Direct Visualization of Single Nuclear Pore Complex Proteins Using Genetically-Encoded Probes for DNA-PAINT

Thomas Schlichthaerle+, Maximilian T. Strauss+, Florian Schueder+, Alexander Auer,
Bianca Nijmeijer, Moritz Kueblbeck, Vilma Jimenez Sabinina, Jervis V. Thevathasan, Jonas Ries,
Jan Ellenberg, and Ralf Jungmann*

Abstract: The nuclear pore complex (NPC) is one of the largest and most complex protein assemblies in the cell and, among other functions, serves as the gatekeeper of nucleocy-toplasmic transport. Unraveling its molecular architecture and functioning has been an active research topic for decades with recent cryogenic electron microscopy and super-resolution studies advancing our understanding of the architecture of the NPC complex. However, the specific and direct visualization of single copies of NPC proteins is thus far elusive. Herein, we combine genetically-encoded self-labeling enzymes such as SNAP-tag and HaloTag with DNA-PAINT microscopy. We resolve single copies of nucleoporins in the human Y-complex in three dimensions with a precision of circa 3 nm, enabling studies of multicomponent complexes on the level of single proteins in cells using optical fluorescence microscopy.

Super-resolution techniques allow diffraction-unlimited fluorescence imaging[1] and with recent advancements, true biomolecular resolution is well within reach.[2] One implementation of single-molecule localization microscopy (SMLM) is called DNA points accumulation in nanoscale topography[2b] (DNA-PAINT), where dye-labeled DNA strands (called "imager" strands) transiently bind to their complements (called "docking" strands) on a target of interest, thus creating the typical "blinking" used in SMLM to achieve super-resolution. While localization precisions down to approximately one nanometer (basically the size of a single dye molecule) are now routinely achievable from a technology perspective, this respectable spatial resolution has yet to be translated to cell biological research. Currently, this is mainly hampered by the lack of small and efficient protein labels. Recent developments of nanobody- or apta-mer-based tagging approaches[3] are providing an attractive route ahead, however both approaches are not yet deploying their full potential either due to limited binder availability (in the case of nanobodies) or lack of large-scale analysis of suitable super-resolution probes (in the aptamer case).

While we are convinced that some of these issues might be resolved in the future, we introduce herein the combination of widely-used, genetically-encoded self-labeling enzymes such as SNAP-tag[4] and HaloTag[5] with DNA-PAINT to enable 1:1 labeling of single proteins in the nuclear pore complex (NPC) using ligand-conjugated DNA-PAINT docking strands. The NPC is responsible for the control of nucleocy-toplasmic transport and a highly complex and sophisticated protein assembly. NPCs contain multiple copies of approx-imately 30 different nuclear pore proteins called nucleoporins (NUPs) and have an estimated total molecular mass of about 120 MDa, placing NPCs among the largest cellular protein complexes.[6] Owing to their diverse function in controlling molecular transport between the nucleus and the cytoplasm, NPCs are a major target for structural biology research with characterization by for example, cryogenic electron micros-copy[6] (cryo-EM) or optical super-resolution techniques.[7] State-of-the-art cryo-EM studies,[8] reaching impressive pseudo-atomic resolution, have advanced our structural understanding in recent years. It is now possible to not only elucidate how NUPs in NPCs are arranged, but also to shed light on how structural changes of NPCs are connected to their dysfunction.[9] However, even with recent advancements in cryo-EM instrumentation, molecular specificity necessary to resolve single NUPs in NPCs proves still elusive, mainly due to the lack of high protein-specific contrast. Fluores-cence-based techniques on the other hand offer exquisite molecular contrast and specificity owing to the use of dye-labeled affinity reagents targeting single protein copies in cells. However, until recently, the necessary resolution to spatially resolve single small proteins in a larger complex has not been achieved because of limitations in labeling (small and efficient probes) and imaging technology (providing sub-10-nm spatial resolution). In order to spatially resolve sub-10-nm distances using SMLM, one needs to obtain a localization precision of circa 4 nm. This is readily achievable with DNA-PAINT, as a comparably large number of photons is available for localization per single binding—that is, blinking—event.

[*] T. Schlichthaerle,[+] Dr. M. T. Strauss,[+] F. Schueder,[+] A. Auer,
Prof. Dr. R. Jungmann
Faculty of Physics and Center for Nanoscience, LMU Munich
Geschwister-Scholl-Platz 1, 80539 Munich (Germany),
and
Max Planck Institute of Biochemistry
Am Klopferspitz 18, 82152 Martinsried (Germany)
E-mail: jungmann@biochem.mpg.de

B. Nijmeijer, M. Kueblbeck, Dr. V. Jimenez Sabinina,
Dr. J. V. Thevathasan, Dr. J. Ries, Dr. J. Ellenberg
Cell Biology and Biophysics Unit, European Molecular Biology
Laboratory (EMBL)
Meyerhofstraße 1, 69117 Heidelberg (Germany)

[+] These authors contributed equally.

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
https://doi.org/10.1002/anie.201905685.

While one can easily reach tens of thousands of photons per blinking event with DNA-PAINT, this is hard to achieve using STORM. Furthermore, the intrinsic resistance of DNA-PAINT to photobleaching enables repetitive localizations with high precision, while in the STORM case, the available photon budget is limited by a few fixed, target-bound fluorophores. Herein, we thus combine DNA-PAINT microscopy with small, genetically-encoded self-labeling enzymes such as SNAP- and HaloTag to overcome limitations in optical super-resolution microscopy.

We present a straightforward protocol to target these tags in a variety of engineered cell lines using the DNA-conjugated ligands benzylguanine (BG) and chloroalkane against SNAP-tag[4] and HaloTag,[5] respectively (Figure 1a and b). We investigate the achievable labeling precision and reduction of linkage error of SNAP-tag and HaloTag, examine their performance in contrast to DNA-conjugated nanobodies against GFP-tagged proteins in single NPCs and further compare them to primary and secondary antibody labeling. Finally, we resolve, for the first time, single copies of NUP96 proteins in the Y-complex of the NPC, spaced only circa 12 nm apart.

To implement genetically-encoded self-labeling tags for DNA-PAINT, we first assayed our ability to use BG-modified docking strands to target SNAP-tags C-terminally fused to NUP96 proteins in U2OS cell lines created by CRISPR/Cas9

engineering.[10] Labeling was performed post-fixation and -permeabilization using standard labeling protocols[7b] adapted for DNA-PAINT (see Online Methods). The resulting 2D DNA-PAINT image is shown in Figure 1c. A zoom-in reveals the expected 8-fold symmetry of NUP96 proteins in the super-resolution micrograph (Figure 1d). We then performed labeling of NUP107-GFP fusion proteins using a DNA-conjugated anti-GFP nanobody[11] and obtained qualitatively similar results (Figure 1e–f, see also Supplementary Figure 1 in the Supporting Information for zoom-outs and comparison to diffraction-limited data).

To evaluate labeling quality and precision in a quantitative manner, we next compared results of more traditional labeling of NUP107-GFP using primary-secondary antibodies to those of NUP96-SNAP, NUP96-Halo, NUP107-SNAP, and NUP107-GFP cell lines targeted with their respective small ligands. The NPC architecture presents a well-suited model to benchmark novel labeling approaches with regards to overall labeling efficiency and limits of spatial resolution, in a sense similar to an in vitro DNA origami calibration standard,[12] but inside a cell. Previous EM studies revealed that NUP96 and NUP107 proteins are present in the Y-complex, which forms the cytoplasmic as well as nuclear NPC double ring arrangement in an 8-fold symmetry. The two double rings are spaced approximately 50 nm apart with each side containing 16 protein copies.[8a,13] The two copies of the proteins in each
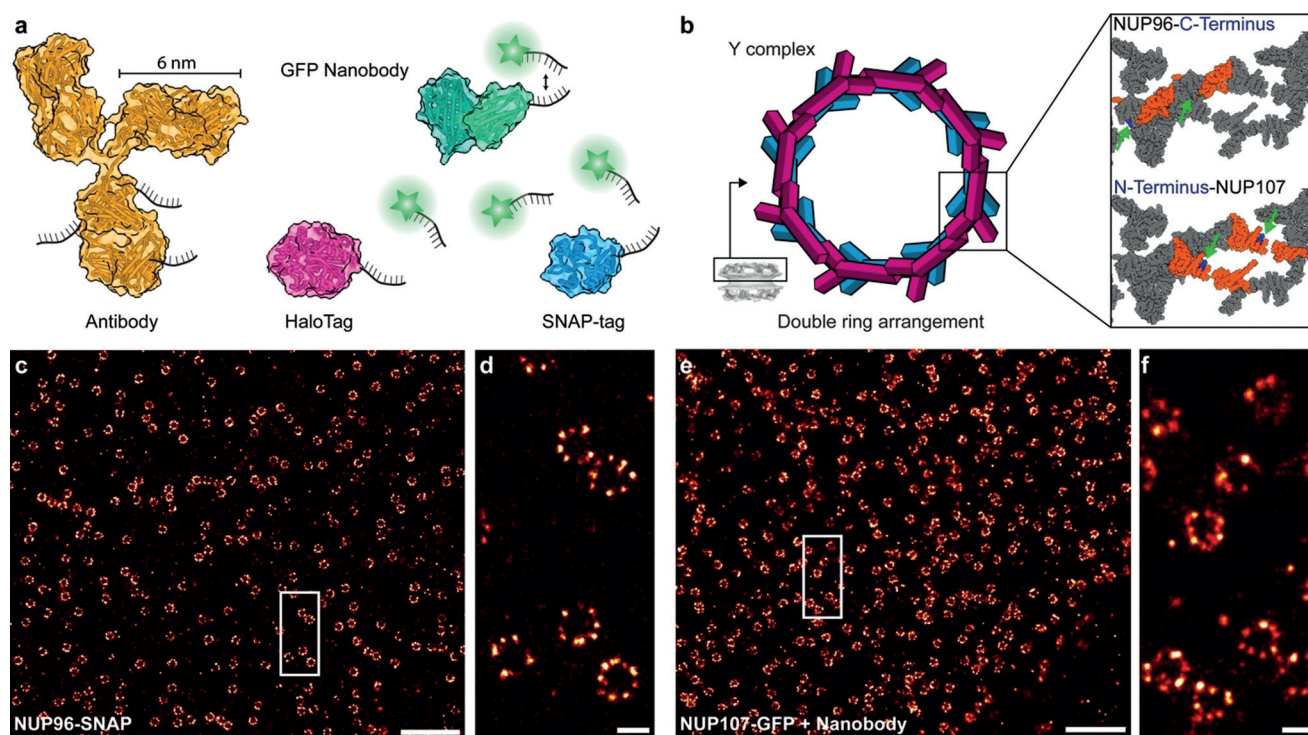

Figure 1. a) Comparison of different labeling probes (secondary antibody: yellow, GFP nanobody: green, HaloTag: magenta, SNAP-tag: blue) conjugated with DNA strands for DNA-PAINT imaging (cartoons are based on protein database (PDB) entries: Secondary antibody (1IGT), GFP nanobody (3K1K), HaloTag (4KAF), SNAP-tag (3KZZ)). Proteins are to scale. b) NPCs contain 16 copies of NUP96 and NUP107 in the cytoplasmic as well as the nuclear ring. Top right: C-terminally-labeled (blue, highlighted by green arrows) NUP96 structure (orange) highlighted in the zoom-in of a symmetry center on the ring (ca. 12 nm apart). Bottom right: N-terminally-labeled (blue, highlighted by green arrows), NUP107 structure (orange) in the zoom-in of a symmetry center on the ring (ca. 12 nm apart). Distances and cartoons derived from PDB entry: Nup(5A9Q). c) DNA-PAINT overview image of NUP96-SNAP in U2OS cells. d) Zoom-in of highlighted area reveals the arrangement of NUP96 in NPCs. e) DNA-PAINT overview image of NUP107-GFP in HeLa cells. f) Zoom-in of highlighted area. Scale bars: 5 µm (c, e), 100 nm (d, f).

symmetry center are arranged in Y-complexes spaced circa 12 nm apart (Figure 1b). In order to quantitatively compare different labeling approaches, we first acquired 2D DNA-PAINT data using identical image acquisition parameters (see Supplementary Tables 1–3 in the Supporting Information for details). Next, we selected single NPC structures in the reconstructed super-resolution image, aligned them on top of each other (that is, the center of the NPC rings, thus creating a sum image) and performed a radial distance measurement over all localizations. This analysis yields two observables for comparison; first, the median fitted ring radius and second, the width of this distribution. Dissimilar fitted radii for the same protein labeled using different tags are a measure for potential systematic biases introduced by a preferential orientation of the labeling probes. The width of the distribution on the other hand is a proxy for label-size-induced linkage error, that is, broader distributions originate from "larger" labeling probes. Our data in Figure 2 provides a quantitative comparison of NUP107-SNAP, NUP107-GFP, NUP96-SNAP, and NUP96-Halo cell lines targeted with their respective DNA-conjugated labeling probes (see also Supplementary Figures 2–6 in the Supporting Information). Furthermore, we compare our results with NUP107-GFP labeled using primary and DNA-conjugated secondary antibodies. We obtained radii of 53.7 ± 13.1 nm for NUP107-SNAP (Figure 2a and Supplementary Figure 2 in the Sup-

porting Information) and 54.6 ± 11.9 nm for NUP107-GFP (nanobody staining) (Figure 2b and Supplementary Figure 3 in the Supporting Information), as well as 55.9 ± 12.6 nm for NUP96-SNAP (Figure 2c and Supplementary Figure 4 in the Supporting Information) and 56.2 ± 10.2 nm for NUP96-Halo (Figure 2d and Supplementary Figure 5 in the Supporting Information), in close overall agreement to earlier EM- and fluorescence-based studies.[7b,8a] For the antibody-stained sample against NUP107-GFP (Figure 2e and Supplementary Figure 6 in the Supporting Information), we obtained a considerably larger radius of 65.9 nm. This could be explained by primary and DNA-conjugated secondary antibodies potentially binding preferentially towards the outside of the NPCs. However, not only did the antibody-stained samples yield a larger apparent NPC radius, also the measured width of the distribution (18 nm) was larger compared to the genetically-encoded tags due to the increased size of the antibodies. In the case of genetically-encoded tags, the width of the distributions is considerably smaller (see also Supplementary Table 4 in the Supporting Information) due to the reduced linkage error to the actual protein location.[3c,e,14]

Next, we sought out to further optimize image acquisition conditions with respect to overall localization precision, sampling of single protein sites, and three-dimensional image acquisition (Supporting Information, Supplementary Figure 7). This allowed us to visualize single copies of NUP96
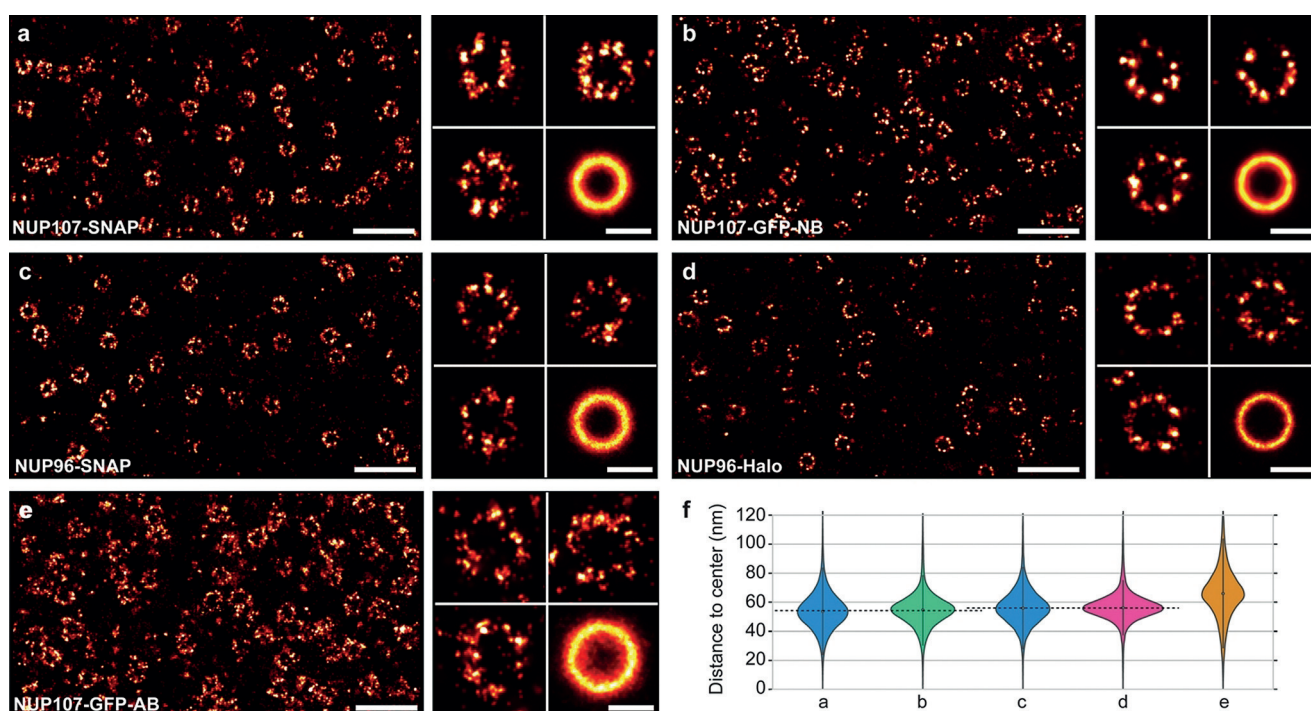

Figure 2. a) NUP107-SNAP overview image (left). Zoom-in to individual NPCs and sum image ($n=398$) (right). b) NUP107-GFP nanobody overview image (left). Zoom-in to individual NPCs and sum image ($n=486$) (right). c) NUP96-SNAP overview image (left). Zoom-in to individual NPCs and sum image ($n=288$) (right). d) NUP96-Halo overview image (left). Zoom-in to individual NPCs and sum image ($n=191$) (right). e) NUP107-GFP-Antibody overview image (left). Zoom-in to individual NPCs and sum image ($n=185$) (right). f) Violin plots of the distances between ring center and localizations. Median radii and standard deviation were obtained for each label: NUP107-SNAP (from **a**, 127 773 fitted localizations, 53.7 ± 13.1 nm radius), NUP107-GFP (from **b**, 219 398 fitted localizations, 54.6 ± 11.9 nm radius), NUP96-SNAP (from **c**, 57 297 fitted localizations, 55.9 ± 12.6 nm), NUP96-Halo (from **d**, 45 143 fitted localizations, 56.2 ± 10.2 nm radius), NUP107-GFP-Anitbody (from **e**, 69 834 fitted localizations, 65.9 ± 17.5 nm). See also Supplementary Table 2 in in the Supporting Information. Scale bars: 500 nm (overviews), 100 nm (individual NPCs and sum image).
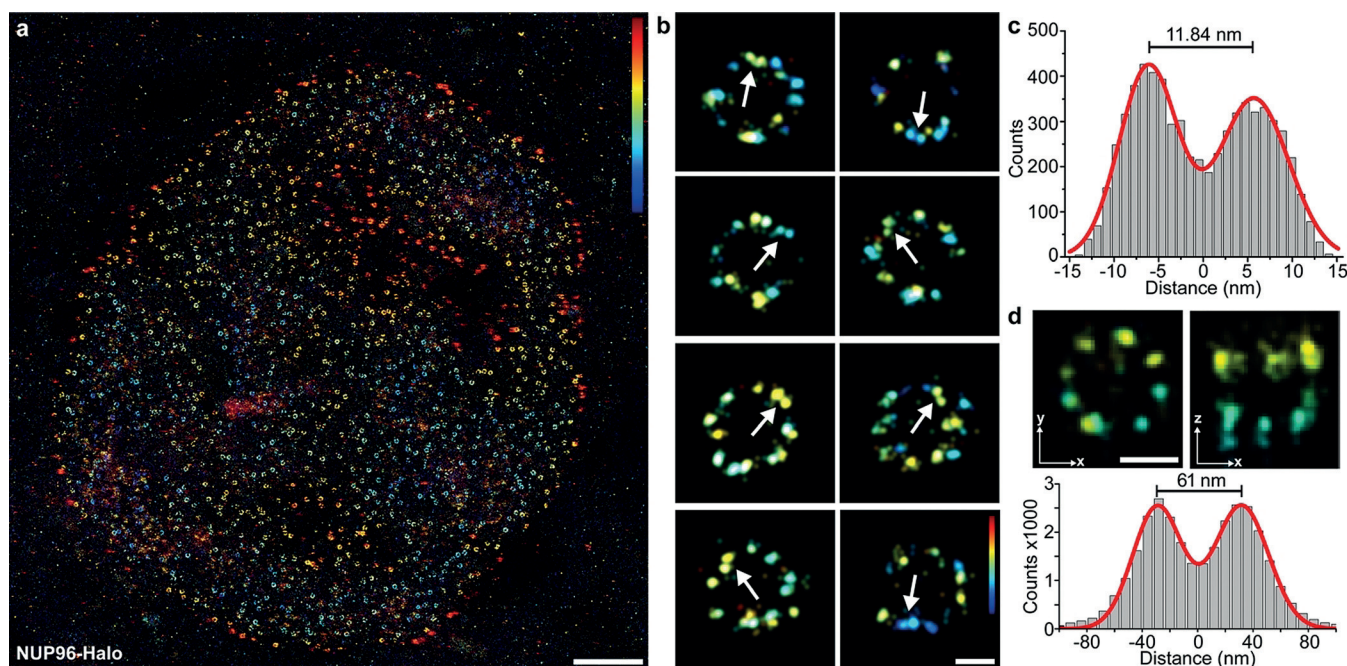
Figure 3. a) Overview image of NUP96-Halo imaged using 3D DNA-PAINT (color indicates height, range: −200 (blue) to 200 nm (red)). b) Selection of single NPCs. Arrows are highlighting two copies of NUP96 proteins in the same symmetry center of the same ring (that is at the same height) spaced ca. 12 nm apart from each other (color indicates height, range: −100 (blue) to 100 nm (red)). c) Cross-sectional histogram of 3D-summed pairs ($n = 45$) of NUP96 proteins in single symmetry centers as highlighted by arrows in **b** yields ca.12 nm distance between single proteins. d) NUP96-Halo-labeled NPCs show the typical eightfold symmetry (*xy*-projection, left) and the organization in nuclear and cytoplasmic rings (*xz*-projection, right). Micrographs represent sum data from aligned NPCs ($n = 31$). Bottom: Cross-sectional histogram of localizations in the xz-projection yields ca. 61 nm separation between cytoplasmic and nuclear rings. Scale bars: 2 μm (**a**), 50 nm (**b, d**).

proteins (Figure 3) using the NUP96-Halo cell line, which we chose based on its superior performance in the 2D study presented above (smallest distribution width). An overview of a typical 3D DNA-PAINT dataset is shown in Figure 3a. Zooming in to some of the NPCs (Figure 3b) reveals distinctive pairs of close-by "localization clouds" (arrows in Figure 3b), which we attribute to single NUP96 proteins. To quantitatively asses the Euclidian distance of the two copies of NUP96 on the two cytoplasmic or nuclear rings of the NPC, we selected about 50 pairs in NPCs, aligned them on top of each other and subsequently carried out particle averaging with Picasso.[2b, 15] We then performed a cross-sectional histogram analysis of the resulting sum image and fitted the distribution with two Gaussian functions (Figure 3c). The fit yields a peak-to-peak distance of about 12 nm, well in agreement with the expected distance of NUP96 proteins on adjacent Y-complexes as derived from EM models.[8a] Furthermore, each peak fit exhibits a standard deviation of only circa 3 nm, highlighting the high localization precision and accuracy achievable with the combination of genetically-encoded tags with DNA-PAINT. Additionally, we measured the separation between the cytoplasmic and the nuclear rings for NUP96-Halo, yielding a distance of about 61 nm (Figure 3d), which we could clearly resolve. The capability to separate the nuclear from the cytoplasmic side of the NPC is a necessity to convince us, that the NUP pairs in each symmetry center (Figure 3b and c) are indeed part of either the nuclear or cytoplasmic rings of the NPC. Furthermore, we obtained qualitatively and quantitatively similar results for

the NUP96-SNAP cell line (Supporting Information, Supplementary Figure 8).

In conclusion, we present an approach to combine DNA-PAINT with genetically-encoded self-labeling tags. This provides a tool to investigate single proteins in higher order protein complexes in cells. However, we could only achieve a relatively modest labeling efficiency of approximately 30% (Supporting Information, Supplementary Table 5). Thus, one of the main challenges in the field remains, which is a route to highly efficient labeling probes (> 90% labeling efficiency) without requiring genetic engineering. Besides the availability of peptide tags combined with nanobodies[16] and small scaffolds like nanobodies,[3a] affimers,[17] darpins,[18] or SOMAmers[19] novel approaches are necessary to tackle this challenge. Probes could include optimized host–guest systems,[20] direct transient binders,[21] or rationally-designed small proteins.[22] However, even with our current labeling efficiency, studies of single proteins in multicomponent complexes are within reach.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

[1] a) S. J. Sahl, S. W. Hell, S. Jakobs, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 685–701; b) M. Sauer, M. Heilemann, *Chem. Rev.* **2017**, *117*, 7478–7509.

[2] a) M. Dai, R. Jungmann, P. Yin, *Nat. Nanotechnol.* **2016**, *11*, 798–807; b) J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, R. Jungmann, *Nat. Protoc.* **2017**, *12*, 1198–1228; c) F. Balzarotti, Y. Eilers, K. C. Gwosch, A. H. Gynna, V. Westphal, F. D. Stefani, J. Elf, S. W. Hell, *Science* **2017**, *355*, 606–612.

[3] a) J. Ries, C. Kaplan, E. Platonova, H. Eghlidi, H. Ewers, *Nat. Methods* **2012**, *9*, 582–584; b) S. Sograte-Idrissi, N. Oleksiievets, S. Isbaner, M. Eggert-Martinez, J. Enderlein, R. Tsukanov, F. Opazo, *Cells* **2019**, *8*, 48; c) F. Opazo, M. Levy, M. Byrom, C. Schafer, C. Geisler, T. W. Groemer, A. D. Ellington, S. O. Rizzoli, *Nat. Methods* **2012**, *9*, 938–939; d) V. Fabricius, J. Lefèbre, H. Geertsema, S. F. Marino, H. Ewers, *J. Phys. D* **2018**, *51*, 474005; e) S. Strauss, P. C. Nickels, M. T. Strauss, V. Jimenez Sabinina, J. Ellenberg, J. D. Carter, S. Gupta, N. Janjic, R. Jungmann, *Nat. Methods* **2018**, *15*, 685–688.

[4] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86–89.

[5] G. V. Los, L. P. Encell, M. G. McDougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Bulleit, K. V. Wood, *ACS Chem. Biol.* **2008**, *3*, 373–382.

[6] M. Beck, E. Hurt, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 73–89.

[7] a) A. Loschberger, S. van de Linde, M. C. Dabauvalle, B. Rieger, M. Heilemann, G. Krohne, M. Sauer, *J. Cell Sci.* **2012**, *125*, 570–575; b) A. Szymborska, A. de Marco, N. Daigle, V. C. Cordes, J. A. Briggs, J. Ellenberg, *Science* **2013**, *341*, 655–658.

[8] a) A. von Appen, J. Kosinski, L. Sparks, A. Ori, A. L. DiGuilio, B. Vollmer, M. T. Mackmull, N. Banterle, L. Parca, P. Kastritis, K. Buczak, S. Mosalaganti, W. Hagen, A. Andres-Pons, E. A. Lemke, P. Bork, W. Antonin, J. S. Glavy, K. H. Bui, M. Beck, *Nature* **2015**, *526*, 140–143; b) D. H. Lin, T. Stuwe, S. Schilbach, E. J. Rundlet, T. Perriches, G. Mobbs, Y. Fan, K. Thierbach, F. M. Huber, L. N. Collins, A. M. Davenport, Y. E. Jeon, A. Hoelz, *Science* **2016**, *352*, aaf1015; c) S. J. Kim, J. Fernandez-Martinez, I. Nudelman, Y. Shi, W. Zhang, B. Raveh, T. Herricks, B. D. Slaughter, J. A. Hogan, P. Upla, I. E. Chemmama, R. Pellarin, I. Echeverria, M. Shivaraju, A. S. Chaudhury, J. Wang, R. Williams, J. R. Unruh, C. H. Greenberg, E. Y. Jacobs, Z. Yu, M. J. de la Cruz, R. Mironska, D. L. Stokes, J. D. Aitchison, M. F. Jarrold, J. L. Gerton, S. J. Ludtke, C. W. Akey, B. T. Chait, A. Sali, M. P. Rout, *Nature* **2018**, *555*, 475–482.

[9] M. Eibauer, M. Pellanda, Y. Turgay, A. Dubrovsky, A. Wild, O. Medalia, *Nat. Commun.* **2015**, *6*, 7532.

[10] a) B. Koch, B. Nijmeijer, M. Kueblbeck, Y. Cai, N. Walther, J. Ellenberg, *Nat. Protoc.* **2018**, *13*, 1465–1487; b) J. V. Thevatha-san, M. Kahnwald, K. Cieśliński, P. Hoess, S. K. Peneti, M. Reitberger, D. Heid, K. C. Kasuba, S. J. Hoerner, Y. Li, Y.-L. Wu, M. Mund, U. Matti, P. M. Pereira, R. Henriques, B. Nijmeijer, M. Kueblbeck, V. J. Sabinina, J. Ellenberg, J. Ries, *bioRxiv* **2019**, https://doi.org/10.1101/582668.

[11] S. S. Agasti, Y. Wang, F. Schueder, A. Sukumar, R. Jungmann, P. Yin, *Chem. Sci.* **2017**, *8*, 3080–3091.

[12] C. Steinhauer, R. Jungmann, T. L. Sobey, F. C. Simmel, P. Tinnefeld, *Angew. Chem. Int. Ed.* **2009**, *48*, 8870–8873; *Angew. Chem.* **2009**, *121*, 9030–9034.

[13] K. Kelley, K. E. Knockenhauer, G. Kabachinski, T. U. Schwartz, *Nat. Struct. Mol. Biol.* **2015**, *22*, 425–431.

[14] T. Schlichthaerle, M. Ganji, A. Auer, O. K. Wade, R. Jungmann, *ChemBioChem* **2019**, *20*, 1032–1038.

[15] F. Schueder, J. Lara-Gutierrez, B. J. Beliveau, S. K. Saka, H. M. Sasaki, J. B. Woehrstein, M. T. Strauss, H. Grabmayr, P. Yin, R. Jungmann, *Nat. Commun.* **2017**, *8*, 2090.

[16] D. Virant, B. Traenkle, J. Maier, P. D. Kaiser, M. Bodenhofer, C. Schmees, I. Vojnovic, B. Pisak-Lukats, U. Endesfelder, U. Rothbauer, *Nat. Commun.* **2018**, *9*, 930.

[17] C. Tiede, A. A. Tang, S. E. Deacon, U. Mandal, J. E. Nettleship, R. L. Owen, S. E. George, D. J. Harrison, R. J. Owens, D. C. Tomlinson, M. J. McPherson, *Protein Eng. Des. Sel.* **2014**, *27*, 145–155.

[18] P. Forrer, H. K. Binz, M. T. Stumpp, A. Pluckthun, *ChemBio-Chem* **2004**, *5*, 183–189.

[19] J. C. Rohloff, A. D. Gelinas, T. C. Jarvis, U. A. Ochsner, D. J. Schneider, L. Gold, N. Janjic, *Mol. Ther. Nucleic Acids* **2014**, *3*, e201.

[20] K. J. Kossmann, C. Ziegler, A. Angelin, R. Meyer, M. Skoupi, K. S. Rabe, C. M. Niemeyer, *ChemBioChem* **2016**, *17*, 1102–1106.

[21] T. Kiuchi, M. Higuchi, A. Takamura, M. Maruoka, N. Watanabe, *Nat. Methods* **2015**, *12*, 743–746.

[22] A. Chevalier, D. A. Silva, G. J. Rocklin, D. R. Hicks, R. Vergara, P. Murapa, S. M. Bernard, L. Zhang, K. H. Lam, G. Yao, C. D. Bahl, S. I. Miyashita, I. Goreshnik, J. T. Fuller, M. T. Koday, C. M. Jenkins, T. Colvin, L. Carter, A. Bohn, C. M. Bryan, D. A. Fernandez-Velasco, L. Stewart, M. Dong, X. Huang, R. Jin, I. A. Wilson, D. H. Fuller, D. Baker, *Nature* **2017**, *550*, 74–79.