Protein Counting

International Edition: DOI: 10.1002/anie.201505664
German Edition: DOI: 10.1002/ange.201505664

A Set of Homo-Oligomeric Standards Allows Accurate Protein Counting

Kieran Finan, Anika Raulf, and Mike Heilemann**

Abstract: *Techniques based on fluorescence microscopy are increasingly used to count proteins in cells, but few stoichiometrically well-defined standards are available to test their accuracy. A selection of bacterial homo-oligomers were developed that contain 10–24 subunits and fully assemble when expressed in mammalian cells, and they can be used to easily validate/calibrate molecular counting methods. The utility of these standards was demonstrated by showing that nuclear pores contain 32 copies of the Nup107 complex.*

Traditional structural biology techniques such as X-ray crystallography have been key to revealing the mechanisms through which small and medium-sized protein complexes function. However, large protein assemblies (>20 nm in diameter), such as nuclear pores, transcription factories, or centrioles, are not accessible to such techniques and are less well understood as a result. In the case of the nuclear pore, cutting-edge experiments have revealed many structural details, but estimates of its mass and subunit stoichiometry differ by factors of two or more.[1,2]

One way to investigate the structural details of large protein assemblies is to determine the copy number of tagged subunits by using fluorescence microscopy. This can be achieved by analyzing step-wise photobleaching[3] or by comparing the total intensity of a structure to those of standards containing known numbers of tags.[4,5]

The combination of such counting strategies with super-resolution techniques[6] that are able to sensitively resolve individual protein complexes holds immense promise. Localization-based super-resolution techniques (e.g., PALM/STORM)[6] have already been used to count proteins;[7–10] however these strategies are complicated by fluorophore photophysics,[11] and positive controls are required to ensure that the counts they produce are accurate (see Text S1 in the Supporting Information).[7–9]

Indeed, a general problem that reduces the robustness of all microscopy-based counting methods is that no reliable and easy-to-use standards are available to calibrate and validate them. The *Escherichia coli* proteins MotB and FliM are present at approximately 22 and 30 copies in active flagellar motors, and have been used as standards for counting experiments;[4,5] however, a nontrivial flagellar-immobilization assay[3] is required to distinguish fully formed motors from the large pool of partially assembled complexes (Text S2).[12] GFP-tagged rotavirus-like particles, created through the co-expression of two viral capsid proteins, contain 120 tagged molecules[4,13] but they require a difficult expression and purification procedure. Constructs containing one to three copies of fluorescent proteins have also been used as standards, but they tend to produce weak signals that can be difficult to detect and quantify[7–9] (Text S1).

To develop easy-to-use and robust protein-counting controls, we sought stable protein complexes that contain known numbers of subunits and exhibit uniform stoichiometry. We hypothesized that well-characterized bacterial homo-oligomers might serve as good standards, and we thus chose the *E. coli* proteins fructose-6-phosphate aldolase (FsaA; a 10-mer), glutamine synthetase (GlnA; a 12-mer), and ferritin (FtnA; a 24-mer), as well as the Dps protein (a 12-mer) of *Thermosynechococcus elongatus* (a thermophilic bacterium) for further study. All of these have previously been structurally characterized by X-ray crystallography[14–17] (Figure 1a), and they are known to assemble stably and completely (Text S3).

We constructed plasmids encoding variants of these proteins that were tagged with the fast-maturing "Ypet" yellow fluorescent protein (YFP; see Text S4 in the Supporting Information), transfected each construct separately into U2OS cells, and imaged the cells by confocal microscopy after fixation. In cells with low expression levels, we observed clear, bright foci with relatively uniform intensities for each of the four homo-oligomeric species (Figure S1 in the Supporting Information).

To ensure that the homo-oligomers had assembled with the expected stoichiometries, we quantified their relative intensities.[5,12] For each YFP-tagged species, we lysed the transfected cells, spread the lysate over coverslips, fixed the homo-oligomers, imaged them by using 3D confocal microscopy (Figure 1b), and determined the average intensity of in-focus complexes (Figure 1c). As expected, we observed an intensity ratio of 10:12:12:24 for FsaA/GlnA/DpsTE/FtnA (Figure 1c), a result that strongly suggests that the oligomers were fully assembled (Texts S5, S6 and Figure S2). Analysis of foci present inside fixed cells revealed similar intensities to those observed with coverslip-bound homo-oligomers (Figure S3).

To further confirm that the oligomers contained the expected number of fluorescent proteins, we used a single-

[*] Dr. K. Finan,[+] A. Raulf, Prof. Dr. M. Heilemann
Single Molecule Biophysics
Institute for Physical and Theoretical Chemistry, Goethe-University Frankfurt
Max-von-Laue-Strasse 7, 60438 Frankfurt am Main (Germany)
E-mail: heilemann@chemie.uni-frankfurt.de
[+] Current address: UCL Medical School
Gower St., London (UK)
E-mail: kieran.finan.13@ucl.ac.uk

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201505664.

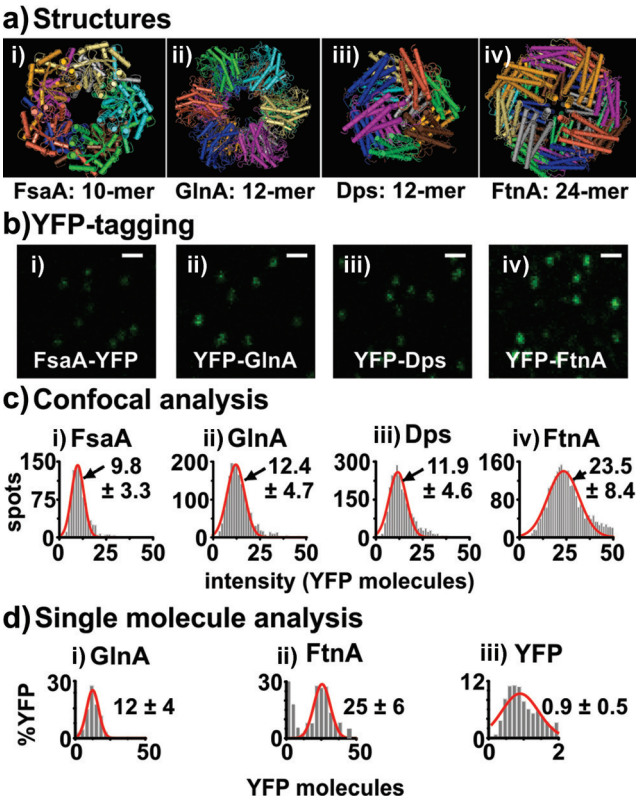

Figure 1. Validation of the homo-oligomeric standards. a) Crystal structures of the bacterial homo-oligomers: i) fructose-6-phosphate aldolase,[15] ii) glutamine synthetase,[16] iii) Dps,[14] and iv) ferritin.[17] Gene abbreviations and subunit numbers are indicated below the images. b) Confocal imaging of tagged homo-oligomers. For each homo-oligomer in (a), a population of U2OS cells was transfected with a plasmid encoding a YFP-tagged variant of the protein and lysed. The lysate was spread over a poly-L-lysine coated coverslip, fixed, and imaged through confocal microscopy z-stacks. In-focus images are shown. Scale bar: 700 nm. c) Quantification of homo-oligomer fluorescence. For each homo-oligomer, the total integrated intensities of in-focus spots in (b) were quantified by using an image-analysis program, and frequency histograms were plotted and fitted with Gaussian functions. The mean values and standard deviations of the fits are indicated. Data was obtained from 993 spots in 16 images for FsaA-YFP, 1858 spots in 16 images for YFP-GlnA, 2671 spots in 15 images for YFP-Dps, and 2950 spots in 16 images for YFP-FtnA. The intensity of one YFP moiety, which was used to covert the intensity scales to molecular units, was calculated by dividing the average intensity of each homo-oligomer by the expected number of YFP molecules and then taking the average of these four values. d) Single-molecule analysis of tagged homo-oligomers. Coverslips prepared as in (b) were imaged using a single-molecule-sensitive microscope with TIRF illumination and the spot total integrated intensities were analyzed. Frequency histograms of i) YFP-GlnA ($n = 193$ spots), ii) YFP-FtnA ($n = 313$ spots), and iii) individual YFP molecule ($n = 2599$ spots) intensities were plotted and fitted as in (c).

molecule-sensitive widefield microscope to compare YFP-FtnA and YFP-GlnA to single YFP molecules. As expected, we found a ratio of approximately 24:12:1 when we compared the average intensities of the three proteins (Figure 1d and Text S6). Intensity distributions were narrow (s.d. < 30%; Figure 1dii, Figure S4). Further analysis of measurement noise revealed that the true variation of homo-oligomer intensities was less than 14% (s.d.), thus indicating that more than 85% of the YFP molecules were fluorescent (Figure S5 and Text S7). Together, our confocal and single-molecule data indicate that the tagged proteins form uniform populations of fully-assembled oligomers when expressed in mammalian cells and can thus be used as standards in counting experiments.

To demonstrate the value of our standards, we used them to distinguish between popular models for nuclear-pore structure. Most theories are based around the multiprotein Nup107 complex, which is thought to have a copy number of 16 or 32 molecules[1,2] (Text S8). To determine which, if any, of these models might be correct, we counted the number of Nup133 molecules (a component of the Nup107 complex) in a pore.

We created a stable cell line in which approximately 95% of the total Nup133 was tagged with YFP (Figure 2a and Figure S6), and imaged these cells in parallel with homo-oligomer-coated coverslips by using confocal microscopy z-stacks (Figure 2bi). We identified individual pores/homo-oligomers by using an algorithm designed to find in-focus and symmetric point-spread functions, while rejecting the asymmetrical spots that arise from adjacent pores (Figure S7). Employing a 'ratiometric' method,[4,5] we used the average intensities of our homo-oligomers, which contained known numbers of tags, to convert the intensities of the Nup133-YFP spots to molecular units. We found that the frequency histogram of the Nup133-YFP spot intensities peaked at approximately 35 molecules (Figure 2bii), which is very close to the 32 molecules predicted by some models of pore structure.[2] Fewer than 5% of spots had intensities less than the average intensity expected for 16 molecules. We noted that the intensity distribution had a "tail" and hypothesized that some spots contained pairs of adjacent pores but had not been filtered out by our algorithm. Fitting the distribution by using two Gaussian peaks confirmed this assumption, yielding one peak at 31 molecules, and a second at about twice that value (58 molecules; Figure 2bii and Figure S2).

To ensure that our previous result had not been influenced by an inability to resolve individual pores, we repeated the Nup133 counting experiment using astigmatism-based 3D single-molecule localization microscopy.[18] In this super-resolution technique, individual fluorophores are stochastically activated, localized to precisions of approximately 10–40 nm, and then driven back into a dark state. After considering various protein labelling strategies, we chose to mark our targets with SNAP tags[19] labelled with Alexa Fluor 647 (Text S9 and Figure S9). In this imaging modality, fluorophores can be localized several times or not at all; however, the average number of localizations per cluster should be roughly proportional to the number of fluorophores (Text S10).

We generated and imaged a U2OS cell line stably expressing Nup133-SNAP, in which more than 90% of the Nup133 protein carried the SNAP tag (Figure 2a). Images revealed well-separated ring-like structures (i.e., pores; Figure 2ci,ii). In order to determine the number of tagged Nup133 molecules present in these pores, we imaged them in parallel with U2OS cells transiently transfected with the

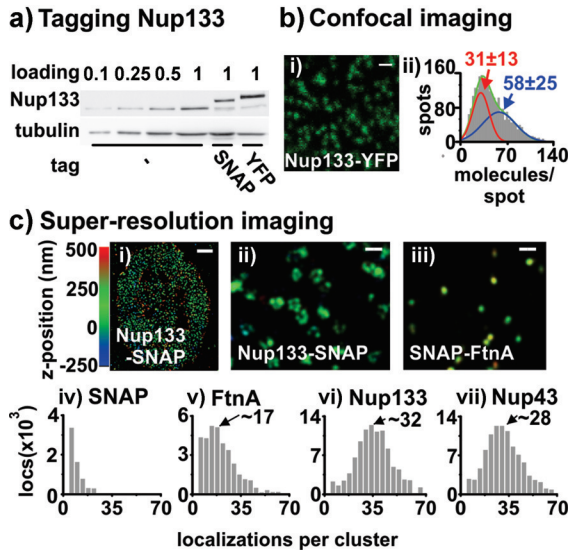

Figure 2. Counting Nup107 complexes in nuclear pores. a) Determining the fraction of Nup133 tagged. U2OS cells stably transfected with plasmids expressing Nup133-SNAP, Nup133-YFP, or no plasmid were lysed and their protein contents subjected to SDS PAGE and western blotting with an anti-Nup133 antibody.. The band corresponding to Nup133-YFP (lane 6) is ca. twice as bright as the 1× Nup133 band (lane 4), thus indicating that the protein has been overexpressed approximately 2-fold. The endogenous Nup133 in the Nup133-YFP cell line (lane 6) produced a band that is roughly as bright as the 0.1× Nup133 lane (lane 1).[20] b) Counting Nup133 molecules by using confocal microscopy. i) The cells stably expressing Nup133-Ypet in (a) were fixed and the basal surfaces of their nuclei were imaged using confocal microscopy *z*-stacks. An in-focus slice is shown. Scale bar: 700 nm. ii) Symmetrical in-focus Gaussian spots were automatically identified and a frequency histogram of their total integrated intensities was plotted and fitted using a two-Gaussian function model (green line; the individual Gaussian functions are plotted in red and blue, mean values and s.d. are indicated)). Data were obtained from 2506 pores in 43 cells. The intensity of a single YFP was determined using the homo-oligomeric standards as in Figure 1c. c) Counting Nup133 and Nup43 molecules using super-resolution microscopy. U2OS cells stably expressing Nup133-SNAP or SNAP-Nup43, or transiently expressing the 24-mer SNAP-FtnA or SNAP alone, were fixed, labelled with Alexa Fluor 647-BG, and imaged using 3D *d*STORM. The z-positions of individual localizations are indicated by their color. i,ii) Images of Nup133-SNAP; scale bars: 3 μm and 200 nm. iii) Image of SNAP-FtnA; scale bar: 200 nm. iv–vii) Analysis of super-resolution images. Clusters of localizations were identified by using the DBSCAN algorithm, and frequency histograms of the number of localizations per cluster were plotted. Peak values are indicated. The 24-mer SNAP-FtnA (2335 clusters analyzed in 4 cells) yielded clusters with an average 17 localizations (range of ca. 14–22), while the Nup107 proteins SNAP-Nup133 (3428 clusters analyzed in 4 cells) and SNAP-Nup43 (3928 clusters analyzed in 4 cells) yielded clusters with averages of ca. 28–32 localizations. To study the SNAP tag alone, 1018 clusters were analyzed in 4 cells.

24-mer SNAP-FtnA (Figure 2c iii). We determined the average number of localizations present in the pores and FtnA homo-oligomers (Figure 2c iii–v) and found that the pores contained roughly 1.3× more localizations than the 24-mer. This result is consistent with the former containing 32 rather than 16 molecules.

To ensure that this result was not affected by different SNAP-fusion constructs binding the Alexa Fluor 647-BG substrate at different frequencies, we generated a cell line stably expressing SNAP-Nup43, another member of the Nup107 complex, along with an shRNA construct targeting the 3′ untranslated region (UTR) of endogenous Nup43 mRNA (Figure S8; more than 80% of the protein tagged). Imaging this strain yielded pores containing approximately 28 localizations, a count similar to the 32 localizations obtained with Nup133. This suggests that the efficiency at which the fluorophore substrate conjugates to the SNAP tags is somewhat independent of the SNAP fusion partner, and further confirms that nuclear pores contain 32 Nup107 complexes.

Our standards can be used to convert the intensity scales of fluorescence microscopy images into molecular units through a simple method that can be used with any microscope capable of producing z-stacks. DNA encoding three out of the four homo-oligomers is available in most labs (in the *E. coli* genome), and the ability to check the relative intensities of multiple standards provides a powerful internal control. The ability of our homo-oligomers to fully assemble appears to be independent of the of the tag with which they are marked (Figure S1) and they should thus be compatible with many types of fluorescence microscopy, including SIM and STED.[6] Our standards will also help to improve existing localization-microscopy-based counting methods[7–9] by aiding users in determining molecular detection efficiencies (Text S1) and by serving as positive controls to help assure nonspecialists that the complex corrections associated with such techniques have functioned correctly.

Acknowledgements

We would like to thank Petra Freund, Petra Gessner, and Lisa Pliess for technical assistance. We are also extremely grateful to Dr. Nicole Wagner and Prof. Sueleyman Erguen (University of Wuerzburg, Germany) for use of the N-STORM microscope, and to Helge Ewers (King's College London) for allowing KF the use of his equipment and reagents while revising the manuscript. We thank Andrea Ilari and Pierpaolo Ceci (National Research Council, Italy) for their gift of the *T. elongatus* Dps ORF. We are grateful for financial support by the Federal Ministry of Science and Education (BMBF, grant 0315262) and the German Science Foundation (DFG, grant EXC115).

Keywords: fluorescence microscopy · nuclear pores · proteins · protein counting · super-resolution imaging

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 12049–12052
Angew. Chem. **2015**, *127*, 12217–12220

[1] F. Alber, et al., *Nature* **2007**, *450*, 695–701.
[2] K. H. Bui, et al., *Cell* **2013**, *155*, 1233–1243.
[3] M. C. Leake, J. H. Chandler, G. H. Wadhams, F. Bai, R. M. Berry, J. P. Armitage, *Nature* **2006**, *443*, 355–358.
[4] J. Lawrimore, K. S. Bloom, E. D. Salmon, *J. Cell Biol.* **2011**, *195*, 573–582.

[5] V. C. Coffman, P. Wu, M. R. Parthun, J. Q. Wu, *J. Cell Biol.* **2011**, *195*, 563 – 572.
[6] L. Schermelleh, R. Heintzmann, H. Leonhardt, *J. Cell Biol.* **2010**, *190*, 165 – 175.
[7] P. Sengupta, T. Jovanovic-Talisman, D. Skoto, M. Renz, S. L. Veatch, J. Lippincott-Schwartz, *Nat. Methods* **2011**, *8*, 969 – 975.
[8] E. M. Puchner, J. M. Walter, R. Kasper, B. Huang, W. A. Lim, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16015 – 16020.
[9] N. Durisic, L. Laparra-Cuervo, A. Sandoval-Alvarez, J. S. Borbely, M. Lakadamyali, *Nat. Methods* **2014**, *11*, 156.
[10] U. Endesfelder, K. Finan, S. J. Holden, P. R. Cook, A. N. Kapanidis, M. Heilemann, *Biophys. J.* **2013**, *105*, 172 – 181.
[11] P. Annibale, M. Scarselli, A. Kodiyan, A. Radenovic, *J. Phys. Chem. Lett.* **2010**, *1*, 1506 – 1510.
[12] N. J. Delalez, G. H. Wadhams, G. Rosser, Q. Xue, M. T. Brown, I. M. Dobbie, R. M. Berry, M. C. Leake, J. P. Armitage, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11347 – 11351.
[13] A. Charpilienne, M. Nejmeddine, M. Berois, N. Parez, E. Neumann, E. Hewat, G. Trugnan, J. Cohen, J. Biol. Chem. **2001**, *276*, 29361 – 29367.
[14] S. Franceschini, P. Ceci, F. Alaleona, E. Chiancone, A. Ilari, *FEBS J.* **2006**, *273*, 4913 – 4928.
[15] S. Thorell, M. Schurmann, G. A. Sprenger, G. Schneider, *J. Mol. Biol.* **2002**, *319*, 161 – 171.
[16] H. S. Gill, D. Eisenberg, *Biochemistry* **2001**, *40*, 1903 – 1912.
[17] T. J. Stillman, P. D. Hempstead, P. J. Artymiuk, S. C. Andrews, A. J. Hudson, A. Treffry, J. R. Guest, P. M. Harrison, *J. Mol. Biol.* **2001**, *307*, 587 – 603.
[18] B. Huang, W. Wang, M. Bates, X. Zhuang, *Science* **2008**, *319*, 810 – 813.
[19] X. Sun, et al., *ChemBioChem* **2011**, *12*, 2217 – 2226.
[20] K. Maeshima, et al., *Nat. Struct. Mol. Biol.* **2010**, *17*, 1065 – 1071.

Received: June 19, 2015
Published online: August 20, 2015

