nature chemistry

ARTICLES

PUBLISHED ONLINE: 17 JULY 2017 | DOI: 10.1038/NCHEM.2811

Cell-permeable nanobodies for targeted immunolabelling and antigen manipulation in living cells

Henry D. Herce[1*†], Dominik Schumacher[2,3*†], Anselm F. L. Schneider[2,3], Anne K. Ludwig[1], Florian A. Mann[2,3], Marion Fillies[4‡], Marc-André Kasper[2,3], Stefan Reinke[2], Eberhard Krause[2], Heinrich Leonhardt[5], M. Cristina Cardoso[1*] and Christian P. R. Hackenberger[2,3*]

Functional antibody delivery in living cells would enable the labelling and manipulation of intracellular antigens, which constitutes a long-thought goal in cell biology and medicine. Here we present a modular strategy to create functional cell-permeable nanobodies capable of targeted labelling and manipulation of intracellular antigens in living cells. The cell-permeable nanobodies are formed by the site-specific attachment of intracellularly stable (or cleavable) cyclic arginine-rich cell-penetrating peptides to camelid-derived single-chain VHH antibody fragments. We used this strategy for the non-endocytic delivery of two recombinant nanobodies into living cells, which enabled the relocalization of the polymerase clamp PCNA (proliferating cell nuclear antigen) and tumour suppressor p53 to the nucleolus, and thereby allowed the detection of protein–protein interactions that involve these two proteins in living cells. Furthermore, cell-permeable nanobodies permitted the co-transport of therapeutically relevant proteins, such as Mecp2, into the cells. This technology constitutes a major step in the labelling, delivery and targeted manipulation of intracellular antigens. Ultimately, this approach opens the door towards immunostaining in living cells and the expansion of immunotherapies to intracellular antigen targets.

Antibodies are powerful biological macromolecules used extensively for a wide range of applications from basic research to some of the most successful medical therapies[1,2]. Important applications include antigen detection in cells and tissues, visualization of protein–protein interactions[3] and potent therapeutic inhibitors[4,5]. However, a central limitation is that antibodies are not able to penetrate living cells, which limits most of their applications to either extracellular targets or fixed and permeabilized tissues. Therefore, intense research is currently being invested towards developing general methods to make antibodies available within living cells[6–8]. It is, in principle, possible to obtain recombinant antibody fragments and antibodies that are optimized for intracellular expression, so-called intrabodies, by transfecting cells; however, the cytosolic expression of antigen-binding proteins is complex and remains challenging because antibodies are often not stable in the reducing environment of the cytosol[6,8–13]. Moreover, as the *in vivo* transfection of recombinant DNA presents multiple serious safety concerns, this approach remains limited to research purposes. Expanding the application of antibodies to the vast number of intracellular targets demands alternative strategies that would enable the permeation of antigen-binding proteins into living cells[14].

Many different methods for the cellular delivery of proteins have been developed within the past decades, among them the use of nanocarriers[15], nanoparticles[16], supercharged proteins[17–19], lipid-mediated delivery systems[20], disruption of the plasma membrane[8,9] and anthrax toxin protective antigen[21]. However, a severe limitation of several of these techniques is that the proteins follow an endocytosis-dependent mode of cellular uptake, and thus the cargo suffers endosomal entrapment and lysosomal degradation. Endosomal escape often requires membrane destabilization caused by lipids, the 'proton sponge effect'[22], or protein unfolding to allow its translocation through transmembrane pores[23]. A prominent strategy for the non-endocytic intracellular delivery of molecules consists of decorating them with cell-penetrating peptides (CPPs) composed of short amino acid sequences rich in arginine amino acids[24–27]. An exceptional aspect of these peptides is that they can directly cross the cell plasma membrane. Although linear arginine-rich peptides enable the efficient direct cellular entry of small molecules, such as short peptides and fluorophores, this pathway is significantly reduced for larger cargoes, such as proteins[28]. Recently, we discovered that employing cyclic arginine-rich peptides significantly enhances this pathway[29]. Cyclic arginine-rich peptides are able to enhance dramatically the non-endocytic cellular uptake kinetics[29] and cargo size[30] relative to their linear counterparts to deliver green fluorescent protein (GFP) (27 kDa) efficiently into living cells using a site-specifically conjugated cyclic TAT (cTAT) peptide[30].

Based on these encouraging results, we envisioned transferring this approach to the design of cell-permeable antibodies. To accomplish this goal, we proposed a general strategy based on two essential modular components: nanobodies (camelid-derived single-chain VHH antibody fragments[31]) for intracellular antigen targeting,

[1]Department of Biology, Technische Universität Darmstadt, Schnittspahnstrasse 10, 64287 Darmstadt, Germany. [2]Chemical Biology Department, Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP), Robert-Rössle-Strasse 10, 13125 Berlin, Germany. [3]Department of Chemistry, Humboldt Universität zu Berlin, Brook-Taylor-Strasse 2, 12489 Berlin, Germany. [4]Max Delbrück Center for Molecular Medicine (MDC), Robert-Rössle-Strasse 10, 13125 Berlin, Germany. [5]Department of Biology II, and Center for Integrated Protein Science Munich, Ludwig Maximilians Universität München, Großhadenerstrasse 2, 82152 Martinsried, Germany. †These authors contributed equally. ‡Present addresses: Charité – Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany. *e-mail: hdherce@gmail.com; dschumacher@fmp-berlin.de; cardoso@bio.tu-darmstadt.de; hackenbe@fmp-berlin.de

and cyclic CPPs (cCPPs) for the transport of the nanobody into living cells. As the size of the CPP cargo is an important parameter for the direct cellular transport of molecules with CPPs[28] we rationalized that nanobodies would be an optimal type of antibody for our goal. VHHs are small (13–14 kDa) antibody fragments that consist of just one variable heavy-chain domain of a heavy-chain antibody[32]. In addition to their reduced size, nanobodies have an increased stability and solubility compared with those of conventional immunoglobulins, which makes them ideal for intracellular applications[33]. These superior properties led to an extensive use of VHHs in molecular and cellular biology[31,33,34]. A second aspect in our design includes the decoration of these nanobodies with the highly efficient CPP cyclic R_{10} (cR_{10}) peptide, compared with the previously employed cTAT peptide[29]. Based on these considerations, we synthesized site-specifically functionalized nanobodies by expressed protein ligation (EPL), and studied their ability to cross the cellular membrane and to bind their antigen in the interior of living cells. Moreover, we tested their ability for the co-transport of antigens as well as their capacity to target, manipulate and label intracellular antigens and their interacting partners.

Results and discussion

Cell-permeable nanobody visualization and quantification strategy. To visualize the cellular uptake and assess the targeted manipulation of antigens mediated by cell-permeable nanobodies in individual living cells without altering or increasing the molecular size of the cargo, we exploited our previous observation that arginine-rich CPPs have a high affinity for RNA and associate strongly with the nucleolus[35,36]. Here we show that cyclic arginine-rich CPPs have a similar tendency to accumulate in the nucleolus (Fig. 1a and Supplementary Fig. 1). Essentially, based on this observation, we aimed to engineer high-affinity GFP nanobodies (GFP-binding proteins (GBPs)) coupled to arginine-rich CPPs and exploit this particular subcellular localization of cyclic arginine-rich CPPs to target and manipulate the respective antigen (GFP) by relocalizing it to the nucleolus. We hypothesized that in cells expressing GFP, in which GFP is mostly excluded from the nucleolus (Fig. 1a), the cellular uptake of these cell-permeable nanobodies would translate in an increased GFP signal at the nucleoli. As illustrated in Fig. 1b, this change in GFP distribution would be a consequence of the arginine-rich CPP affinity for RNA within the nucleolus. The CPP would bind nucleolar RNA and consequently relocalize the GFP bound to the GFP-binding nanobody. The nucleolus is the largest subnuclear structure and can be recognized using transmitted light microscopy to facilitate its detection, visualization and quantification. This assay simultaneously probes whether the nanobodies undergo a free cytosolic uptake, being able to reach the nucleus, and whether they remain functional within living cells after being decorated by cCPPs. To demonstrate the broad applicability of this assay, we used two different high-affinity GFP-targeting nanobodies with different binding epitopes, GBP4 (**1**) and GBP1 (**2**) (Fig. 2a and Supplementary Fig. 2)[37]. On conjugation with arginine-rich CPPs they would enable an immediate readout of the cell permeation and intact function after transduction into GFP-expressing cells (Fig. 1b). Furthermore, this readout allows for the automated quantification of cellular uptake (as described in Supplementary Fig. 42).

Synthesis of cell-permeable nanobodies. As the antigen-binding region of the nanobodies is located close to their N terminus (Fig. 2a), we envisioned the C terminus to be the best-suited site for the attachment of cCPPs with a minimal impact on the nanobodies' function. Hence, we aimed to generate cell-permeable nanobodies that are specifically functionalized with a single CPP. An ideal strategy for the C-terminal modification of proteins with

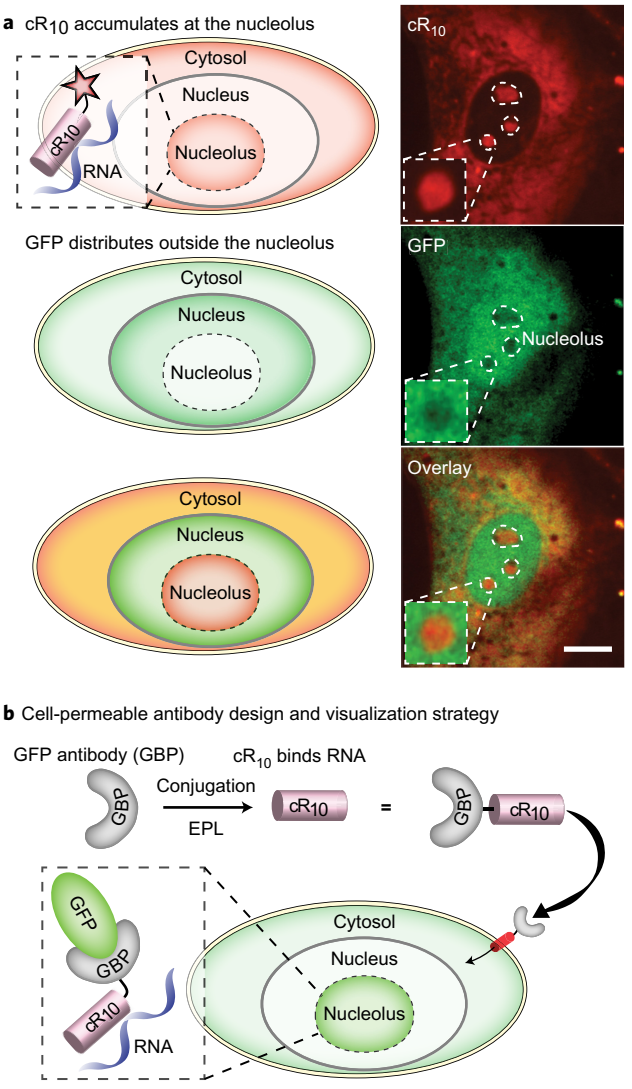


Figure 1 | Strategy to visualize the successful transduction of functional cell-permeable nanobodies. a, Arginine-rich peptides, such as cR_{10} (shown here labelled with tetramethylrhodamine (TAMRA)), have a high affinity for RNA and this drives their distribution primarily at the cytosol and nucleolus, whereas GFP alone is mostly excluded from the nucleolus. 3T3 cells incubated with the TAMRA-labelled peptide cR_{10} for 1 h show that the peptide is accumulating in the nucleolus and GFP remains in the nucleus and cytosol. Scale bar, 10 μm. **b**, Coupling cR_{10} to a nanobody would make it cell permeable. Provided the nanobody remains functional and as the antigen of the nanobody is GFP, the cellular uptake of this cell-permeable nanobody would result in a clear change in the subcellular distribution of GFP as it becomes accumulated into the nucleolus.

chemical probes is EPL[38–41]. Here the protein is expressed as an intein fusion protein that enables the formation of an active thioester. This thioester can be further captured by a synthetic peptide that carries an N-terminal cysteine, and consequently allows the addition of almost any unnatural group to the protein's C terminus via a native peptide bond.

We started our investigation by exploiting the semisynthesis of GBP4 (**1**) and planned to transfer the respective findings of the cCPP attachment to GBP1 (**2**). As we were not able to obtain sufficient amounts of GBP–cCPPs using internal Cys residues as the ligation sites (Supplementary Fig. 3), we proposed the expression of full-length GBP4 (**1**) and GBP1 (**2**) as intein fusions (**1A** and **2A**) to allow the ligation with a synthetic Cys peptide and thereby

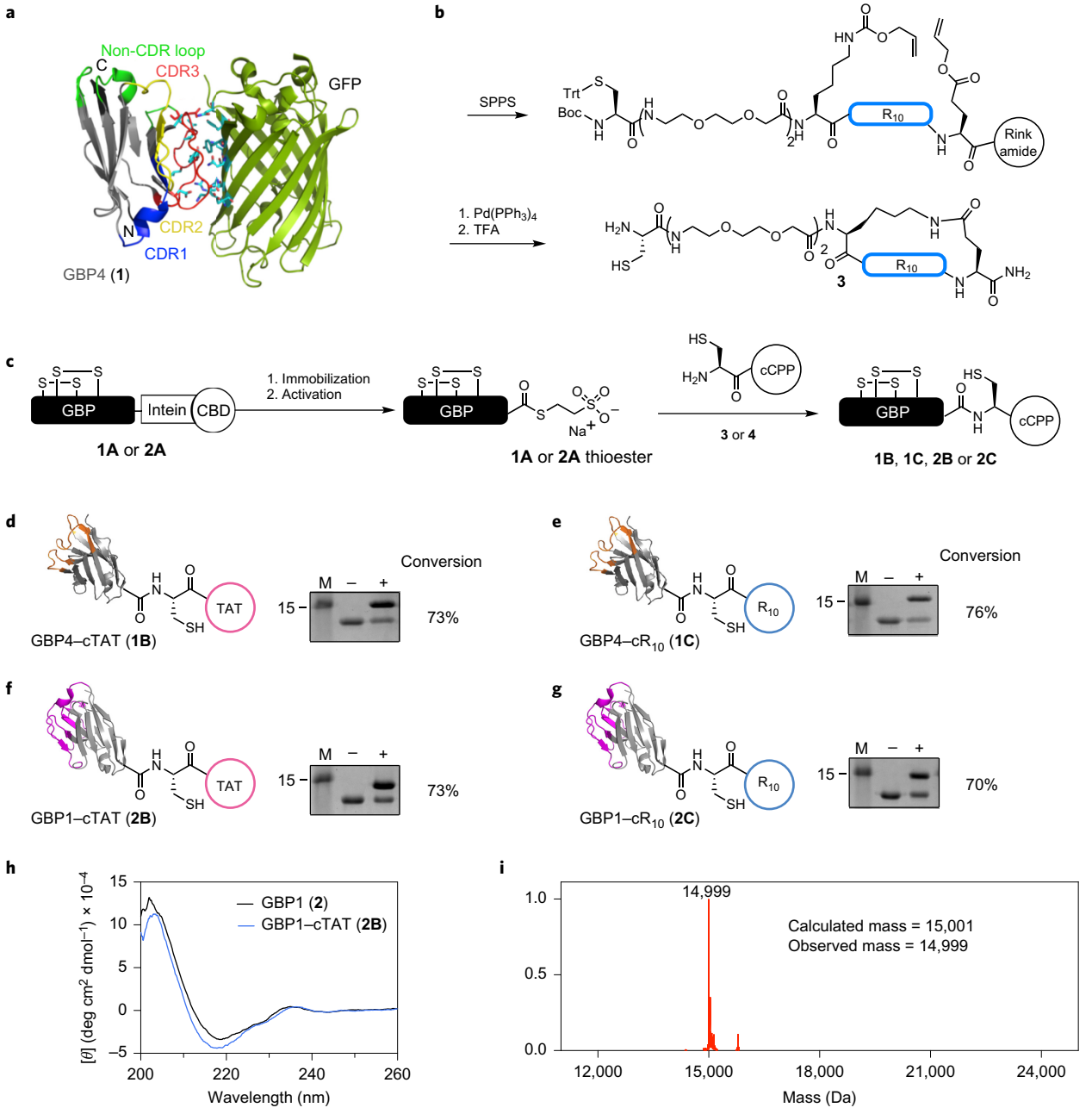


Figure 2 | Synthesis of cell-permeable antigen-binding proteins. a, X-ray structure of GBP4 (**1**) binding to its antigen GFP. The variable antigen-binding CDRs (complementary determining regions) 1–3 are highlighted in blue, yellow and red, respectively. The loops of the conserved nanobody framework that are the most distal to the antigen-binding interface and thus best for site-specific functionalization are highlighted in light green (PDB ID: 3G9A)[37] and GFP is drawn in dark green. **b**, Scheme for the synthesis of Cys–cR_{10} (**3**). Trt, trityl. **c**, Synthetic strategy of GBP–cCPPs. The full length nanobody (**1A** or **2A**) is expressed as an intein–CBD fusion and is reacted with the Cys-containing CPP (**3** or **4**) by MESNA-induced ligation. **d**–**g**, EPL of Cys–cCPPs to GBP4 (**1**) (**d** and **e**) and GBP1 (**2**) (**f** and **g**). The gel shifts in SDS–PAGE show an efficient conjugation (conversions of the NCL reactions are determined by SDS–PAGE analysis based on the ratio of conjugated and unconjugated nanobodies. +, EPL mixture that includes Cys–cCPP; –, EPL reaction that lacks Cys–cCPP; M, marker (kDa)). **h**, CD spectrum that compares the secondary structure of recombinant GBP1 (**2**) and GBP1–cTAT (**2B**). **i**, Deconvoluted electrospray ionization–mass spectrometry (ESI–MS) spectrum of GBP1–cTAT (**2B**).

place an additional Cys at the C terminus of the nanobody. For this, we synthesized cR_{10} (**3**) and cTAT (**4**) peptides equipped with an N-terminal Cys residue for native chemical ligation (NCL) using a slightly adapted version of our previously published solid-phase peptide-synthesis (SPPS) protocol (2.6% yield for **3** (Fig. 2b and Supplementary Scheme 2) and 4.8% for **4** (Supplementary Scheme 3))[30]. The EPL strategy is outlined in Fig. 2c. The subsequent NCL of both peptides **3** and **4** to GBP4 (**1**) gave high yields of the desired products (73% GBP4–cTAT (**1B**) and 76% GBP4–cR_{10} (**1C**) (Fig. 2d,e)). To obtain high intein cleavage and ligation yields for GBP1 (**2**), three alanine residues were placed between the nanobody and the intein. Thereby, we could synthesize GBP1–cTAT (**2B**) and GBP1–cR_{10} (**2C**) conjugates with conversion rates of 73 and 70% (Fig. 2f,g). To elucidate the impact of cyclization of the CPPs, we also synthesized linear R_{10} (**5**) and TAT (**6**) peptides and performed EPL to GBP1 (**2**) using the same strategy as for the cyclic peptides to give GBP1–linTAT (**2D**) and GBP1–linR10 (**2E**) (Supplementary Fig. 4). The expression of GBP–linTAT and

GBP–linR_{10} fusions was not efficient, which makes the introduction of CPPs to the nanobodies by EPL essential (Supplementary Fig. 5).

Circular dichroism (CD) spectra confirmed the conserved secondary structure of the functionalized nanobodies compared to the recombinantly expressed wild-type **1** and **2** (Fig. 2h and Supplementary Fig. 6). Moreover, all the conjugates were characterized using mass spectroscopy (MS) analysis (Fig. 2i, Supplementary Fig. 31 for **2B** and Supplementary Figs 30 and 33–39 for **1B**, **1C**, **2A** and **2C**–**2E**). Disulfide bond formation within the nanobodies synthesized by EPL was shown using selective maleimide coupling and 5,5′-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent) (Supplementary Fig. 7). To validate the binding to the antigen and potential modulation of the antigen's fluorescence spectra, as reported before, microscale thermophoresis and a previously reported GFP-fluorescence *in vitro* assay were used. The results are shown in Supplementary Figs 8–10 and reveal the full functionality and intact GFP binding of the CPP conjugates[34,37].

Live-cell uptake of fully functional antigen-recognizing proteins. To detect and quantify the cellular permeation of functional nanobodies, as described in Fig. 1, we generated a mammalian cell line (mouse 3T3 fibroblasts) that stably expresses GFP (Supplementary Information). In this cell line, GFP is present in the cytosol and most prominently in the nucleus (because of the addition of a nuclear localization signal) and is mostly excluded from the nucleolus (Fig. 3a). When the 3T3 GFP cells were incubated with different GFP cell-permeable nanobodies, their uptake changed the intracellular distribution of GFP, which resulted in a higher distribution of GFP in the nucleolus than in the rest of the cell. We incubated the cells with 20 μM of GBP1–cR_{10} (**2C**) and this, indeed, resulted in a consistent and significant increase in the GFP distribution in the nucleolus (Fig. 3b), which was further validated by two different nucleolar markers (Supplementary Fig. 11). Next, we used this readout to compare the efficiency of cTAT to that of cR_{10} peptides in enabling the uptake of nanobodies. To assess the relative uptake efficiencies of these cell-permeable nanobodies, we measured the cellular uptake of both GFP-binding nanobodies **1** and **2** (GBP4 and GBP1), each ligated to the two cyclic arginine-rich CPPs, **3** and **4** (cTAT and cR_{10}). The 3T3 GFP cells were incubated with the cell-permeable nanobody (20 μM) for one hour at room temperature, and were subsequently imaged. The relocalization of GFP to the nucleolus was used to quantify the number of positive cells that display a successful uptake (Fig. 3c for GBP1_cTAT (**2B**) and GBP1_cR_{10} (**2C**) and Fig. 3d for GBP4_cTAT (**1B**) and GBP4_cR_{10} (**1C**)). Both nanobodies consistently showed a threefold increase in the number of cells that take up the nanobody when coupled to cR_{10} relative to cTAT. These findings confirm that the enhanced efficiency of cR_{10} peptides, which has been shown before for small molecules, is preserved when coupled to larger proteins, which makes them more suitable for an efficient nanobody delivery[29].

Targeted manipulation of antigens in living cells. Next, we asked if the cell-permeable nanobodies would bind and relocalize functional GFP-labelled proteins within living cells. For this purpose, we performed cellular uptake studies in a human cervical cancer cell line (HeLa) that expressed GFP–PCNA (**8**)[42] (PCNA, proliferating cell nuclear antigen) fusion proteins by live-cell microscopy. PCNA serves as a loading platform for several repair and replication factors and is recruited to subnuclear sites of DNA during the DNA synthesis phase (S phase)[43], whereas it is mostly excluded from the RNA-containing nucleoli (Fig. 4a). As arginine-rich CPPs accumulate mostly at the nucleolus (Fig. 1a), we hypothesized that after the cellular delivery of the GBP1–cR_{10} (**2C**), binding to GFP–PCNA would occur in the nucleus followed by its relocalization to the nucleolus (Fig. 4a). As shown in Fig. 4b (and Supplementary Fig. 12), GFP–PCNA was relocalized to the nucleolus during both the gap phase (G phase) and the S phase of the cell cycle. In this case, the cells were incubated with 10 μM of the cell-permeable nanobody GBP1–cR_{10} (**2C**) for one hour, then washed and imaged. Moreover, after 24 hours the cells were still viable and able to accomplish DNA synthesis (assayed by the cell's ability to incorporate the thymidine analogue BrdU into newly synthetized DNA *in vivo*) and the PCNA fusion remained bound to the cell-permeable nanobody GBP1–cR_{10} (**2C**), which indicates a high and long-term stability of the cell-permeable nanobodies inside the living cells (Fig. 4c). Hence, we could show that not only intracellular GFP (Fig. 3) but also GFP-fusion proteins, which participate in physiological processes, could be manipulated, opening up the possibility to study cellular (patho)physiology. Consequently, we focused on this assay to quantify further the uptake efficiency of different cell-permeable nanobodies delivered under different conditions.

To probe the difference between linear and cCPPs, we compared their uptake into HeLa cells (Supplementary Fig. 13, one hour incubation, 10 μM, 25 °C). The cyclic arginine-rich peptides are consistently more efficient in cellular uptake than the linear versions and the R_{10} peptides are more efficient than the TAT peptides. Next, four concentrations were tested, 2.5, 5, 10 and 20 μM of the GBP1–cR_{10} (**2C**, one hour incubation, 25 °C in HEPES buffer (Supplementary Fig. 13)). This cell-permeable nanobody was chosen as it proved to be taken up the most efficiently. The uptake efficiency reached a plateau at 10 μM. At this concentration, almost all the cells took up the GBP1–cR_{10}. To test whether the cellular uptake was energy independent, the uptake of GBP1–cR_{10} (**2C**) at 4 °C was compared with the uptakes at 25 and 37 °C (one hour incubation, 10 μM in HEPES buffer). At all the temperatures, the cellular uptake was observed in over 90% of the cells (Supplementary Fig. 13). Moreover, the cellular uptake of GBP1–cR_{10} (**2C**) at 10 and 20 μM was studied in the presence of cell-culture media (DMEM) plus serum. In this case, the concentration of the nanobody had to be increased to reach similar results to those when a HEPES buffer was used. Supplementary Fig. 13 and Supplementary Movie 1 show a time lapse of the uptake of GBP1–cR_{10} (**2C**) in the presence of DMEM plus serum at 20 μM and 37 °C. Within the first ten minutes of incubation, almost 30% of the cells had taken up the cell-permeable nanobody, consistent with a non-endocytic mode of uptake. That we assay for antigen relocalization furthermore reinforces that the nanobody is not encapsulated in cytosolic vesicles. Finally, we added TAMRA-labelled cR_{10} to unconjugated GBP1 and performed uptake experiments. Although the cR_{10} is efficiently taken up and localized at the nucleolus of the cells, GBP1 was not internalized at all (Supplementary Fig. 14). These results establish that CPP conjugation to the nanobody is essential for cellular uptake. Furthermore, cCPPs are taken up more efficiently than their linear counterparts, their uptake is non-endocytic, concentration- and time-dependent and can take place even in the presence of serum in the media.

Encouraged by this these promising results, we next tested whether we could also relocalize both the antigen and an interacting protein simultaneously.

Visualization of protein–protein interactions in living cells. Protein–protein interactions are essential to all processes in living cells. However, most methods to study these interactions rely on *in vitro* biochemical assays, which disregard the innate complexity of living cells. The fluorescent three-hybrid (F3H) strategy is a versatile approach to study protein–protein interactions in living cells and is based on the GBP nanobody coupled to an intracellular localization sequence[3]. This fusion–protein complex

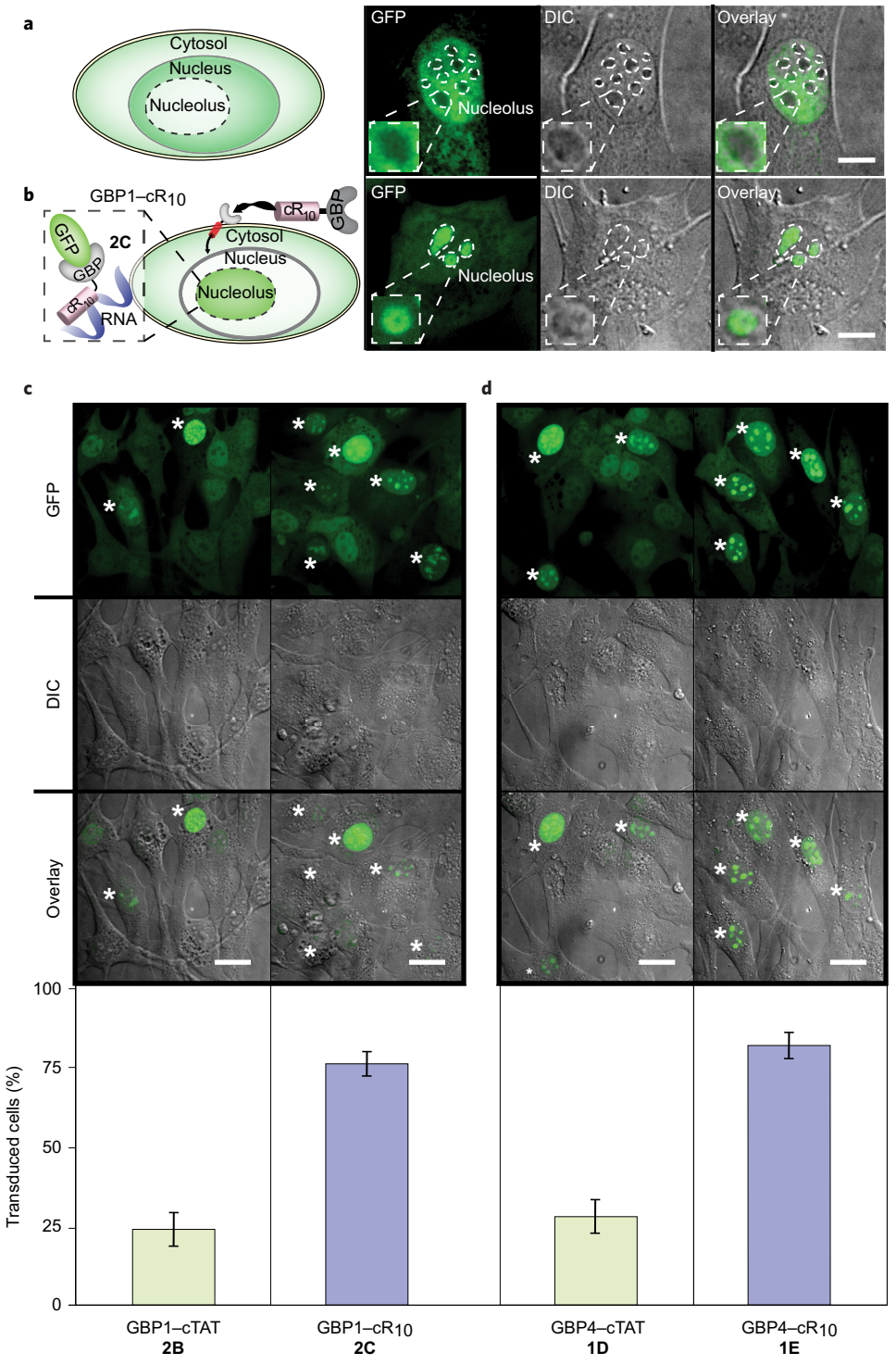


Figure 3 | Nanobodies efficiently transduce into living cells when coupled to cR_{10} (3), and remain functional and able to bind and relocalize intracellular GFP to the nucleolus. a, 3T3 cells that express GFP are distributed outside the nucleolus. **b**, 3T3 cells incubated with 20 μM of GBP1–cR_{10} (**2C**) for 1 h show that the complex is able to permeate into the living cell where it relocalizes GFP into the nucleolus. Scale bar, 10 μm. **c**,**d**, The uptake efficiency of **2B**, **2C**, **1B** and **1C** was measured by counting the proportion of cells with GFP relocalized to the nucleolus. Cells that display relocalized GFP at the nucleolus were scored as positive (indicated by asterisks). The bar plot shows the percentage of positive cells imaged in each case. Clearly, both nanobodies when coupled to cR_{10} show a threefold increase in cells that take up the conjugates compared with cTAT, quantified by the number of cells that show nucleolar accumulation of GFP (based on the average of 350 cells from three biological replicates). Error bars, s.d. Scale bars, 20 μm.

recruits GFP-tagged proteins along with other fluorescently tagged interacting partners to specific regions, or structures, within cells, which leads to a powerful way to visualize protein–protein interactions in living cells. However, this strategy requires the transfection of cells with the plasmid coding for the nanobody fused to a localization sequence. We hypothesized that the cell-permeable nanobody GBP1–cR_{10} (**2C**) could expand the standard F3H method and so enable the immediate detection of

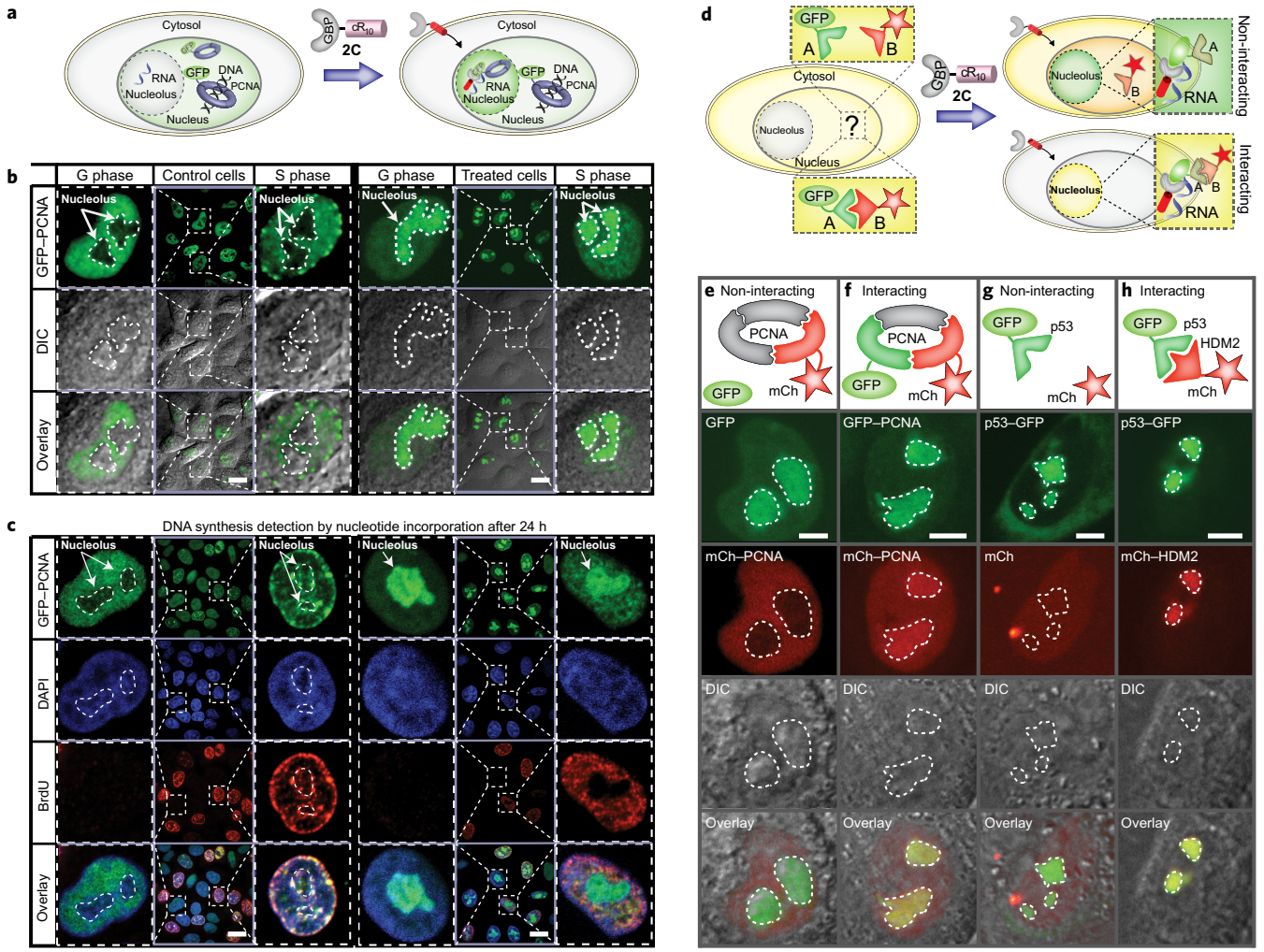


Figure 4 | The cell-permeable GFP-binding antibody 2C targets and relocalizes GFP-labelled functional proteins along with other interacting proteins, which facilitates the detection of protein–protein interactions in living cells. a, Scheme of the relocalization to the nucleolus of GFP–PCNA. **b**, In the absence of GBP1–cR_{10} (**2C**), GFP–PCNA is distributed outside the nucleolus at all stages of the cell cycle. The cell-permeable antibody GBP–cR_{10} (**2C** (10 μM)) transduces into living cells and can relocalize GFP–PCNA to the nucleolus at all stages of the cell cycle. Scale bars, 20 μm. **c**, Halogenated nucleotide incorporation (BrdU) in cells undergoing DNA synthesis 24 h after cell-permeable antibody delivery. Cells remain viable and the cell-permeable nanobody is stable, as after 24 h it remains bound and keeps relocalizing GFP–PCNA to the nucleolus. Scale bars, 20 μm. **d**, Use of cell-permeable nanobodies to facilitate the detection of protein–protein interactions in living cells. If the proteins do not interact, only the GFP-tagged protein is relocalized to the nucleolus by GBP1–cR_{10} (**2C** (10 μM)), whereas if the proteins interact both proteins become relocalized to the nucleolus. **e**, GFP is relocalized to the nucleolus by GBP1–cR_{10} (**2C**), but mCherry–PCNA is not relocalized, because it does not interact with GFP. **f**, PCNA oligomerizes to form a homotrimer; consequently, in cells that express GFP–PCNA and mCherry–PCNA, the PCNA trimers contain a mix of the two fluorescently labelled proteins, GBP1–cR_{10} (**2C**) relocalizes trimers that contain GFP–PCNA along with mCherry–PCNA. **g**, p53-GFP is relocalized to the nucleolus by GBP1–cR_{10} (**2C**), whereas non-interacting proteins, such as mCherry, are not. **h**, HDM2 binds and blocks the tumour suppressor protein p53, **2C** relocalizes p53–GFP and the interacting protein regulator mCherry–HDM2 to the nucleolus. Scale bars (**e**–**h**), 5 μm.

protein–protein interactions without the need to transfect the cells. As shown in Fig. 4d, the cell-permeable nanobody would drive the relocalization to the nucleolus of the GFP-tagged protein of interest and potentially bring along other interacting proteins labelled with a different fluorophore. Accordingly, if the two proteins are non-interacting, only the GFP-tagged protein would be recruited to the nucleolus. In contrast, if the proteins interact the second protein would be recruited to the nucleolus and display a similar redistribution of its fluorescent signal.

To test this hypothesis, we incubated HeLa cells that expressed fusion proteins tagged with mCherry as well as GFP tagged with GBP1–cR_{10} (**2C**). After the successful cellular uptake of GBP1–cR_{10} (**2C**) and binding to its antigen GFP, protein–protein complexes should be relocalized to the nucleolus. As a first proof of concept, we applied this to cells that expressed PCNA proteins in different colour fusions, which form homotrimers (Fig. 4e,f). When GFP was uncoupled from PCNA, only GFP was recruited to the nucleolus on the uptake of GBP1–cR_{10} (**2C**), whereas the non-interacting protein (mCherry–PCNA) was found mostly excluded from the nucleolus (Fig. 4e). When PCNA was fused to GFP, mCherry–PCNA interacted with GFP–PCNA to form a trimer, and both fusion proteins were successfully co-recruited to the nucleolus (Fig. 4f).

Next, we applied this strategy to visualize one of the most important protein–protein interactions in cancer, the interaction between the tumour suppressor p53 (nicknamed the 'guardian of the genome') and HDM2. In this case, we used the human osteosarcoma cell line U2OS, because these cells express wild-type p53

suppressed by an overexpression of *HDM2*, and thereby constitute a relevant cell line to study this particular protein–protein interaction. Figure 4g shows that GBP1–cR_{10} (**2C**) transported p53–GFP to the nucleolus and the non-interacting protein mCherry remained freely distributed over the cell. On the other hand, when mCherry was fused to HDM2, both proteins relocalized to the nucleolus, which indicates that protein–protein interactions between HDM2 and p53 were taking place (Fig. 4h).

These results demonstrate that intracellular antigens and their interacting partners can be manipulated efficiently and relocalized by the cell-permeable nanobodies.

Delivery of antigen proteins by cell-permeable nanobodies. We next probed whether cR_{10}-containing nanobodies would enable the cellular delivery of purified antigen-containing proteins from the extracellular milieu into living cells (Fig. 5a). Therefore, we co-incubated GBP1–cR_{10} (**2C**) and purified eGFP (**7**) and studied the cellular uptake of this large complex (43 kDa) in mouse 3T3 and human HeLa cells. As shown in Fig. 5b, eGFP (**7** (20 μM)) was successfully co-transported into both cell lines by GBP1–cR_{10} (**2C** (20 μM)). The characteristic strong nucleolus accumulation indicated that the antigen was bound to GBP1–cR_{10} (**2C**) after the cellular uptake of the complex.

We next asked whether this approach could be used to deliver full-length proteins labelled with GFP into living cells. To elucidate this, we used PCNA tagged with GFP (**8** (Supplementary Fig. 15)), which, as shown in Fig. 4b, when bound by GBP1–cR_{10} (**2C**) accumulates at the nucleolus. As depicted in Fig. 5c, we analysed whether GBP1–cR_{10} would be able to co-transport GFP-tagged PCNA (**8** (Supplementary Fig. 15)) into living cells. Indeed, co-incubating 20 μM of GBP1–cR_{10} (**2C**) with 20 μM of GFP–PCNA resulted in the accumulation of GFP–PCNA at the nucleolus of HeLa cells, which shows that GBP1–cR_{10} is able to mediate the transport into living cells of GFP–PCNA (Fig. 5d). In this experiment, we used a HeLa cell line that stably expressed mCherry–PCNA[42] to probe simultaneously the specificity for GFP of **2C**. As can be seen in Fig. 5d, the mCherry-tagged PCNA was excluded from the nucleolus, whereas the GFP–PCNA fusion was recruited to the nucleolus. Accordingly, we conclude that GBP1–cR_{10} (**2C**) is capable of the unambiguous transport of GFP-tagged proteins into living cells.

Next, we asked whether the use of cell-permeable nanobodies would allow the co-transport of other full-length proteins. For this, we chose the protein Mecp2, whose gene is mutated in patients with Rett syndrome. Rett syndrome is a devastating neurological disease in which neurons lack functional Mecp2 proteins[44,45]. Experiments in mice showed that symptoms of Rett syndrome can be reversed by systemic viral delivery of the functional gene[46]. Accordingly, these data provide a promising therapeutic approach to treat Rett patients by making the functional *MECP2* gene available within living cells. However, viral therapies in humans present several technical and safety challenges, such as the uncontrolled delivery of a gene into the host genome, the potential induction of cancer and immune responses to the viral vector[47]. Even though proteins sometimes also elicit an immune response, alternative methods that could directly supply functional proteins are desirable and novel approaches for the delivery of Mecp2 should be developed. Given this, we produced and purified the fusion protein Mecp2–GFP (**9** (Supplementary Fig. 15)) and incubated it (20 μM) with the GBP1–cR_{10} (**2C** (20 μM)) for 30 minutes to allow the binding of the cell-permeable antibody to its antigen (GFP). The complex was then added to mouse 3T3 cells and incubated for one hour before the cells were washed and imaged. Mouse cells, such as 3T3, present well-defined DNA-dense structures composed of tandem arrays of major satellite DNA, to which Mecp2 binds with a very strong affinity[48]. Therefore, we hypothesized

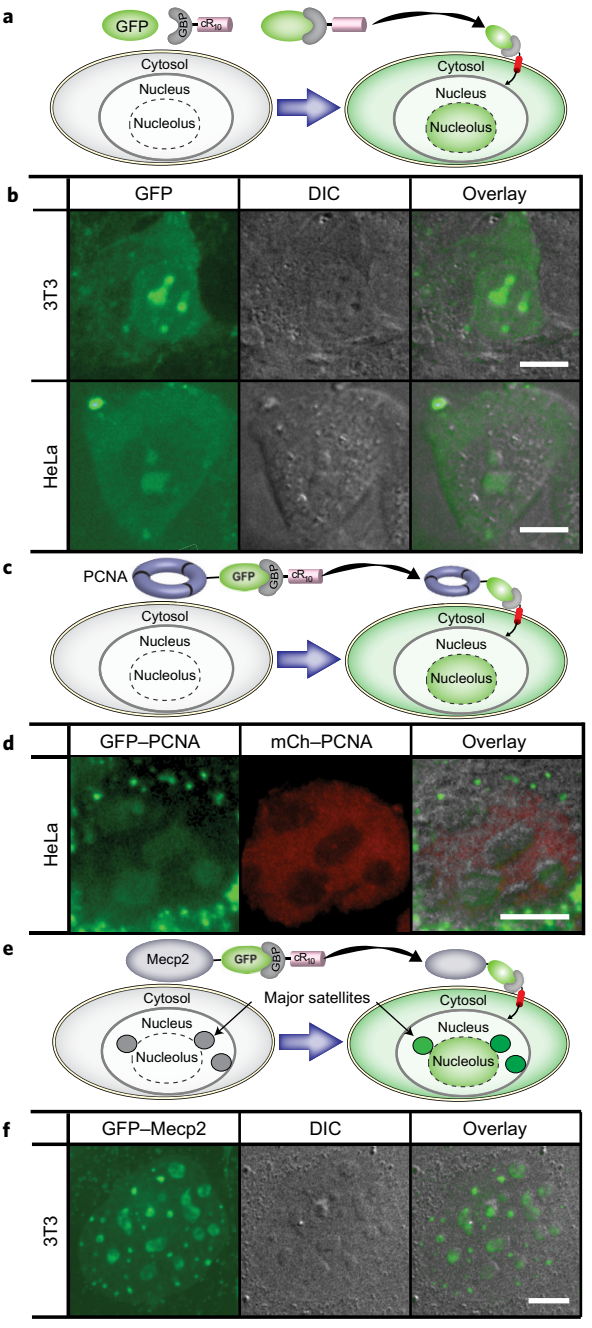


Figure 5 | The cell-permeable antibody is able to bind its antigen and transport it into living cells. a, The scheme shows the extracellular addition of eGFP (**7** (20 μM)) and GBP1-cR_{10} (**2C** (20 μM)); the GFP antibody binds eGFP and the complex enters into living cells. **b**, A 3T3 cell and a HeLa cell into which eGFP has been co-transported. The accumulation of eGFP in the nucleolus indicates that it remains bound to GBP1-cR_{10} (**2C**). **c**,**d**, GFP-labelled PCNA (**8** (20 μM)) is co-transported by GBP1-cR_{10} (**2C** (20 μM)) into a HeLa cell stably expressing mCherry-PCNA and accumulates at the nucleolus, whereas the mCherry-tagged PCNA remains excluded from the nucleolus. **e**,**f**, GFP-labelled Mecp2 (**9** (20 μM)) is co-transported by GBP1-cR_{10} (**2C** (20 μM)) and binds its natural target, the chromocentres, located in the nucleus. Scale bars, 10 μm.

that, after the Mecp2–GFP cellular uptake triggered by the cell-permeable nanobody, the Mecp2 fusion would be transported to the nucleolus and, in addition, bind to the major satellite repeats (depicted in Fig. 5e). As shown in Fig. 5f, we observed that the Mecp2–GFP fusion (**9**), besides displaying nucleolar distribution,

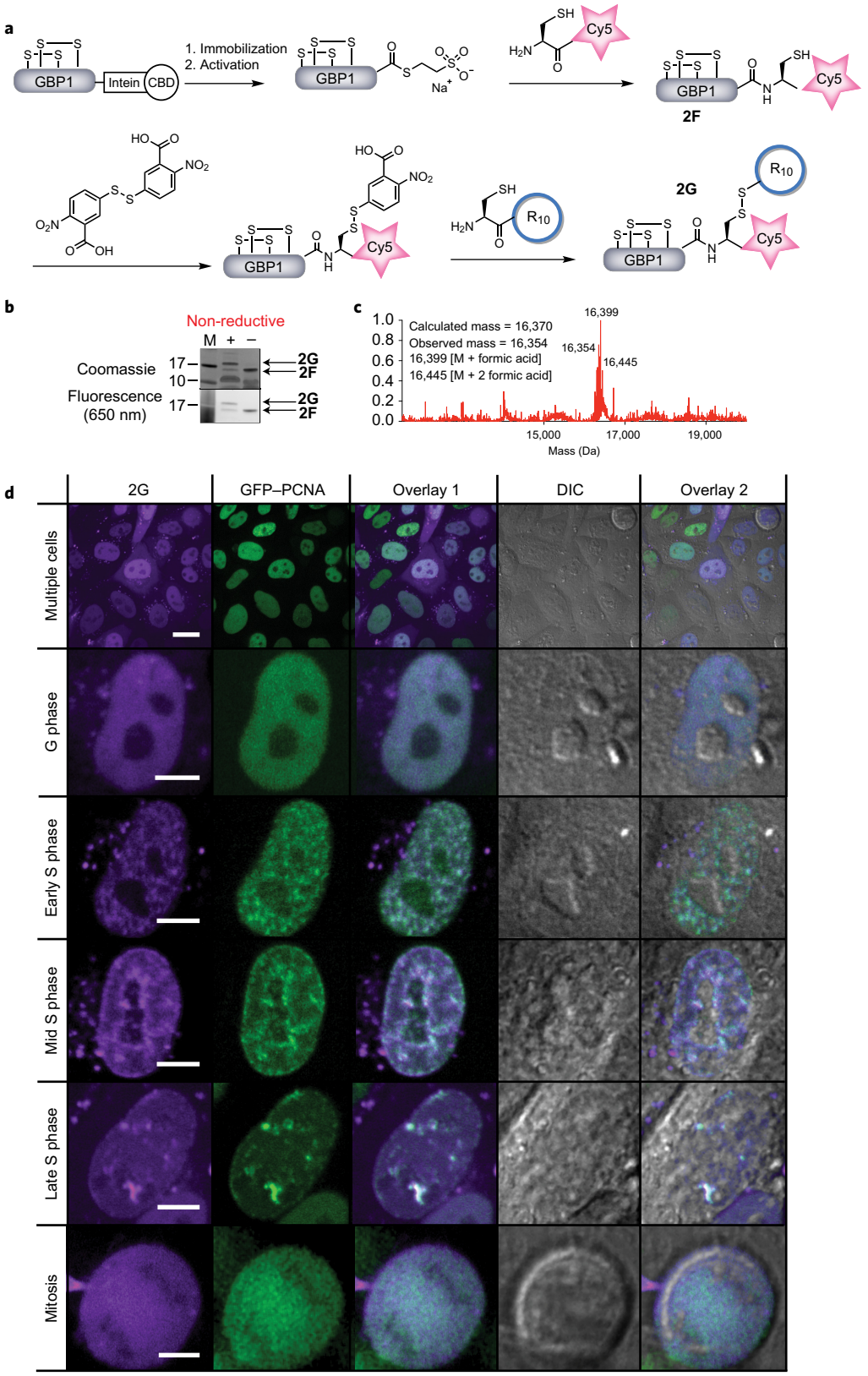


Figure 6 | Cleavable cell-permeable nanobodies for immunostaining in living cells. a, Synthetic strategy for the generation of fluorescently labelled cell-permeable nanobodies with a cleavable CPP. **b**, Non-reductive SDS–PAGE of the reaction mixture after EPL and cCPP coupling. The gel shifts in SDS–PAGE show efficient conjugation yields (yields of the NCL reaction are determined by SDS–PAGE analysis and based on the ratio of conjugated and unconjugated nanobody. +, mixture that includes Cys-cR_{10} (**3**); –, EPL reaction that lacks Cys-cR_{10} (**3**); M, marker (kDa)). **c**, Deconvoluted ESI–MS spectrum of GBP1–Cy5–S–S–cR_{10} (**2G**). **d**. The cell-permeable nanobody (**2G** (5 μM)) enters a cell and labels its nuclear antigen (GFP–PCNA) without relocalizing it to the nucleolus, which indicates an efficient cellular uptake and intracellular cleavage of the cR_{10} peptide **3** from the functional nanobody. Scale bars, 20 μm.

was, indeed, able to reach and localize at its natural target, the major satellite DNA repeats (Supplementary Figs 16 and 17).

Although uptake for the complexed eGFP, GFP–PCNA and Mecp2–GFP proteins was detected in just 2–10% of cells, these results indicate that cR_{10}-functionalized nanobodies can serve as versatile delivery vehicles for the transduction of large recombinant proteins at defined concentrations.

Immunostaining in living cells using intracellularly cleavable cell-permeable nanobodies. Up to this point we had shown that a stable covalent bond between the nanobody and the cCPP cR_{10} led to a nucleolar redistribution of the antigen. Although this could be advantageous for some applications, such as targeting specific proteins to the nucleolus, as demonstrated for the visualization of protein–protein interactions, it is generally not desirable to alter the position of the antigen for immunostaining purposes. Therefore, we synthesized a fluorescently labelled cell-permeable nanobody in which the cCPP moiety is conjugated via an intracellularly cleavable disulfide bond. For this, we performed a NCL of the GBP1–intein–CBD fusion (**2A**) to a Cy5-containing peptide (**10**) to yield GBP1–Cy5 (**2F**). Afterwards, we took advantage of the additional C-terminal cysteine residue at the ligation site and incubated **2F** with Ellman's reagent, and thereby transformed the cysteine into an electrophilic disulfide. On the addition of the N-terminal cysteine that contained cR_{10} peptide **3**, the disulfide-linked nanobody GBP1–Cy5–S–S–cR_{10} (**2G**) was obtained with a high conversion rate of 80% (Fig. 6a,b; Fig. 6c and Supplementary Figs 40 and 41 give the MS analysis; Supplementary Fig. 6 gives CD spectra).

With this conjugate in hand, we probed whether the cleavable cR_{10} moiety facilitates efficient cellular uptake and antigen labelling without nucleolar relocalization. For this, we incubated HeLa cells that stably expressed GFP–PCNA with **2G**. As shown in Fig. 6d, the Cy5-labelled cell-permeable nanobody **2G** was efficiently taken up by the cells and labelled its antigen in the nucleus, which indicates an efficient reductive cleavage of the cR_{10} peptide **3**. The cleavable cell-permeable nanobody accumulated primarily at the nucleus and was able to label DNA replication sites. This contrasted significantly with the distribution of the cleavable cell-permeable nanobody in cells that stably express mCherry–PCNA (Supplementary Fig. 18). In this case, the nanobody was located more homogeneously in the cytosol and nucleus, and, as expected, did not co-localize with DNA replication sites.

Based on the intracellular fluorescence intensity, we found that the intracellular concentration of the cell-permeable nanobody reached an average of 0.54 ± 0.01 μM relative to the extracellular concentration (Supplementary Fig. 19). At this low micromolar concentration in the extracellular milieu (5 μM) and relatively short incubation time (one hour), the intracellular concentrations exhibited a significant variability from cell to cell, consistent with the stochastic nature of the cellular binding and the particular membrane composition of the cells.

In summary (Supplementary Movie 2), we have shown, for the first time, a modular synthetic strategy for the generation of functional cell-permeable nanobodies that leads to their efficient non-endocytic cellular uptake. We showed that we can use these cell-permeable nanobodies to label and manipulate antigens and their interacting partners in living cells. Furthermore, we demonstrated that we can use them to transport their respective antigens and antigen-coupled proteins into living cells. Depending on the synthesis strategy, cCPPs were conjugated via stable amide or intracellularly cleavable disulfide bonds. Our optimized EPL protocol allowed us to conjugate cR_{10} (**3**) and cTAT (**4**) peptides efficiently to two different nanobodies. Incubating living cells with low micromolar concentrations of the newly designed conjugates was sufficient to obtain a cellular uptake of up to 95% of the cells within an hour of incubation, which resulted in immediate intracellular antigen binding, as shown by fluorescence co-localization or the relocalization of intracellularly expressed GFP to the nucleolus. The transduction efficiencies of cR_{10}-containing nanobodies were three-times higher compared with those of the cTAT derivatives, which makes cR_{10} peptides more suitable for the design of cell-permeable proteins. Furthermore, the highly efficient cR_{10} conjugate GBP1–cR_{10} (**2C**) enabled us to co-transport GFP (27.5 kDa), GFP–PCNA (63 kDa) and the therapeutically relevant Mecp2–GFP fusion protein (83 kDa) into cells. This establishes the cell-permeable nanobodies as general tools for the controlled cellular uptake of recombinant proteins and opens up new possibilities for the delivery of therapeutically relevant proteins. Moreover, the cell-permeable antigen-binding proteins allowed us to visualize protein–protein interactions in living cells and further expand the previously reported fluorescent three-hybrid assay. Finally, the development of a cleavable cell-permeable nanobody allowed the staining of antigens in living cells. We propose that the site-specific conjugation of cCPPs is well suited for the production of other cell-permeable nanobodies and constitutes a very promising tool for immunostaining and immunomanipulation in the interior of living cells.

Methods

General synthesis of cell-permeable nanobodies by EPL. The dialysed lysate with intein–CBD fusion proteins (2 ml) was loaded on 1 ml of pre-equilibrated (3 × 4 ml equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity flow through (three times). The solid phase was washed six times with 5 ml of equilibration buffer, flushed with 1 ml of cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA (sodium 2-mercaptoethanesulfonate), pH 8.5) and incubated for 18 h at room temperature with 1 ml of cleavage buffer, which included 512 or 437 nmol of peptide **3** or **4**, respectively. The ligation mixture was eluted from the chitin beads and the resin washed six times with 500 μl of washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). Product-containing fractions were pooled, incubated for 2 h at 4 °C with BioBeads (0.2 g ml^{-1} (BioRad)) and the peptide excess removed by desalting columns (Zeba Spin, 7 kDa cutoff, 5 ml (Thermo Fisher Scientific)). The solution was concentrated (Amicon Ultra-0.5, 10,000 MWCO (Merck)) to 200 μl and rebuffered via dilution/spin (seven additions of 300 μl) to 1× HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5).

Cell-permeable antibody delivery. Cells were seeded at 50% confluence in a tissue-culture-treated 384-well optical glass bottom plate (ThermoFisher) 24 h before peptide treatment. Unless specified otherwise, the cellular uptake assays were done by washing twice with a HEPES buffer and replacing the buffer solution with the HEPES buffer with the cell-permeable antibody added at a final concentration of 20 μM. The sample was incubated for 1 h, washed with cell culture media, taken to the microscope and imaged. To establish the proportion of cells that contained the cell-permeable antibody, we computed the number of cells that displayed accumulation of GFP or GFP-tagged proteins at the nucleolus relative to the total number of imaged cells[29]. The nucleus and nucleoli areas were visualized using the differential interference contrast (DIC) channel. Each experiment was repeated at least three times and the average and the standard error plotted. Cell viability was assessed with the DIC images used to detect the cell morphology throughout the experiments. To evaluate further the viability of the cells after uptake of the cell-permeable antibody we used two approaches: (1) the cell division was monitored for 16 h (imaged every 30 min) and (2) in cells that permanently expressed fluorescently tagged PCNA, we monitored the DNA synthesis by simultaneously visualizing DNA replication patterns and the incorporation of modified nucleotides (Supplementary Section 2.2 gives more details).

Data availability. The authors declare that the main data supporting the findings of this study are available within the article and its Supplementary Information. Extra data are available from the corresponding authors on request.

Received 17 May 2016; accepted 24 May 2017;
published online 17 July 2017

References

1. Chames, P., Van Regenmortel, M., Weiss, E. &, Baty, D. Therapeutic antibodies: successes, limitations and hopes for the future. *Br. J. Pharmacol.* **157,** 220–233 (2009).
2. Leavy, O. Therapeutic antibodies: past, present and future. *Nat. Rev. Immunol.* **10,** 297 (2010).

3. Herce, H. D., Deng, W., Helma, J., Leonhardt, H. & Cardoso, M. C. Visualization and targeted disruption of protein interactions in living cells. *Nat. Commun.* **4,** 2660 (2013).
4. Courtete, J. *et al.* Suppression of cervical carcinoma cell growth by intracytoplasmic codelivery of anti-oncoprotein E6 antibody and small interfering RNA. *Mol. Cancer. Ther.* **6,** 1728–1735 (2007).
5. Smith, K. G. & Clatworthy, M. R. FcγRIIB in autoimmunity and infection: evolutionary and therapeutic implications. *Nat. Rev. Immunol.* **10,** 328–343 (2010).
6. Freund, G. *et al.* Targeting endogenous nuclear antigens by electrotransfer of monoclonal antibodies in living cells. *MAbs* **5,** 518–522 (2013).
7. Kimura, H., Hayashi-Takanaka, Y., Stasevich, T. J. & Sato, Y. Visualizing posttranslational and epigenetic modifications of endogenous proteins *in vivo*. *Histochem. Cell. Biol.* **144,** 101–109 (2015).
8. Marschall, A. L., Frenzel, A., Schirrmann, T., Schungel, M. & Dubel, S. Targeting antibodies to the cytoplasm. *MAbs* **3,** 3–16 (2011).
9. Biocca, S., Ruberti, F., Tafani, M., Pierandrei-Amaldi, P. & Cattaneo, A. Redox state of single chain Fv fragments targeted to the endoplasmic reticulum, cytosol and mitochondria. *Biotechnology* **13,** 1110–1115 (1995).
10. Desplancq, D. *et al.* Targeting the replisome with transduced monoclonal antibodies triggers lethal DNA replication stress in cancer cells. *Exp. Cell. Res.* **342,** 145–158 (2016).
11. Marschall, A. L., Dubel, S. & Boldicke, T. Specific *in vivo* knockdown of protein function by intrabodies. *MAbs* **7,** 1010–1035 (2015).
12. Sato, Y. *et al.* Genetically encoded system to track histone modification *in vivo*. *Sci. Rep.* **3,** 2436 (2013).
13. Worn, A. & Pluckthun, A. Stability engineering of antibody single-chain Fv fragments. *J. Mol. Biol.* **305,** 989–1010 (2001).
14. Heng, B. C. & Cao, T. Making cell-permeable antibodies (transbody) through fusion of protein transduction domains (PTD) with single chain variable fragment (scFv) antibodies: potential advantages over antibodies expressed within the intracellular environment (intrabody). *Med. Hypotheses* **64,** 1105–1108 (2005).
15. Gu, Z., Biswas, A., Zhao, M. & Tang, Y. Tailoring nanocarriers for intracellular protein delivery. *Chem. Soc. Rev.* **40,** 3638–3655 (2011).
16. Kamaly, N., Xiao, Z., Valencia, P. M., Radovic-Moreno, A. F. & Farokhzad, O. C. Targeted polymeric therapeutic nanoparticles: design, development and clinical translation. *Chem. Soc. Rev.* **41,** 2971–3010 (2012).
17. McNaughton, B. R., Cronican, J. J., Thompson, D. B. & Liu, D. R. Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. *Proc. Natl Acad. Sci. USA* **106,** 6111–6116 (2009).
18. Fuchs, S. M. & Raines, R. T. Arginine grafting to endow cell permeability. *ACS Chem. Biol.* **2,** 167–170 (2007).
19. Cronican, J. J. *et al.* Potent delivery of functional proteins into mammalian cells *in vitro* and *in vivo* using a supercharged protein. *ACS Chem. Biol.* **5,** 747–752 (2010).
20. Zelphati, O. *et al.* Intracellular delivery of proteins with a new lipid-mediated delivery system. *J. Biol. Chem.* **276,** 35103–35110 (2001).
21. Liao, X., Rabideau, A. E. & Pentelute, B. L. Delivery of antibody mimics into mammalian cells via anthrax toxin protective antigen. *ChemBioChem* **15,** 2458–2466 (2014).
22. Khalil, I. A., Kogure, K., Akita, H. & Harashima, H. Uptake pathways and subsequent intracellular trafficking in nonviral gene delivery. *Pharmacol. Rev.* **58,** 32–45 (2006).
23. Krantz, B. A., Trivedi, A. D., Cunningham, K., Christensen, K. A. & Collier, R. J. Acid-induced unfolding of the amino-terminal domains of the lethal and edema factors of anthrax toxin. *J. Mol. Biol.* **344,** 739–756 (2004).
24. Brock, R. The uptake of arginine-rich cell-penetrating peptides: putting the puzzle together. *Bioconjug. Chem.* **25,** 863–868 (2014).
25. Duchardt, F., Fotin-Mleczek, M., Schwarz, H., Fischer, R. & Brock, R. A comprehensive model for the cellular uptake of cationic cell-penetrating peptides. *Traffic* **8,** 848–866 (2007).
26. Herce, H. D., Garcia, A. E. & Cardoso, M. C. Fundamental molecular mechanism for the cellular uptake of guanidinium-rich molecules. *J. Am. Chem. Soc.* **136,** 17459–17467 (2014).
27. Ter-Avetisyan, G. *et al.* Cell entry of arginine-rich peptides is independent of endocytosis. *J. Biol. Chem.* **284,** 3370–3378 (2009).
28. Tunnemann, G. *et al.* Cargo-dependent mode of uptake and bioavailability of TAT-containing proteins and peptides in living cells. *FASEB J.* **20,** 1775–1784 (2006).
29. Lattig-Tunnemann, G. *et al.* Backbone rigidity and static presentation of guanidinium groups increases cellular uptake of arginine-rich cell-penetrating peptides. *Nat. Commun.* **2,** 453 (2011).
30. Nischan, N. *et al.* Covalent attachment of cyclic TAT peptides to GFP results in protein delivery into live cells with immediate bioavailability. *Angew. Chem. Int. Ed.* **54,** 1950–1953 (2015).
31. Helma, J., Cardoso, M. C., Muyldermans, S. & Leonhardt, H. Nanobodies and recombinant binders in cell biology. *J. Cell. Biol.* **209,** 633–644 (2015).
32. Hamers-Casterman, C. *et al.* Naturally occurring antibodies devoid of light chains. *Nature* **363,** 446–448 (1993).
33. Rothbauer, U. *et al.* Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat. Methods* **3,** 887–889 (2006).
34. Schumacher, D. *et al.* Versatile and efficient site-specific protein functionalization by tubulin tyrosine ligase. *Angew. Chem. Int. Ed.* **54,** 13787–123791 (2015).
35. Martin, R. M. *et al.* Principles of protein targeting to the nucleolus. *Nucleus* **6,** 314–325 (2015).
36. Herce, H. D., Rajan, M., Lattig-Tunnemann, G., Fillies, M. & Cardoso, M. C. A novel cell permeable DNA replication and repair marker. *Nucleus* **5,** 590–600 (2014).
37. Kirchhofer, A. *et al.* Modulation of protein properties in living cells using nanobodies. *Nat. Struct. Mol. Biol.* **17,** 133–138 (2010).
38. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **266,** 776–779 (1994).
39. Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: a general method for protein engineering. *Proc. Natl Acad. Sci. USA* **95,** 6705–6710 (1998).
40. Hackenberger, C. P. & Schwarzer, D. Chemoselective ligation and modification strategies for peptides and proteins. *Angew. Chem. Int. Ed.* **47,** 10030–10074 (2008).
41. Muir, T. W. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72,** 249–289 (2003).
42. Chagin, V. O. *et al.* 4D visualization of replication foci in mammalian cells corresponding to individual replicons. *Nat. Commun.* **7,** 11231 (2016).
43. Leonhardt, H. *et al.* Dynamics of DNA replication factories in living cells. *J. Cell Biol.* **149,** 271–280 (2000).
44. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23,** 185–188 (1999).
45. Rett, A. On an until now unknown disease of a congenital metabolic disorder. *Krankenschwester* **19,** 121–122 (1966).
46. Guy, J., Gan, J., Selfridge, J., Cobb, S. & Bird, A. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **315,** 1143–1147 (2007).
47. Thomas, C. E., Ehrhardt, A. & Kay, M. A. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4,** 346–358 (2003).
48. Brero, A. *et al.* Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *J. Cell. Biol.* **169,** 733–743 (2005).

Acknowledgements

We thank K. K. Hassanin and A. Lehmkuhl for excellent technical assistance. We are grateful to J. Hewing and A. Krella for the generation and characterization of MaSat fusion and PCNA expression constructs, respectively, and R. Kühne for providing the pGEX4T1eGFP plasmid. Furthermore, we thank J. Helma for his great support during the nanobody cloning and expression. This work was supported by grants from the Deutsche Forschungsgemeinschaft (SPP1623) to M.C.C. (CA 198/8-2), C.P.R.H. (HA 4468/9-1) and H.L. (LE 721/13-2), the Einstein Foundation Berlin (Leibniz-Humboldt Professorship) and the Boehringer-Ingelheim Foundation (Plus 3 award) to C.P.R.H., the Fonds der Chemischen Industrie to C.P.R.H. and to D.S. (Kekulé fellowship) and A.F.L.S (Chemiefonds fellowship) and the Nanosystems Initiative Munich to H.L.

Author contributions

H.D.H. and D.S. contributed equally. M.C.C., H.L. and C.P.R.H. designed and conceived the project. H.D.H. conceived and performed the cellular uptake experiments, the relocalization-based visualization assay, the uptake of recombinant GFP and Mecp2–GFP, PCNA relocalization and the modified F3H assay and microscale thermophoresis measurements to determine the binding constant of functionalized nanobodies. D.S. designed and optimized the cell-permeable nanobody synthesis, cloned and expressed GBP–intein–CBD fusions, established the refolding protocol, performed the EPL and analysed all the constructs (MS, CD, binding to GFP), synthesized the linear, cyclic and cleavable CPPs, generated double-functionalized nanobodies and performed eGFP expression and purification. A.F.L.S. generated the $GBP1_{1\text{-}117}A_3$–intein–CBD fusion, established a purification strategy, performed EPLs and synthesized cCPPs. A.K.L. purified recombinant proteins and performed some cellular uptake experiments as well as RNA isolation and RNA-binding assays. F.A.M. optimized the EPL conditions and synthesized cCPPs. M.F. generated and characterized the cell lines with the permanent expression of GFP and its fusions. M.-A.K. synthesized Cy5. S.R. performed the cloning and initial testing of the GBP–intein–CBD fusions. E.K. contributed to the matrix-assisted laser desorption ionization measurements. H.L. provided the nanobodies. H.D.H. and D.S. wrote the manuscript supported by M.C.C., C.P.R.H., F.A.M., A.F.L.S. and A.K.L.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to H.D.H., D.S., M.C.C. and C.P.R.H.

Competing financial interests

The authors declare no competing financial interests.

nature chemistry

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCHEM.2811

In the format provided by the authors and unedited.

Cell-permeable nanobodies for targeted immunolabeling and antigen manipulation in living cells

Henry D. Herce[1†*], Dominik Schumacher[2,3†*], Anselm F. L. Schneider[2,3], Anne K. Ludwig[1], Florian A. Mann[2,3], Marion Fillies[4], Marc-André Kasper[2,3], Stefan Reinke[2], Eberhard Krause[2], Heinrich Leonhardt[5], M. Cristina Cardoso[1*], and Christian P. R. Hackenberger[2,3*]

[1]Technische Universität Darmstadt, Department of Biology, Schnittspahnstrasse 10, 64287 Darmstadt (Germany)

[2]Leibniz-Institut für Molekulare Pharmakologie (FMP), Chemical Biology Department, Robert-Rössle-Strasse 10, 13125 Berlin (Germany)

[3]Humboldt Universität zu Berlin, Department of Chemistry, Brook-Taylor-Str. 2, 12489 Berlin (Germany)

[4]Max Delbrück Center for Molecular Medicine (MDC), Robert-Rössle-Strasse 10, 13125 Berlin (Germany); present address: Charité - Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin (Germany)

[5]Ludwig Maximilians Universität München, Department of Biology II, and Center for Integrated Protein Science Munich, Großhadenerstr. 2, 82152 Martinsried (Germany)

[†] These authors contributed equally

[*]e-mail: hdherce@gmail.com; dschumacher@fmp-berlin.de; cardoso@bio.tu-darmstadt.de; hackenbe@fmp-berlin.de

1 SUPPLEMENTARY RESULTS 4

1.1 Cyclic arginine-rich peptides bind RNA driving their recruitment to the nucleolus 4

1.2 Binding epitopes of GBP1 4

1.3 Expression of nanobody-intein fusions to target internal cysteines as ligation sites 5

1.4 EPL of linear R_{10} and TAT peptides to GBP1 (2) 7

1.5 Expression of recombinant nanobodies with linear CPPs 7

1.6 CD-spectra of recombinant and cell-penetrating nanobodies 8

1.7 Formation of intracellular disulfides of nanobodies after functionalization by EPL 9

1.8 Nanobody conjugates retain their original GFP-fluorescence modulation properties ... 10

1.9 Estimation of antigen binding affinity of cell-permeable nanobody detected by microscale thermophoresis 12

1.10 The cell-permeable nanobody recruits GFP to the nucleolus as shown by a live-cell nucleolar marker and by immunostaining in fixed cells 13

1.11 The cell-permeable nanobody recruits GFP tagged PCNA to the nucleolus as shown by a live-cell nucleolar staining and by immunostaining in fixed cells 14

1.12 Cellular uptake quantification of the cell-permeable nanobody 15

1.13 Control experiment with cR_{10} and unconjugated GBP1 16

1.14 GFP tagged PCNA (8) and GFP tagged Mecp2 (9) 17

1.15 Control cell line stably expressing Mecp2-GFP 17

1.16 Cell-permeable nanobody delivery of Mecp2-GFP into a mouse cell line expressing a protein that labels major satellite DNA 18

1.17 Distribution of the cleavable cell-permeable nanobody in human HeLa cells expressing mCherry-PCNA 19

1.18 Intracellular concentration estimation of the cell-permeable nanobody 20

1.19 Tables of synthesized constructs, cells and plasmids 21

2 SUPPLEMENTARY MATERIAL AND METHODS 24

2.1 General Information 24

2.1.1 Analytical UPLC-MS 24
2.1.2 Preparative HPLC 24
2.1.3 High-resolution mass spectra (HRMS) 24
2.1.4 Protein MS 24
2.1.5 Protein concentration 24
2.1.6 Column chromatography 25
2.1.7 NMR 25
2.1.8 Reagents and solvents 25
2.1.9 SPPS 25
2.1.10 SDS-PAGE gel analysis 25
2.1.11 Heparin based separation of CPP-nanobodies for MS analysis. 25

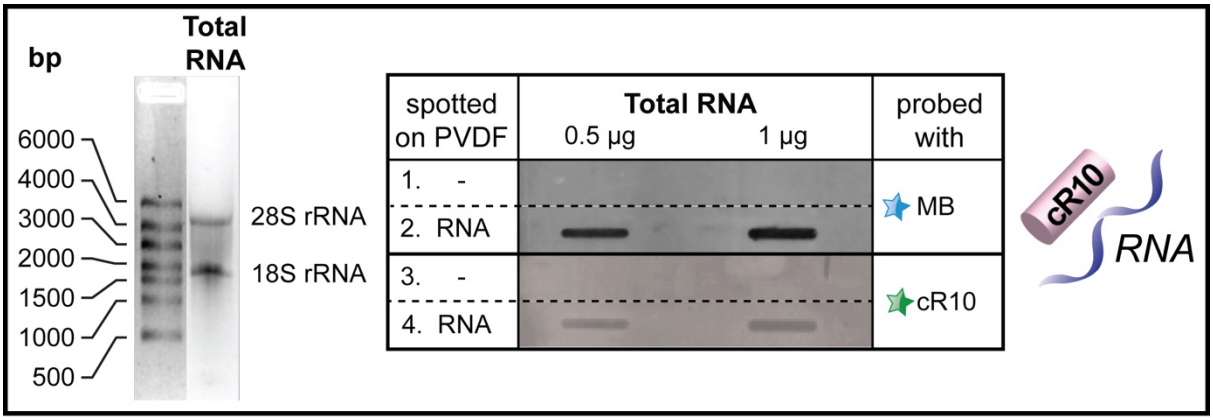
2.2 Experimental section 26
2.2.1 Expression and purification of proteins 26
2.2.1.1 Plasmids 26
2.2.1.2 Expression of recombinant GBP4 (**1**) 27
2.2.1.3 Expression of recombinant GBP1 (**2**) 27
2.2.1.4 eGFP (**7**) 28
2.2.1.5 GFP-tagged PCNA (**8**) 29
2.2.1.6 GFP-tagged Mecp2 (**9**) 29
2.2.1.7 Expression of recombinant nanobodies with linear CPPs 30
2.2.1.8 Expression of truncated and full-length GBP4-intein fusion proteins **S2** and **1A** 30
2.2.1.9 Expression of full-length GBP1-A3-intein fusion protein **2A** 30
2.2.2 RNA purification and *in vitro* RNA binding assay 30
2.2.3 Peptide Synthesis 31
2.2.3.1 Synthesis of 5,6-FAM-cyclic R_{10} (**S1**) 31
2.2.3.2 Synthesis of $GBP4_{97-133}$-5,6-FAM (**S3**) 32
2.2.3.3 Synthesis of cyclic R_{10} (**3**) 33
2.2.3.4 Synthesis of cyclic TAT (**4**) 34
2.2.3.5 Synthesis of linear R_{10} peptide **5** 35
2.2.3.6 Synthesis of linear TAT peptide **6** 36
2.2.3.7 Synthesis of Cy5-peptide **10** 37
2.2.4 Intein-cleavage experiments 37
2.2.5 Synthesis 38
2.2.5.1 1,2,3,3-tetramethyl-3H-indolium iodide (**S5**) 38
2.2.5.2 1-(5-Carboxypentyl)-2,3,3-trimethyl-3H-indolium iodide (**S6**) 38
2.2.5.3 1,3,3-trimethyl-2-((1E,3E)-4-(N-phenylacetamido)buta-1,3-dien-1-yl)-3H-indolium chloride (**S7**) 38
2.2.5.4 Cy5-COOH **S8** 39
2.2.6 Expressed protein ligation 40
2.2.6.1 GBP4-5,6-FAM **S4** 40
2.2.6.2 GBP4-cTAT **1B** 40
2.2.6.3 GBP1-cTAT **2B** 41
2.2.6.4 GBP1-linTAT **2D** 42
2.2.6.5 GBP4-cR_{10} **1C** 43
2.2.6.6 GBP1-cR_{10} **2C** 45
2.2.6.7 GBP1-$linR_{10}$ **2E** 46
2.2.6.8 GBP1-Cy5-S-S-cR_{10} **2G** 47
2.2.7 Mammalian cell culture 48
2.2.8 Development of cell lines stably expressing nuclear GFP (and GFP-tagged Mecp2) 48
2.2.8.1 Expression vectors: 48
2.2.8.2 Host cell lines 49
2.2.8.3 Transfection and selection of stably expressing cell lines 49
2.2.9 GFP fluorescence *in vitro* assay 49
2.2.10 Circular dichroism 49
2.2.11 Ellman's test for the detection of free cysteines 50
2.2.12 Live cell confocal microscopy 50
2.2.13 Nucleotide incorporation after cell-permeable antibody delivery 50
2.2.14 Nucleolar labeling 50
2.2.15 Automatic computerized cellular uptake quantification 51
2.2.16 Intracellular nanobody quantification 52
2.2.17 Protein-protein interactions - Fluorescence three Hybrid method (F3H) 52
2.2.18 Cell-permeable nanobody interaction with its antigen using microscale thermophoresis 53

2.3 Protein sequences 54
2.3.1 GBP4 (**1**) 54
2.3.2 GBP4-int-CBD (**1A**) 54
2.3.3 GBP1 (**2**) 54
2.3.4 GBP1-int-CBD (**2A**) 54

3 REFERENCES 55

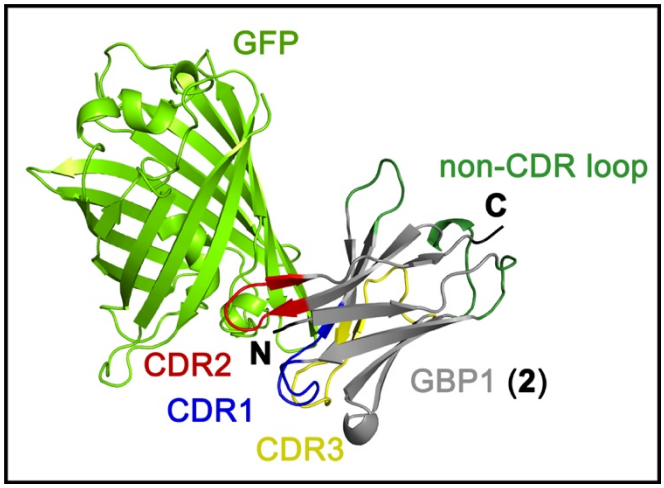
1 Supplementary Results

1.1 Cyclic arginine-rich peptides bind RNA driving their recruitment to the nucleolus



Supplementary Figure 1 | cR_{10} binds RNA, which drives its accumulation to the nucleolus. Total RNA used for the *in vitro* RNA binding assay separated by size showing the absence of genomic DNA contamination, as well as characteristic bands for the 28S and 18S rRNA, which are primarily synthesized and localized in nucleoli. Rows 1 and 2 of the slot blot (representative from two independent experiments) show a methylene blue (MB) stained PVDF membrane in the absence (1) and presence (2) of different amounts of total RNA (0.5 µg and 1 µg) and were used as a loading control. The rest of the blot shows the binding of fluorescently tagged cR_{10} **S1** (see section 2.2.3.1 and Supplementary Fig. 23) (3 and 4) in the absence (3) and presence (4) of total RNA.

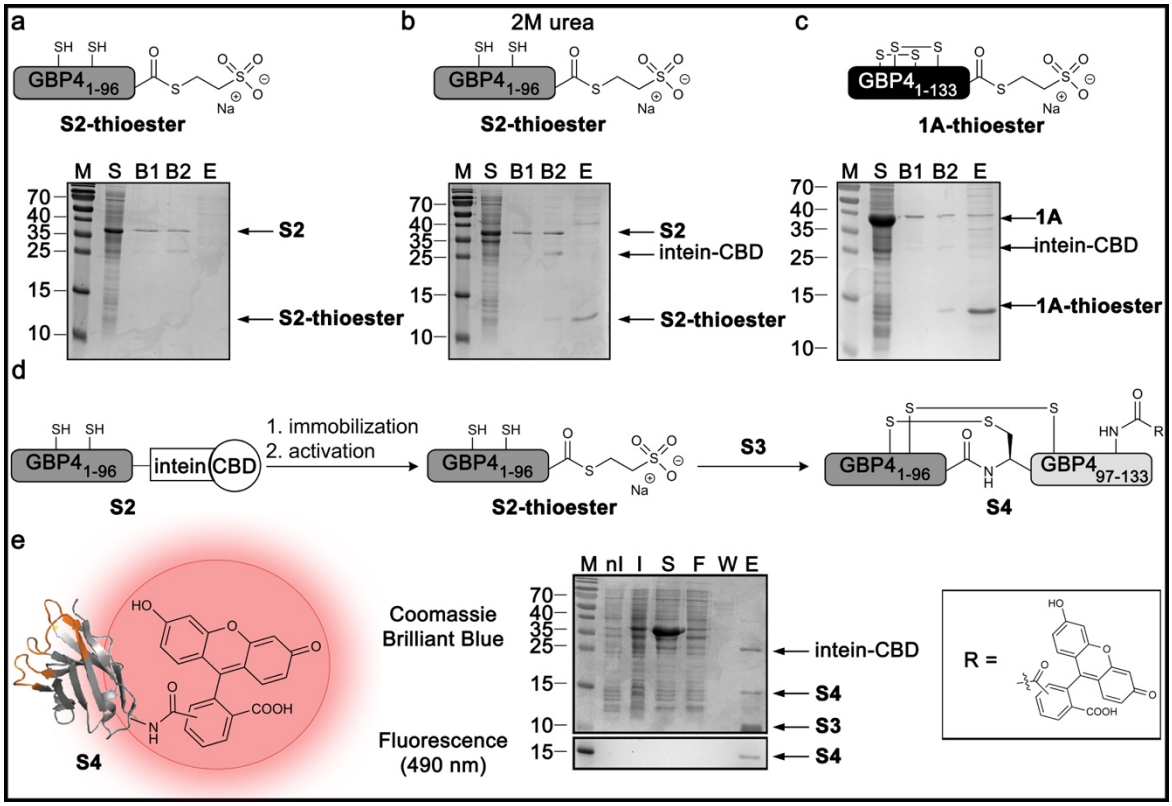
1.2 Binding epitopes of GBP1



Supplementary Figure 2 | Crystal structures of GBP1 (**2**) binding GFP. The variable antigen binding CDRs (complementary determining regions) 1-3 are highlighted in blue, yellow and red. The loops of the conserved nanobody framework that are most distal from the antigen binding interface and thus best for site-specific functionalization are highlighted in dark green. GFP is highlighted in light green (PDB ID: 3K1K)[1].

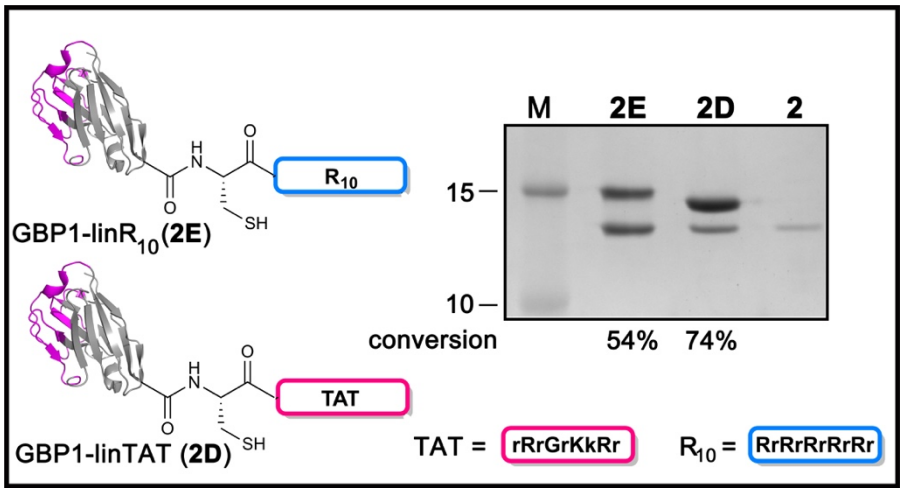
1.3 Expression of nanobody-intein fusions to target internal cysteines as ligation sites

First, we elucidated to use the internal cysteines of GBP4 (**1**) as ligation site for expressed protein ligation (EPL). For this, we genetically fused a truncation of $GBP4_{1-96}$ to the intein-CBD domain (**S2**). However, the expression of **S2** resulted in the formation of inclusion bodies and required the resolubilization by guanidinium-HCl, urea and stepwise dialysis based refolding (see section 2.2.1.8). The soluble protein (4mL) was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer. To elucidate the efficiency of thioester formation we induced cleavage of the immobilized fusion protein by incubating the beads for 18 h at RT either with cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM Sodium 2-mercaptoethanesulfonate (MESNA), pH 8.5) or cleavage buffer containing 2M urea (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, 2M urea, pH 8.5) (see Supplementary Fig. 3a and b). We could detect activated thioester (**S2-thioester**) in the presence of 2M urea (see Supplementary Fig. 3b). However, initial test ligations to a fluorescently labeled Cys-$GBP4_{97-133}$-peptide (**S3**, synthesized as described in section 2.2.3.2, Supplementary Scheme 1 and Fig. 24) to form fluorescently labeled GBP4 (**S4**) were highly inefficient (see Supplementary Fig. 3d and e). Since low amounts of functionalized product generated by this strategy would hamper further cell experiments, we employed the use of the full length $GBP4_{1-137}$-intein-CBD fusion (**1A**). Once again the expression resulted in insoluble protein (see section 2.2.1.8 for refolding details). However, this time we could generate high amounts of soluble thioester **1A-thioester** upon intein cleavage (see Supplementary Fig. 3c).



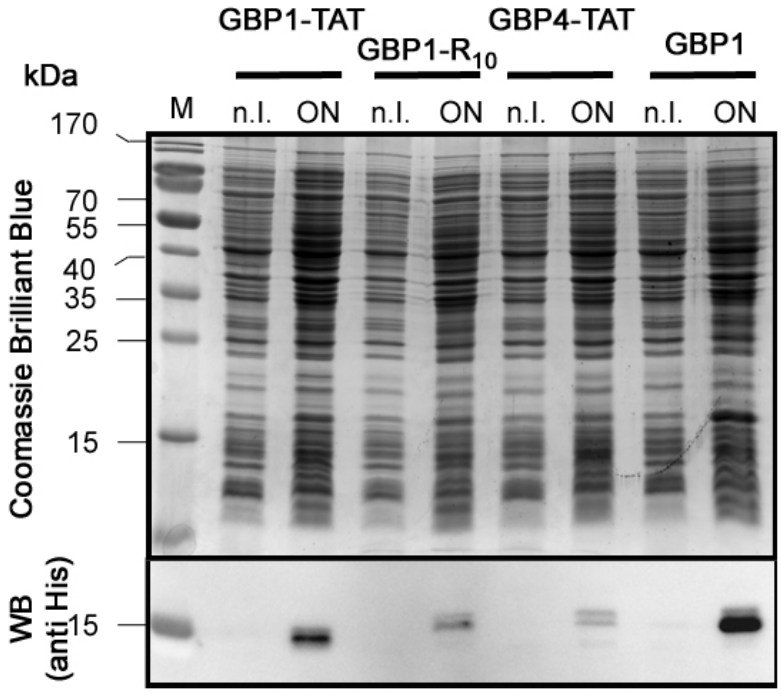
Supplementary Figure 3 | Optimization of the EPL strategy for the synthesis of cell-permeable nanobodies. To test and compare the efficiencies of the strategies for EPL, we performed intein cleavage experiments of **S2** and **1A**. **a**, Cleavage of the $GBP4_{1-96}$-intein-CBD (**S2**) shows low amounts of soluble **S2-thioester** formation. **b**, Cleavage of the $GBP4_{1-96}$-intein-CBD (**S2**) shows low amounts of soluble **S2-thioester** formation even at slightly denaturing conditions. **c**, In contrast to this, soluble thioester of the full length GBP4 **1A-thioester** is generated in high amounts. The same amount of chitin-beads and eluate was boiled with SDS-sample buffer and applied to SDS-PAGE. Excess of soluble protein was added to the chitin-beads to ensure complete loading of the beads (M: marker in kDa; S: soluble protein; B1: GBP-intein-CBD (**S2** or **1A**) loaded on chitin-beads; B2: chitin-beads after 2h incubation with MESNA; E: Elution after 2 h incubation with MESNA). **d**, Test ligations of a carboxyfluorescein labeled Cys-$GBP4_{97-133}$ peptide **S3** (see section 2.2.3.2 for synthesis details) to $GBP4_{1-96}$ expressed as an intein-CBD-fusion protein **S2**. **e**, SDS-PAGE analysis revealed low amounts of fluorescently labeled semisynthetic $GBP4_{1-133}$ **S4** (M: marker in kDa; nI: uninduced *E. coli* cells; I: whole cells induced with IPTG, 18 h at 18 °C; S: Partly refolded $GBP4_{1-97}$-intein-CBD fusion 1B in 2 M urea; F: flowthrough of chitin-column; W: wash fraction; E: NCL to 5,6-FAM-$GBP4_{97}$-133 **S3**).

1.4 EPL of linear R_{10} and TAT peptides to GBP1 (**2**)



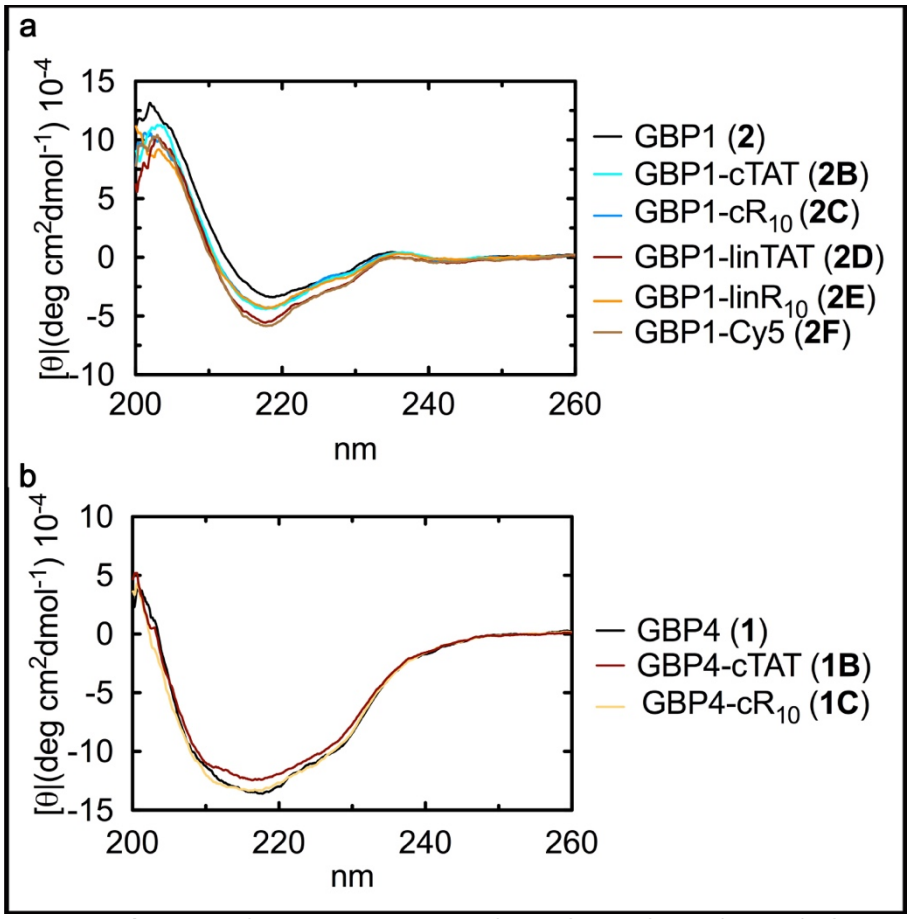
Supplementary Figure 4 | EPL of GBP1-intein-CBD (**2A**) with Cys-linear R_{10} **5** and TAT **6** peptides. Gel shifts in SDS-PAGE show efficient conjugation (conversions of the NCL reactions are given in % and were determined by SDS-PAGE analysis based on the ratio of conjugated and unconjugated nanobody. M: marker in kDa). For synthetic details of peptide **5** see Supplementary section 2.2.3.5, for peptide **6** see Supplementary section 2.2.3.6. For EPL reactions and MS analysis of nanobody conjugate **2E** see Supplementary section 2.2.6.7 and Supplementary Fig. 38-39**,** for conjugate **2D** see Supplementary section 2.2.6.4 and Supplementary Fig. 33-34.

1.5 Expression of recombinant nanobodies with linear CPPs



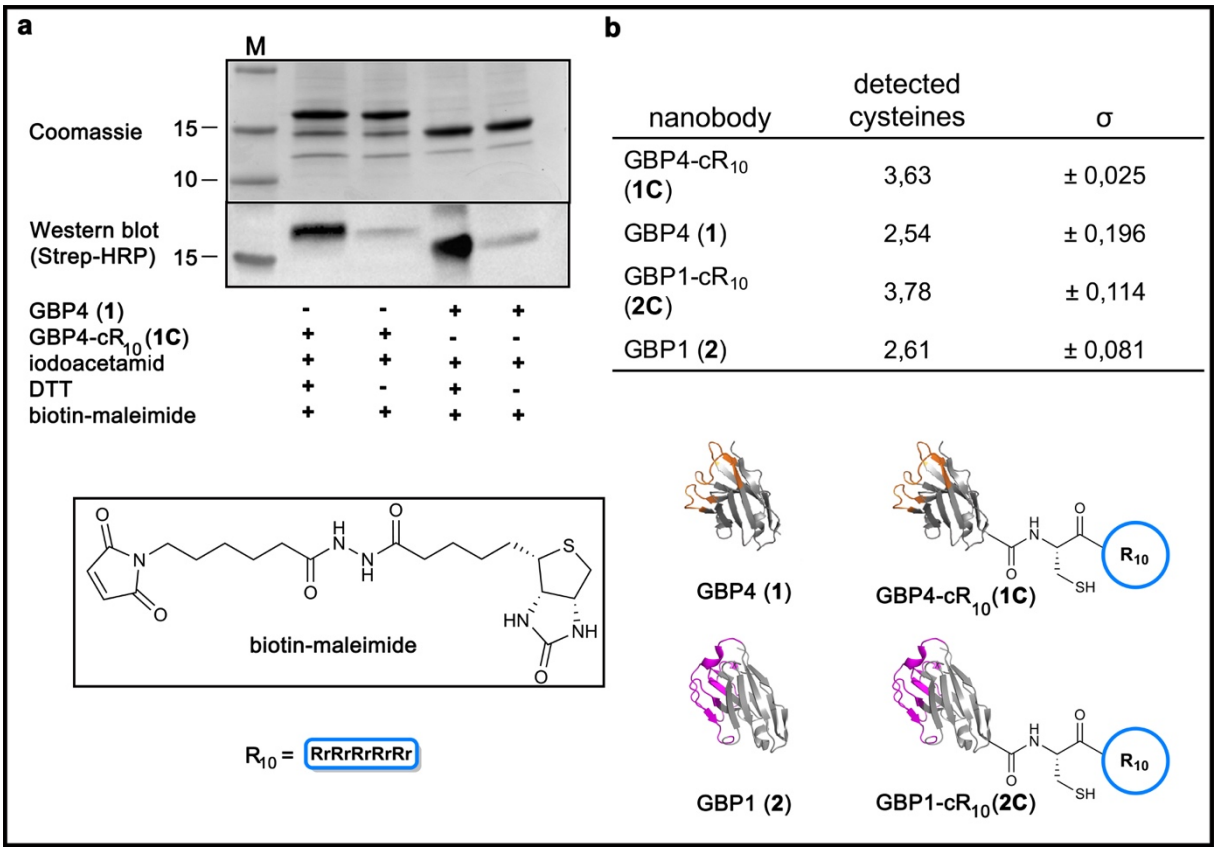
Supplementary Figure 5 | Expression of linear CPP nanobody fusions. Nanobody-CPP fusion expression constructs were generated by standard molecular biology techniques resulting in N-terminally 6xHis tagged GBP1-TAT, GBP1-R_{10}, GBP4-TAT. The linear CPP nanobody fusions were expressed using the same conditions as wild type nanobodies (GBP1 exemplarily shown). As shown by SDS-PAGE of whole cell samples and anti-His Western Blot, yields for linear CPP fusions are significantly lower compared to GBP1 and purification with Ni-NTA resin was not successful (M: marker in kDa; n.I.: uninduced E. coli cells; ON: whole cells induced with IPTG, 18 h at 18 °C).

1.6 CD-spectra of recombinant and cell-penetrating nanobodies



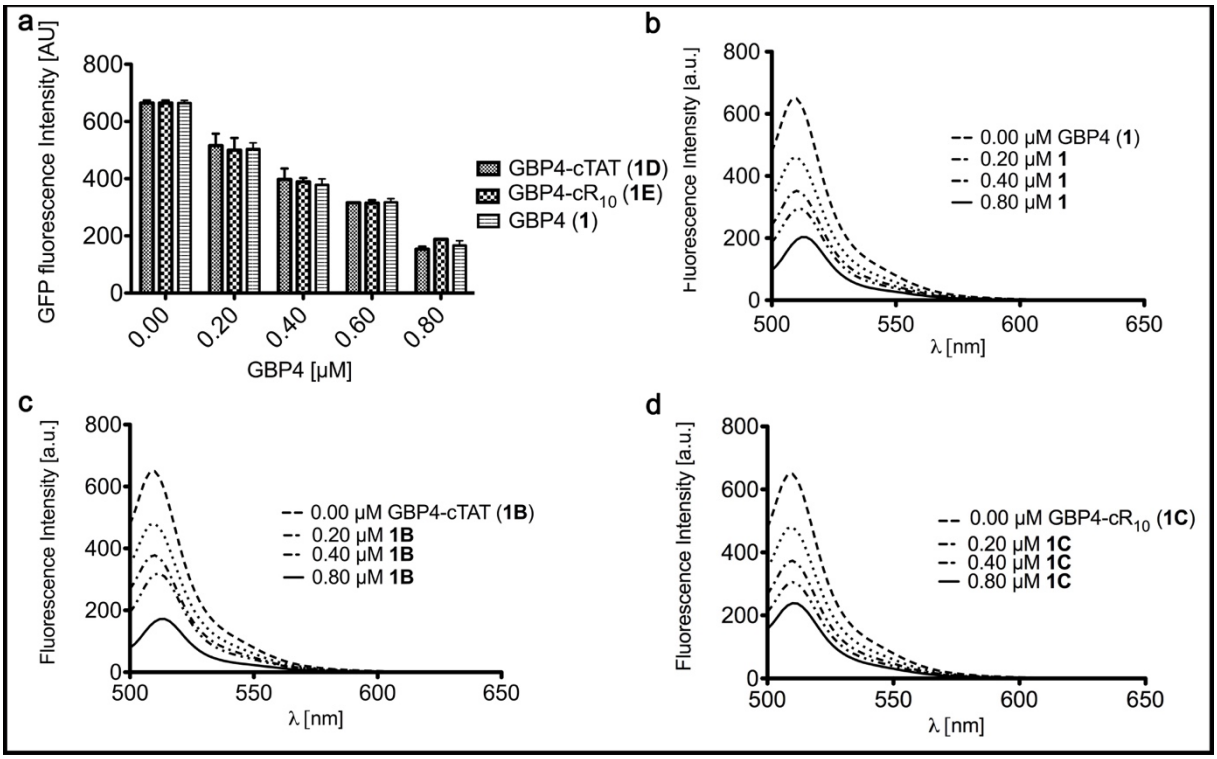
Supplementary Figure 6 | Far-UV CD-spectra obtained for **a,** GBP1-CPPs (**2B**-**2E**), GBP1-Cy5 (**2F**) and recombinant GBP1 (**2**) and **b,** GBP4-CPPs (**1B** and **1C**) and recombinant GBP4. Similar spectra suggest similar secondary structures of functionalized nanobodies and the corresponding recombinant wild type.

1.7 Formation of intracellular disulfides of nanobodies after functionalization by EPL

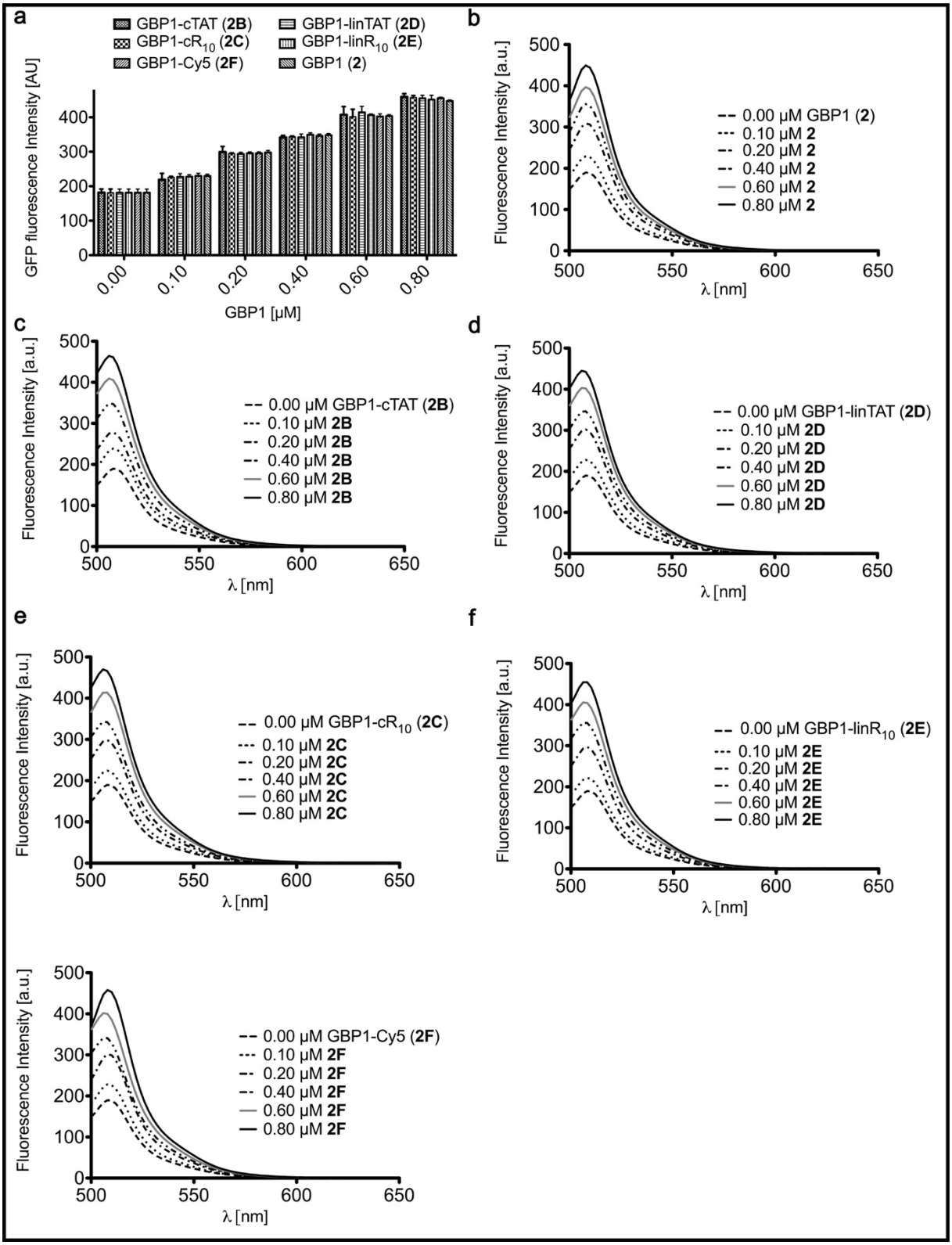


Supplementary Figure 7 | The presence of disulfides in nanobodies synthesized by EPL was elucidated. **a**, For this, GBP4-cR$_{10}$ **1C** and recombinant GBP4 (**1**) were treated with iodoacetamide for 1h at 30 °C to alkylate all free cysteines present in the molecule. Afterwards, excess iodoacetamide was removed using desalting columns (Zeba^{TM} Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). Half of the sample was incubated in 200 mM DTT at 30 °C for 30 minutes. Excess DTT was removed using desalting columns (Zeba^{TM} Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). DTT treated and not reduced protein was incubated with 10 eq. biotin-maleimide (Sigma Aldrich) and samples analyzed by SDS-PAGE and Western Blot using a Strep-HRP conjugate. Selective biotin labeling of DTT reduced samples indicates that disulfides are formed in recombinantly expressed GBP4 (**1**) and EPL synthesized GBP4-cR$_{10}$ (**1C**). **b**, The amount of free cysteine was determined using Ellman's reagent (see Supplementary Section 2.2.11 for details). Values for EPL derived nanobodies are increased by one cysteine compared to recombinantly expressed nanobodies. This additional cysteine corresponds to the cysteine incorporated by EPL to the nanobodies C-terminus. The mean value and standard deviation (SD) of three replicate reactions is shown.

1.8 Nanobody conjugates retain their original GFP-fluorescence modulation properties

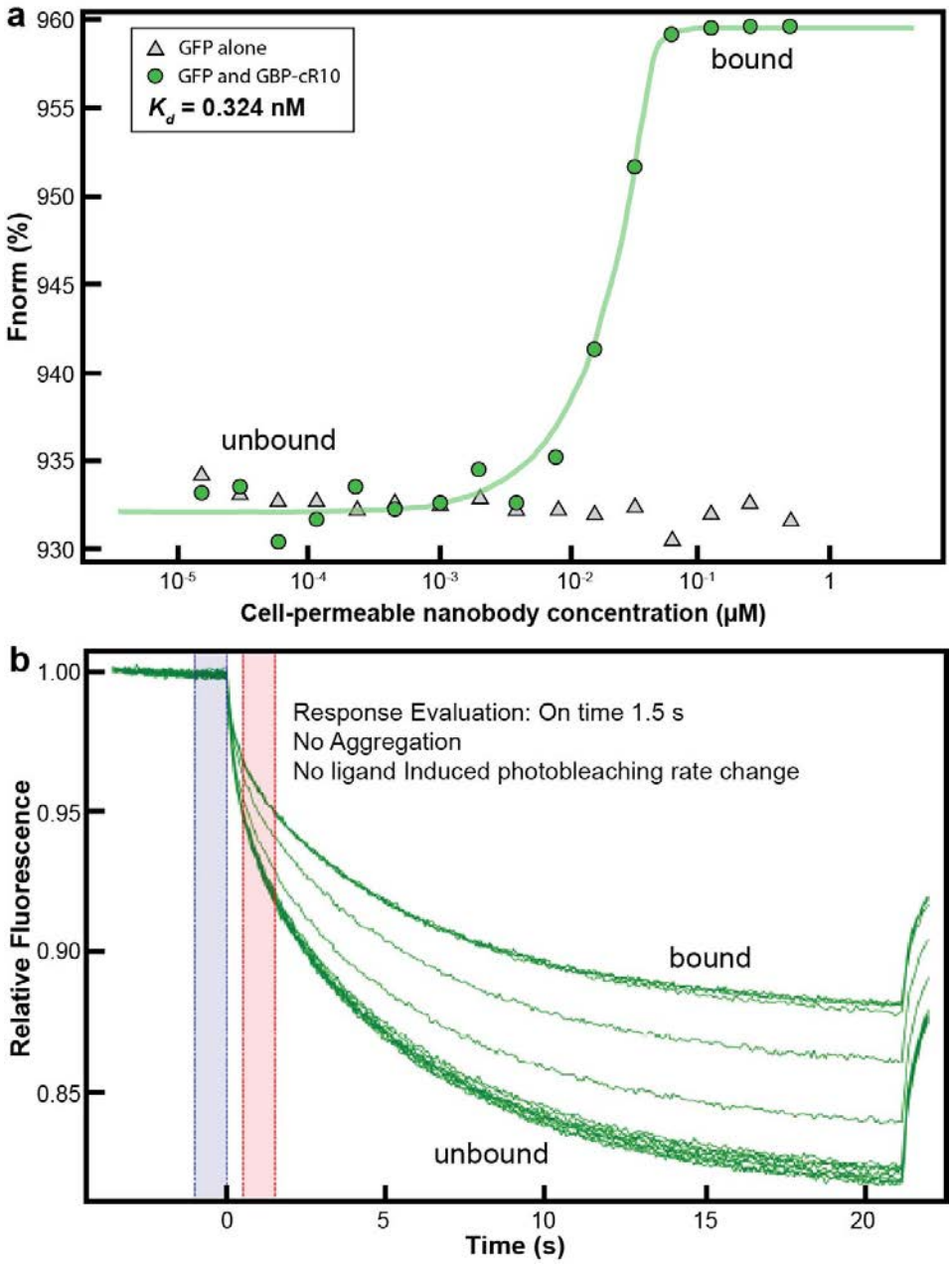


Supplementary Figure 8 | GBP4 conjugates retain their GFP binding properties. The presented binding assay was performed in accordance to a known protocol in literature and is based on the concentration dependent reduction of GFP-fluorescence upon binding of GBP4[1, 2]. 0.5 µM GFP in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) was incubated with varying concentrations [0.00 – 0.80 µM) of GBP4 and GBP4-cCPP respectively. **a,** Similar decrease indicates comparable GFP-binding capacities of functionalized and recombinant nanobodies. The mean value and standard deviation (SD) of three replicate reactions is shown. **b-d,** GFP emission spectra upon titration of **1** (**b**), **1D** (**c**) and **1E** (**d**) to a solution of 0.5 µM GFP.



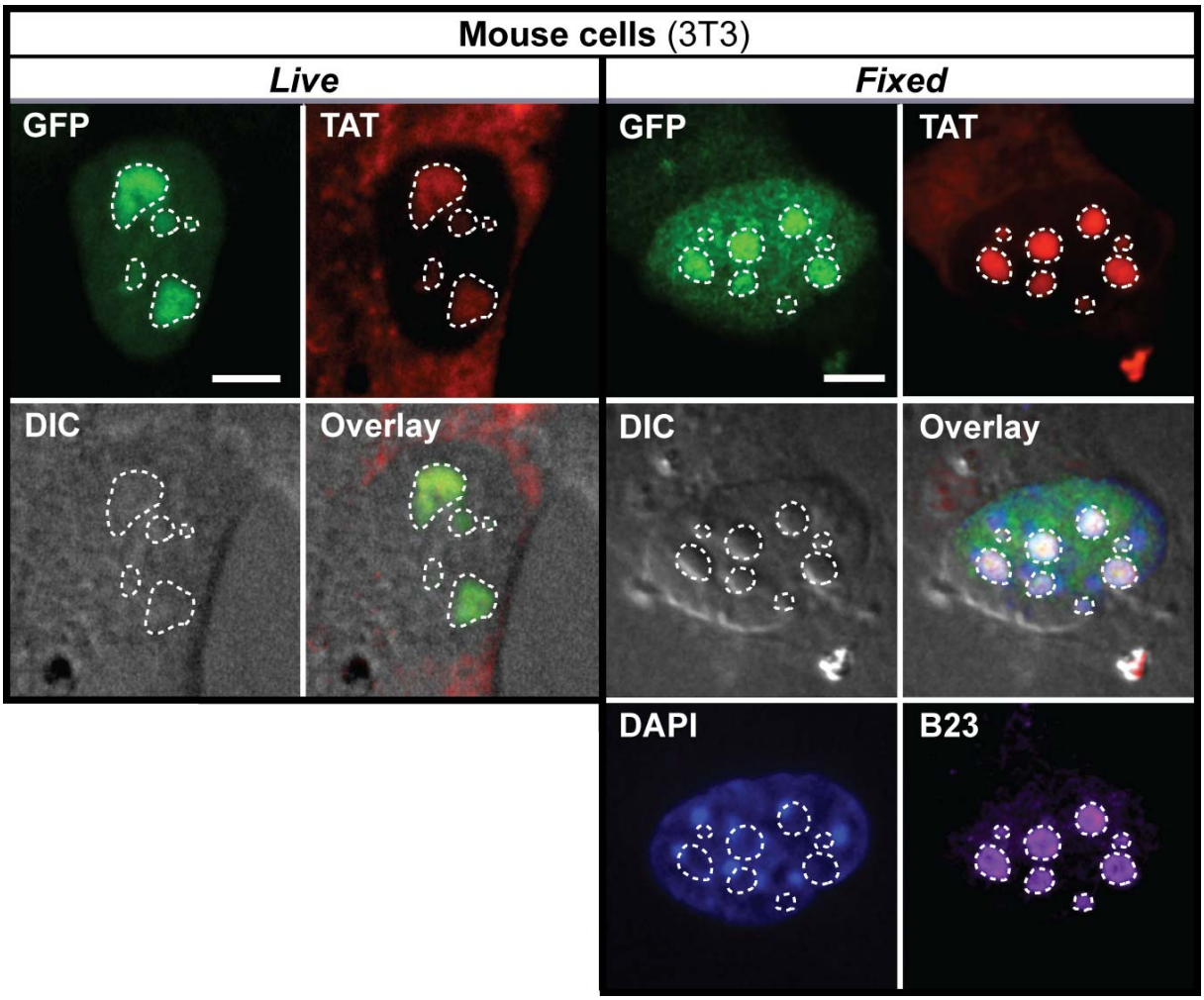
Supplementary Figure 9 | GBP1 conjugates retain their GFP binding properties. The presented binding assay was performed according to a known protocol in literature and is based on the concentration dependend increase of GFP-fluorescence upon binding of GBP1[1, 2]. 0.5 µM GFP in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) was incubated with varying concentrations [0.00 – 0.80 µM) of GBP1 (**2**), GBP1-Cy5 (**2F**) and GBP1-CPP (**2B**-**2E**) respectively. **a**, Similar increase of fluorescence indicates comparable GFP-binding capacities of functionalized and recombinant nanobodies. The mean value and standard deviation (SD) of three replicate reactions is shown. **b-g**, GFP emission spectra upon titration of **2** (**b**), **2B** (**c**) **2D** (**d**), **2C** (**e**), **2E** (**f**) and **2F** (**g**) to a solution of 0.5 µM GFP.

1.9 Estimation of antigen binding affinity of cell-permeable nanobody detected by microscale thermophoresis



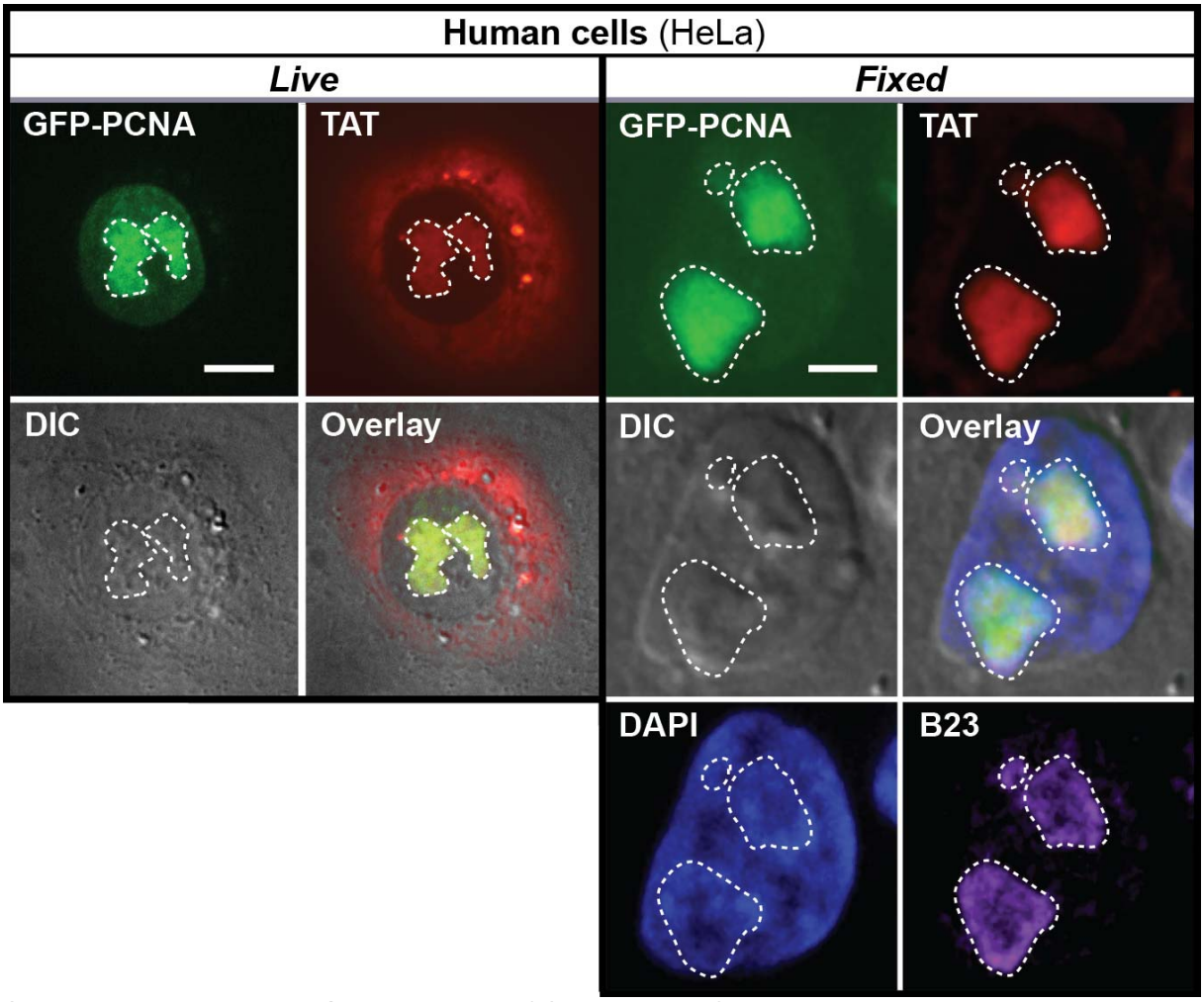
Supplementary Figure 10 | The cell-permeable nanobody GBP1-cR_{10} (**2C**) binds with subnanomolar affinity to its antigen (GFP). To determine the affinity of the binding reaction, the cell-permeable nanobody GBP1-cR_{10} (**2C**) was serially titrated while the fluorescent antigen (GFP) was kept at a constant concentration using a protocol previously reported for GBP1[3]. **a,** Dose response curve of GFP alone (control) and in the presence of GBP1-cR_{10} (**2C**). **b,** Microscale thermophoresis traces for GFP in the presence of GFP-cR_{10} (**2C**). The obtained K_d value is in the same range to that published for GBP1 [1, 3].

1.10 The cell-permeable nanobody recruits GFP to the nucleolus as shown by a live-cell nucleolar marker and by immunostaining in fixed cells



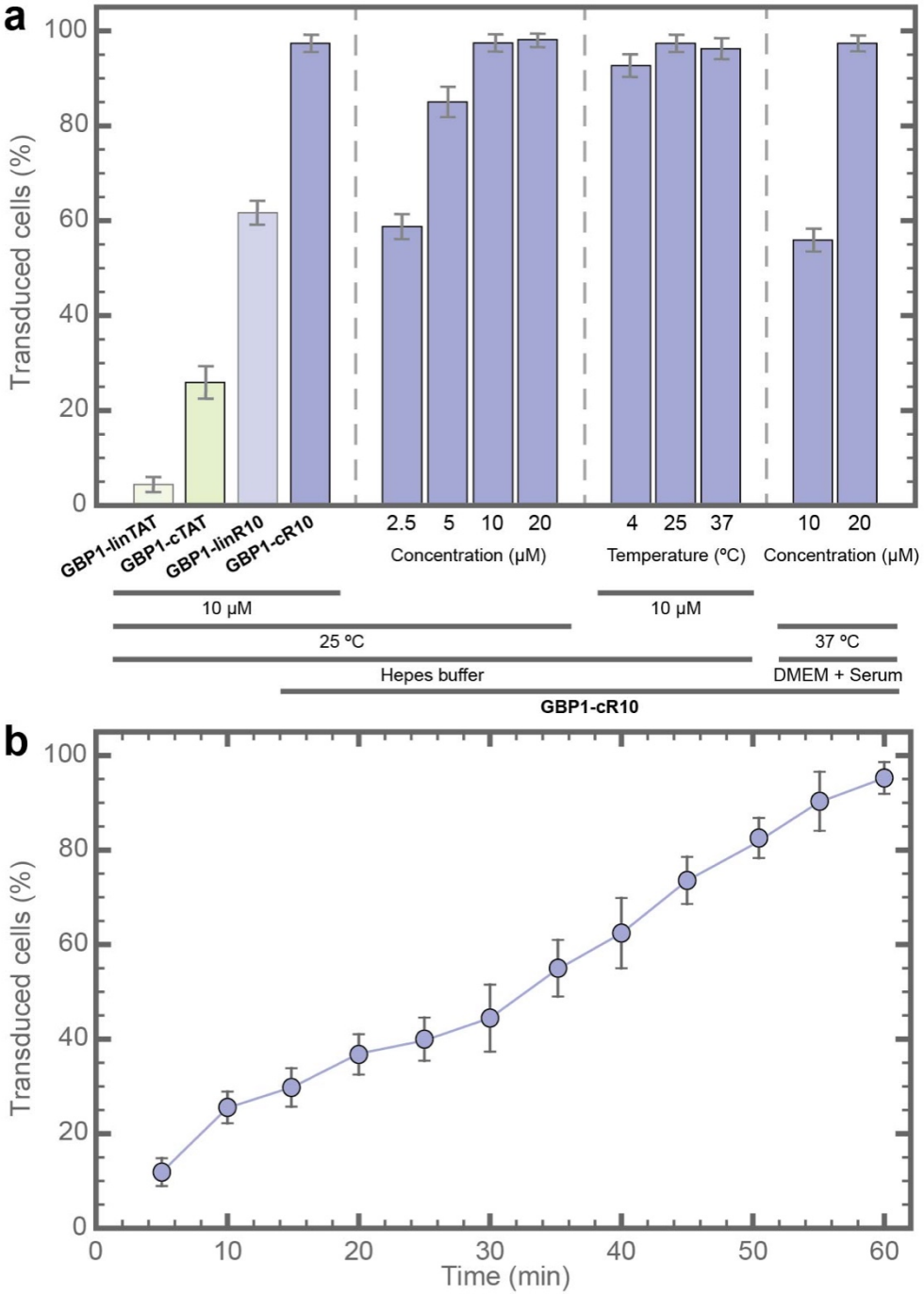
Supplementary Figure 11 | Redistribution of GFP within the nucleolus mediated by the cell-permeable nanobody GBP1-cR$_{10}$ (**2C**) in mouse cells. In live cells (left) we used a linear TAT peptide labeled with TAMRA (as a nucleolar marker for living cells[4]). After incubating the cells with 10 μM of GBP1-cR$_{10}$, GFP colocalizes with the TAT peptide showing that the GFP has been recruited to the nucleolus. On the right panel, the cells are fixed and co-stained with DAPI and an antibody for B23 (a protein found primarily at the nucleolus associated with nucleolar ribonucleoprotein structures and nucleic acids[5]). Scale bars: 5 μm.

1.11 The cell-permeable nanobody recruits GFP tagged PCNA to the nucleolus as shown by a live-cell nucleolar staining and by immunostaining in fixed cells



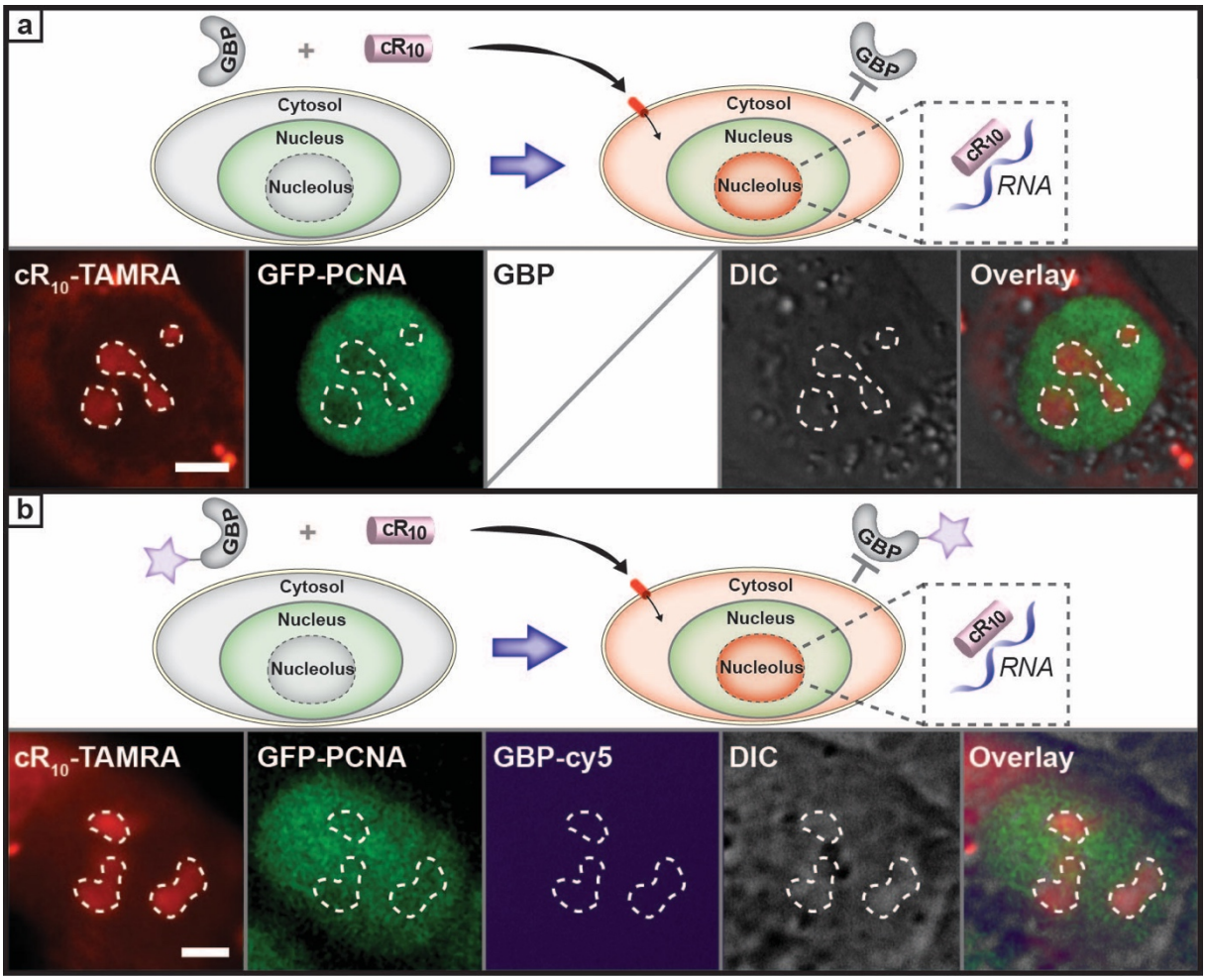
Supplementary Figure 12 | Redistribution of GFP tagged PCNA within the nucleolus mediated by the cell-permeable nanobody GBP1-cR$_{10}$ (2C) in human cells. In live cells (left) we used a linear TAT peptide labeled with TAMRA (proposed to be a nucleolar marker for living cells[4]). After incubating the cells with 10 µM of GBP1-cR$_{10}$, GFP tagged PCNA colocalizes with the TAT peptide showing that the GFP tagged PCNA has been recruited to the nucleolus. On the right panel the cells are fixed and co-stained with DAPI and an antibody for B23 (a protein found primarily at the nucleolus associated with nucleolar ribonucleoprotein structures and nucleic acids[5]). Scale bars: 5 µm.

1.12 Cellular uptake quantification of the cell-permeable nanobody



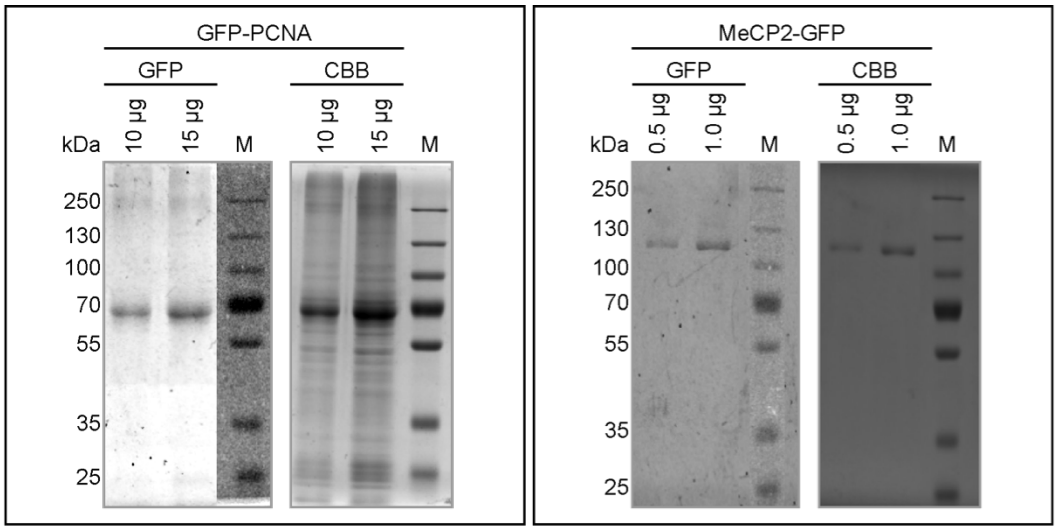
Supplementary Figure 13 | Cellular uptake quantification of the cell-permeable nanobodies in living cells. **a,** Cellular uptake evaluation of GBP1-linTAT (**2D**), GBP1-cTAT (**2B**), GBP1-linR_{10} (**2E**) and GBP1-cR_{10} (**2C**), (1 h incubation,10 μM, 25 ºC in 5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). The cyclic arginine-rich peptides are consistently more efficient than the linear versions and the R_{10} peptides are more efficient than the TAT peptides. Next, four concentrations were tested, 2.5, 5, 10, and 20 μM of the GBP1-cR_{10} (1 h incubation, 25 ºC in Hepes buffer). We observed that at 10 μM most of the cells show uptake of the cell-permeable nanobody. To show that the uptake is energy independent, the uptake of GBP1-cR_{10} at 4 ºC is compared with the uptake at 25 and 37 ºC, (1 h incubation, 10 μM in Hepes buffer). Finally, the cellular uptake of GBP1-cR_{10} at 10 and 20 μM is studied in regular cell culture media (DMEM) plus serum. In this case the concentration had to be increased to reach similar effects as when incubated with Hepes buffer. **b,** Time lapse uptake of GBP1-cR_{10} in the presence of cell culture media (DMEM) plus serum, 20 μM and 37 ºC (see Movie 1).

1.13 Control experiment with cR_{10} and unconjugated GBP1



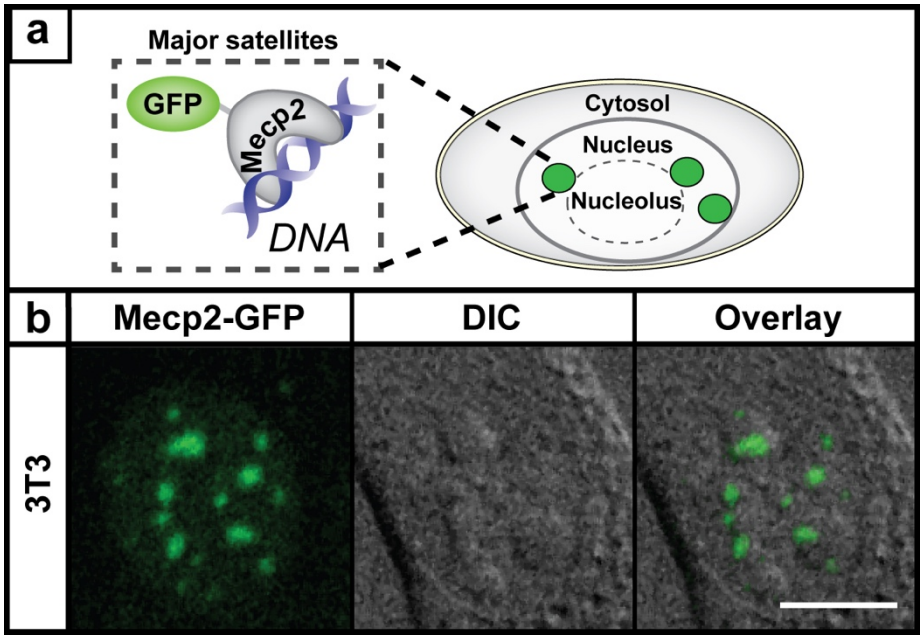
Supplementary Figure 14. | Covalent attachment of the nanobody to the cyclic cell-penetrating peptide is required for cellular permeation followed by manipulation or labeling of the antigen. **a-b,** Cells where incubated for 1 h in the presence of 10 μM of labeled cR_{10} (**S1**) and 10 μM of the nanobody (GBP) either unlabeled (**a**, **2**) or labeled (**b**, **2F** see Supplementary Section 2.2.6.8). In both cases we did not detect permeation of the nanobody, manipulation and/or labeling of the antigen. Scale bars 5 μM.

1.14 GFP tagged PCNA (**8**) and GFP tagged Mecp2 (**9**)



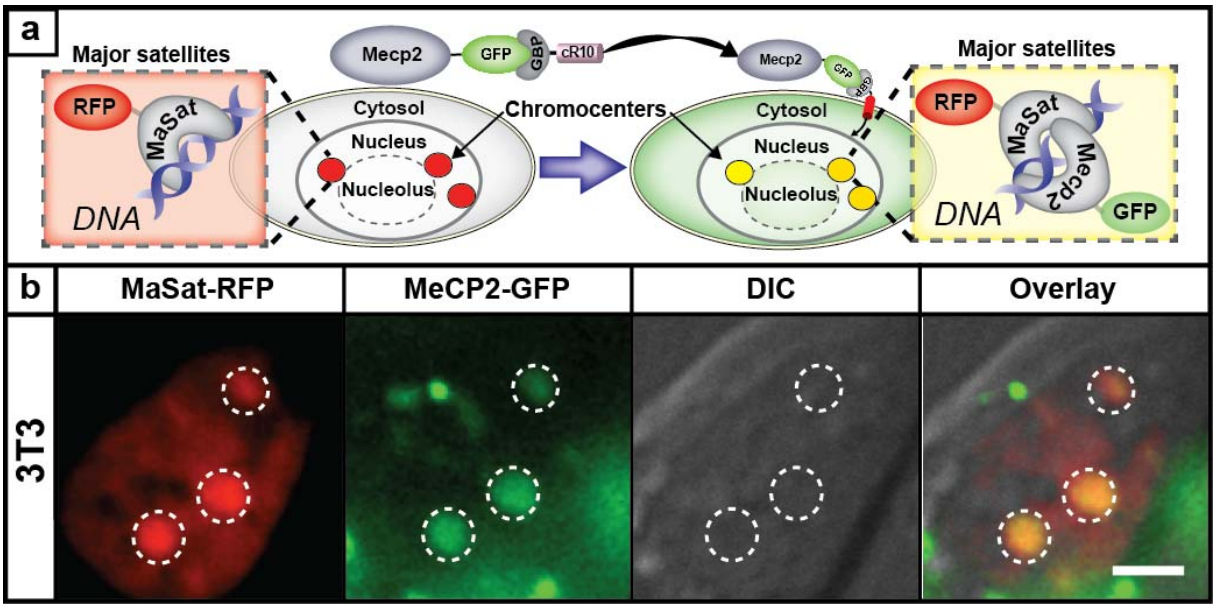
Supplementary Figure 15 | Separation of purified, GFP-tagged PCNA (**8**) and Mecp2 (**9**) via electrophoresis through a denaturing polyacrylamide gel visualized using a fluorescent imager and by Coomassie Brilliant Blue (CBB) staining. Purified GFP tagged PCNA has the expected size of 63 kDa. Purified Mecp2-GFP runs at the height of 110 kDa (predicted size of 83 kDa). This apparent discrepancy is a result of the very high arginine/lysine content of Mecp2, which makes it run anomalously in SDS-PAGE. M: PageRuler™ Plus Prestained Protein Ladder (Thermo Fisher Scientific, Dreieich, Germany).

1.15 Control cell line stably expressing Mecp2-GFP



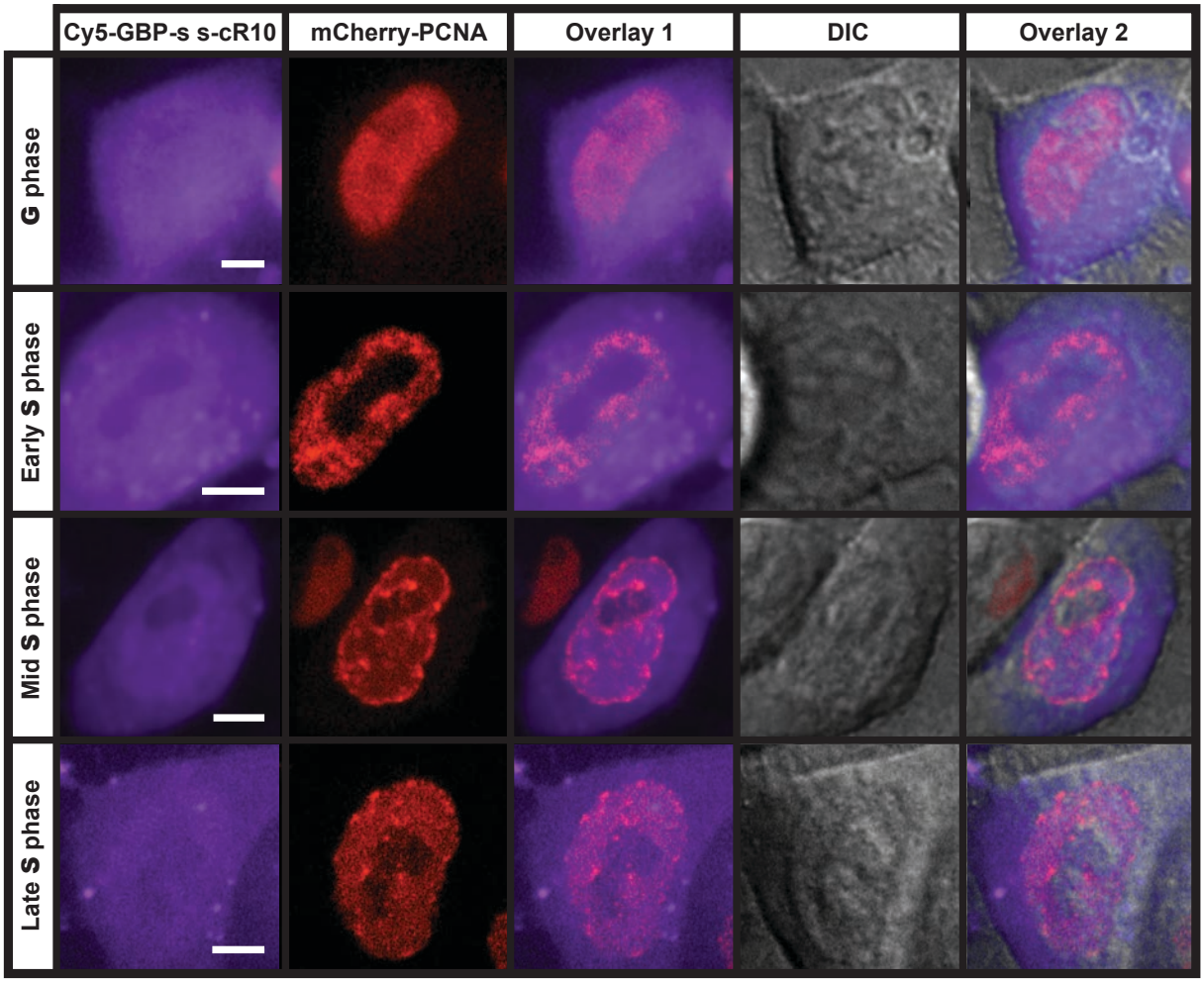
Supplementary Figure 16 | Mecp2-GFP binds and labels major satellite repeats. **a**, Scheme depicting the recruitment of Mecp2-GFP at major satellite repeats. **b**, Fluorescent image showing that in the absence of GBP1-cR_{10} (**2C**), Mecp2-GFP binds exclusively to major satellite repeats and does not accumulate at the nucleolus visualized by differential interference contrast (DIC) microscopy. Scale bar: 10 µm.

1.16 Cell-permeable nanobody delivery of Mecp2-GFP into a mouse cell line expressing a protein that labels major satellite DNA



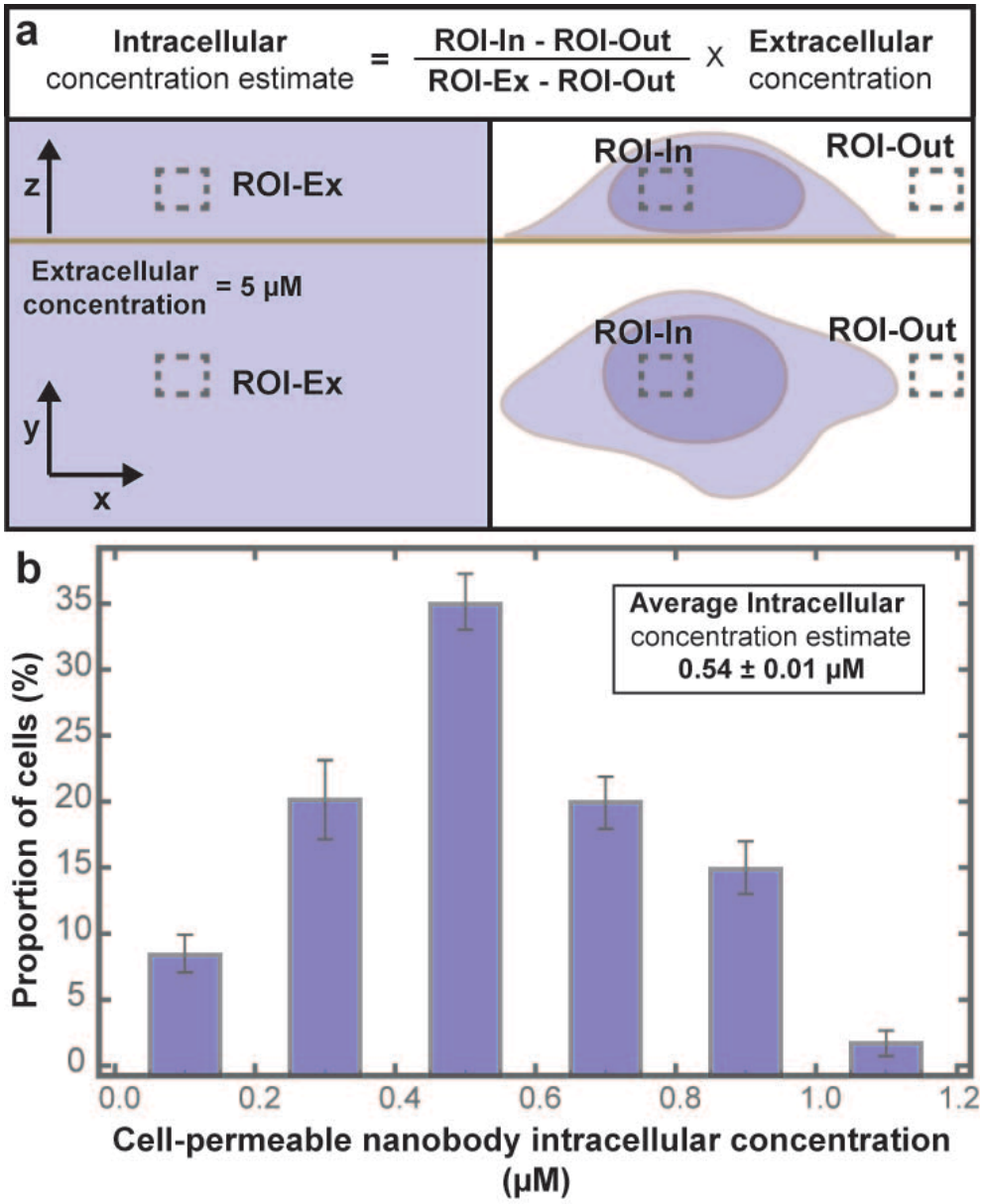
Supplementary Figure 17 | Mecp2-GFP colocalizes with the mouse major satellite protein fused to RFP (MaSat-RFP), showing that it binds and labels major satellite repeats. **a**, Scheme depicting the recruitment of MaSat-RFP and Mecp2-GFP at major satellite repeats. **b**, Fluorescent image showing that Mecp2-GFP, delivered by GBP1-cR_{10} (2C), colocalizes with MaSat-RFP at major satellite DNA repeats. Scale bar: 5 μm.

1.17 Distribution of the cleavable cell-permeable nanobody in human HeLa cells expressing mCherry-PCNA



Supplementary Figure 18 | Distribution of the cleavable cell-permeable nanobody GBP1-ss-cR$_{10}$ (**2G**) in HeLa Kyoto cells stably expressing mCherry-PCNA. Scale bar: 5 µm.

1.18 Intracellular concentration estimation of the cell-permeable nanobody



Supplementary Figure 19 | Intracellular concentration quantification of the fluorescently labeled cell-permeable nanobody GBP1-ss-cR$_{10}$ (**2G**). **a**, Scheme depicting the method employed to estimate the intracellular concentration of the nanobody. We quantified the average fluorescence intensity of the cell-permeable nanobody (5 µM) in an empty well (ROI-Ex). Then the cells where incubated in the presence of the cell-permeable nanobody (5 µM) for 1 h, washed and imaged. The background (ROI-Out) and the average intracellular fluorescence intensity in individual cells (ROI-In) were determined. These values were plugged into the equation shown to estimate the intracellular concentration. **b**, Histogram of the percentage of cells showing intracellular concentrations in interval ranges of 0.2 µM, between 0 and 1.2 µM. The inset shows the average intracellular concentration of all the cells and the standard error.

1.19 Tables of synthesized constructs, cells and plasmids

Supplementary Table 1: Constructs used synthesized and/or purified in this study

Number	Name
1	GBP4
1A	GBP4-intein-CBD fusion
1B	GBP4-cTAT
1C	GBP4-cR_{10}
2	GBP1
2A	GBP1-intein-CBD fusion
2B	GBP1-cTAT
2C	GBP1-cR_{10}
2D	GBP1-linTAT
2E	GBP1-linR_{10}
2F	GBP1-Cy5
2G	GBP1-Cy5-S-S-cR_{10}
3	Cys-cR_{10} peptide
4	Cys-cTAT peptide
5	Cys-linR_{10} peptide
6	Cys-linTAT peptide
7	H_6-eGFP
8	His-NLS-GFP-PCNA
9	Mecp2-GFP
10	Cy5-peptide
S1	*5,6*-FAM-cR_{10}
S2	$GBP4_{1-96}$-intein-CBD fusion
S3	$GBP4_{97-133}$-*5,6*-FAM peptide
S4	GBP4-*5,6*-FAM
S5	1,2,3,3-tetramethyl-3H-indolium iodide
S6	1-(5-Carboxypentyl)-2,3,3-trimethyl-3H-indolium iodide
S7	1,3,3-trimethyl-2-((1E,3E)-4-(N-phenylacetamido)buta-1,3-dien-1-yl)-3H-indolium chloride
S8	Cy5-COOH
S9	GBP4-5,6-FAM

Supplementary Table 2: Cells used in this study

Name	Species	Celly type	Reference	Used in Figure
NIH Flp-In 3T3	Mus musculus	fibroblast cells	Invitrogen (R761-07)	5b, 5f, S17
NIH Flp-In 3T3 NGFP	Mus musculus	fibroblast cells	this study	1, 3 and S11
NIH Flp-In 3T3 Mecp2-GFP	Mus musculus	fibroblast cells	this study	S16
HeLa Kyoto	Homo sapiens	cervical adenocarcinoma cells	Landry et al., 2013[6]	5b
HeLa Kyoto GFP-PCNA	Homo sapiens	cervical adenocarcinoma cells	Chagin et al., 2016[7]	4b-c, 6, S12, S13, S19 and S42
HeLa Kyoto Cherry-PCNA	Homo sapiens	cervical adenocarcinoma cells	Chagin et al., 2016[7]	4e-f, 5d and S18
U2OS 2-6-3	Homo sapiens	osteosarcoma cells	Janicki et al., 2004[8]	4g-h
Sf9	Spodoptera frugiperda	-	Gibco (11496-015)	5f, S15 and S17
BL21(DE3)	Escherichia coli	-	Studier and Moffatt, 1986[9]	5d and S15
JM109	Escherichia coli	-	Kirchhofer et al., 2010[1]	S20
T7 Express	Escherichia coli	-	New England Biolabs	S2 and S4

Supplementary Table 3: Plasmids used in this study

Name	No.	Protein	Promoter	Resistance*	Expression in	Reference	Used in Figure
pRHGPCNA	1068	His-NLS-GFP-PCNA (**8**)	T7	Amp	BL21(D3) bacteria	this study	5d and S15
pFBMecp2G	1571	Mecp2-GFP (**9**)	Polyhedrin	Amp/Gent	Sf9 insect cells	Jost et al., 2011[10]	5f, S15 and S17
pEGFP-N1	713	eGFP	CMV	Kan/Neo	Mammalian cells	Clontech (GenBank Accession #U55762)	4e
pENeGFPCNAL2Amut	653	NLS-eGFP-PCNA	CMV	Amp	Mammalian cells	Leonhardt et al., 2000[11]	4f
p53-GFP	2468	p53-GFP	CMV	Kan/Neo	Mammalian cells	Herce et al., 2013[12]	4g-h
pmCherry	2387	mCherry	CMV	Kan/Neo	Mammalian cells	this study	4g
pCAG-mCh-HD (NTD)	2339	mCherry-HDM2 (aa 1-128)	CAG	Kan/Neo	Mammalian cells	Herce et al., 2013[12]	4h
pMaSat-mRFP	2063	MaSat-mRFP	CMV	Kan/Neo	Mammalian cells	this study	S17
pEF5/FRT/V5-D-TOPO	-	-	EF-1alpha	Amp/Hyg	Mammalian cells	Invitrogen (K603501)	-
pUB-BSD-TOPO	-	-	hUbC	Amp/Bsd	Mammalian cells	Invitrogen (K512-20)	-
pFRT-B-NGFP	1656	NLS-GFP	EF-1alpha	Amp/Bsd	Mammalian cells	this study	1, 3 and S11
pFRT-B-Mecp2G	1233	Mecp2-GFP	EF-1alpha	Amp/Bsd	Mammalian cells	this study	S16
pET22b	-	GBP1 (**2**),	T7	Amp	T7 Express bacteria	this study	S4, S21-S22
pET22b-GBP1-TAT	-	GBP1-TAT	T7	Amp	T7 Express bacteria	this study	S4
pET22b-GBP1-R_{10}	-	GBP1-R_{10}	T7	Amp	T7 Express bacteria	this study	S4
pET22b-GBP4-TAT	-	GBP4-TAT	T7	Amp	T7 Express bacteria	this study	S4
pHEN6-GBP4	-	GBP4 (**1**)	T7	Amp	JM109 bacteria	Kirchhofer et al., 2010[1]	S20
pTXB1-GBP4_1-96	-	GBP4 (aa 1-96)-int-CBD (**S2**)	T7	Amp	T7 Express bacteria	this study	S2
pTXB1-GBP4	-	GBP4-int-CBD (**1A**)	T7	Amp	T7 Express bacteria	this study	2
pTXB1-GBP1	-	GBP1-int-CBD (**2A**)	T7	Amp	T7 Express bacteria	this study	2
pET28a-eGFP	-	eGFP (**7**)	T7	Amp	BL21(D3) bacteria	this study	5

*Amp: Ampicillin, Gen: Gentamicin, Neo: Neomycin, Kan: Kanamycin, Hyg: Hygromycin, Bsd: Blasticidin

2 Supplementary Material and Methods

2.1 General Information

2.1.1 Analytical UPLC-MS

UPLC-UV traces were obtained on a Waters H-class instrument equipped with a Quaternary Solvent Manager, a Waters autosampler and a Waters TUV detector connected to a 3100 mass detector with an Acquity UPLC-BEH C18 1.7 µm, 2.1x 50 mm RP column with a flow rate of 0.6 mL/min (Water Corp., USA). The following gradient was used: Method A: (A = H_2O + 0.1% TFA, B = MeCN + 0.1% TFA) 5-95% B 0-3 min, 95% B 3-5 min. UPLC-UV chromatograms were recorded at 220 nm.

2.1.2 Preparative HPLC

Preparative HPLC was performed either on a Gilson PLC 2020 system (Gilson Inc., WI, Middleton, USA) using a Nucleodur C18 HTec Spum column (Macherey-Nagel GmbH & Co. Kg, Germany, 100 A, 5 m, 250 mm x 32 mm, 30 mL/min, Method B) or on a Shimadzu SCL-8A system with a Shimadzu C-R4AX controler (Shimadzu Corp., Kyoto, Japan) using a Nucleodur C18 column (Macherey-Nagel GmbH & Co. Kg, Germany, 100 A, 5 m, 250 mm x 46 mm, 1 mL/min, Method C).
The following gradients were used: Method B: (A = H_2O + 0.1% TFA, B = MeCN + 0.1% TFA) flow rate 30 mL/min, 5-40% B 0-60 min. Method C: (A = H_2O + 0.1% TFA, B = 80% MeCN/20 % H_2O + 0.1% TFA) flow rate 1 mL/min, 5-95% B 0-40 min.

2.1.3 High-resolution mass spectra (HRMS)

High-resolution mass spectra (HRMS) were measured on an Acquity UPLC system and a LCT Premier™ (Waters Corp., USA) time-of-flight mass spectrometer with electrospray ionization using water and acetonitrile (10-90% gradient) with 0.1% formic acid as eluent.

2.1.4 Protein MS

Protein-MS was measured on an Acquity UPLC system and a LCT Premier™ (Waters Corp., USA) time-of-flight mass spectrometer with electrospray ionization using water and acetonitrile (10-90% gradient) with 0.1% formic acid as eluent. Deconvolution was performed using *MaxEnt 1*.

2.1.5 Protein concentration

The concentration of proteins was determined by absorption spectroscopy measurements at 280 nm using the extinction coefficient of the respective protein and/or using a BCA protein assay (Thermo Fisher Scientific, USA). The mean value of at least three individual measurements was taken. In the case of mixed samples containing unconjugated and conjugated nanobody, the concentration was corrected by the ratio of Coomassie stained protein bands. For the “co-transported” proteins (GFP-tagged PCNA and Mecp2) protein concentrations were measured using the Pierce 660 nm protein assay reagent (Thermo Fisher Scientific, USA) and for GFP-tagged PCNA, further evaluated using a BSA calibration standard loaded and stained on the same gel.

2.1.6 Column chromatography

Column chromatography was performed on silica gel (Acros Silica gel 60 Å, 0.035-0.070 mm).

2.1.7 NMR

NMR spectra were recorded with a Bruker Ultrashield 300 MHz spectrometer (Bruker Corp., USA) at ambient temperature. The chemical shifts are reported in ppm relative to the residual solvent peak.

2.1.8 Reagents and solvents

Reagents and solvents were, unless stated otherwise, commercially available as reagent grade and did not require further purification. Resins and Fmoc-protected amino acids were purchased from IRIS BioTech (Germany) or Novabiochem (Germany). TAMRA labeled linear TAT and cR_{10} were purchased from Biosyntan GmbH (Germany).

2.1.9 SPPS

Solid-phase peptide synthesis (SPPS) was either carried out manually or with a Tribute-UV peptide synthesizer (Protein Technologies, USA) via standard Fmoc-based conditions. Fmoc deprotection was done using 25% piperidine in DMF (3 x 10 minutes) and couplings were performed with at least 5 eq. of HOBt, HBTU and DIPEA or 5 eq. of HATU and DIPEA for 45 min in DMF[13].

2.1.10 SDS-PAGE gel analysis

Coomassie or fluorescently stained bands in SDS-Gels were imaged using a ChemiDoc imaging system (BioRad, USA) and band intensities analysed by the software ImageLab (BioRad, USA).

2.1.11 Heparin based separation of CPP-nanobodies for MS analysis.

Crude NCL mixtures were dialysed to 10 mM Tris-HCl, 50 mM NaCl, pH 8.0 and highly positively charged CPP-containing nanobodies immobilized on heparin agarose beads equilibrated with 10 mM Tris-HCl, pH 8.0. The beads were washed with 20 eq. 10 mM Tris-HCl pH 8.0 and bound protein eluted with 10 mM Tris-HCl, 500 mM NaCl, pH 8.0 and finally dialysed against 0.5 mM HEPES, 14 mM NaCl, 0.25 mM KCl, 0.5 mM glycine, pH 7.5. Purified conjugates were directly used for MS analysis.

2.2 Experimental section

2.2.1 Expression and purification of proteins

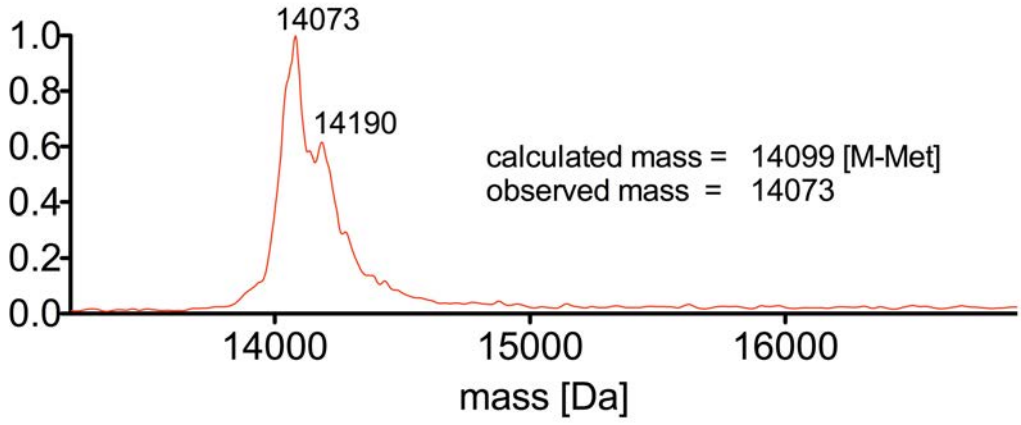
2.2.1.1 Plasmids

All plasmids used are summarized in Supplementary Table 3.
Mammalian expression constructs coding for NLS-eGFP-PCNA, p53-GFP, as well as mCherry-HDM2 (aa 1-128) have been described previously[11, 12].
The mammalian expression construct pEGFP-N1 was purchased from Clontech (GenBank Accession #U55762).
For construction of the mammalian expression vector coding for MaSat-mRFP, Mecp2 was replaced from pMecp2mR[14] by CMV-MaSat of pMaSat-GFP[15] using SnaBI and AgeI restriction endonuclease sites.
For construction of pmCherry, eGFP was removed from pEGFP-N1 and replaced by mCherry of pFRT-B-CPCNA[7] using BsrGI and BamHI restriction endonuclease sites.
Additional plasmids for the generation of stable cell lines (pEF5/FRT/V5-D-TOPO, pUB-BSD-TOPO, pFRT-B-NGFP and pFRT-B-Mecp2G) are described in section 2.2.8.1 and Supplementary Table 3.
The insect cell expression construct coding for the GFP-tagged rat Mecp2 (**9**) full-length protein was previously described[10, 16].
For construction of the bacterial expression construct pRHGPCNA coding for the His-NLS-GFP-tagged human PCNA (**8**) full-length protein, TAT-p27 was replaced from pTATp27WT[17] by NLS-GFP-PCNA of pENeGFPCNAL2mut[11] using BamHI and EcoRI/MfeI restriction endonuclease sites.
For construction of the bacterial expression constructs coding for GBP4-intein-CBD fusion (**1A**) and GBP1-intein-CBD fusion (**2A**), GBP1 and GBP4 were cloned from pHEN6GBP1[1] and pHEN6GBP4[1] into pTXB1(New England Biolabs) using XhoI and SapI restriction endonuclease sites. The GBP1 construct included an Ala_3 spacer.
For construction of the bacterial expression constructs coding for GBP1 (**2**) and nanobody-linear CPP fusion (GBP1-TAT, GBP1-R_{10}, GBP4-TAT), GBP1-linCPP or GBP4-linCPP was amplified from pHEN6GBP1[1] and pHEN6GBP4[1] by PCR with respective primers (see below) and cloned into pET22b using NcoI and EcoRI restriction endonuclease sites resulting in nanobodies with an pelB leader-sequence, N-terminal 6xHis tag and C-terminal CPPs (TAT: CKRRRGRKKRRE, R_{10}: CKRRRRRRRRRRE). Primer:
GBP1-R_{10} forward:
GGGGCCATGGCCCACCATCATCACCATCATGATGTGCAGCTGGTGGAGT,
GBP1-R_{10} reverse:
CCCCGAATTCTTATTCGCGACGGCGACGGCGGCGACGGCGACGGCGACGTTT
GCATGAGGAGACGGTGACC
GBP1-TAT forward:
GGGGCCATGGCCCACCATCATCACCATCATGATGTGCAGCTGGTGGAGT
GBP1-TAT reverse:
CCCCGAATTCTTATTCGCGACGTTTTTTGCGGCCACGGCGACGTTTGCATGAGG
AGACGGTGACC
GBP4-TAT forward:
GGGGCCATGGCCCACCATCATCACCATCATGATGTGCAGCTGCAGGAGT,
GBP4-TAT reverse:
CCCCGAATTCTTATTCGCGACGTTTTTTGCGGCCACGGCGACGTTTGCATGAGG
AGACGGTGACC

For construction of the bacterial expression constructs coding for eGFP, DNA encoding GFP was amplified from pGEX4T1eGFP (provided by Ronald Kühne) by PCR with respective primers (see below) and cloned into pET28a using NdeI and XhoI restriction endonuclease sites. Primer:
eGFP forward:
GGGGCCCATATGGGATCAATTCAGATG
eGFP reverse:
GGGGCTCGAGTATTACTTGTACAGCTCGTC
The bacterial expression construct coding for the GBP4 (**1**) was previously described[1].

*2.2.1.2 Expression of recombinant GBP4 (**1**)*

GBP4 (**1**) was expressed in *E. coli* (JM109) (see Supplementary Table 2). Cells were induced with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and incubated at 18 °C for 18 h. Lysis was performed in presence of lysozyme (100 µg/mL), DNAse (25 µg/mL) and PMSF (2 mM) followed by sonication (Branson® Sonifier; 16 x 8sec, 20% Amplitude) and debris centrifugation at 20.000 g for 30 min. The protein was purified with an Äkta FPLC system using a 5 mL His-Trap (GE Healthcare, USA) column, peak fractions were concentrated to 2 mL using Amicon filter columns (cut-off 3 kDa; (Merck Millipore, Germany) and subjected to size exclusion chromatography using a Superdex 75 column (GE Healthcare, USA). Peak fractions were pooled and protein aliquots were shock-frozen and stored at -80 °C. For MS analysis, the protein was precipitated by acetone, resolubilized in ddH_2O and subjected to MALDI measurements. Spectrum is shown below (Supplementary Fig. 20).

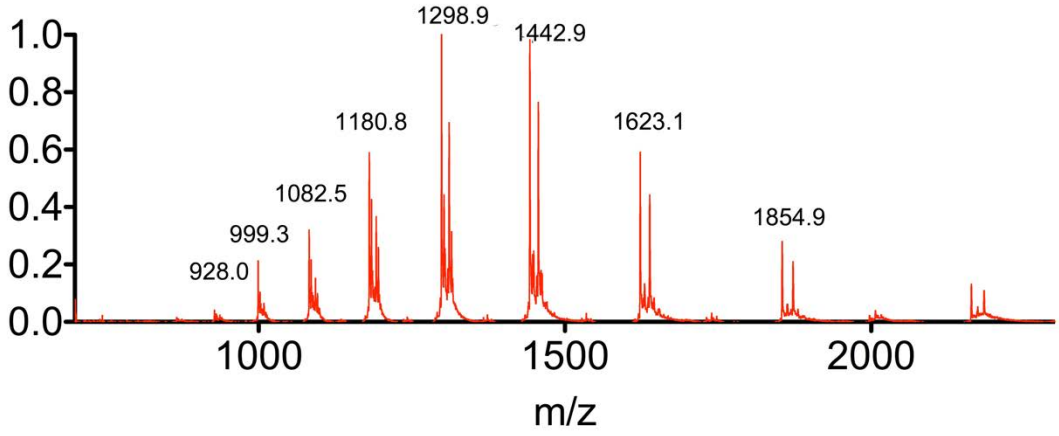


Supplementary Figure 20 | MALDI of GBP4 (**1**).

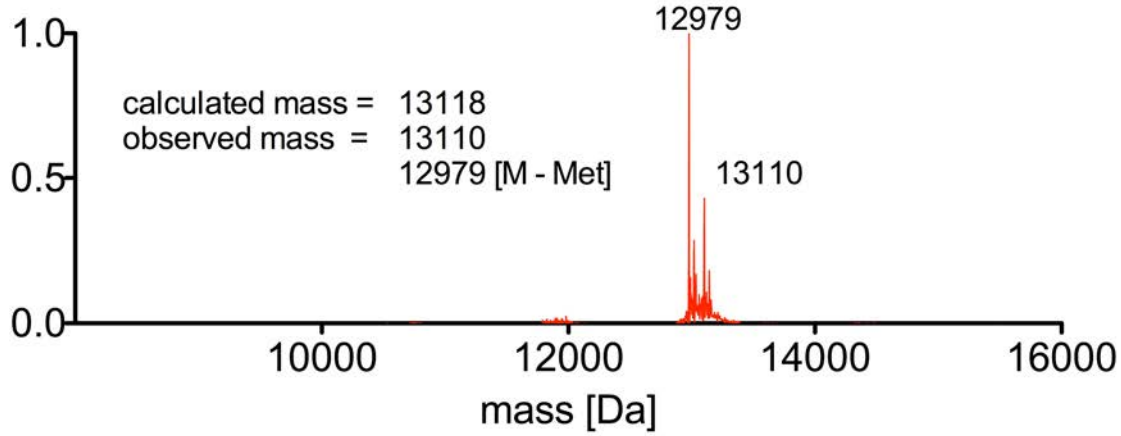
*2.2.1.3 Expression of recombinant GBP1 (**2**)*

GBP1 was expressed in *E. coli* (T7 Express) (see Supplementary Table 2). Cells were induced with 0.1 mM IPTG and incubated at 18 °C for 18 h. Lysis was performed in presence of lysozyme (100 µg/mL), DNAse (25 µg/mL) and PMSF (2 mM) followed by sonication (Branson® Sonifier; 16 x 8sec, 20% Amplitude) and debris centrifugation at 20.000 g for 30 min. The protein was purified with a BioRad NGC system (BioRad, USA) using a 5 mL HisTrap FF (GE Healthcare, USA) column,

peak fractions were collected and subjected to size exclusion chromatography using a Superdex 75 column (GE Healthcare, USA). Peak fractions were pooled and protein aliquots were shock-frozen and stored at -80 °C. The protein was analyzed by ESI-MS. Spectra are shown below (Supplementary Fig. 21 and 22).



Supplementary Figure 21 | ESI-MS of GBP1 (**2**).



Supplementary Figure 22 | Deconvoluted mass of GBP1 (**2**).

*2.2.1.4 eGFP (**7**)*

The H_6-eGFP protein was expressed in *E. coli* BL21(DE3) (see Supplementary Table 2) using media containing 100 µg/mL ampicillin (LB_{AMP}). Cells were grown at 37°C, 180 rpm until OD_{600} reached approximately 0.8, induced with 0.3 mM IPTG and incubated at 18°C for 19 h. Lysis was performed in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) using a high pressure homogenizer (Microfluidics LM10 Microfluidizer) and debris centrifuged at 20,000 g for 30 min. The protein was purified with a BioRad NGC system (BioRad, USA) using a 5 mL HisTrap FF (GE Healthcare, USA) column, the protein eluted with elution buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4,1.8 mM KH2PO4 and 500 mM imidazole), peak fractions were collected and desalted to PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) using a HiPrep 26/10 Desalting column (GE Healthcare, USA). Thrombin (1 µL/mL; Thrombin restriction grade, Merck Millipore, Germany) was added to the protein fractions and incubated

for 16 h at 16°C. The protein was concentrated to 1 mL using Vivaspin 20 (cut-off 19 kDa; Merck Millipore, Germany) and subjected to a final size exclusion chromatography in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 and 1.8 mM KH_2PO_4) using a Superdex 75 10/300 GL column and (GE Healthcare, USA). Peak fractions were pooled, 0.1 mM PMSF added and aliquots were shock-frozen and stored at -80 °C.

*2.2.1.5 GFP-tagged PCNA (**8**)*

BL21(DE3) (see Supplementary Table 2) cells overexpressing His-NLS-GFP-tagged PCNA (pRHGPCNA) were pelleted and incubated at -20˚C for 2.5 h. Thawed pellets were resuspended in extraction buffer (250 mM sodium phosphate, 1.5 M sodium chloride, pH 8) containing 0.75 mg/mL Lysozyme (Sigma-Aldrich, USA). After incubation on ice for 90 min, cells were disrupted by sonication and three alternate freeze and thaw steps in liquid nitrogen and water, respectively. Following addition of 10 µg/mL RNaseA (Qiagen, Germany) and 10 µg/mL DNaseI (Sigma-Aldrich, USA), lysates were incubated for 1 h on ice and centrifuged (12,000 g, 4˚C) for 20 min. Supernatant was removed from cell debris, supplemented with protease inhibitors PMSF (100 µM, Carl Roth, Germany), Pepstatin A (1 µM, Sigma Aldrich, USA) and E64 (10 µM, AppliChem, Germany) and immobilized to TALON Metal Affinity Resin (Clontech Laboratories, Inc., USA) by incubation at 4˚C for 3 h. Immobilized proteins were washed 3x with wash buffer (50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 8) and 3x in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4), both containing protease inhibitors as described above. For elution of the His-NLS-GFP-tagged PCNA proteins, beads were incubated with elution buffer (50 mM sodium phosphate, 300 mM sodium chloride, 150 mM imidazole, pH 8) supplemented with protease inhibitors as described above for 10 min on ice. After centrifugation (800 rpm, 4˚C, 3 min), the eluate was separated from the beads. Elution buffer was exchanged to 1x PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) using Amicon Ultra centrifugal filter units (Sigma-Aldrich, USA). Purified proteins were analyzed by SDS-PAGE (see Supplementary Fig. 15).

*2.2.1.6 GFP-tagged Mecp2 (**9**)*

Sf9 insect cells (Invitrogen, Paisley PA4 9RF, UK) (see Supplementary Table 2) used for protein production were cultivated and transfected as previously described[16]. GFP-tagged Mecp2 proteins were purified from Sf9 insect cells as previously described[10, 16] with following exceptions: The re-suspension buffer was supplied with protease inhibitors in following concentrations: E64: 10 µM (AppliChem, Germany), Pepstatin A: 1 µM (Sigma-Aldrich, USA) and PMSF: 100 µM (Carl Roth, Germany). Proteins were eluted from GFP binder sepharose beads by the addition of 4 M $MgCl_2$, pH 4.4 and subsequent incubation on ice for 10 min. Elution buffer was exchanged to 1x PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) using Amicon Ultra centrifugal filter units (Sigma-Aldrich, USA). Purified proteins were analyzed by SDS-PAGE (see Supplementary Fig. 15).

2.2.1.7 Expression of recombinant nanobodies with linear CPPs

Proteins were expressed in *E. coli* (T7 Express) (see Supplementary Table 2). Cells were induced with 0.1 mM IPTG and incubated at 18 °C for 18 h. Lysis was performed in presence of Lysozyme (100 µg/mL), DNAse (25 µg/mL) and PMSF (2 mM) followed by sonication (Branson® Sonifier; 16 x 8sec, 20% Amplitude) and debris centrifugation at 20.000 g for 30 min. Lysates were analyzed by SDS-PAGE and anti-His tag Western Blot (see Supplementary Fig. 5).

*2.2.1.8 Expression of truncated and full-length GBP4-intein fusion proteins **S2** and **1A***

E. coli T7 Express (see Supplementary Table 2) was transformed with the expressions vectors described above. Bacterial expression was induced by 0.4 mM IPTG and carried out for 20 h at 18°C, 180 rpm. The cells were collected and lyzed using a high pressure homogenizer (Microfluidics LM10 Microfluidizer) in 40 mL lysis buffer (20 mM Tris-HCl, pH 8.5, 0.5 M NaCl). The pellet containing insoluble protein/inclusion bodies was suspended and stirred in 10 mL breaking buffer (1 h, 4°C). The solution was centrifuged (20,000 g, 4°C, 20 min) and the supernatant loaded into a dialysis bag (cutoff 3000 kDa). The solution was dialyzed consecutively against renaturation buffers A, B, C, D and, in case of full refolding, E (see Supplementary Table 4) for 3 h each, centrifuged for removal of impurities or incorrectly folded protein and directly used for further experiments (intein-cleavage/CPP-attachment).

Supplementary Table 4: Refolding buffers

Lysis buffer	20 mM Tris-HCl, pH 8.5, 0.5 M NaCl
Breaking buffer	20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 7 M guanidine-HCl
Renaturation Buffer A	20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 8 M urea
Renaturation Buffer B	20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 6 M urea
Renaturation Buffer C	C 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 4 M urea
Renaturation Buffer D	20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 2 M urea
Renaturation Buffer E	20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 0.1 mM ox. glutathione, 1 mM red. glutathione

*2.2.1.9 Expression of full-length GBP1-A3-intein fusion protein **2A***

E. coli T7 Express (see Supplementary Table 2) transformed with the vector described above was incubated with 0.1 mM IPTG for 20 h at 18°C /180 rpm. The cells were collected and lysed by sonication (Branson sonifier, amplitude 30%, 2 min, pulse on 2 sec, pulse of 1 sec, max temp: 14°C) in 10 mL lysis buffer (20 mM Tris-HCl, pH 8.5, 0.5 M NaCl). The cell lysate was directly used for further experiments.

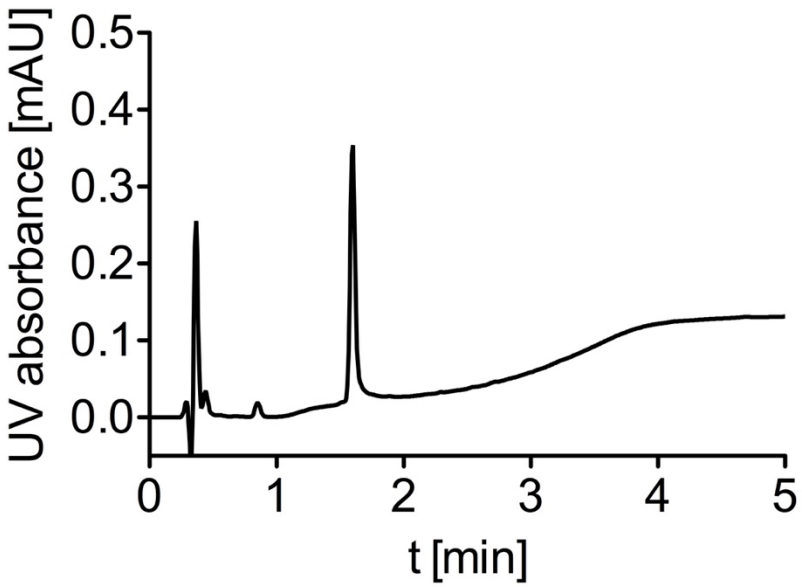
2.2.2 RNA purification and *in vitro* RNA binding assay

Purification of total RNA (see Supplementary Fig. 1) from human embryonic kidney (HEK) cells, as well as the *in vitro* RNA binding assay were peformed as previously described[18].

2.2.3 Peptide Synthesis

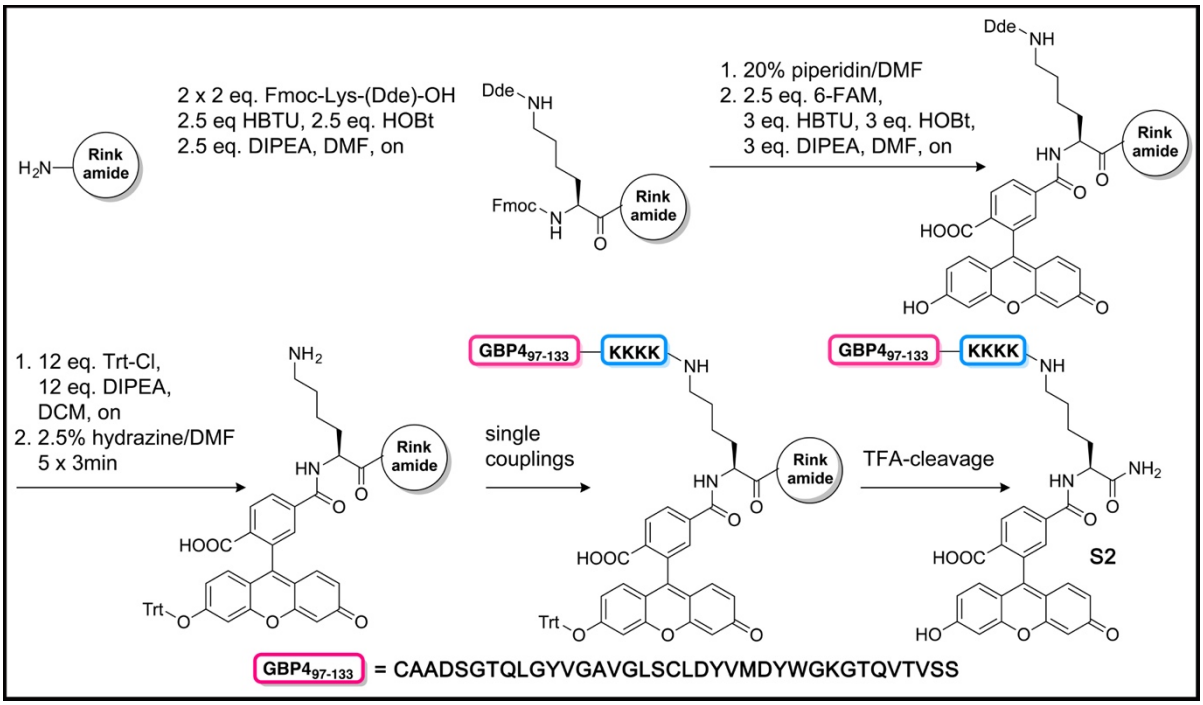
*2.2.3.1 Synthesis of 5,6-FAM-cyclic R_{10} (**S1**)*

The linear peptide of the sequence (Fmoc)(Trt)C(PEG)$_2$(Alloc)KRrRrRrRrRr(OAll)E (Trt = trityl) was synthesized on a Rink Amide resin (0.1 mmol, 0.14 mmol/g). Upper case letters correspond to L-, lower case letters to D-amino acids. After Fmoc deprotection (25% piperidine in DMF), 5,6-FAM was coupled using HATU (2 eq. each) for 3 h. The Alloc- and OAll protecting groups were removed using $Pd(PPh_3)_4$ (11.6 mg, 0.1 eq.) and $PhSiH_3$ (25 eq., 270.5 mg, 308 µL) in dry DCM for 30 minutes at ambient temperature under argon atmosphere. To remove the Pd catalyst afterwards, the resin was washed additionally with 0.2 M DIPEA/DMF. The cyclization of the peptide was carried out using 1 eq. HATU/2 eq. DIPEA in 8 mL DMF for 2 h at rt. After washing and drying the peptide was cleaved from the solid support (five hours in 8 mL 95 % TFA, 2.5 % TIS, 2.5 % DTT), TFA evaporated via N_2-stream and the peptide precipitated in 40 mL diethyl ether. The product was purified using preparative HPLC using method B to yield a white trifluoroacetate (9.9 mg, 2.86 µmol, yield 2.8%, molar mass (peptide) = 2570 Da, molar mass (TFA_{11}-salt) = 3824 Da) in good purity (see Supplementary Fig. 23). HRMS: m/z: 857.1304 $[M+3H]^{3+}$ (calcd. m/z: 857.1301).



Supplementary Figure 23 | UPLC-UV purity of the peptide **S1**

2.2.3.2 *Synthesis of GBP4$_{97-133}$-5,6-FAM (**S3**)*

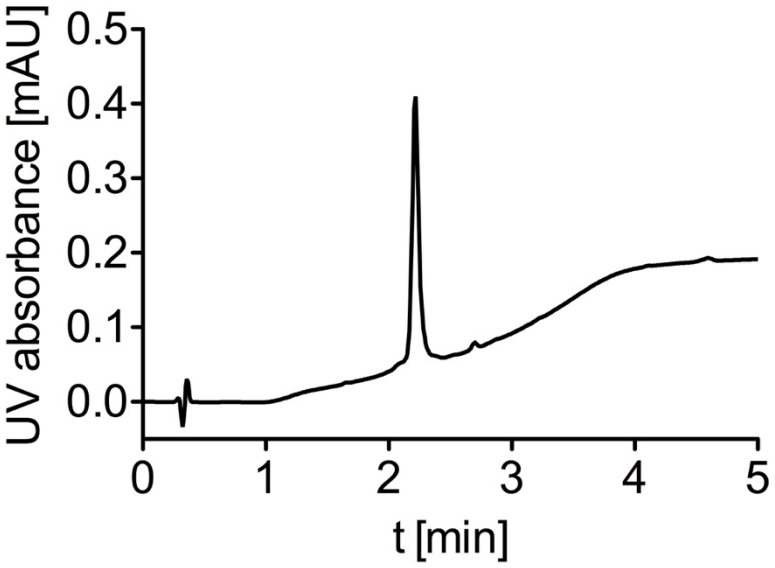


Supplementary Scheme 1 | Synthesis of peptide **S3**.

The 5,6-FAM labeled C-terminal nanobody peptide **S3** was synthesized by standard Fmoc-based chemistry in a linear synthesis on an Activotec peptide synthesizer as outlined in Supplementary Scheme 1.

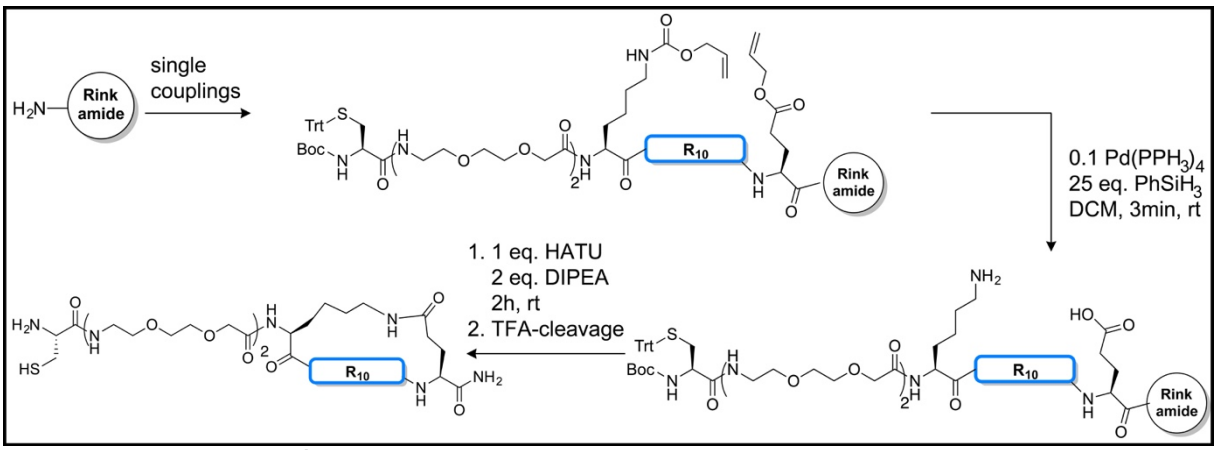
In the first step, N_α-Fmoc-N_ε-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl]-L-lysine was coupled to a Rink Amide resin (scale: 0.2 mmol, subst: 0.71 mmol/g) using standard coupling conditions (2 eq. AA, 2.5 eq. HBTU, 2.5 eq. HOBt, 2.5 eq. DIPEA in DMF). Following Fmoc-deprotection (20% piperidine/DMF), 5,6-carboxyfluorescein (5,6-FAM) was attached to the α-NH_2 in a single coupling step (2.5 eq. 5,6-FAM, 3 eq. HBTU, 3 eq. HOBt, 3 eq. DIPEA, (DMF), o.n.). For the assembly of the C-terminal nanobody fragment, the carboxyfluorescein's phenol was protected using 12 eq. triphenylmethyl chloride (Trt-Cl) and 12 eq. DIPEA in DCM (1 h) followed by deprotection of the ε-NH_2 (2.5% hydrazine/DMF, 5x 3 min). The free ε-NH_2-group was then used for the linear assembly of the growing peptide chain using double couplings with 5 eq. of each amino acid and 10 eq. of HBTU, HOBt and DIPEA after the introduction of a KKKK-spacer to improve the solubility of the C-terminal nanobody fragment.

After washing and drying, the peptide was cleaved from the solid support (3 h in 8 mL 95% TFA, 2% TIS, 2% Thioanisol, 1% DTT). The cleavage cocktail was evaporated via N_2-stream and the peptide precipitated in 40 mL diethyl ether. The peptide was purified using preparative HPLC (method C) to yield a white trifluoroacetate (37.3 mg, 7.77 µmol, yield 3.9%, molar mass (peptide) = 4800.23 Da in good purity (see Supplementary Fig. 24). HRMS: m/z: 1200.5753 $[M+4H]^{4+}$ (calcd. m/z: 1200.5759).



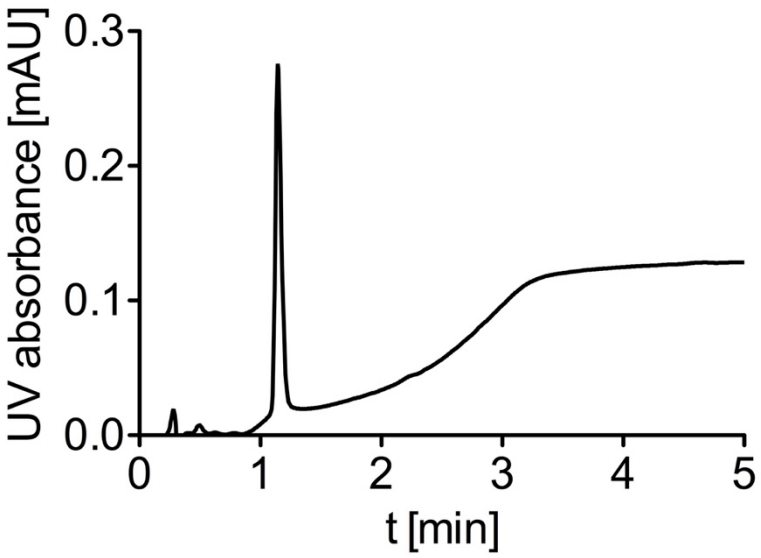
Supplementary Figure 24 | UPLC-UV purity of the peptide **S3.**

*2.2.3.3 Synthesis of cyclic R_{10} (**3**)*



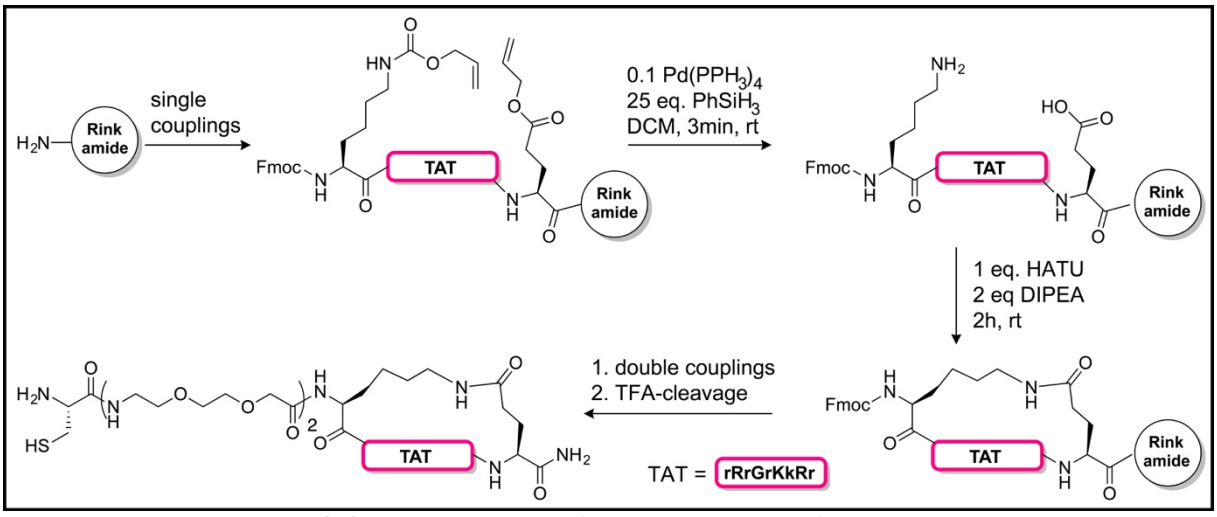
Supplementary Scheme 2 | Synthetic scheme for the synthesis of a circular cell-penetrating R_{10} peptide **3**.

The synthesis of peptide **3** is outlined in Supplementary Scheme 2. A linear peptide of the sequence (Boc)(Trt)C(PEG)$_2$(Alloc)KRrRrRrRrRr(OAll)E was synthesized on a Rink amide resin (0.1 mmol, 0.14 mmol/g). Upper case letters correspond to L-, lower case letters to D-amino acids. The couplings were performed using 5 eq. of amino acid. The Alloc- and OAll protecting groups were removed using $Pd(PPh_3)_4$ (11.6 mg, 0.1 eq.) and $PhSiH_3$ (25 eq., 270.5 mg, 308 µL) in dry DCM for 30 min at ambient temperature under argon atmosphere. To remove the Pd catalyst afterwards, the resin was washed additionally with 0.2 M DIPEA/DMF. The cyclization of the peptide was carried out using 1 eq. HATU and 2 eq. DIPEA in 8 mL DMF for 2 h at RT. After washing and drying the peptide was cleaved from the solid support (5 h in 8 mL 95% TFA, 2.5% TIS, 2.5% DTT), TFA evaporated via N_2-stream and the peptide precipitated in 40 mL diethyl ether. The peptide was purified using preparative HPLC method B to yield a white trifluoroacetate (9.9 mg, 2.86 µmol, yield 2.8%, molar mass (peptide) = 2212 Da, molar mass (TFA_{11}-salt) = 3466 Da) in good purity (see Supplementary Fig. 25). HRMS: m/z: 737.7731 $[M+3H]^{3+}$ (calcd. m/z: 737.7739).



Supplementary Figure 25 | UPLC-UV purity of the cR_{10} peptide **3**.

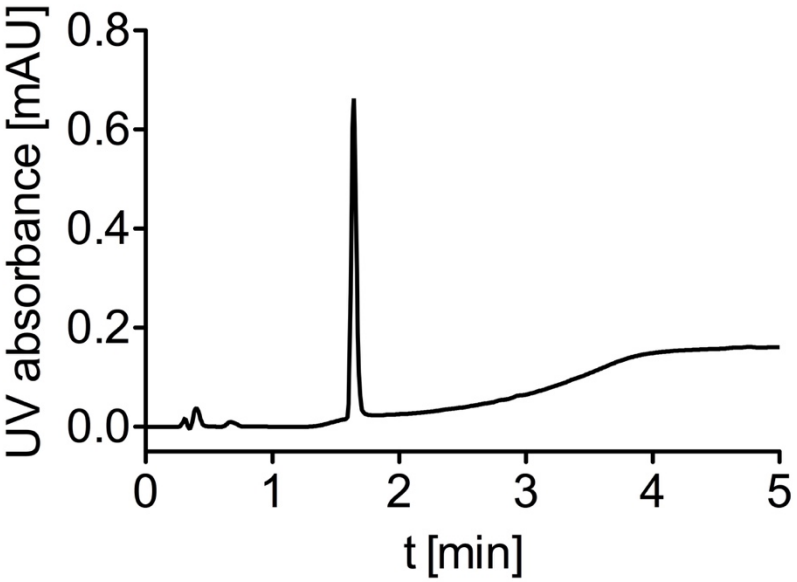
*2.2.3.4 Synthesis of cyclic TAT (**4**)*



Supplementary Scheme 3 | Synthetic scheme for the synthesis of a circular cell-penetrating TAT peptide **4**.

The synthesis of peptide **4** is outlined in Supplementary Scheme 3. First, a linear peptide of the sequence (Fmoc)(Alloc)KrRrGrKkRr(OAll)E was synthesized on a Rink amide resin (0.1 mmol, 0.14 mmol/g). Upper case letters correspond to L-, lower case letters to D-amino acids. The couplings were performed using 5 eq. of amino acid. Alloc- and OAll protecting groups were removed using $Pd(PPh_3)_4$ (58 mg, 0.5 eq.) in a mixture of $CHCl_3$/AcOH/NMM in a ratio of 37/2/1 for 2 h at ambient temperature under argon atmosphere. To remove the Pd catalyst afterwards, the resin was washed additionally with 0.2 M DIPEA/DMF. The cyclization of the peptide was carried out using 1 eq. HATU and 2 eq. DIPEA in 8 mL DMF for 2 h at RT, followed by capping. After Fmoc-removal (2x 20 min piperidine/DMF [1:4]), a linker consisting of two repeats of [2-[2-aminoethoxy]ethoxy]acetic acid was introduced before Boc-Cys was coupled to the N-terminus of this linker. These couplings were each performed twice using 5 eq. of HOBt/DIC activated acid in DMF. After washing and drying the peptide was cleaved from the solid support (4 h in 8 mL 95% TFA,

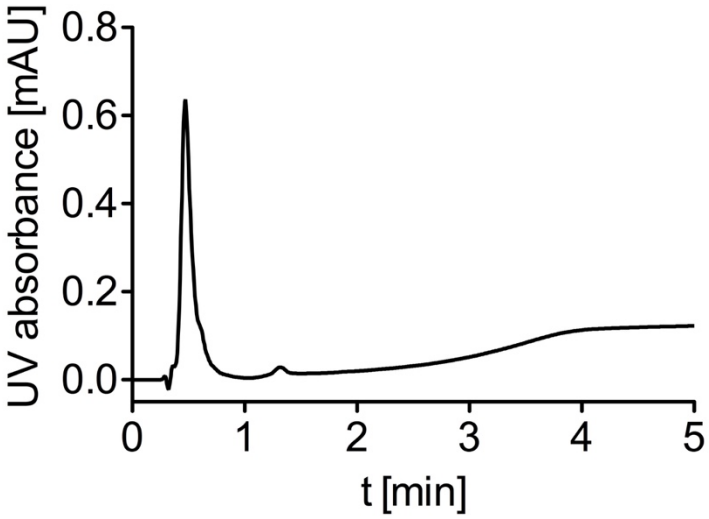
2.5% TIS, 2.5% DTT), TFA evaporated via N_2-stream and the peptide precipitated in 40 mL diethy lether. The peptide was purified using preparative HPLC method B to yield a white trifluoroacetate (14 mg, 4.78 µmol, yield 4.8%, molar mass (peptide) = 1901 Da, molar mass (TFA_8-salt) = 2927 Da) in good purity (see Supplementary Fig. 26). HRMS: m/z: 634.0493 $[M+3H]^{3+}$ (calcd. m/z: 634.0429).



Supplementary Figure 26 | UPLC-UV purity of the cTAT peptide **4**.

*2.2.3.5 Synthesis of linear R_{10} peptide **5***

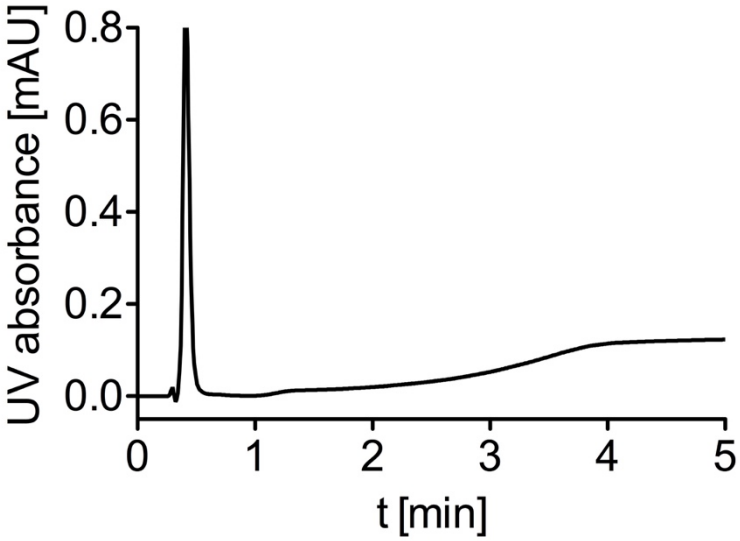
The peptide of the sequence (Boc)(Trt)C$(PEG)_2$KrRrGrKkRrE was synthesized on a Rink Amide resin (0.1 mmol, 0.14 mmol/g). Upper case letters correspond to L-, lower case letters to D-amino acids. The couplings were performed using 5 eq. of amino acid. After washing and drying the peptide was cleaved from the solid support (5 h in 8 mL 95% TFA, 2.5% TIS, 2.5% DTT), TFA evaporated via N_2-stream and the peptide precipitated in 40 mL diethy lether. The peptide was purified using preparative HPLC method B to yield a white trifluoroacetate (25 mg, 11.2 µmol, yield 11%, molar mass (peptide) = 2229 Da, molar mass (TFA_{11}-salt) = 3483 Da) in good purity (see Supplementary Fig. 27). HRMS: m/z: 744,1214 $[M+3H]^{3+}$ (calcd. m/z: 744.2298).



Supplementary Figure 27 | UPLC-UV of the linear R_{10} peptide **5**.

*2.2.3.6 Synthesis of linear TAT peptide **6***

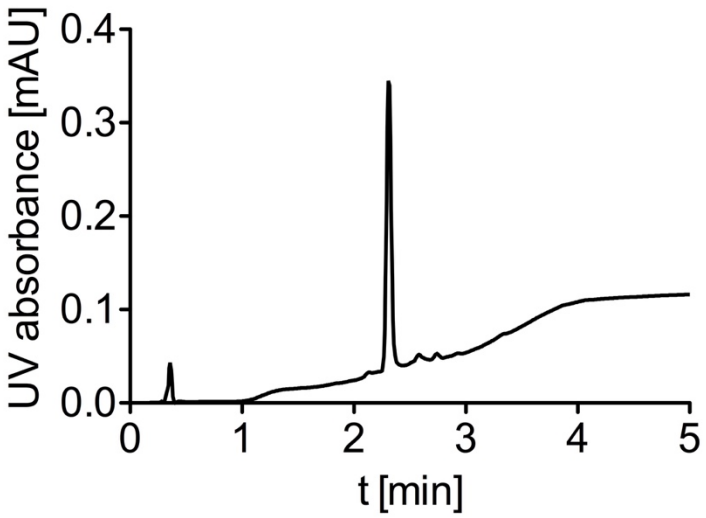
The peptide of the sequence (Boc)(Trt)C(PEG)$_2$KrRrGrKkRrE was synthesized on a Rink Amide resin (0.1 mmol, 0.14 mmol/g). Upper case letters correspond to L-, lower case letters to D-amino acids. The couplings were performed using 5 eq. of amino acid. After washing and drying the peptide was cleaved from the solid support (5 h in 8 mL 95% TFA, 2.5% TIS, 2.5% DTT), TFA evaporated via N_2-stream and the peptide precipitated in 40 mL diethyl ether. The peptide was purified using preparative HPLC method B to yield a white trifluoroacetate (30 mg, 15.63 µmol, yield 16%, molar mass (peptide) = 1919 Da, molar mass (TFA_{11}-salt) = 2945 Da) in good purity (see Supplementary Fig. 28). HRMS: m/z: 640,3884 $[M+3H]^{3+}$ (calcd. m/z: 640.8433).



Supplementary Figure 28 | UPLC-UV of the linear TAT peptide **6**.

*2.2.3.7 Synthesis of Cy5-peptide **10***

A linear peptide of the sequence (Boc)(Trt)C(PEG)$_2$(Alloc)KG was synthesized on a Rink Amide resin (0.1 mmol, 0.54 mmol/g). The couplings were performed using 5 eq. of amino acid. The Alloc protecting group was removed using Pd(PPh$_3$)$_4$ (11.6 mg, 0.1 eq.) and PhSiH$_3$ (25 eq., 270.5 mg, 308 µL) in dry DCM for 30 min at ambient temperature under Argon atmosphere. To remove the Pd catalyst afterwards, the resin was washed additionally with 0.2 M DIPEA/DMF. 2 eq. Cy5-COOH **S8** (for synthetic details see Supplementary section 2.2.5) was coupled using 2 eq. HATU/2 eq. DIPEA in 8 mL DMF for 2h at RT. After washing and drying the peptide was cleaved from the solid support (5 h in 8 mL 95% TFA, 2.5% TIS, 2.5% DTT), TFA evaporated via N_2-stream and the peptide precipitated in 40 mL diethylether. The peptide was purified using preparative HPLC method C to yield a white trifluoroacetate (60 mg, 56 µmol, yield 56%, molar mass (peptide) = 1062 Da) in good purity (see Supplementary Fig. 29). HRMS: m/z: 1062.5767 $[M+H]^+$ (calcd. m/z: 1062.5824).



Supplementary Figure 29 | UPLC-UV of the Cy5-peptide **10**.

2.2.4 Intein-cleavage experiments

0.5 mL of the dialyzed lysate of **S2 or 1A** was loaded on 125 µL pre-equilibrated (3x 1 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 1 mL equilibration buffer, flushed with 125 µL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 125 µL cleavage buffer for 2 h at RT.

2.2.5 Synthesis

*2.2.5.1 1,2,3,3-tetramethyl-3H-indolium iodide (**S5**)*

The compound was synthesized according to a previously published procedure by microwave irradiation of 2,3,3-trimethlyindolenine together with 1.2 eq. methyl iodide for 20 min at 120 °C[19].

^{1}H NMR (600 MHz, DMSO-d_6) δ = 7.93 – 7.88 (m, 1H), 7.85 – 7.80 (m, 1H), 7.66 – 7.57 (m, 2H), 3.97 (s, 3H), 2.77 (s, 3H), 1.53 (s, 6H).
^{13}C NMR (151 MHz, DMSO) δ = 195.98, 142.07, 141.57, 129.29, 128.79, 123.26, 115.09, 53.91, 34.69, 21.69, 14.12.

Data is in accordance with literature values[20].

*2.2.5.2 1-(5-Carboxypentyl)-2,3,3-trimethyl-3H-indolium iodide (**S6**)*

A 10-ml microwave vessel equipped with a stirring bar was charged with 500 mg 2,3,3-trimethlyindolenine (3.140 mmol, 1.2 eq.), 510 mg 6-bromohexanoic acid (2.615 mmol, 1.0 eq.) and 2 ml nitromethane. The mixture was irradiated for 1 h at 100 °C, cooled to room temperature and poored into 100 ml of cold ether. The precipitate was collected by centrifugation, dissolved in 2 ml of methanol, again precipitated with cold ether and collected. The purple oil was dried under high vacuum to give a purple solid that was sufficiently pure for the next steps.

^{1}H NMR (300 MHz, DMSO-d_6) δ = 8.01 (dd, *J*=6.2, 2.9, 1H), 7.86 (dd, *J*=5.7, 3.1, 1H), 7.61 (dd, *J*=5.8, 3.1, 2H), 4.48 (t, *J*=7.7, 2H), 2.88 (s, 3H), 2.22 (t, *J*=7.1, 2H), 1.91 – 1.78 (m, 2H), 1.54 (s, 6H), 1.63 – 1.34 (m, 4H).
^{13}C NMR (75 MHz, DMSO) δ = 196.58, 174.39, 141.92, 141.09, 129.42, 128.98, 123.61, 115.61, 54.24, 47.60, 33.45, 27.03, 25.46, 24.10, 22.08, 14.34.

Data is in accordance with literature values[21].

*2.2.5.3 1,3,3-trimethyl-2-((1E,3E)-4-(N-phenylacetamido)buta-1,3-dien-1-yl)-3H-indolium chloride (**S7**)*

The compound was synthesized according to a previously published procedure from 1,2,3,3-tetramethyl-3*H*-indolium iodide and 1.1 eq. malonaldehyde dianilide hydrochloride in a 1:1 mixture of acetic acid and acetic anhydride[21].

^{1}H NMR (600 MHz, DMSO-d_6) δ = 8.88 (d, *J*=13.2, 1H), 8.50 (dd, *J*=15.1, 11.2, 1H), 7.78 (d, *J*=7.4, 1H), 7.72 (d, *J*=7.9, 1H), 7.68 – 7.50 (m, 5H), 7.45 (d, *J*=7.6, 2H), 6.84 (d, *J*=15.1, 1H), 5.52 (dd, *J*=13.1, 11.3, 1H), 3.81 (s, 3H), 2.05 (bs, 3H), 1.70 (s, 6H).
^{13}C NMR (151 MHz, DMSO) δ = 181.02, 170.13, 156.83, 143.29, 142.28, 138.32, 130.88, 130.06, 129.20, 129.10, 128.89, 128.64, 123.15, 119.45, 114.58, 113.27, 112.25, 51.61, 33.65, 26.11, 23.71.

Data is in accordance with literature values[21].

*2.2.5.4 Cy5-COOH **S8***

2-((1E,3E,5Z)-5-(1-(5-carboxypentyl)-3,3-dimethylindolin-2-ylidene)penta-1,3-dienyl)-1,3,3-trimethyl-3H-indolium iodide

The compound was synthesized according to a previously published procedure from 1,3,3-trimethyl-2-((1*E*,3*E*)-4-(*N*-phenylacetamido)buta-1,3-dien-1-yl)-3*H*-indolium chloride and 1.1 eq. 1-(5-Carboxypentyl)-2,3,3-trimethyl-3*H*-indolium iodide in pyridine[21].

^{1}H NMR (300 MHz, Chloroform-*d*) δ = 8.61 (s, 1H), 8.06 (t, *J*=12.9, 2H), 7.40 – 7.30 (m, 4H), 7.25 – 7.05 (m, 4H), 6.77 (t, *J*=12.4, 1H), 6.27 (dd, *J*=17.9, 13.6, 2H), 4.02 (t, *J*=7.5, 2H), 3.66 (s, 3H), 2.55 – 2.26 (m, 2H), 1.87 – 1.63 (m, 4H), 1.73 (s, 6H), 1.72 (s, 6H), 1.61 – 1.43 (m, 2H).
^{13}C NMR (75 MHz, $CDCl_3$) δ = 173.29, 172.75, 153.40, 142.58, 141.81, 141.10, 140.86, 128.58, 128.53, 126.13, 125.06, 122.21, 122.08, 110.51, 110.43, 110.40, 103.87, 103.48, 49.32, 49.20, 44.12, 34.29, 32.11, 28.05, 27.92, 26.81, 26.17, 24.40.

Data is in accordance with literature values[21].

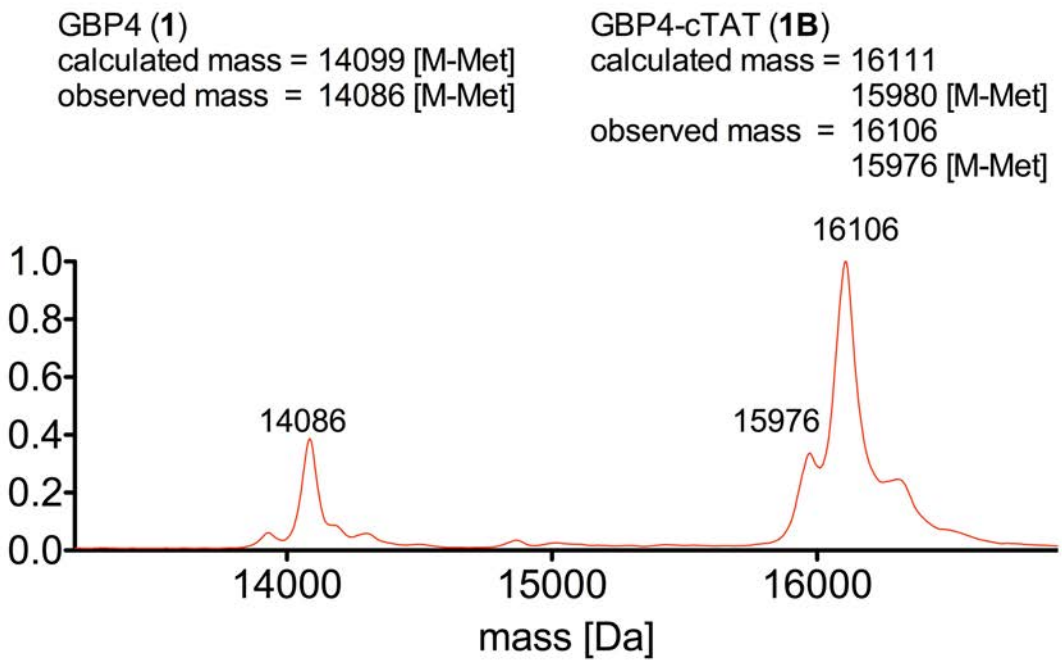
2.2.6 Expressed protein ligation

*2.2.6.1 GBP4-5,6-FAM **S4***

2 mL of partly refolded $GBP4_{1-96}$-intein-CBD (**S2**) in 2 M urea were loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 2 M urea, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, 2 M urea, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg $GBP4_{97-133}$-5,6-FAM (**S3**, 512 nmol, 0.512 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed with 6x 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, 2 M urea, pH 8.5). Product containing fractions were pooled and 2x dialyzed against 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 0.1 mM ox. glutathione, 1 mM red. glutathione for 3 h. The product was analyzed by SDS-PAGE (Supplementary Fig. 3).

*2.2.6.2 GBP4-cTAT **1B***

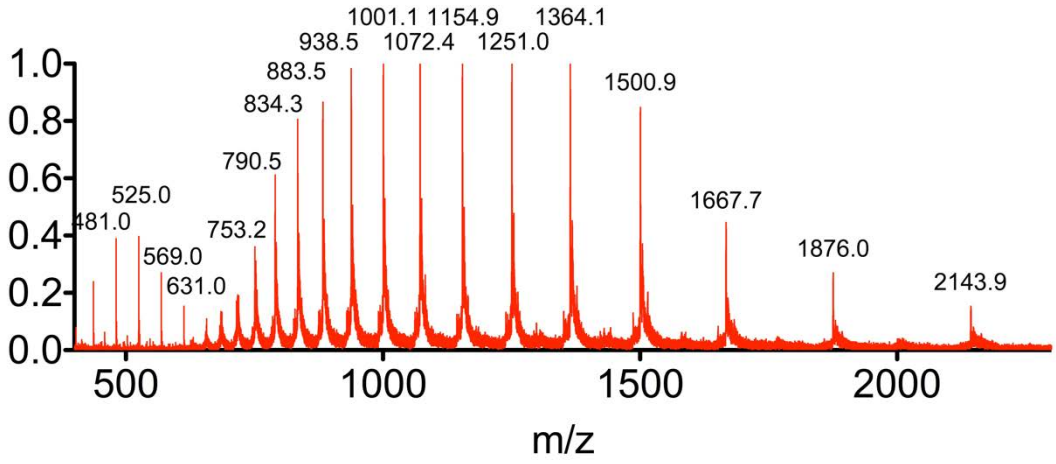
2 mL of the dialyzed lysate of **1A** was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg cTAT peptide **4** (512 nmol, 0.512 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed 6x with 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). Product containing fractions were pooled, incubated for 2 h at 4°C with BioBeads (0.2 g/mL, BioRad, USA) and the peptide excess removed by desalting columns (Zeba™ Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). The solution was concentrated (Amicon Ultra-0.5, 10000 CO, Merck KGaA, Germany) to 200 µL and rebuffered via dilution/spin (7x addition of 300 µL) to 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). For MS analysis, the reaction mixture was precipitated by acetone, resolubilized in ddH_2O and subjected to MALDI measurements. Spectra are shown below (Supplementary Fig. 30).



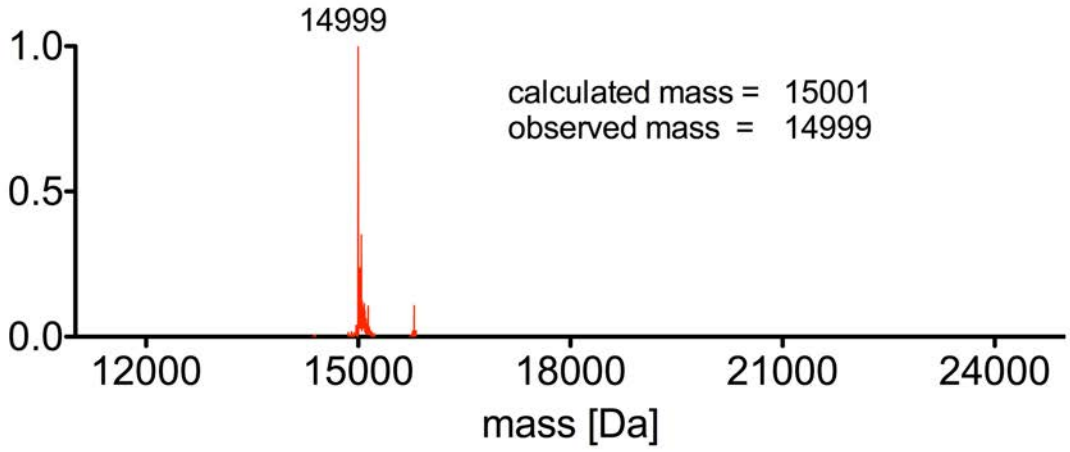
Supplementary Figure 30 | MALDI of the GBP4-cTAT reaction mixture after EPL.

*2.2.6.3 GBP1-cTAT **2B***

4 mL of the clear lysate of **2A** was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg cTAT-CPP **4** (512 nmol, 0.512 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed 6x with 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). Product containing fractions were pooled, incubated for 2 h at 4°C with BioBeads (0.2 g/mL, BioRad, USA) and the peptide excess removed by desalting columns (Zeba^{TM} Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). The solution was concentrated (Amicon Ultra-0.5, 10000 CO, Merck KGaA, Germany) to 200 µL and rebuffered via dilution/spin (7x addition of 300 µL) to 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). The product was analyzed by ESI-MS. ESI-MS are shown below (Supplementary Fig. 31 and 32).



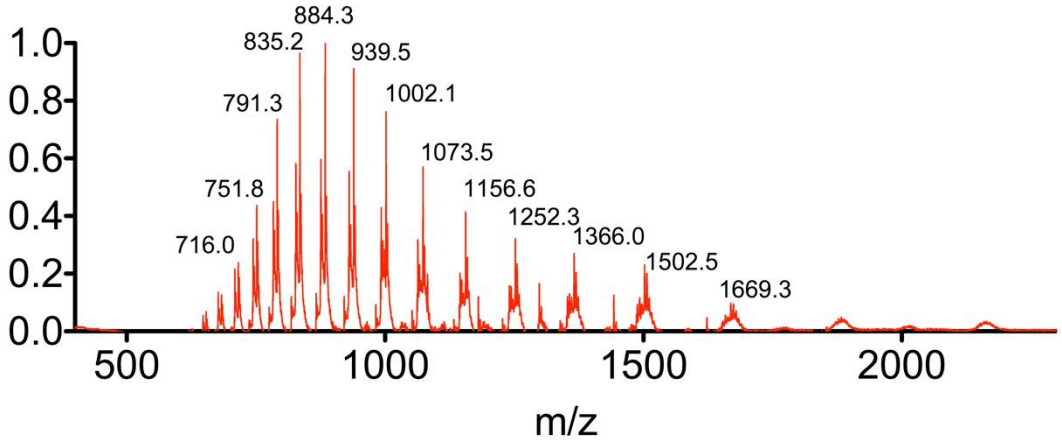
Supplementary Figure 31 | ESI-MS of GBP1-cTAT (**2B**).



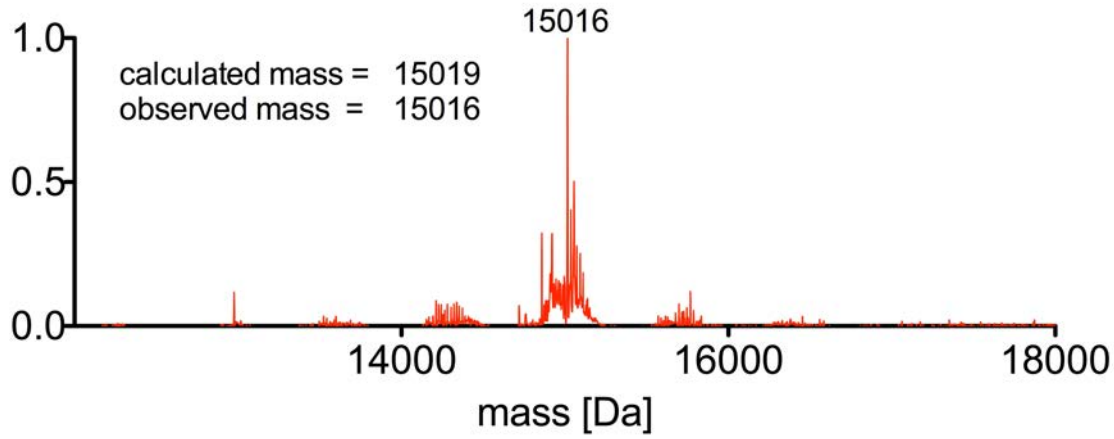
Supplementary Figure 32 | Deconvoluted mass of GBP1-cTAT (**2B**).

*2.2.6.4 GBP1-linTAT **2D***

4 mL of the clear lysate was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg linear TAT-CPP **6** (437 nmol, 0.43 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed 6x with 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). Product containing fractions were pooled, incubated for 2 h at 4°C with BioBeads (0.2 g/mL, BioRad, USA) and the peptide excess removed by desalting columns (Zeba^{TM} Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). The solution was concentrated (Amicon Ultra-0.5, 10000 CO, Merck KGaA, Germany) to 200 µL and rebuffered via dilution/spin (7x addition of 300 µL) to 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). The product was analyzed by ESI-MS. ESI-MS are shown below (Supplementary Fig. 33 and 34).



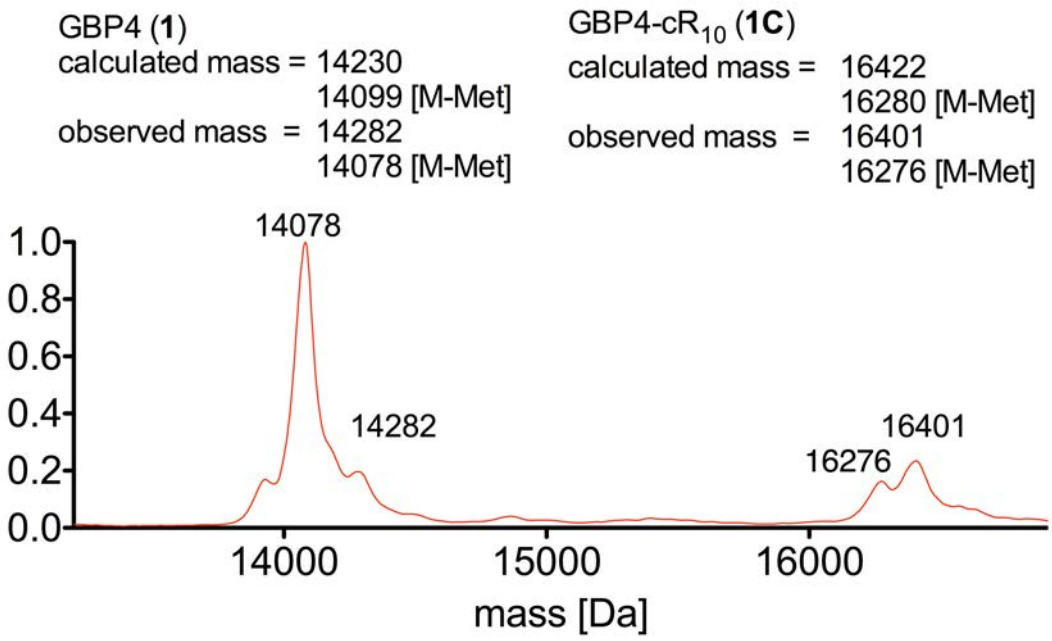
Supplementary Figure 33 | ESI-MS of GBP1-linTAT (**2D**).



Supplementary Figure 34 | Deconvoluted mass of GBP1-linTAT (**2D**).

2.2.6.5 GBP4-cR$_{10}$ ***1C***

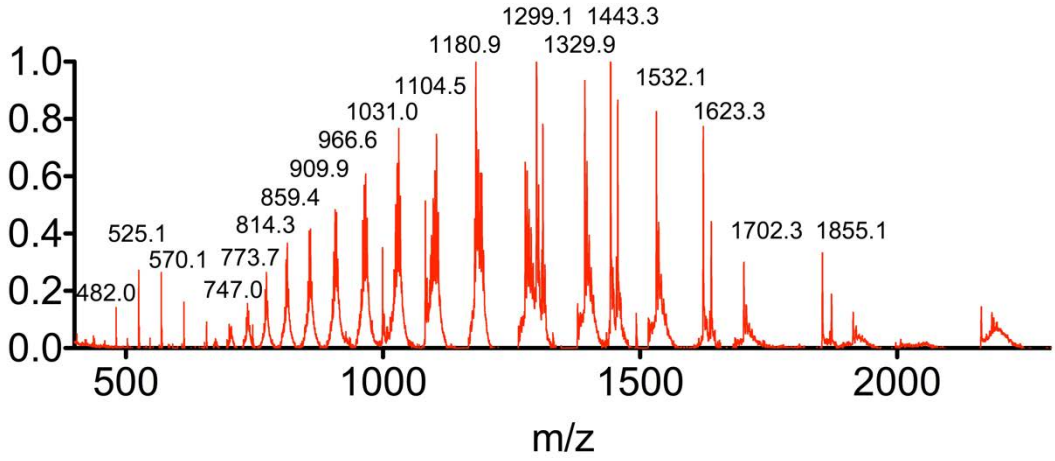
2 mL of the dialyzed lysate was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg cR$_{10}$-CPP **3** (437 nmol, 0.43 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed 6x with 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). Product containing fractions were pooled, incubated for 2 h at 4°C with BioBeads (0.2 g/mL, BioRad, USA) and the peptide excess removed by desalting columns (Zeba^{TM} Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). The solution was concentrated (Amicon Ultra-0.5, 10000 CO, Merck KGaA, Germany) to 200 µL and rebuffered via dilution/spin (7x addition of 300 µL) to 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). For MS analysis, the reaction mixture was precipitated by acetone, resolubilized in ddH_2O and subjected to MALDI measurements. Spectra are shown below (Supplementary Fig. 35).



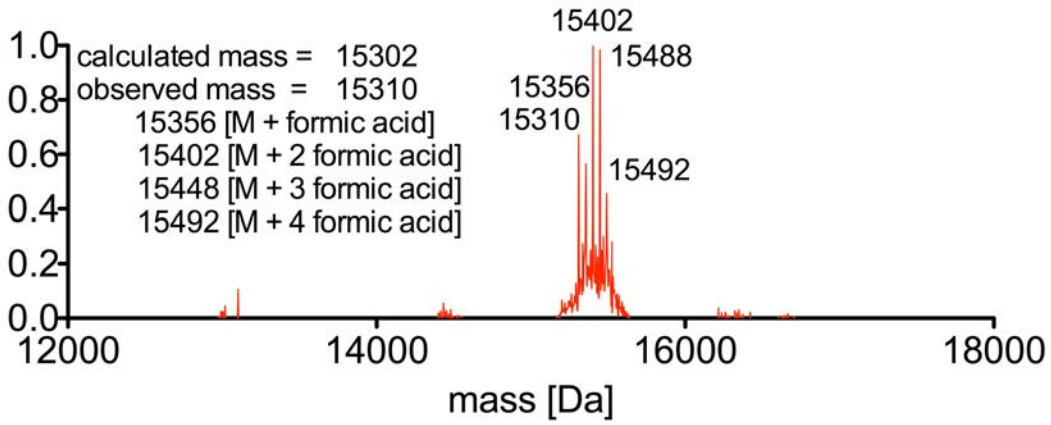
Supplementary Figure 35 | MALDI of the GBP4-cR_{10} (**1C**) reaction mixture after EPL. The starting material GBP4 has a higher tendency to ionize compared to the cR_{10} conjugate (see SDS-PAGE in Fig. 2e).

2.2.6.6 *GBP1-cR$_{10}$* ***2C***

4 mL of the clear lysate was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg cR$_{10}$-CPP **3** (437 nmol, 0.43 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed 6x with 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). Product containing fractions were pooled, incubated for 2 h at 4°C with BioBeads (0.2 g/mL, BioRad, USA) and the peptide excess removed by desalting columns (Zeba^{TM} Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). The solution was concentrated (Amicon Ultra-0.5, 10000 CO, Merck KGaA, Germany) to 200 µL and rebuffered via dilution/spin (7x addition of 300 µL) to 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). The product was analyzed by ESI-MS. Spectra are shown below (Supplementary Fig. 36 and 37).



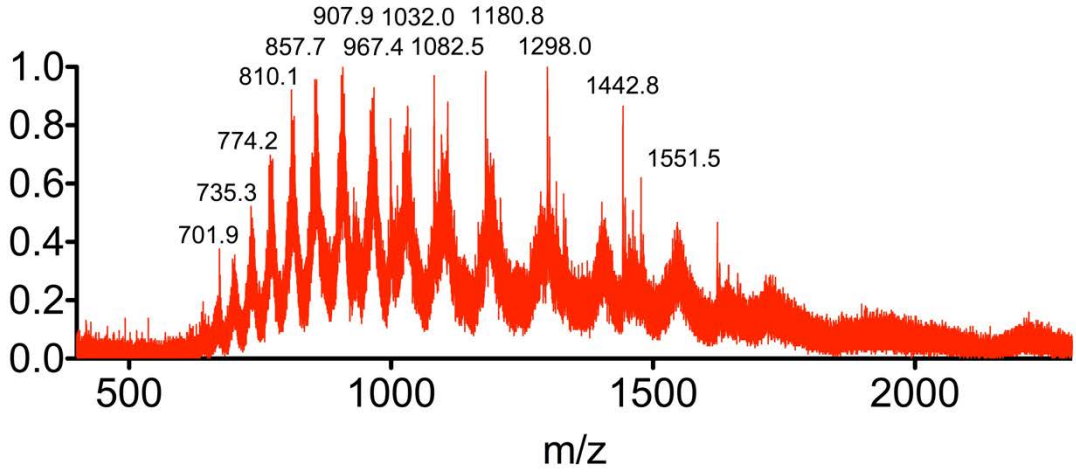
Supplementary Figure 36 | ESI-MS of GBP1-cR$_{10}$ (**2C**).



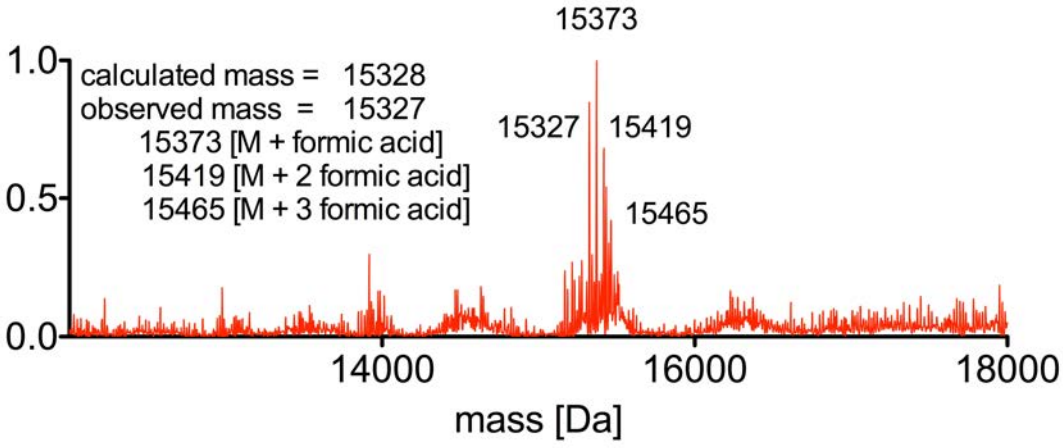
Supplementary Figure 37 | Deconvoluted mass of GBP1-cR$_{10}$ (**2C**).

2.2.6.7 *GBP1-linR$_{10}$* ***2E***

4 mL of the clear lysate was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg linear R_{10}-CPP **5** (437 nmol, 0.43 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed 6x with 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). Product containing fractions were pooled, incubated for 2 h at 4°C with BioBeads (0.2 g/mL, BioRad, USA) and the peptide excess removed by desalting columns (Zeba^{TM} Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). The solution was concentrated (Amicon Ultra-0.5, 10000 CO, Merck KGaA, Germany) to 200 µL and rebuffered via dilution/spin (7x addition of 300 µL) to 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). The product was analyzed by ESI-MS. Spectra are shown below (Supplementary Fig. 38 and 39).



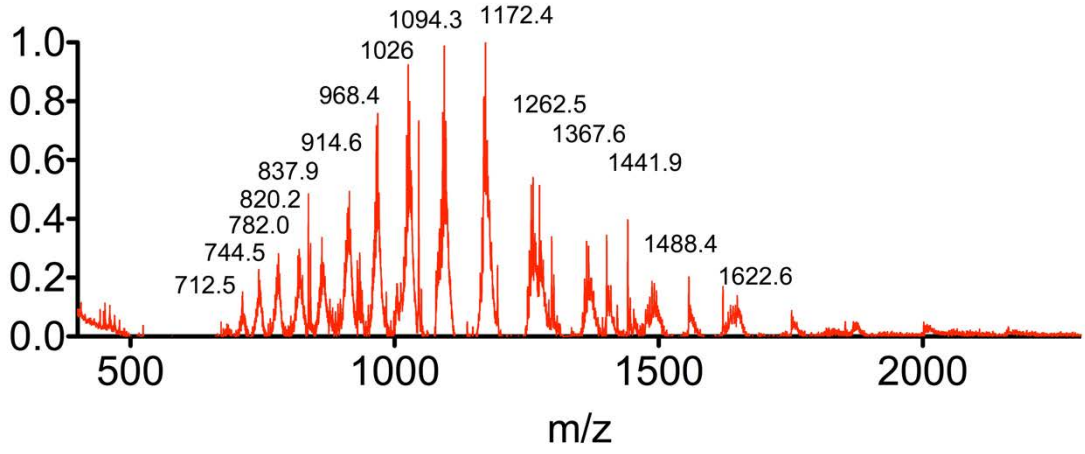
Supplementary Figure 38 | ESI-MS of GBP1-linR$_{10}$ (**2E**).



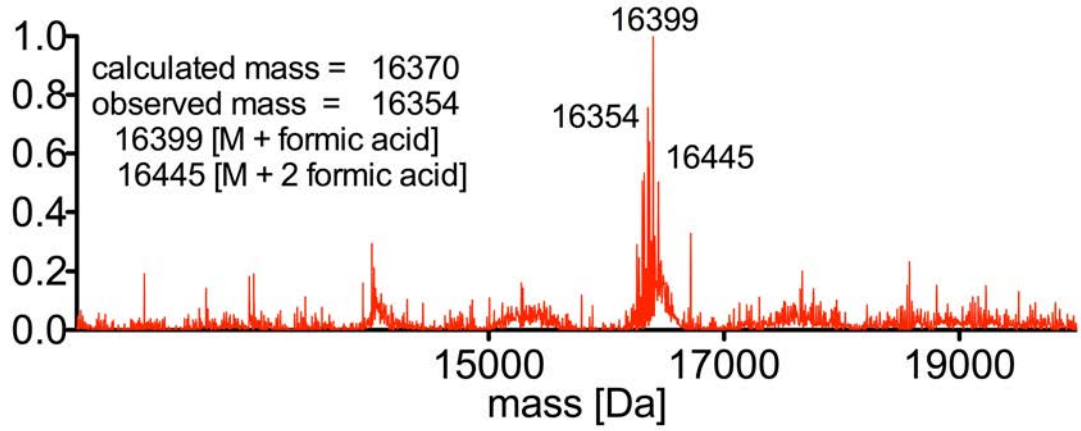
Supplementary Figure 39 | Deconvoluted mass of GBP1-linR$_{10}$ (**2E**).

2.2.6.8 GBP1-Cy5-S-S-cR$_{10}$ ***2G***

4 mL of the clear lysate was loaded on 1 mL pre-equilibrated (3x 4 mL equilibration buffer, 20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, pH 8.5) chitin beads via gravity-flow through (3x). The solid phase was washed 6x with 5 mL equilibration buffer, flushed with 1 mL cleavage buffer (20 mM Tris-HCl, 0.5 M NaCl, 0.1% TritonX100, 1 mM EDTA, 100 mM MESNA, pH 8.5) and incubated with 1 mL cleavage buffer including 1.5 mg Cy5-peptide **10** (1412 nmol, 1.41 mM) for 18 h at RT. The ligation mixture was eluted from the chitin beads and the resin washed 6x with 500 µL washing buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.5). GBP1-Cy5 (**2F**) containing fractions were pooled, incubated for 2 h at 4°C with BioBeads (0.2 g/mL, BioRad, USA) to remove the excess Cy5-peptide. **2F** in 200 µL 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 8.5, 1 mg/mL) was activated with 10 eq. Ellman's reagent for 1h at ambient temperature. Excess of Ellman's reagent was removed by dialysis (dialysis membrane Spectra/Por® - MWCO 10000) against 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 9.0) over night at 4°C. 1.5 mg of cR$_{10}$-CPP **3** (437 nmol, 0.43 mM) was added and incubated for 18 h at ambient temperature. The peptide excess removed by desalting columns (Zeba™ Spin, 7 kDa cutoff, 5 mL, Thermo Fisher Scientific Inc., USA). The solution was concentrated (Amicon Ultra-0.5, 10000 CO, Merck KGaA, Germany) to 200 µL and rebuffered via dilution/spin (7x addition of 300 µL) to 1x HEPES buffer (5 mM HEPES, 140 mM NaCl, 2.5 mM KCl, 5 mM glycine, pH 7.5). The product was analyzed by ESI-MS. Spectra are shown below (Supplementary Fig. 40 and 41).



Supplementary Figure 40 | ESI-MS of GBP1-Cy5-S-S-cR$_{10}$ (**2G**).



Supplementary Figure 41 | Deconvoluted mass of GBP1-Cy5-S-S-cR$_{10}$ (**2G**).

2.2.7 Mammalian cell culture

Human HeLa cells[7], U2OS 2-6-3 human osteosarcoma cells and mouse NIH Flp-In 3T3 fibroblast cells (see Supplementary Table 2) were cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal calf serum, 50 µg/mL gentamicin and 2 mM glutamine. Cells expressing GFP-tagged proteins were supplemented once a week with 2.5 µg/mL blasticidin for selection. Both cell lines were grown at 37° C in a humidified atmosphere with 5% CO_2. Media components were purchased from Sigma-Aldrich, Germany.

2.2.8 Development of cell lines stably expressing nuclear GFP (and GFP-tagged Mecp2)

2.2.8.1 Expression vectors:

For the generation of the expression vectors (see Supplementary Table 3) pFRT-B-NGFP and pFRT-B-Mecp2G, pEF5/FRT/V5-D-TOPO (Invitrogen, USA) was used as backbone plasmid. This plasmid contains a human EF-1alpha promoter to drive expression of the gene of interest and includes a FRT (Flp-recombinase target) site and an ATG-less gene coding for hygromycin resistance for selection in mammalian cells. The hygromycin gene was excised by cutting with BstZ17I and StyI restriction endonucleases and ligated to a similarly cut PCR amplified fragment containing an ATG-less gene coding for blasticidin resistance, obtained from the pUB-Bsd vector (Invitrogen, USA) with the following primers:

5' tag aat cct tgg aat tca agg cca agc ctt tgt ctc a 3'

5' tac atc gta tac gga att cag aca tga taa gat aca ttg 3'

In subsequent steps, NLS-GFP, as well as rat Mecp2-GFP[14] cDNAs were cloned downstream of the EF-1alpha promoter. The resulting plasmid constructs contain the EF-1alpha promoter driving expression of the NLS-GFP or the Mecp2-GFP cDNAs and the FRT site followed by an ATG-less blasticidin resistance gene.

2.2.8.2 Host cell lines

Mouse (NIH Swiss) embryonic fibroblasts (Flp-In^{Tm}-3T3) were purchased from Invitrogen (Carlsbad, USA). These cells contain a FRT recombination site stably inserted into the genome. The construct comprises a SV40 promoter driving expression of an open reading frame composed of an ATG start codon followed by a FRT site and a *lacZeo* fusion gene.

2.2.8.3 Transfection and selection of stably expressing cell lines

Flp-In^{Tm}-3T3 cells (described above) were co-transfected with a plasmid coding for the Flp recombinase (pOG44; Invitrogen, California, USA) and pFRT-B-NGFP or pFRT-B-Mecp2G, respectively using TransFectin^{Tm} lipid reagent (Bio-Rad, USA) according to the manufacturer's instructions. Upon co-transfection, the Flp recombinase expressed from the pOG44 plasmid catalyzes the homologous recombination of the FRT containing plasmids into the genomically integrated FRT site in the host cells. This results in the insertion of the ATG-less blasticidin resistance gene in frame with the ATG initiation codon in the host cells' genome leading to expression of the functional blasticidin resistance gene and concomitant integration of the EF-1alpha promoter followed by the NLS-GFP or the Mecp2-GFP cDNAs (see also Chagin et al., 2016)[7]. This leads to disruption of the expression of the *lacZeo* fusion gene. Positive clones of cells were selected by adding 2.5 µg/mL blasticidin S (InvivoGen; California, USA) to the culture medium.

Stable expression of the nuclear GFP, as well as the GFP-tagged Mecp2 was evaluated using fluorescence microscopy and found to be homogenous through the cell population. In addition, the subcellular distribution of the proteins was validated by fluorescence microscopy. The nuclear GFP fluorescence was correlated with phase contrast images of the cells. Moreover, co-staining the Mecp2-GFP expressing cells with the DNA dye DAPI confirmed the enrichment of Mecp2 at nuclear sites corresponding to pericentric heterochromatin[14].

2.2.9 GFP fluorescence *in vitro* assay

0.5 µM GFP in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) was incubated with varying concentrations [0.00 – 0.80 µM) of nanobody and change in fluorescence emission detected. GFP fluorescence spectra were recorded with a Jasco FP-6500 spectrometer (Jasco Research Ltd., Canada) using an excitation wavelength of 488 nm.

2.2.10 Circular dichroism

CD spectra were recorded with a Jasco J-720 spectrometer (Jasco Research Ltd., Canada) using a 0.1-cm-pathlength cuvette. Sample concentrations were typically 10 µM protein in 20 mM sodium phosphate buffer, 150 mM NaF (pH 7.2). All spectra were recorded between 250 and 200 nm at 25 °C and 8 scans averaged without smoothing. The spectrum of the buffer was subtracted from all protein samples.

2.2.11 Ellman's test for the detection of free cysteines

Free cysteines within recombinantly expressed and EPL derived nanobodies were determined using Ellman's reagent[22]. For this, 50 µL DTNB (Dithionitrobenzoic acid) stock solution (50 mM sodium acetate and 2 mM DTNB) were added to 840 µL of molecular grade biology water and 100 µL of 1M Tris-HCl pH 8.0. 1 µL of protein sample was added to 99 µL to DTNB reagent, incubated at ambient temperature for 5 minutes and optical absorbance measured at 412 nm using a TECAN Safire multiplate reader (Tecan Group, Switzerland). The molarity of the free sulfhydryl groups was determined by dividing the absorbance by 13600 $M^{-1}cm^{-1}$.

2.2.12 Live cell confocal microscopy

Confocal images were collected using an Ultra VIEW VoX spinning disc system (Perkin Elmer) on a Nikon Ti microscope equipped with an oil immersion Plan Apochromat VC ×60/1.45 NA (pixel size in XY = 111 nm, Z-step = 0.3 – 1 µm), with laser lines at 488 nm, 561 nm, a differential interference contrast (DIC) setup, and a temperature, humidity and CO_2 incubation control (ACU control, Olympus).

2.2.13 Nucleotide incorporation after cell-permeable antibody delivery

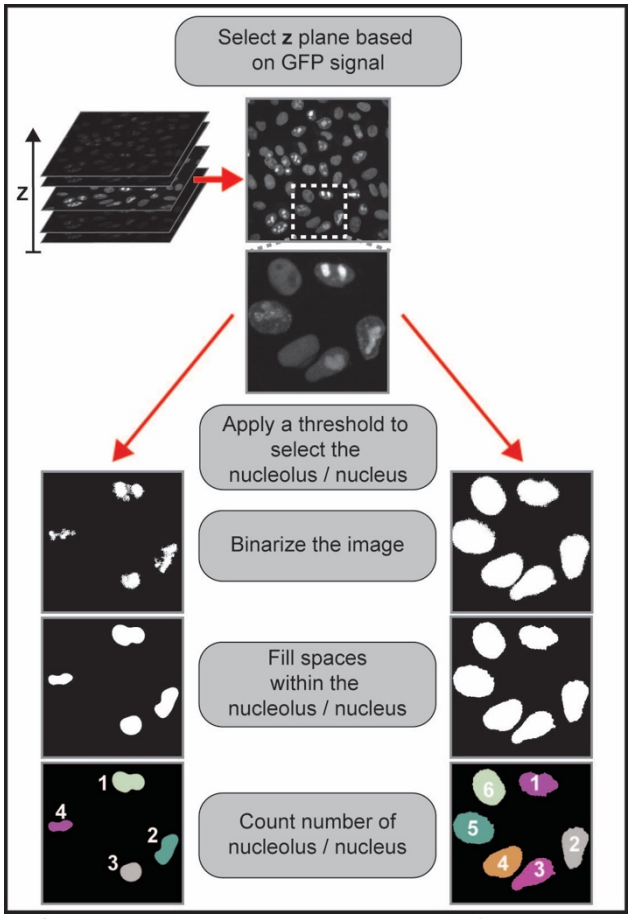
To analyze replication patterns after antibody delivery, we made use of an established nucleotide analog pulse labeling and immunostaining for the incorporated nucleotide analog[7] with the modifications described below. Cells were pulse labeled with 5 µM BrdU for 20 min at 37 ˚C. Cells were fixed for immunofluorescence in 4% formaldehyde for 10 min at RT and permeabilized with 0.5% Triton X-100 for 15 min at RT.
After blocking in 4% BSA/PBS for 1 h at RT, incorporated BrdU was recognized by rabbit anti-BrdU antibody (1:500, Biomol GmbH, Hamburg, Germany) in conjunction with 0.05 U/µL DNaseI (Sigma-Aldrich, St. Louis, MO, USA) for 1 h at 37˚C in 2% BSA/100 mM sodium acetate/5 mM magnesium sulfate. Cells were then washed in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) containing 0.01% Tween and 1 mM EDTA to stop DNaseI digestion. Following incubation with the secondary Alexa Fluor 594 conjugated goat anti rabbit IgG antibody (1:500, The Jackson Laboratory, Bar Harbor, USA) in 4% BSA for 1 h at RT, cells were washed 3x for 5 min and 1x for 30 min in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) containing 0.01% Tween. DNA was counterstained with 1 µg/mL DAPI for 10 min at RT, before cells were washed 1x in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) containing 0.01% Tween, 1x in PBS (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7 mM KCl and 137 mM NaCl, pH 7.4) and 1x in ddH_2O and mounted in Mowiol 4-88 embedding medium (Sigma-Aldrich, St. Louis, MO, USA).

2.2.14 Nucleolar labeling

To confirm the relocalization of GFP to the nucleolus driven by the cellular uptake of GBP1-cR_{10}, we performed colocalization[23] studies with the TAT-TAMRA peptide (previously proposed as a nucleolar marker for living cells[4, 18]) and in fixed cells using an antibody for nucleophosmin/B23, a protein present in the granular component of the nucleolus and fibrillarin localized in the dense fibrillar component of the nucleolus.

2.2.15 Automatic computerized cellular uptake quantification

The cell-permeable nanobody nucleolar relocalization assay allows the implementation of methods to automatically quantify cellular uptake. The specific implementation of a computerized method would naturally depend on multiple variables, in particular the expression levels and distribution of GFP, or GFP-fusion proteins, within the cells. To illustrate this approach, we developed a simple code to automatically quantify the number of cells displaying uptake of the cell-permeable nanobody and apply it in the quantifications shown in Movie 1 and Supplementary Fig. 13. In Supplementary Fig. 42 is shown the flow code used for the automatic segmentation and counting of the number of cells that display a relocalization of GFP tagged PCNA to the nucleolus. First, based on the GFP fluorescence signal a plane was selected from a z-stack of images based on the maximum GFP fluorescent signal. Next, to segment and count the number of cells and cells with signal relocalized to the nucleolus we applied the Otsu's method[24] to automatically determine a threshold necessary to subtract the background signal. This threshold was applied and the image was binarized. The disconnected pixels or areas within the nucleus were joined by filling the spaces by applying a morphological dilation of the image using a disk matrix with a radius of one pixel followed by a morphological erosion using a disk matrix of 4 pixels. Next, the connected regions within the final binary image were automatically counted (as matrix blocs) by performing a connected components analysis. To count the cells displaying nucleolar accumulation, a morphological binarization was applied in which a threshold 25 % higher than the Otsu's threshold used to segment the nuclei was used, followed by the same steps to connect the nucleolar fractions within the individual nucleus and count the number of cell displaying GFP nucleolar relocalization. This processes was automatically done for each time point in the relocalization computation done for the time-lapse relocalization assays shown in Supplementary Fig. 13b and Movie 1.



Supplementary Figure 42 | Flow code and steps implemented for the automatic segmentation of the nucleolus and nucleus.

2.2.16 Intracellular nanobody quantification

To quantify the intracellular concentration of the cell-permeable nanobody we used the fluorescently labeled cell-permeable nanobody GBP1-ss-cR$_{10}$ (**2G**). As shown in Supplementary Fig. 19a, we quantified the average fluorescence intensity of 5 µM of the cell-permeable nanobody in an empty well (ROI-Ex). Then the cells where incubated in the presence of 5 µM of the cell-permeable nanobody for 1 h, washed and imaged. The background (ROI-Out) and the average intracellular fluorescence intensity in individual cells (ROI-In) were determined. These values were plugged into the equation shown in Supplementary Fig. 19a to estimate the intracellular concentration based on the intracellular fluorescence signal.

2.2.17 Protein-protein interactions - Fluorescence three Hybrid method (F3H)

Protein-protein interactions were studied using PCNA fluorescent fusions and p53/Hdm2 fluorescent fusions (Supplementary Table 3).

To visualize the oligomerization of PCNA using the F3H, HeLa Kyoto cells stably expressing mCherry-PCNA[7] were used and transiently transfected with GFP tagged PCNA as described before[25].

U2OS cell lines were transiently transfected using PEI (poly-ethyleneimine; 1 mg/ml in ddH_2O, neutralized with HCl). For the transfection, 200 µl DMEM serum-free with 12 µl of PEI was vortexed for 2 min and added to 200 µl serum-free DMEM with 4 µg DNA. After incubating the DNA-PEI mixture for 15 min at room temperature, the

solution was added to the cells drop wise and incubated overnight at 37 °C and 5% CO_2.

Next, the cells were incubated for 1 h in the presence of the cell-permeable nanobody GBP1-cR_{10} (**2C**), washed and imaged.

2.2.18 Cell-permeable nanobody interaction with its antigen using microscale thermophoresis

To validate that the cell-permeable nanobody GBP1-cR_{10} (**2C**) retains a strong GFP binding affinity by microscale thermophoresis we used the Monolith NT.115 (Nano Temper Technologies GmbH) as described previously for GBP1[3]. The GBP-cR_{10} was titrated from 500 nM (a concentration higher than the expected dissociation constant K_d = 0.59 nM)[1] down to sub stoichiometric concentrations with respect to GFP (kept at 50 nM). As control similar measurements were done in the absence of GBP-cR_{10}. The samples were loaded in 100 µm diameter glass capillaries. An infrared laser was used to create spatial temperature increase of 25 µm diameter. This rise in temperature induced a spatial concentration distribution change that was visualized using the GFP fluorescence signal. The thermophoresis of the proteins was measured by measuring the change in concentration between the initial state and the steady state. The fluorescence is monitored continuously before and after laser-induced heating, showing the homogeneous distribution of GFP, and the changes induced after infrared-laser heating. At low nanobody concentrations, the profile in Supplementary Fig. 10 reflects the thermophoresis of the unbound state, whereas at high concentrations it reflects the bound state of the cell-permeable nanobody (GBP1-cR_{10}) and its antigen (GFP). The fraction bound (f) is a linear superposition of the bound and unbound states $F_{norm} = f \cdot F_{bound} + (1-f) \cdot F_{unbound}$. From this binding curve, the dissociation constant K_d was calculated.

2.3 Protein sequences

2.3.1 GBP4 (**1**)

MADVQLQESGGGSVQAGGSLRLSCAASGDTFSSYSMAWFRQAPGKECELVSNIL
RDGTTTYAGSVKGRFTISRDDAKNTVYLQMVNLKSEDTARYYCAADSGTQLGYVA
VGLSCLDYVMDYWGKGTQVTVSSA

2.3.2 GBP4-int-CBD (**1A**)

MADVQLQESGGGSVQAGGSLRLSCAASGDTFSSYSMAWFRQAPGKECELVSNIL
RDGTTTYAGSVKGRFTISRDDAKNTVYLQMVNLKSEDTARYYCAADSGTQLGYVG
AVGLSCLDYVMDYWGKGTQVTVSSACITGDALVALPEGESVRIADIVPGARPNSDN
AIDLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVP
TLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHH
RDPDAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATG
LTGLNSGLTTNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPA
LWQLQ

2.3.3 GBP1 (**2**)

MADVQLVESGGALVQPGGSLRLSCAASGFPVNRYSMRWYRQAPGKEREWVAGM
SSAGDRSSYEDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYYCNVNVGFEYWG
QGTQVTVSSAAA

2.3.4 GBP1-int-CBD (**2A**)

MADVQLVESGGALVQPGGSLRLSCAASGFPVNRYSMRWYRQAPGKEREWVAGM
SSAGDRSSYEDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYYCNVNVGFEYWG
QGTQVTVSSAAACITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPV
LADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDY
AVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTD
GRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGLTTNPGV
SAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ

3 References

1. Kirchhofer A, Helma J, Schmidthals K, Frauer C, Cui S, Karcher A*, et al.* Modulation of protein properties in living cells using nanobodies. *Nat. Struct. Mol. Biol.* 2010, **17**(1)**:** 133-138.

2. Schumacher D, Helma J, Mann FA, Pichler G, Natale F, Krause E*, et al.* Versatile and Efficient Site-Specific Protein Functionalization by Tubulin Tyrosine Ligase. *Angew. Chem. Int. Ed.* 2015, **54**(46)**:** 13787-13791.

3. Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* 2010, **1:** 100.

4. Martin RM, Tunnemann G, Leonhardt H, Cardoso MC. Nucleolar marker for living cells. *Histochem. Cell Biol.* 2007, **127**(3)**:** 243-251.

5. Shaw PJ, Jordan EG. The Nucleolus. *Annu. Rev. Cell. Dev. Biol.* 1995, **11:** 93-121.

6. Landry JJ, Pyl PT, Rausch T, Zichner T, Tekkedil MM, Stutz AM*, et al.* The genomic and transcriptomic landscape of a HeLa cell line. *G3 (Bethesda)* 2013, **3**(8)**:** 1213-1224.

7. Chagin VO, Casas-Delucchi CS, Reinhart M, Schermelleh L, Markaki Y, Maiser A*, et al.* 4D Visualization of replication foci in mammalian cells corresponding to individual replicons. *Nat. Commun.* 2016, **7:** 11231.

8. Janicki SM, Tsukamoto T, Salghetti SE, Tansey WP, Sachidanandam R, Prasanth KV*, et al.* From silencing to gene expression: real-time analysis in single cells. *Cell* 2004, **116**(5)**:** 683-698.

9. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 1986, **189**(1)**:** 113-130.

10. Jost KL, Rottach A, Milden M, Bertulat B, Becker A, Wolf P*, et al.* Generation and characterization of rat and mouse monoclonal antibodies specific for MeCP2 and their use in X-inactivation studies. *PLoS One* 2011, **6**(11)**:** e26499.

11. Leonhardt H, Rahn HP, Weinzierl P, Sporbert A, Cremer T, Zink D*, et al.* Dynamics of DNA replication factories in living cells. *J. Cell Biol.* 2000, **149**(2)**:** 271-280.

12. Herce HD, Deng W, Helma J, Leonhardt H, Cardoso MC. Visualization and targeted disruption of protein interactions in living cells. *Nat. Commun.* 2013, **4:** 2660.

13. Chan WC. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*. Oxford University Press, 2000.

14. Brero A, Easwaran HP, Nowak D, Grunewald I, Cremer T, Leonhardt H*, et al.* Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *J. Cell. Biol.* 2005, **169**(5)**:** 733-743.

15. Lindhout BI, Fransz P, Tessadori F, Meckel T, Hooykaas PJ, van der Zaal BJ. Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucleic Acids Res.* 2007, **35**(16)**:** e107.

16. Becker A, Allmann L, Hofstatter M, Casa V, Weber P, Lehmkuhl A*, et al.* Direct homo- and hetero-interactions of MeCP2 and MBD2. *PLoS One* 2013, **8**(1)**:** e53730.

17. Nagahara H, Vocero-Akbani AM, Snyder EL, Ho A, Latham DG, Lissy NA*, et al.* Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27Kip1 induces cell migration. *Nat. Med.* 1998, **4**(12)**:** 1449-1452.

18. Martin RM, Ter-Avetisyan G, Herce HD, Ludwig AK, Lattig-Tunnemann G, Cardoso MC. Principles of protein targeting to the nucleolus. *Nucleus* 2015, **6**(4)**:** 314-325.

19. Owens EA, Bruschi N, Tawney JG, Henary M. A microwave-assisted and environmentally benign approach to the synthesis of near-infrared fluorescent pentamethine cyanine dyes. *Dyes Pigm.* 2015, **113:** 27-37.

20. Kiyose K, Hanaoka K, Oushiki D, Nakamura T, Kajimura M, Suematsu M*, et al.* Hypoxia-sensitive fluorescent probes for in vivo real-time fluorescence imaging of acute ischemia. *J. Am. Chem. Soc.* 2010, **132**(45)**:** 15846-15848.

21. Simmons RL, Yu RT, Myers AG. Storable arylpalladium(II) reagents for alkene labeling in aqueous media. *J. Am. Chem. Soc.* 2011, **133**(40)**:** 15870-15873.

22. Ellman GL. Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* 1959, **82**(1)**:** 70-77.

23. Herce HD, Casas-Delucchi CS, Cardoso MC. New image colocalization coefficient for fluorescence microscopy to quantify (bio-)molecular interactions. *J. Microsc.* 2013, **249**(3)**:** 184-194.

24. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans. Sys., Man., Cyber.* 1979, **9**(1)**:** 62-66.

25. Herce HD, Rajan M, Lattig-Tunnemann G, Fillies M, Cardoso MC. A novel cell permeable DNA replication and repair marker. *Nucleus* 2014, **5**(6)**:** 590-600.