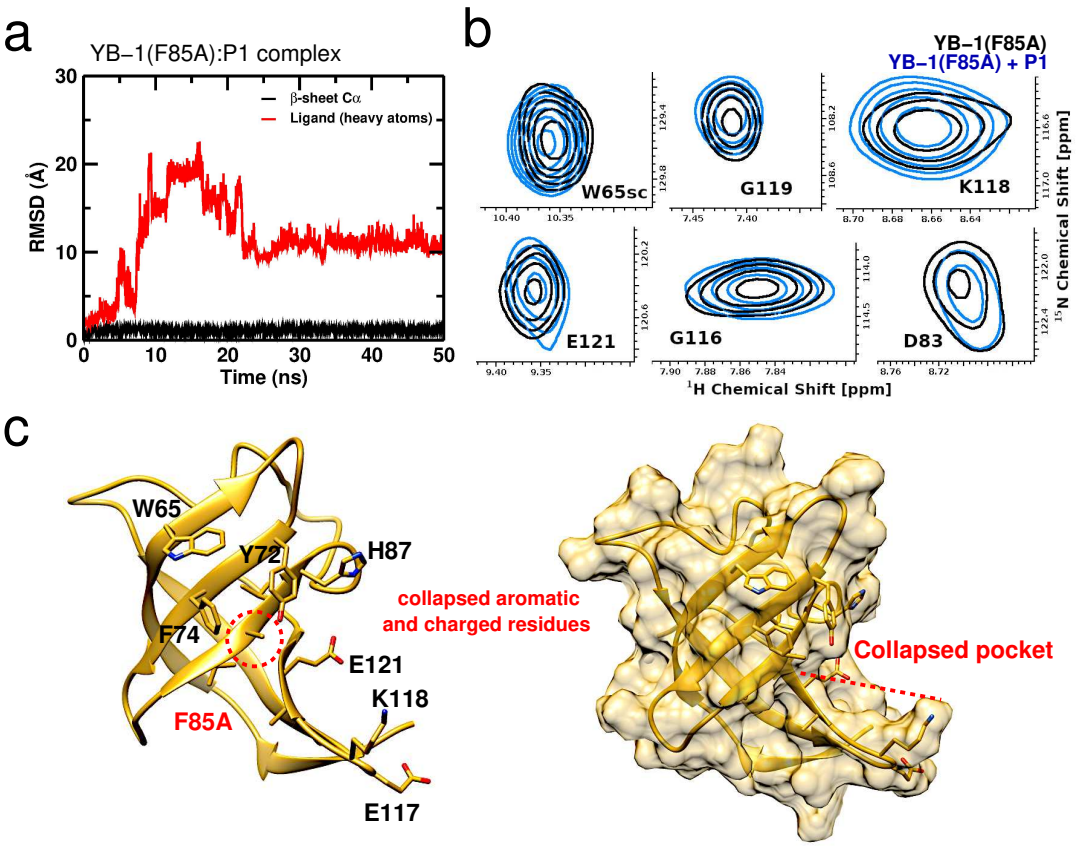


Figure 6–Figure supplement 2. Mapping the effect of F85A mutation on P1 binding to YB-1 by MD and NMR. (a) Root mean square deviation (RMSD) of atomic positions measured along 50 ns of MD simulation of the YB-1(F85A):P1 complex, for the backbone of YB-1(F85A) (black) and P1 heavy atoms (red). The protein structure is stable across the simulation, while the ligand leaves the binding site after 1 ns. (b) Zoom on pocket residues from 2D ^{1}H-^{15}N-SOFAST-HMQC spectra of YB-1(F85A) in the absence (black) and presence (blue) of P1. (c) YB-1(F85A) structural analysis by MD simulations. 3D structure extracted from the simulation shows, compared to WT (Figure 1-Figure supplement 1), a different side chain structural rearrangement of aromatic and polar residue following F85A mutation. The residues in question are labeled and the F85A mutation is indicated by a dashed red circle. The surface representation on the right hand-side emphasizes the collapsing pocket (indicated by a red dashed line).

1501

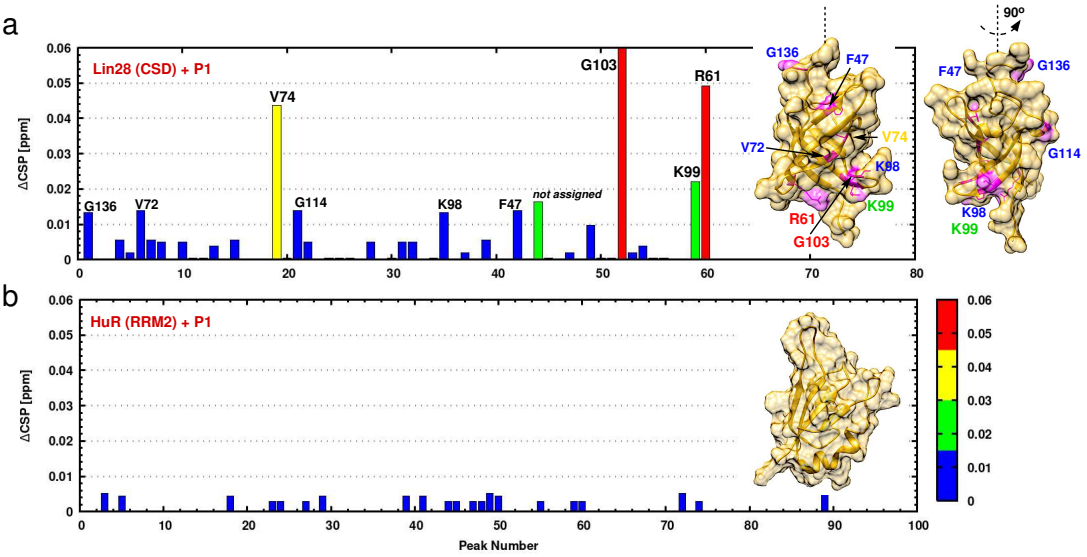


Figure 6–Figure supplement 3. Normalized chemical shift changes of the ^{1}H-^{15}N NMR spectra of ^{15}N-labeled CSD of LIN28 (a) and ^{15}N-labeled RRM2 of HuR (b) in the presence of P1. The color code of ΔCSP for both panels is indicated on the right-hand side of panel (b). Peak numbers are ordered with respect to their ^{1}H ppm value. The insets show the 3D structures of each of the studied proteins. For LIN28, in panel (a), residues representing a significant CSP are highlighted in magenta and labeled with the same color code as the corresponding ΔCSPs. Only peaks showing significant ΔCSPs were assigned. LIN28(CSD) resonance assignments were taken from the data kindly provided by the laboratory of Dr. Piotr Sliz, Harvard Medical School, USA.

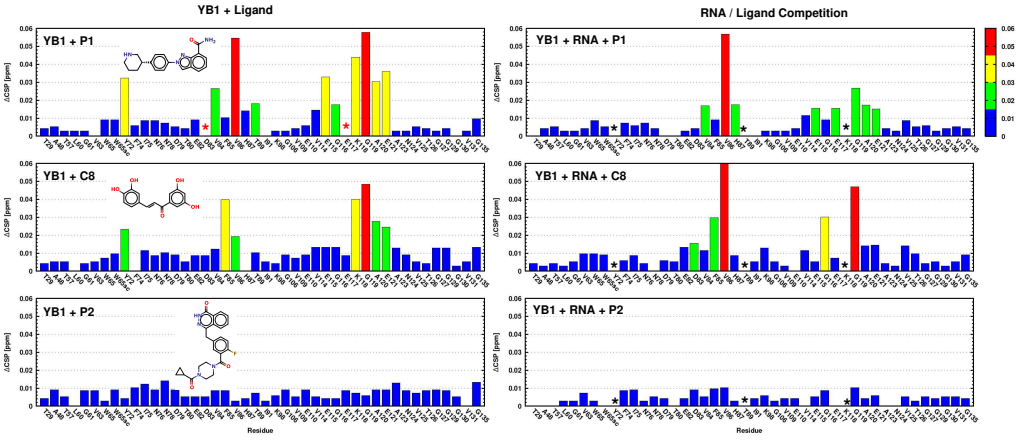


Figure 7–Figure supplement 1. Mapping Ligand/RNA competition on binding YB-1 using NMR. Histograms show normalized chemical shift changes of the ^{1}H-^{15}N NMR spectra of ^{15}N-labeled YB-1 in the absence (left panels) and in the presence of 5-nt long poly(C) RNA (right panels), while in the presence of P1 (upper panels), C8 (middle panels) and P2 (lower panels). The color code of ΔCSP is indicated on the top right-hand side. The insets show the chemical structures of the tested molecules P1, C8 and P2. Red and black asterisks indicate disappearing residues in the presence of P1 and RNA, respectively.

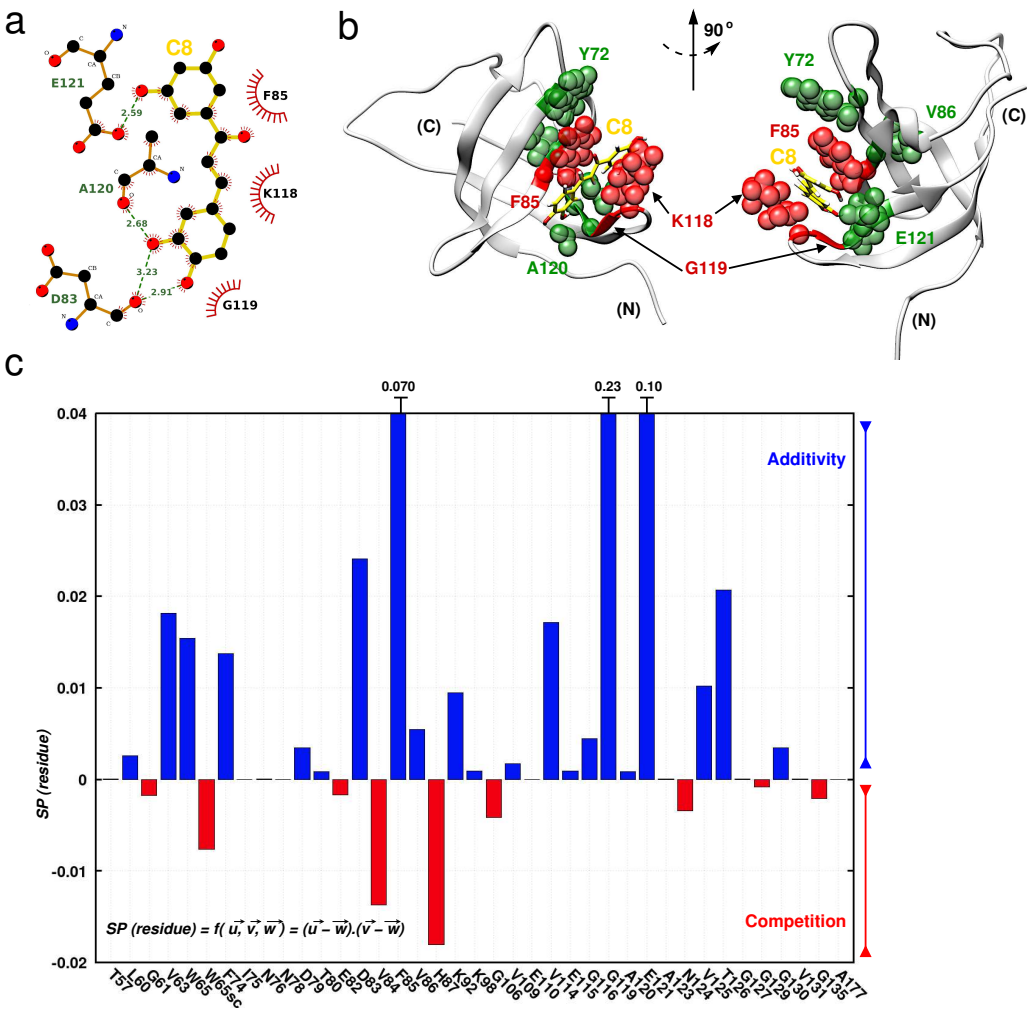


Figure 7–Figure supplement 2. Investigating C8 binding to YB-1 in the presence of RNA by MD and NMR. (a) 2D interaction diagram between C8 (gold) and YB-1. (b) 3D representation of the YB-1:C8 complex. The structure used in both representations (2D and 3D) is of the zero-energy complex obtained from 200 ns MD simulations. The protein is represented in light grey cartoon, C8 in gold sticks. Residues implicated in the binding and showing significant CSPs in NMR are represented with spheres: residues with high and intermediate CSPs are in red, and green respectively (see Figure 7-Figure supplement 1 for CSP values). (c) Histogram showing the scalar product of pair displacement vectors $SP(residue)$ calculated for each residue based on CSPs extracted from ^{1}H-^{15}N NMR spectra of ^{15}N-labeled YB-1 in the absence and/or absence of C8 and/or C5 RNA. $SP(residue) > 0$ means that vectors $\vec{w}$, $\vec{v}$ and $\vec{u}$ move in the same direction, and thus an additive effect is observed in the NMR spectra. $SP(residue) < 0$ means that vectors $\vec{w}$, $\vec{v}$ and $\vec{u}$ move in the opposite direction, and thus a competitive effect is observed.

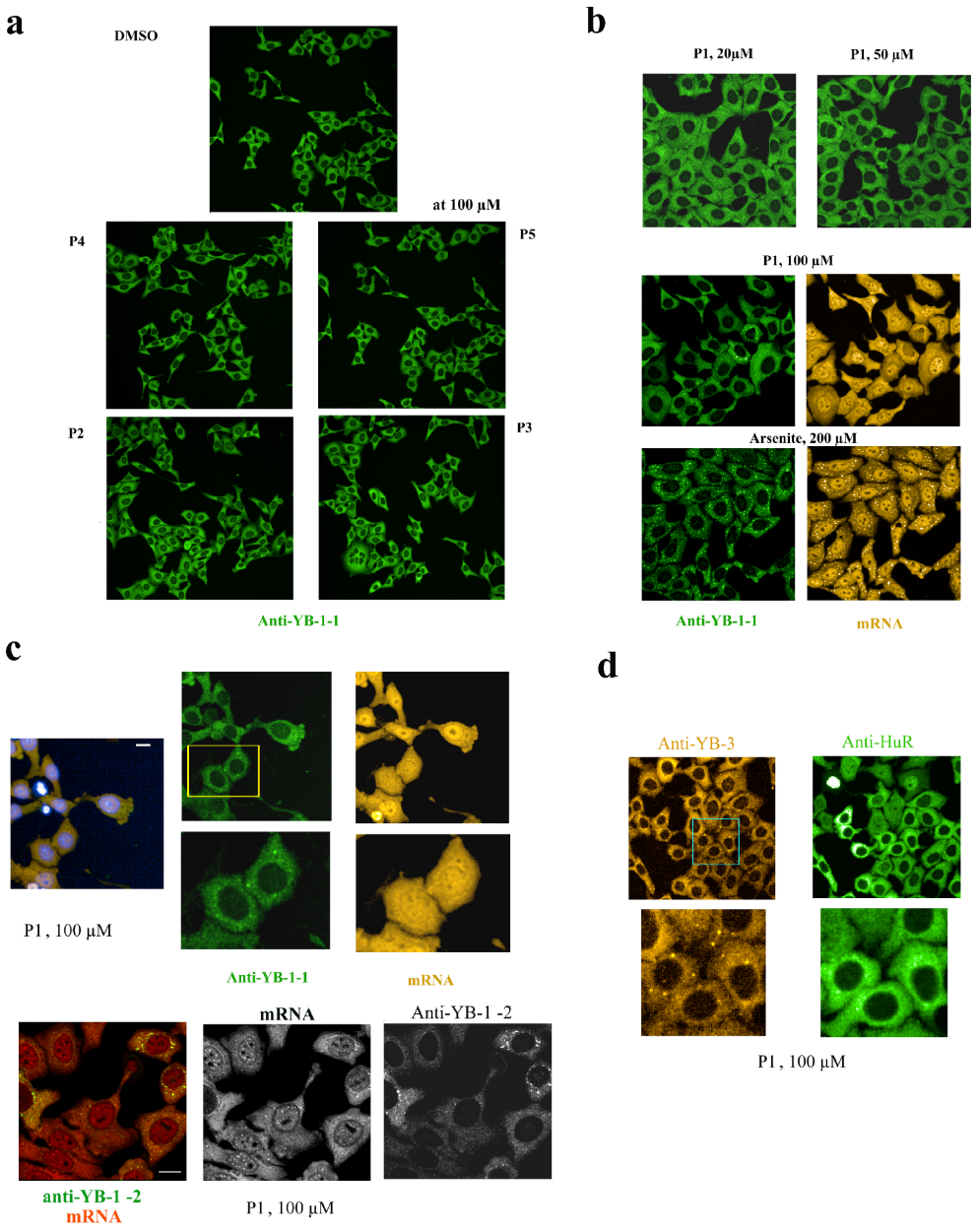
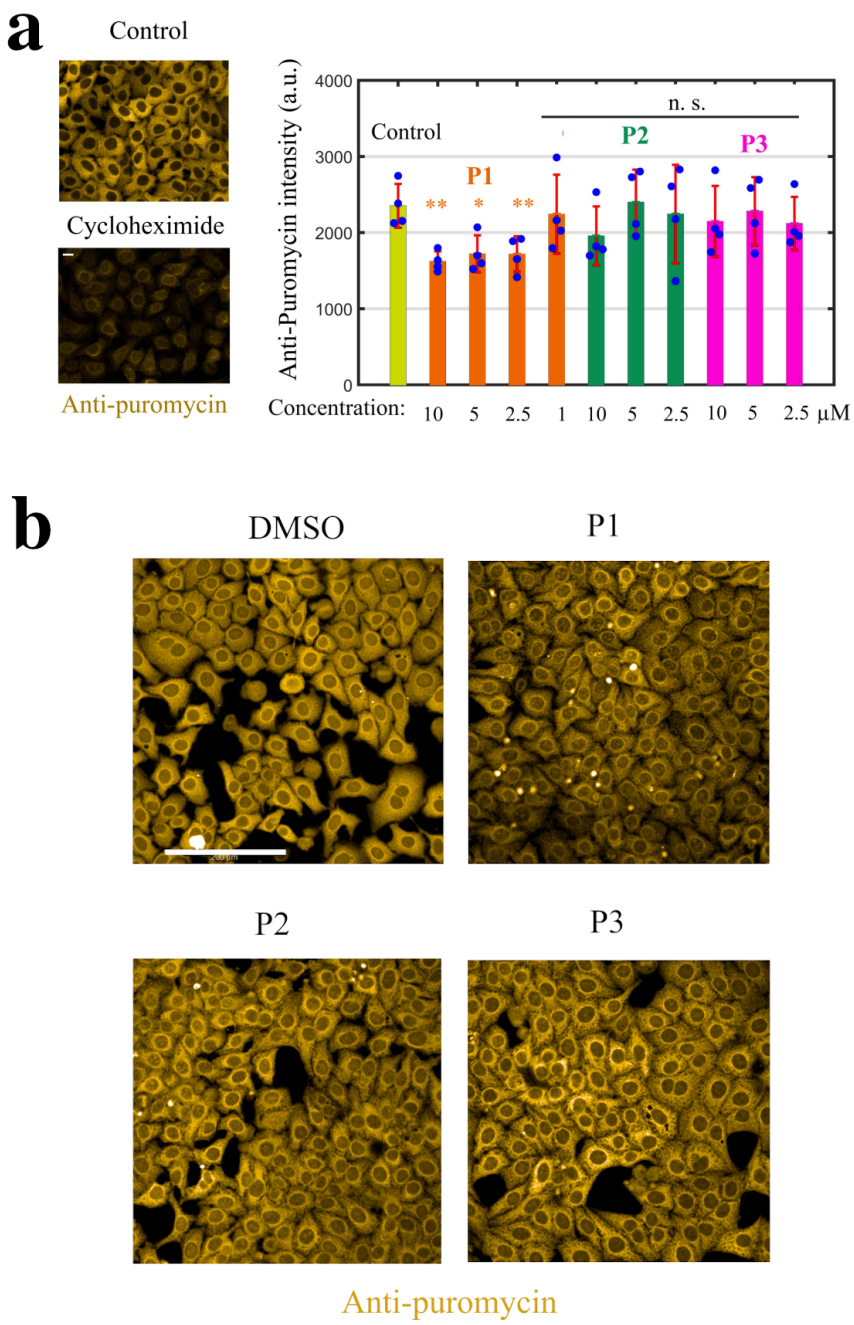
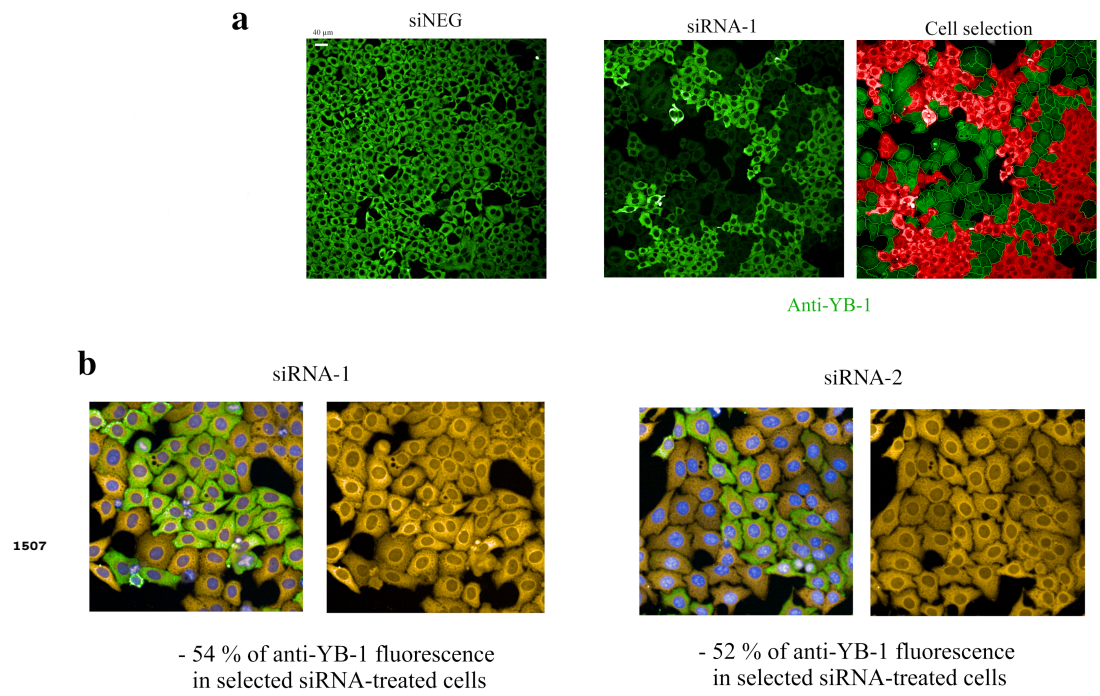


Figure 8–Figure supplement 1. Cytoplasmic YB-1-rich granules in HeLa cells at high P1 concentrations (≥ 20 μM). (a) Representative images of HeLa cells exposed to PARP-1 inhibitors (P4, P5, P2, P3 at 100 μM) during 2h. No YB-1-rich granules are observed. (b) Representative images of HeLa cells exposed to P1 (at 20, 50 and 100 μM) or Arsenite (at 200 μM) for 2 h at indicated concentrations. P1 and Arsenite lead to the appearance of YB-1-rich granules (fluorescent dots). The critical P1 concentration to observe the appearance of YB-1-rich granules under experimental condition is 50 μM. In Arsenite-treated cells, YB-1-rich granules are stress granules (lower left panel) owing to their highly enriched non-polysomal mRNA (lower right panel). On the other hand, YB-1-rich granules detected in P1-treated cells (middle left panel) were poorly enriched in mRNA (middle right panel). (c) The YB-1-rich granules of P1-treated cells can be observed with two different YB-1 antibodies (anti-YB-1-1 and anti-YB-1-2). Green, YB-1; orange/red, mRNA. (d) same as (c) with anti-YB-3 and anti-HuR antibodies. Note the marked presence of YB-3 in the cytoplasmic granules.



1506

Figure 8–Figure supplement 2. Monitoring puromycin incorporation in HeLa cells. (a) Left panel: Cycloheximide (Cyclo, 10 μg.ml^{-1}, 1 h), a translation inhibitor, prevents puromycin incorporation. Right panel: Histogram showing the efficiency of puromycin incorporation after indicated treatment for 2h. **, p<0.01; *, p<0.05; n.s., non-significant; t-test with two tails (n=4 samples *per* condition). (b) Anti-puromycin fluorescence images of HeLa cells exposed to 10 μM of DMSO, P1, P2, and P3 for 2 h; these cells were briefly exposed to puromycin 10 min prior to fixation.



1507

Figure 8–Figure supplement 3. Detection of puromycin incorporation in HeLa cells. (a) Anti-fluorescence images of HeLa cells after treatment with a negative control (siNEG, upper panel) or treatment with siRNA targeting YB-1 transcripts (siRNA-1, lower left panel). Under our experimental conditions, siRNAs lead to the appearance of clusters of HeLa cells expressing low levels of YB-1 that coexists with clusters of HeLa cells expressing normal levels of YB-1 (dim green vs fluorescent green in the lower left panel). Using the HARMONY software, we measured the mean intensity of anti-YB-1 fluorescence in the cytoplasm and selected only cells expressing YB-1 at a low level for further analysis of the puromycin incorporation level (green, lower right panel; red cells are cells expressing normal levels of YB-1 and thus were not selected). (b) Representative HeLa images obtained to measure the incorporation of puromycin in siRNA treated cells. The two siRNAs used in this study (siRNA-1 and -2) decreased the expression level of YB-1 to a significant extent.

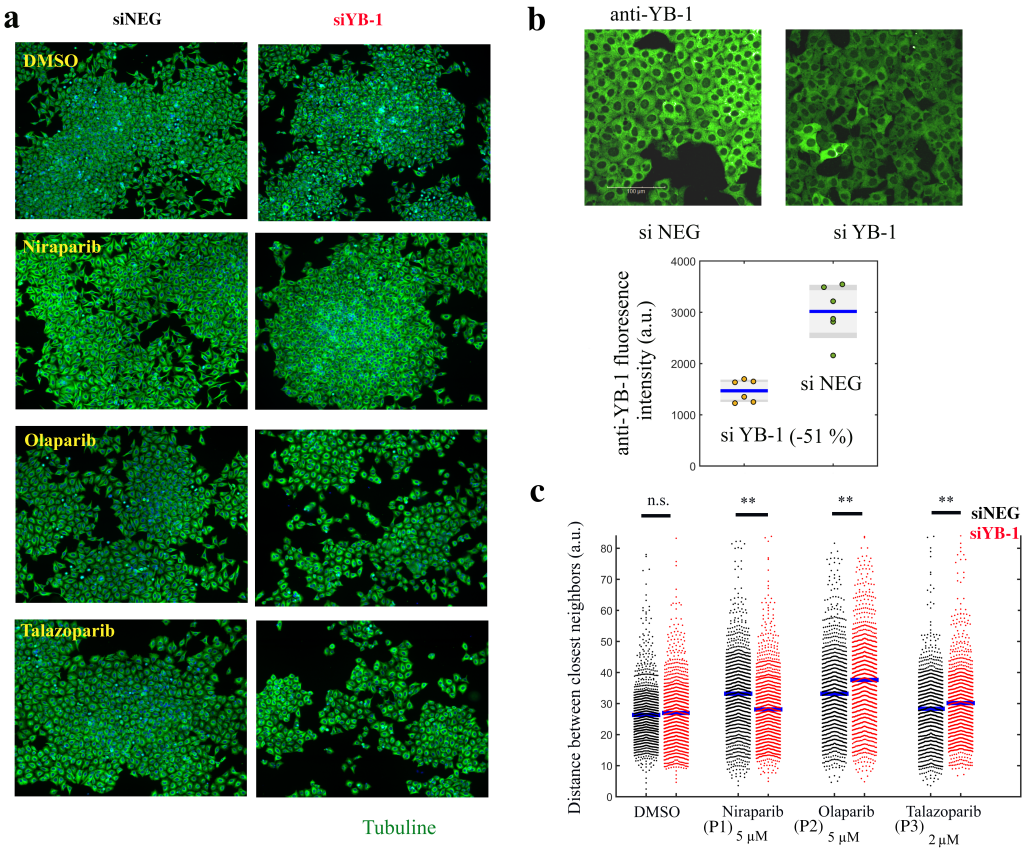


Figure 8–Figure supplement 4. (a) Representative cell images after indicated treatments for 48 h (staining: anti-tubulin). (b) Analysis of YB-1 level in HeLa cells after treatment with siRNA-1 or siNEG. siRNA treatment, under such conditions, allows to significantly repress YB-1 expression in most cells (c) Automatic measurement of the distance between closest neighbors at the single cell level after indicated treatments. $p < 0.01$; two-tailed t-test.

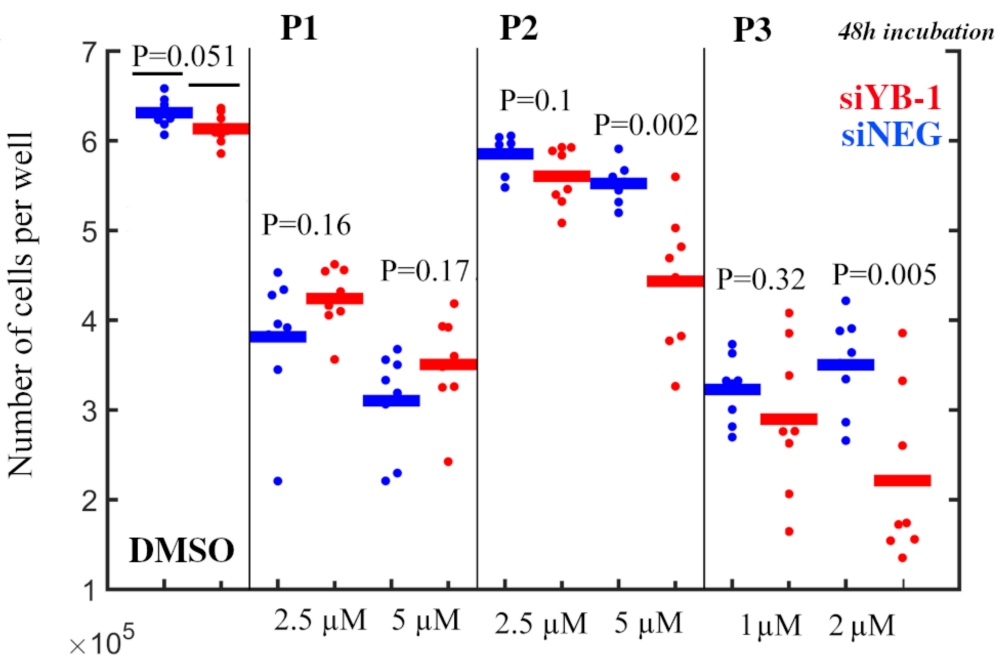


Figure 8–Figure supplement 5. Number of cells *per* well after indicated treatment (48 h) for siRNA- or siNEG-treated cells (siRNA-1). *t*-test with two tails (octuplates). siNEG vs siYB-1.