TECHNICAL BRIEF

A universally applicable process for preparing stoichiometrically 1:1 labelled functional proteins

Achim Friedrich, Jörg D. Hoheisel, Jens-Peter Knemeyer and Nicole Marmé

Division of Functional Genome Analysis, German Cancer Research Center (DKFZ), Heidelberg, Germany

A universally applicable labelling and purification process was established to prepare biologically active proteins with a stoichiometric 1:1 ratio of attached dye-label. The dye-label is linked to a specific DNA sequence, which acts as a barcode-like tag for affinity purification. The DNA-dye tag is covalently bound to the target protein, which is present in excess to assure the binding of not more than one dye per molecule. Affinity purification occurs at magnetic beads that are functionalized with oligonucleotides that are complementary to the DNA-tag of the labelled proteins but for one or two mismatches. Washing removes all unbound, unlabelled molecules. The labelled protein is subsequently released by the addition of a fully complementary oligonucleotide. This process allows a gentle purification of a protein fraction that has exactly one label attached to each molecule under conditions that preserve protein structure.

Received: April 1, 2011
Revised: June 21, 2011
Accepted: July 1, 2011

Keywords:
Protein function word count / Protein labelling / Single-molecule analysis / Technology

Technical advances in single-molecule fluorescence techniques and the ambition of quantification in biological research demand stoichiometrically labelled biomolecules, especially proteins [1]. Besides direct chemical modification [2], a broad variety of protein-labelling procedures have emerged over the past decades, e.g. fusion proteins [3], enzymatic protein ligation [1, 4], snap tags [5], or labelling by antibodies [6, 7]. All of them have particular advantages but also disadvantages. None of them is generally applicable without affecting the biological activity of the labelled molecule. Fusion proteins, for example, show poor spectroscopic properties and high molecular weight. Thus, they are often inappropriate for single-molecule applications. Ligation techniques are limited to proteins with a compatible amino acid sequence at the enzymatic ligation site. From the spectroscopic point of view, organic fluorophores provide the best characteristics concerning absorption coefficient, fluorescent quantum yield, blinking, and photostability [8]. They have a low molecular weight, which minimizes negative effects on biological activity and well-established labelling protocols are available. Consequently, they are used in most single-molecule fluorescence experiments. However, standard labelling protocols are inappropriate for generating stoichiometric labelling, a fact that complicates quantification significantly [9, 10]. Therefore, additional calculations of the brightness of fluorescent signals or cumbersome techniques like stepwise photobleaching measurements have to be applied for quantitative statements [11, 12].

To overcome these limitations, we report here a universal method for preparing stoichiometrically 1:1 dye-labelled proteins. As a model system, BSA was labelled with the fluorescent dye Atto520. The method is based on a dye that is attached to a specific DNA sequence. The hydrophilic DNA-strand is hydrated and therefore increases the dye-protein distance. Consequently, the dye molecule interferes less with hydrophobic protein portions, with the result that protein structure and activity are less affected. By contrast, standard labelling with organic dyes often interacts with hydrophobic protein regions and often causes a loss of native structure [2]. The DNA-dye label is covalently linked to the ε-amino group of the amino acid lysine, which is the

Correspondence: Dr. Jörg D. Hoheisel, Functional Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
E-mail: j.hoheisel@dkfz.de
Fax: +49-6221-424682

Abbreviations: DNA-D, displacement strand oligonucleotide; **DNA-SP,** solid-phase strand oligonucleotide

numerically most available one on protein surfaces. The attachment occurs randomly at different protein surface locations. To avoid the binding of multiple dyes per protein, the amount of label and protein was adjusted appropriately. The linkage itself is done by the chemoselective crosslinker pair succinimidyl 4-hydrazinonicotinate acetone hydrazone (SANH) and succinimidyl 4-formylbenzoate (SFB), which is forming a stable hydrazone bond between DNA-tag and protein [2, 13]. Such conditions result in a considerable fraction of unlabelled proteins. The challenge is the separation of the labelled from the unlabelled proteins. Owing to the small mass difference between both fractions, size exclusion techniques are not feasible. Although a separation via hydrophobic interaction chromatography or related methods is possible, distinct conditions are required for every protein, which often also reduce biological activity.

In contrast, the method presented here is universally applicable (Fig. 1). A detailed protocol can be found in the Supporting Information. The fluorescent label is attached to a specific DNA sequence (tag strand: DNA-T), which binds to the protein together with the fluorophore. By hybridization to a solid-phase material, such as magnetic beads, which is functionalized with an oligonucleotide of almost complementary sequence (solid-phase strand, DNA-SP), the labelled proteins can be purified. By removal of the supernatant, unlabelled proteins are separated from the labelled fraction. The labelled proteins are subsequently released from the solid phase by adding another oligonucleotide (displacement strand, DNA-D), which forms a more stable DNA duplex with the surface-bound DNA and thus replaces the DNA that is attached to the protein (see Supporting Information for DNA sequences and duplex stabilities). The purification fractions of such a process are also shown in Fig. 1, demonstrating the efficiency of the separation process. Analysis of the amounts of fluorophore, which were found in the washing and elution fractions or remained on the beads, revealed that the whole process provides recovery rates of more than 90% of the labelled protein. The oligonucleotides were designed in terms of GC-content and sequence complexity to achieve such a good yield. The binding and release process are intimately linked. While the loss of material during the washing could be reduced by using DNA-sequences that bind stronger, concomitantly the yield of recovery would be reduced, since the competition reaction would be affected negatively. To circumvent this, recovery from the solid support could be done by DNA denaturation, for example by use of low salt concentration or heat. However, this often results also in denaturation

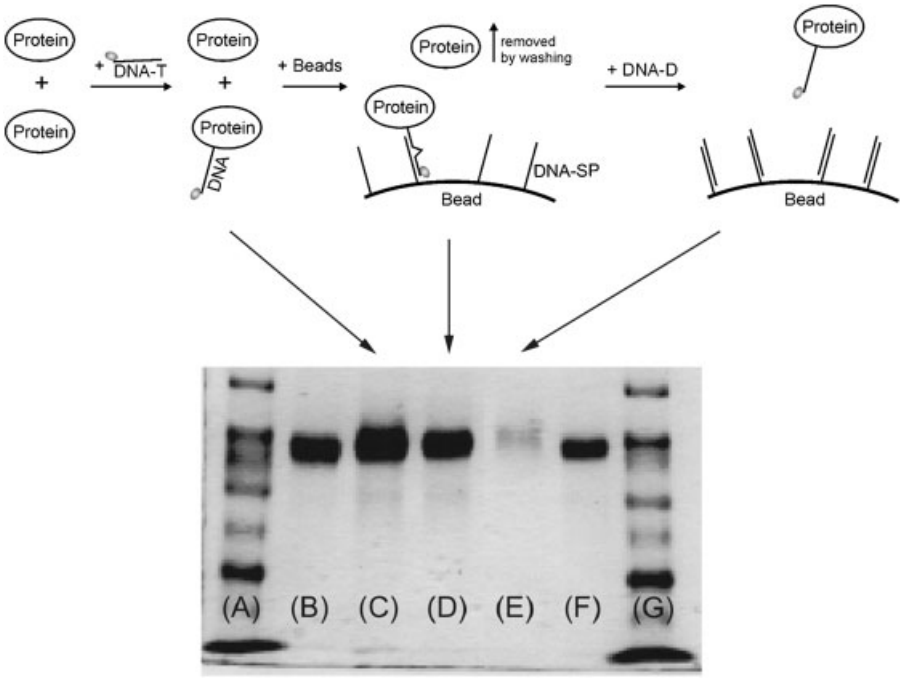

Figure 1. Illustration of the labelling and purification procedure. (Top) After the covalent labelling reaction between the dye-DNA-tag (DNA-T) and an excess of protein, the labelled protein sub-fraction is immobilized on magnetic beads. The beads are functionalized with a DNA of almost complementary sequences (DNA-SP). This enables a gentle removal of the unlabelled protein fraction by washing. Subsequently, DNA sequences (DNA-D) are added, which are the exact complement to the bead sequences, thus displacing the DNA-labelled proteins from the solid phase. (Bottom) Protein fractions obtained during a purification process. The protein obtained at different steps was separated in an SDS-gel electrophoresis and stained with Coomassie-blue. Molecular weight size markers ran in lanes (A) and (G). Unmodified BSA is shown in (B) and (F). In (C), the crude labelling mixture is shown. The protein isolated from the washing supernatant (D) and the purified product recovered from the beads (E) exhibit the expected mobility difference, caused by the T-DNA-fluorophore tag attached to the product.

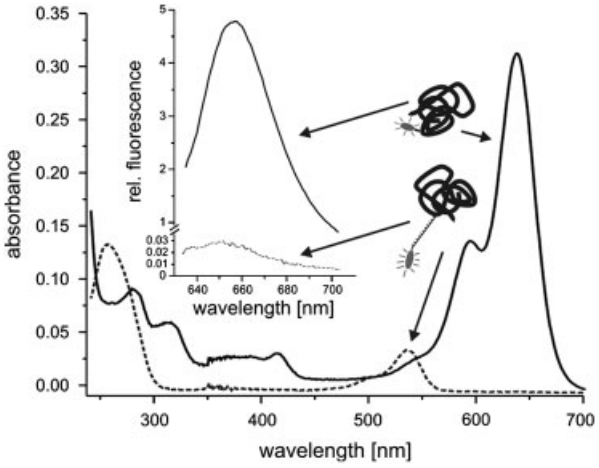

Figure 2. Analysing purification quality. Absorption spectra of the molecules in the supernatant (solid line) and the purified product after elution from the beads (dashed line) are shown. The reaction had been spiked with protein that was labelled directly with the fluorescent dye Atto633 rather than via a DNA-tag. In the inset, the fluorescence spectra of the two molecules types at an excitation at 633 nm are shown.

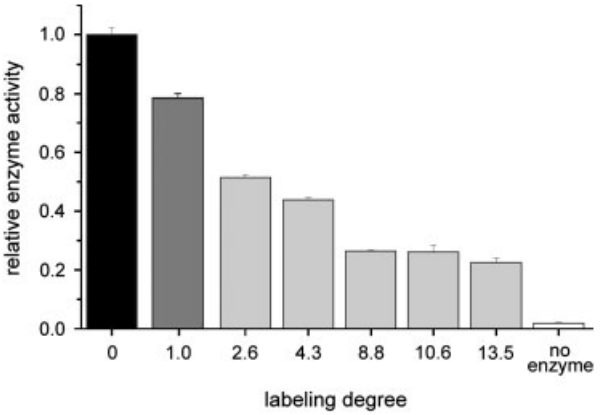

Figure 3. Relative enzyme activity of carboxypeptidase A. The activity was measured on unlabelled (black bar) and stoichiometrically labelled (grey bar) molecules. Also, the activity was determined with enzyme that had been randomly labelled to different degrees (light-grey bars). The amount of enzyme was adjusted to the same level in all assays.

of the proteins. In any case, even better recovery may have been possible by extending the sequence of the DNA-D into the oligo(dA) sequence of the DNA-SP, thereby providing a nucleation point for enhanced hybrid invasion. With an overall yield of 90%, however, such an adaptation was felt to be unnecessary. The yield of the whole labelling and purification process is 3–5% of the overall protein used. This is mainly the result of the fact that an excess of protein is used during the initial labelling step in order to avoid the attachment of more than one dye molecule per protein.

The high specificity of the purification procedure is demonstrated in Fig 2. The labelling reaction mixture was spiked with protein that had been labelled with the red-fluorescent dye Atto633 without a DNA-tag. Compared with the DNA-labelled protein, a five-fold excess of the spike material was present. Whereas the majority of the DNA-tag labelled protein was part of the final product fraction, almost no signal of the spiked BSA can be detected in the purified product. All this protein was found in the supernatant. From these data, it can be calculated that in the normal process $<0.1\%$ of the final product is unlabelled protein, demonstrating the efficiency of the purification protocol.

The low impact of both the fluorophore-DNA-label and the purification procedure on biological activity was demonstrated by comparing the cleavage rates of labelled and unlabelled enzyme. We used two enzymes: *Escherichia coli* β-galactosidase and carboxypeptidase A. After enzymatic cleavage with *E. coli* β-galactosidase, the substrate *o*-nitrophenyl-β-D-galactopyranoside exhibits an increased absorbance at 410 nm. Carboxypeptidase A hydrolyses hippuryl-L-phenylalanin to hippuric acid and phenylalanin, a reaction that can be followed by a change in adsorption at 254 nm. The cleavage rate of both enzymes carrying a fluorophore-DNA-label was about 80% that of an unlabelled enzyme due to the spin column purification (Supporting Information Fig. S1). In contrast, both enzymes exhibited significantly lower activity when labelled fluorescently by standard *N*-hydroxysuccinimide-chemistry; increasing labelling ratios resulted in decreasing cleavage rates. Owing to the Gaussian distribution of random labelling, usually label to protein ratios of five and higher are used in order to avoid a large percentage of unlabelled protein. With a ratio above five, the cleavage rates of both enzymes decreased below 30% (e.g. Fig. 3).

In conclusion, we present a broadly applicable labelling method for the preparation of stoichiometrically 1:1 labelled proteins. Besides achieving stoichiometry, a label ratio of one is bound to preserve the structure and biological activity of the labelled molecules better than the attachment of more dye molecules. If required, however, more than one label could be introduced subsequently in a controlled manner by means of different DNA sequences. Moreover, the system could be used for label molecules other than fluorophores, for instance for the incorporation of biotin. The method could have considerable impact on the quality of quantitative experiments in many fields, ranging from intracellular single-molecule applications to protein microarray analyses. Also, as discussed in the introduction, studies by means of fluorescence spectroscopy on the interaction of different proteins or protein isoforms could be quantified and related limitations in the analysis of posttranslational protein modifications could be circumvented, for example.

The work was funded by the European Union as part of the projects ProteomeBinders, AffinityProteome, and Affinomics.

Achim Friedrich was supported by a fellowship of the Helmholtz Graduate School for Cancer Research.

The authors have declared no conflict of interest.

References

[1] Becker, C. F. W., Seidel, R., Jahnz, M., Bacia, K. et al., C-terminal fluorescence labelling of proteins for interaction studies on the single-molecule level. *Chembiochemistry* 2006, *7*, 891–895.

[2] Hermanson, G. T., *Bioconjugate Techniques*, Elsevier, Heidelberg 2008.

[3] Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., Prasher, D. C., Green fluorescent protein as a marker for gene expression. *Science* 1994, *263*, 802–805.

[4] Muralidharan, V., Muir, T. W., Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods* 2006, *3*, 429–438.

[5] Keppler, A., Kindermann, M., Gendreizig, S., Pick, H. et al., Labelling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. *Methods* 2004, *32*, 437–444.

[6] Amitani, I., Baskin, R. J., Kowalczykowski, S. C., Visualization of Rad54, a chromatin remodeling protein, translocating on single DNA molecules. *Mol. Cell* 2006, *23*, 143–148.

[7] Dillingham, M. S., Wallace, M. I., Protein modification for single molecule fluorescence microscopy. *Org. Biomol. Chem.* 2008, *6*, 3031–3037.

[8] Resch-Genger, U., Grabolle, M., Cavaliere-Jaricot, S., Nitschke, R., Nann, T., Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* 2008, *5*, 763–775.

[9] Hwang, L. C., Wohland, T., Single wavelength excitation fluorescence cross-correlation spectroscopy with spectrally similar fluorophores: resolution for binding studies. *J. Chem. Phys.* 2005, *122*, 114708.

[10] Kosturko, L. D., Maggipinto, M. J., Korza, G., Lee, J. W. et al., Heterogeneous nuclear ribonucleoprotein (hnRNP) E1 binds to hnRNP A2 and inhibits translation of A2 response element mRNAs. *Mol. Biol. Cell* 2006, *17*, 3521–3533.

[11] Moertelmaier, M., Brameshuber, M., Linimeier, M., Schütz, G. J., Stockinger, H., Thinning out clusters while conserving stoichiometry of labelling. *Appl. Phys. Lett.* 2005, *87*, 263903.

[12] Ruprecht, V., Brameshuber, M., Schütz, G. J., Two-color single molecule tracking combined with photobleaching for the detection of rare molecular interactions in fluid biomembranes. *Soft Matter* 2010, *6*, 568–581.

[13] Kozlov, I. A., Melnyk, P. C., Stromsborg, K. E., Chee, M. S. et al., Efficient strategies for the conjugation of oligonucleotides to antibodies enabling highly sensitive protein detection. *Biopolymers* 2004, *73*, 621–630.

