Evaluation of two novel tag-based labelling technologies for site-specific modification of proteins

Aline Tirat, Felix Freuler, Thomas Stettler, Lorenz M. Mayr, Lukas Leder

Novartis Institutes for Biomedical Research, Discovery Technologies, CH-4056 Basel, Switzerland

Received 28 November 2005; received in revised form 23 January 2006; accepted 23 January 2006

Available online 28 February 2006

Abstract

Modern drug discovery strongly depends on the availability of target proteins in sufficient amounts and with desired properties. For some applications, proteins have to be produced with specific modifications such as tags for protein purification, fluorescent or radiometric labels for detection, glycosylation and phosphorylation for biological activity, and many more. It is well known that covalent modifications can have adverse effects on the biological activity of some target proteins. It is therefore one of the major challenges in protein chemistry to generate covalent modifications without affecting the biological activity of the target protein. Current procedures for modification mostly rely on non-specific labelling of lysine or cysteine residues on the protein of interest, but alternative approaches dedicated to site-specific protein modification are being developed and might replace most of the commonly used methodologies. In this study, we investigated two novel methods where target proteins can be expressed in E. coli with a fusion partner that allows protein modification in a covalent and highly selective manner. Firstly, we explored a method based on the human DNA repair protein O6-alkylguanine-DNA alkyltransferase (hAGT) as a fusion tag for site-directed attachment of small molecules. The AGT-tag (SNAP-tag™) can accept almost any chemical moiety when it is attached to the guanine base through a benzyl group. In our experiments we were able to label a target protein fused to the AGT-tag with various fluorophores coupled to O6-benzylguanine. Secondly, we tested in vivo and in vitro site-directed biotinylation with two different tags, consisting of either 15 (AviTag™) or 72 amino acids (BioEase™ tag), which serve as a substrate for bacterial biotin ligase birA. When birA protein was co-expressed in E. coli, biotin was incorporated almost completely into a model protein which carried these recognition tags at its C-terminus. The same findings were also obtained with in vitro biotinylation assays using pure birA independently over-expressed in E. coli and added to the biotinylation reaction in the test tube. For both biotinylation methods, peptide mapping and LC–MS proved the highly site-specific modification of the corresponding tags. Our results indicate that these novel site-specific labelling reactions work in a highly efficient manner, allow almost quantitative labelling of the target proteins, have no deleterious effect on the biological activity and are easy to perform in standard laboratories.

© 2006 Elsevier B.V. All rights reserved.

Keywords: O6-alkylguanine-DNA alkyltransferase; Biotinylation; SNAP-tag™; BioEase™; AviTag™; Site-specific; Protein modification

1. Introduction

Modification of proteins without affecting their biological function remains one of the main challenges faced by protein chemists. Current methods are mostly based on random chemical modification of lysine or cysteine residues on the protein of interest, which usually results in an inhomogeneous protein with sometimes compromised activity. In order to preserve the biological function as much as possible and to obtain a more homogeneously-labelled protein, several novel methods were developed for site-specific modification. Target proteins can be expressed in E. coli with a fusion partner that allows covalent and highly selective modification. We evaluated two of these new tag-based technologies that have superior properties and a broad range of applications.

Abbreviations: AEBSF, 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride; AGT, O6-alkylguanine-DNA alkyltransferase; BG, O6-benzylguanine; BPL, biotin protein ligase; birA, gene for bacterial biotin ligase; DTT, dithiothreitol; E1, ubiquitin activating enzyme; E2, ubiquitin conjugating enzyme; hAGT, O6-alkylguanine-DNA alkyltransferase from human; Hsu, herbesidin; HRP, horseradish peroxidase; IPTG, isopropyl-β-D-thiogalactoside; LC–MS, liquid chromatography-mass spectrometry; MBP, maltose binding protein; NTA, nitrilotriacetic acid; PEG, polyethylene glycol; PVDF, polyvinylidene fluoride; SDS-PAGE, sodium dodecyl sulfate-polycrylamide gel electrophoresis; SUMO, small ubiquitin-related modifier; TAMRA, tetramethylrhodamin; Ub, ubiquitin

Note: This paper was presented at the Challenging Proteins Workshop in Paris, October 17–18, 2005.

Corresponding author. Tel.: +41 61 324 92 31.
E-mail addresses: lukas.leder@novartis.com (L. Leder).
Fig. 1. Reaction scheme of the SNAP-tag. The active site Cys145 residue of the SNAP-tag (engineered form of human AGT) reacts with synthetic 6-benzylguanine substrates carrying a desired label in the para position of the benzyl ring. This reaction results in free guanine and the covalently attached benzyl derivative of the label.

The first method relies on a small protein of about 180 amino acids, called “SNAP-tagTM”. It can be used for covalent and specific linkage of almost any chemical substance to a protein of interest both in vitro and within living cells [1]. The SNAP-tagTM technology is based on the human form of the DNA repair protein Oβ-alkylguanine-DNA alkyltransferase (AGT). The physiological function of AGT is the removal of alkyl groups from guanine bases of DNA to avoid mutations during cell division. AGT undergoes a classical suicide reaction as the cysteine residue in its active site forms a covalent and stable thioether bond with the alkyl group, allowing the release of the de-alkylated, original guanine base [1,2]. This covalent and site-specific modification of AGT was studied in great detail by Prof. Johnsson’s group at EPFL Lausanne, Switzerland. They observed that AGT can accept any chemical compound when it is linked to the guanine base through a benzyl group (Fig. 1). This rather broad substrate specificity of AGT and the fact that the gene for human AGT (hAGT) can be expressed as a soluble protein in various host cells led to the development of the SNAP-tagTM technology. Directed evolution experiments and protein engineering with hAGT resulted in a smaller enzyme with much faster reaction kinetics and devoid of interactions with other biomolecules such as double stranded DNA [3]. The reaction mechanism of the SNAP-tagTM system (Fig. 1) allows the covalent and specific attachment of target proteins to any desired compound or solid surface. This technology, now commercially available from the company Covalys (Witterswil, Switzerland), offers a large number of uses such as in vivo or in vitro labelling, immobilization on a solid phase, or purification of AGT fusion proteins in a functional form [1,4].

The second technique is based on the natural incorporation of biotin molecules into certain proteins. The enzyme biotin protein ligase (BPL) attaches biotin to the biotin carrier domain of certain carboxylases or decarboxylases. BPL catalyzes the post-translational formation of an amide bond between the carboxyl group of biotin and the e-amino group of a specific lysine residue located within the biotin carrier domain in a two step, ATP-dependent reaction (Fig. 2). The biotin carrier domain, usually found at the C-terminus of the (de)carboxylases, contains the specific biotinylated lysine residue within a highly conserved Ala-Met-Lys-Met recognition motif [5]. Used as fusion tags at the C-terminus, the N-terminus or even within the target protein, these biotin carrier domains allow the in vivo or in vitro site-directed biotinylation of recombinant proteins. Together with the avidin/streptavidin technology, they provide various applications for the biotinylated tagged proteins such as detection, purification or attachment to surfaces [6,7]. These tags, called BioEaseTM and AviTagTM, were developed by the companies Invitrogen and Avidity, respectively. The BioEaseTM tag is a 72 amino acid domain derived from the C-terminus

![BioEase](https://example.com/bioease.png)

![AviTag](https://example.com/avitag.png)

Fig. 2. Principle of site-specific biotinylation techniques. The upper panel depicts the reaction scheme for biotinylation catalyzed by biotin protein ligase (BPL) on their natural protein substrates, such as carboxylases. Part of the 72 amino acid BioEase sequence and the full 15 amino acid AviTag peptide are shown in the lower part of the figure. Grey residues correspond to the consensus biotinylation site and the underlined lysines located 35 and 6 amino acids upstream of the C-terminus for BioEase and AviTag, respectively are biotinylated by BPL.
of the *K. pneumoniae* oxalate decarboxylase α subunit [8]. Biotin is covalently attached to this particular subunit, and peptide sequencing has identified a single biotin binding site at lysine 561, which is located 35 residues upstream of the C-terminus (Fig. 2). The entire 72 amino acid domain is required for recognition by the cellular biotinylation enzyme. The AviTag™ contains a 15 amino acid sequence that was discovered by phage display optimization starting from the *E. coli* biotin carboxyl carrier protein (BCCP) subunit of acetyl CoA carboxylase [9,10]. This combinatorial approach was used to identify peptide sequences showing the highest rates of in vivo attachment of biotin. Based on the results of these experiments it was possible to deduce a consensus biotinylation recognition sequence which is displayed in Fig. 2. Interestingly, the two Met residues adjacent to the target lysine residue in the native sequence have been replaced by Gin and Ile. This peptide with a minimal length of 15 aa proved to be highly capable in mimicking the biotin acceptor function of the much larger protein domain normally recognized by the BPL [10]. When fused to a heterologously expressed protein, the BioEase™ and AviTag™ target sequences are recognized and efficiently biotinylated in vivo by the biotin protein ligase birA in *E. coli* [11] or by a similar system in mammalian cells [6]. This method is also suitable for in vitro modifications with purified biotin protein ligase birA which can be easily expressed in *E. coli* biotinylated in vivo by the biotin protein ligase birA in *E. coli* and AviTag™ target sequences are recognized and efficiently biotinylated in vivo by the biotin protein ligase birA in *E. coli* [11] or by a similar system in mammalian cells [6]. This method is also suitable for in vitro modifications with purified biotin protein ligase birA which can be easily expressed in *E. coli* [12].

The SNAP-tag™, BioEase™ and AviTag™ were evaluated as targets for site-specific modification of model proteins in the context of tool production for assay development and HTS. For this purpose, the ubiquitin conjugating enzyme Rad6B was cloned with the SNAP-tag™ fused to its N- or C-terminus. The BioEase™ and AviTag™ sequences were each fused to the C-terminus of the SUMO conjugating enzyme Ubc9. The tagged Rad6B was expressed and purified to perform in vitro labelling with fluorescent benzylguanine derivatives. BioEase™, and AviTag™-tagged Ubc9 were produced and in vivo biotinylated in *E. coli* by co-expression of birA. In parallel, birA and non-biotinylated BioEase™, and Avi-tagged Ubc9 were independently expressed in *E. coli* and purified to perform in vitro biotinylation assays. Finally, modified Rad6B and Ubc9 were characterized by LC-MS and their enzymatic activity was tested by electrophoretic band shift assays with Ub and SUMO as the binding partner, respectively.

### 2. Materials and methods

#### 2.1. Cloning of the expression constructs

Most of the expression constructs were based on the Gateway™ technology from Invitrogen. The proprietary ENTRY vector consisted of a pUC derived backbone with the promoted gene for gentamicin acetyltransferase, and an engineered nucleotide sequence coding for a PreScission (LEVLFQGP) cleavage site located just downstream of the att1 recombination site (Freuler et al., unpublished data). The plasmid for the SNAP-tag™ mutant version [3] was obtained from Covalys, Witterwil, Switzerland and the plasmid containing the BioEase™ was supplied by Invitrogen, Basel, Switzerland. The coding sequences for the two tags were then amplified with “sticky end PCR” described by Zeng [13]. The isolated PCR fragments were then inserted into the EcoRI/NotI restriction sites of the ENTRY vector. In the case of the 15 aa AviTag™ sequence, synthetic oligonucleotides coding for this sequence were directly ligated into the ENTRY vector again by using the EcoRI/NotI restriction sites. The sequences for full-length Rad6B (Swissprot: P23567) and Ubc9 (Swissprot: P63279) were amplified by PCR as described in [14]. The inserts were ligated into the BamHI/NotI restriction sites of the corresponding Gateway ENTRY vector either carrying the SNAP-Taq™ for Rad6B or the BioEase™ and AviTag™ for Ubc9. The correctness of the different ENTRY vectors with the fusion proteins was confirmed by double-strand DNA sequencing. In order to generate the expression plasmids the ENTRY vectors were recombined with engineered proprietary DESTINATION plasmids according to the protocol of the manufacturer Invitrogen. These plasmids are derived from the pET vectors (T7 promoter) and carry a N-terminal MBP-His6-S-tag sequence (MBP = maltose binding protein from *E. coli*) or a GB1-His6-S-tag sequence (GB1 = 56 amino acid stretch derived from protein G from *Streptococcus sp*) upstream of the at1 recombination site. This step resulted in four different recombinant expression plasmids: MBP-His6-S-AGT-Rad6B, MBP-His6-S-Rad6B-AGT, MBP-His6-S-Ubc9-AviTag™, and GB1-His6-S-Ubc9-BioEase. The PreScission cleavage site was always located downstream of the S-tag and just upstream of the Rad6B or Ubc9 fusion protein. The gene for the *E. coli* endogenous biotin holoenzyme ligase (birA) was amplified with sticky-end PCR and ligated into the pET26 expression vector (Novagen, Lucerne, Switzerland) by using the NcoI/NotI restriction sites. We generated an expression vector for birA flanked by a T7 promoter, a lac operator, a T7 terminator and the repressor lacI. This cassette was then transferred to the vector pACYC184 (New England Biolabs, Allschwill, Switzerland), which has an alternative origin of replication allowing the co-expression of the protein in *E. coli*. The same procedure was also repeated with pBADara (Invitrogen) in order to have an expression vector under the control of the arabinoose promoter. Finally, birA was also cloned into the initial Gateway™ ENTRY vector with the PreScission site downstream of the at1 recombination site. This birA ENTRY vector was then recombined again with the DESTINATION plasmid carrying the MBP-His6-S tags.

#### 2.2. Protein expression and purification for in vitro labelling and biotinylation assays

For the production of N- and C-terminally AGT-tagged Rad6B, birA, Ubc9-BioEase and Ubc9-Avi, cells of the *E. coli* expression strain BL21(DE3) Tuner (Novagen) were transformed with the corresponding expression plasmids. The cells were grown at 37 °C in Terrific Broth (TB, complemented with 0.1 M MOPS buffer pH 7.0) containing 30 μg/ml kanamycin in 21 shaker flasks at 200rpm to an OD595 of 0.8. In the case of birA expression, the TB medium was complemented with 20 μM β-biotin (Sigma) in order to enhance the expression and the solubility of the BPL [12]. After induction with 0.1 mM isopropyl-β-thiogalactoside (IPTG) the cells were further
incubated for 16 h at 20 °C and harvested by centrifugation. 30 min prior to harvesting, lysozyme (Roche Diagnostics, Rotkreuz, Switzerland) was added to a final concentration of 50 μg/ml culture. The cells were resuspended in lysis buffer (50 mM Na-phosphate, 300 mM NaCl, 20 mM imidazole pH 8.0) and lysed by two passages on a French Press (EmulsiFlex, Avestin Inc.), after 0.5 mM Pefabloc® (Roche, Basel, Switzerland) was added to a final concentration of 50 μM benzamidine (Merk, Darmstadt, Germany) had been added. The tagged Rad6B fusion proteins used for in vitro labelling assays (MBP-His6-S-AGT/Rad6B and MBP-His6-S-Rad6B-AGT) were purified over a HiTrap HP 5 ml column (GE Healthcare, Freiburg, Germany) installed on an AKTA 100 Explorer (GE Healthcare) chromatography system. After loading of the lysates, the column was washed with five column volumes of lysis buffer and the His6-tag containing fusion proteins were eluted by increasing the concentration of imidazole to 300 mM. The pooled fractions containing the target protein were further incubated with PreScission protease (GE Healthcare) overnight at 4 °C to cleave of the MBP-His6-S fusion tag. Separation of the cleaved tag was achieved by reverse HiTrap chromatography following the same procedure as described above. Since the proteins found in the flow-through were not pure enough (about 80%), the reverse HisTrap protein pool was concentrated and loaded onto a Superdex 75 HiLoad 26/60 size exclusion (GE Healthcare) column equilibrated with 1 × PBS pH 7.4. BirA and Ub9 proteins used for in vitro biotinylation assays (MBP-His6-S-birA, GB1-His6-S-Ubc9/BioEase and MBP-His6-S-Ubc9-Avi) were purified on an AKTA Explorer 100 system equipped with automated protein purification AKTA 3D Kit (GE Healthcare). This set-up allowed successive loading, cleavage and purification of the proteins following the sequence: Ni-cholate affinity – column cleavage with PreScission protease – gel filtration. The affinity step was performed on a 5 ml HiTrap HP column and the size exclusion chromatography on a Superdex 75 HiLoad 16/60 column. All protein preparations were characterized by mass spectrometry and found to contain target protein with the correct mass in a purity >90%.

2.3. SNAP-tag™ site-specific in vitro labelling procedure

After expression and purification, AGT-Rad6B and Rad6B-AGT were labelled following the SNAP-tag™ principle depicted in Fig. 1. The AGT-tagged proteins were pre-incubated at a final concentration of 10 μM in 100 mM HEPES buffer (pH 7.5) supplemented with 1 mM DTT for 10 min at room temperature in order to fully reduce the active cysteine residue of AGT involved in the labelling reaction. A 1.5-fold excess of benzylguanine derivative (15 μM) was added to the mixture, which was further incubated for one hour at room temperature. SDS-PAGE and RP-HPLC analysis performed at different incubation times were used to follow the incorporation over the course of the reaction. After labelling, the AGT-tagged proteins were purified by loading the reaction mixture onto either a gel filtration (Superdex 75 16/60, GE Healthcare) or a desalting column (PD10 Sephadex G 25M, GE Healthcare). Fractions were checked by SDS-PAGE and/or RP-HPLC, those containing the labelled AGT-tagged Rad6B were pooled and further characterized.

Once labelled and purified, the AGT-tagged Rad6B proteins were characterized by mass spectrometry, peptide map and a SDS-PAGE band shift assay. In order to confirm the specificity of the modification reaction, labelled proteins were digested by addition of trypsin (1.50 w/w protease/substrate, Roche Diagnostics) and digestion mixtures were incubated overnight at 37 °C. To check the digestion efficiency, mixtures were analyzed by RP-HPLC on a C8 Reprosil column (Morvay Analytics, Basel, Switzerland) before being further analyzed by LC–MS. The functional activity of the labelled ubiquitin conjugating enzyme Rad6B was finally verified by an SDS-PAGE band shift assay. After ATP-dependent activation by the ubiquitin activating enzyme E1, the formation of a thioester bond between the C-terminal carboxylate of Ub and the active site cysteine of Rad6B was monitored by SDS-PAGE analysis under reducing and non-reducing conditions. Based on an in-house protocol [14], the reaction was run at room temperature for 10 min with in a volume of 10 μl. The mixture contained 1 μg ubiquitin activating enzyme GST-E1, 1 μg Rad6B and 2 μg Ub in 50 mM HEPES pH 8.0, 50 mM NaCl, 5 mM MgCl2, 4 mM ATP and 0.25 mM DTT. The reaction was stopped by addition of 2 × NuPAGE sample buffer (Invitrogen, Basel, Switzerland) with or without 100 mM DTT. Samples were loaded and run at 200 V constant voltage for 35 min on a 4–12% Bis-Tris NuPAGE gel. The fluorescent bands on the gel were detected with a gel documentation system by using a rhodamine filter. The gel was finally stained with Coomassie "Simply Blue SafeStain" solution (Invitrogen).

2.4. BioEase™ and AviTag™ site-specific in vitro biotinylation procedure

The in vitro biotinylation experiments were performed according to Cull and Schatz [15]. Prior to use, proteins were dialyzed against 50 mM Tris–HCl pH 8.0. The reaction mixture (volume of 1 ml) consisted of 40 μM BioEase™ or AviTag™-tagged Ubc9 and 1.50 w/w (enzyme/substrate) birA in 50 mM Tris–HCl pH 8.0, 10 mM MgCl2, 10 mM ATP and 50 μM ubiquitin. The mixture was incubated for 4 h at 30 °C with shaking at 400 rpm in an Eppendorf Thermomixer (Hamburg, Germany). In order to follow the progress of the biotinylation reaction, 10 μl samples (after 30, 60, 120 and 240 min of incubation) were loaded onto a 4–12% Bis-Tris NuPAGE gel. After electrophoresis the gels were blotted onto PVDF membranes (Invitrogen) at 30 V constant voltage for 1 h by using Xcell II Blot module and NuPAGE Transfer buffer (Invitrogen). Membranes were blocked for 30 min in 5% Blottng Grade Blocker Non-Fat Dry Milk (BioRad, Basel, Switzerland) dissolved in 1 × PBS pH 7.4 and for further 30 min in 1 × PBST (1 × PBS pH 7.4, 0.05% Tween20) supplemented with 1% BSA (Sigma). The Streptavidin-horseradish peroxidase conjugate (Sigma) was directly diluted 1:1000 in the BSA blocking solution. Membranes were incubated for 1 h with the conjugate and washed in 1 × PBST. The blots were finally developed with 4-chloro-l-naphtol as HRP substrate. In parallel, 450 μl samples (after 60 and 240 min of incubation) were loaded onto a HiTrap HP 5 ml column.
desalting column (Amersham Biosciences) which allowed the separation of biotinylated Ubc9 from free biotin. Fractions containing the biotinylated BioEase\textsuperscript{TM} or Avi-tagged Ubc9, were subjected to LC–MS analysis in order to determine the degree of modification.

2.5. In vivo biotinylation

Small-scale expressions and Western blot analysis

Ubc9-BioEase and Ubc9-Avi were cloned and expressed in E. coli with the GB1- and MBP-His6-S tag at their N-terminus, respectively. Several expression conditions were tested in small-scale samples consisting of 2 ml bacterial cultures in 48 polypropylene well blocks. Luria Bertani (LB) medium supplemented with 30 mg/ml kanamycin and 34 mg/ml chloramphenicol (when the biA plasmid was co-expressed) was inoculated with overnight pre-cultures (1:10 dilution) of BL21(DE3) Tuner (Novagen) and Ubc9-Avi eluted from the column during the gradient. Once in vivo biotinylated and purified, the BioEase\textsuperscript{TM} and AviTag\textsuperscript{TM} fusion proteins were cleaved with PreScission protease and further purified by cation exchange chromatography (Resource S 6 ml; GE Healthcare) The column had been equilibrated with 30 mM Na-phosphate, 30 mM Na-formate and 60 mM Na-acetate buffer at pH 8.0. The proteins were eluted with a linear salt gradient ranging from 0 to 1 M NaCl within 20 column volumes. As expected, the GB1- and MBP-His6-S tags appeared in the flow-through, whereas cleaved Ubc9-BioEase and Ubc9-Avi eluted from the column.

2.6. In vivo site-specific biotinylated protein expression, purification and characterization

For the production of in vivo biotinylated Ubc9-BioEase and Ubc9-Avi, cells of the E. coli BPL birA and further co-transformed with the co-expression plasmids encoding GB1-His6-S-Ubc9-BioEase or MBP-His6-S-Ubc9-Avi, respectively. For expression, we used LB medium supplemented with 30 mg/ml kanamycin, 34 mg/ml chloramphenicol and 50 mg/ml tetracycline. Cells were grown in shaker flasks at 200 rpm until they reached an OD\textsubscript{600} of 0.8. After induction with 0.1 mM IPTG the cells were further incubated for 5 h at 25 °C. Harvesting of the cells, lysis of cell pellets and the first affinity chromatography step purification were performed as described above. After the HisTrap column, the BioEase\textsuperscript{TM} and AviTag\textsuperscript{TM} fusion proteins were cleaved with PreScission protease and further purified by cation exchange chromatography (Resource S 6 ml; GE Healthcare) The column had been equilibrated with 30 mM Na-phosphate, 30 mM Na-formate and 60 mM Na-acetate buffer at pH 8.0. The proteins were eluted with a linear salt gradient ranging from 0 to 1 M NaCl within 20 column volumes. As expected, the GB1- and MBP-His6-S tags appeared in the flow-through, whereas cleaved Ubc9-BioEase and Ubc9-Avi eluted from the column during the gradient. The functional activity of the biotinylated SUMO conjugating enzyme Ubc9 was finally verified by a band shift assay. After ATP-dependent activation by the heterodimeric SUMO activating enzyme SAE1/SAE2\textsuperscript{[16]}, the reaction was incubated overnight at 37°C and further analyzed by LC–MS. The AGT substrate provided for the reaction was [3H]SUMO (PerkinElmer) and 2 mg/ml lysozyme (Roche Diagnostics) were added to the wells. The plate was incubated for 20 min with mixing (200 rpm) and soluble fractions were collected in a receiver plate by vacuum filtration. The same filter plate wells were filled with 170 μl of denaturing buffer (8 M urea, 50 mM Tris-Cl, 100 mM NaCl, 0.05% Tween20, pH 8.0). After 15 min incubation with mixing (200 rpm), vacuum filtration was repeated and non-soluble fractions were collected in a second receiver plate. Soluble and insoluble fractions were finally loaded on SDS-PAGE gels and processed as described above. The biotin incorporation of the proteins in the soluble lysates was additionally checked by Western blot with the streptavidin-HRP conjugate for detection of biotinylated proteins.

3. Results and discussion

3.1. Site-specific labelling with SNAP-tag\textsuperscript{TM} technology using fluorescent BG derivatives

We were able to express and purify Rad6B fused to the SNAP-tag\textsuperscript{TM} without any major impact of hAGT on expression levels or solubility. The two fusion constructs AGT-Rad6B and Rad6B-AGT could be specifically modified following the SNAP-tag\textsuperscript{TM} principle (Fig. 1). The AGT substrate provided by Covalys was a TAMRA fluorophore linked to benzylguanidine through a polyethylene glycol (PEG) spacer[1,16]. Using this substrate, both N- and C-terminally tagged Rad6B constructs were efficiently labelled. The final yield was about 1 mg of purified TAMRA-AGT-Rad6B or Rad6B-AGT-TAMRA when starting with 2 mg non-modified protein. Fig. 3 shows the results for the modification and characterization from AGT-Rad6B with BG-PEG-TAMRA. Results for the other construct (Rad6B-AGT) were very similar (data not shown). LC–MS characterization of both proteins demonstrated that only one
TAMRA-PEG molecule had been incorporated in both variants of the AGT/Rad6B fusion proteins. In the case of Rad6B-AGT-TAMRA, about 20% of the material consisted of C-terminally truncated forms. It has been described that covalently modified AGT undergoes a slight conformational change, leading to increased susceptibility towards proteolysis especially at the C-terminal part of the enzyme [17]. We therefore conclude that the AGT tag should preferably be fused to the N-terminus of the target protein to avoid generation of these C-terminally heterogeneous forms. Based on the natural AGT reaction, the cysteine residue 145 (numbering of native AGT) forms the covalent thioether bond with benzylguanine substrates [1]. In order to verify that the modification really occurred at this active cysteine residue, a peptide map was generated by digestion with trypsin. Results of RP-HPLC analysis (Fig. 3) showed that only one labelled fragment was produced and LC–MS analysis confirmed that this fragment was corresponding to the digested peptide containing the cysteine 145 of AGT.

We were able to prove the high specificity of labelling by using the SNAP-tag™ technology. Modification of recombinant proteins with the SNAP-tag technology offers several advantages. Firstly, the reaction is easily performed by mixing the AGT tagged target protein with the benzylguanine derivative substrate within the appropriate buffer. Moreover, only a slight excess of labelling reagent was necessary for the modification (1.5 times molar excess) whereas 3–10-fold molar excess is usually needed for other labelling procedures. Finally, less than one hour of incubation time was enough to achieve a nearly 100% incorporation of the label, due to the specificity and fast kinetics of the reaction between the AGT tag and its substrate. In comparison to other labelling protocols that require several hours or even overnight incubation, the SNAP-tag™ technology saves time and clearly diminishes the exposure of the protein to potentially unfavourable conditions. The enzymatic activity of modified AGT-Rad6B was tested with a band shift assay to check if the labelled AGT fusion tag was interfering with the function of the ubiquitin conjugating enzyme Rad6B. The AGT tagged Rad6B was still able to form the thioester bond with the C-terminus of ubiquitin in the presence of the ubiquitin activating enzyme E1 and ATP as depicted in Fig. 4. Since the thioester is disrupted upon addition of DTT [18], the reaction mixtures were analysed under reducing and non-reducing conditions. GST-E1, AGT/Rad6B labelled with TAMRA-PEG and ubiquitin showed the expected enzymatic activity in the gel-based thioester bond formation assay. The Rad6 ubiquitin complex was present under non-reducing conditions (lanes 2 and 8 in Fig. 4), whereas
addition of DTT to the sample buffer completely disrupted the thioester bond of the complex (lanes 2 and 8 in Fig. 4). This experiment clearly demonstrated that Rad6B fully retained its functional activity with the AGT tag fused to its C-terminus. A second AGT substrate consisting of an Alexa555 fluorophore bound to benzylguanine without any spacer [16] was also provided by Covalys and tested in new labelling experiments with AGT-Rad6B. This alternative substrate gave very similar results compared to the TAMRA substrate underlining the broad versatility of the SNAP-tag™ technology (data not shown).

### 3.2. In vivo biotinylation small-scale expressions analysis

For exploration of the related BioEase™ and AviTag™ techniques for site-specific biotinylation, we selected the SUMO conjugating enzyme Ubc9 as a model protein. Ubc9 was cloned into bacterial expression vectors with the biotinylation recognition tags fused to its C-terminus. A second AGT substrate consisting of an Alexa555 fluorophore bound to benzylguanine without any spacer [16] was also provided by Covalys and tested in new labelling experiments with AGT-Rad6B. This alternative substrate gave very similar results compared to the TAMRA substrate underlining the broad versatility of the SNAP-tag™ technology (data not shown).

#### 3.3. In vivo biotinylation large-scale experiments by co-expression of birA from pACYC184 vector

As the previously obtained results were only of qualitative nature, production of the proteins was performed with larger cell cultures to further investigate biotinylation efficiency and specificity of each tag. GB1-His-S-Ubc9-BioEase or MBP-His-S tagged Ubc9-Avi. The two different N-terminal tags were chosen because they were most suited for separation after cleavage of the fusion protein with PreScission protease. Expression of the plasmids was done in parallel in BL21(DE3) Tuner cells and in an identical strain previously transformed with a birA (E. coli) gene expression vector containing the usual T7lac promoter and a chloramphenicol resistance marker. LB medium was either supplemented or not with 50 μM biotin and cells were either induced or not by addition of 0.1 mM IPTG. The biotinylation efficiency in each expression condition was determined with Western blot analysis. As shown in Fig. 5, the highest biotin incorporation was obtained for both constructs when E. coli birA had been co-expressed, the cells had been induced with IPTG and 50 μM biotin had been added to the LB medium (lane 1 on left and right panels). Nevertheless, we also observed a reasonable degree of biotinylation without co-expression of the birA enzyme (lanes 5 on left and right panels). Obviously, endogenous biotin protein ligase of E. coli was present in sufficient amounts to catalyze the in vivo biotinylation reaction. In contrast, the corresponding experiments performed without additional biotin in the medium (lanes 2 and 6 on left and right panels) demonstrated clearly that protein modification was lowered. The bacteria were probably not able to cover the needs for in vivo biotinylation of our target protein by endogenous synthesis of biotin.

![Fig. 4. Activity test of labelled AGT-tagged Rad6B with an SDS-PAGE band shift assay. Panel A depicts the reaction scheme between Ub, E1 and E2s. Panel B shows the SDS-PAGE analysis of activity experiments performed with reaction mixtures containing GST-E1 (~150 kDa), Alexa488-Ub (~9 kDa) made in-house [14] and labelled Rad6B-AGT or AGT/Rad6B (~40 kDa). Lanes M: Novex Mark 12 molecular weight marker (Invitrogen), lanes 1 and 2: activity test of Rad6B-AGT-TAMRA with Alexa488-Ub, lanes 3 and 4: Rad6B-AGT-TAMRA reaction mixture without Alexa488-Ub, lanes 5 and 6: Alexa488-Ub reaction mixture without Rad6B-AGT-TAMRA, lanes 7 and 8: activity test of TAMRA-AGT-Rad6B with Alexa488-Ub, lanes 9 and 10: TAMRA-AGT-Rad6B reaction mixture without Alexa488-Ub. Lanes with odd numbers: mixtures with 100 mM DTT (disruption of the conjugate between Rad6B and Ub) and lanes with even numbers: mixtures without DTT.](image-url)
responds to the expected value for biotinylated Ubc9-Avi and indicates a biotin incorporation of almost 100%. The higher efficiency of the AviTag™ compared to the BioEase™ tag is probably linked to its optimized birA recognition sequence. As described previously the 15 amino acids sequence of the AviTag was derived from phage display experiments and is therefore quite different from the biotin ligase interaction motifs present in natural biotin carrier domains like the BioEase™ tag. We also proved the high specificity of the modification with the peptide map obtained after digestion of both tagged proteins with LysC protease. LC–MS analysis of digestion mixtures identified the expected biotinylated peptides with a molecular weight of 4'449.50 Da for Ubc9-BioEase and 2'669.34 Da for Ubc9-Avi. These results demonstrated that biotinylation occurred only at the designated lysine residues 202 (BioEase™) and 174 (AviTag™) within the birA recognition sequences. In order to check if the biotinylated BioEase™ and Avi™ fusion tags were interfering with the function of SUMO conjugating enzyme Ubc9, the enzymatic activity was tested with electrophoretic band shift assays. The tagged Ubc9 proteins were still able to form a thioester bond with the C-terminus of SUMO in presence of the heterodimeric SUMO activating...

Fig. 5. Small-scale experiments of BioEase™ and Avi™-tagged Ubc9 analysed by Western blot. The left and right panels correspond to the blot analysis of GB1-His-S-Ubc9-BioEase and MBP-His-S-Ubc9-Avi experiments, respectively. Lane M: SeeBlue® Pan2 Novex molecular weight marker, lanes 1 and 2: induced BL21(DE3) Tuner cells co-transformed with birA plasmid, lanes 3 and 4: non-induced BL21(DE3) Tuner cells co-transformed with birA plasmid, lanes 5 and 6: induced BL21(DE3) Tuner cells, lanes 7 and 8: non-induced BL21(DE3) Tuner cells and lane +: biotinylated SUMO control. Lanes with odd numbers correspond to cultures made in presence of biotin in LB medium and lanes with even numbers to cultures made in absence of biotin in LB medium.

Fig. 6. Activity tests of in vivo biotinylated BioEase™ and Avi™-tagged Ubc9 with an SDS PAGE band shift assay. Panel A depicts the reaction between SUMO, SAE1/SAE2 and Ubc9. Panel B shows the SDS-PAGE analysis of activity experiments performed with reaction mixtures containing SAE1/SAE2 heterodimer (~39 and ~71 kDa, respectively), SUMO (~11 kDa) and Ubc9-His (~19 kDa) made in-house and biotinylated Ubc9-BioEase or Ubc9-Avi (~26 and ~20 kDa, respectively). Lanes M: Novex Mark 12 molecular weight marker (Invitrogen), lanes 1 and 6: positive control activity test with Ubc9-His, lanes 2 and 7: activity test of biotinylated Ubc9-Avi, lanes 3 and 8: biotinylated Ubc9-Avi reaction mixture without SAE1/SAE2, lanes 4 and 9: activity test of biotinylated Ubc9-BioEase, lanes 5 and 10: biotinylated Ubc9-BioEase reaction mixture without SAE1/SAE2. Lanes 1–5: mixtures with 100 nM DTT (disruption of the conjugate between Ubc9 and SUMO) and lanes 6–10: mixtures without DTT.
enzyme SAE1/SAE2 and ATP (Fig. 6). Since the thioester is disrupted upon addition of DTT (analogous to Ub-E2 complexes [18]), all reaction mixtures were analysed under reducing and non-reducing conditions. Biotinylated Ubc9-BioEase and Ubc9-Avi showed the expected enzymatic activity in the gel-based thioester bond formation assay: the Ubc9/SUMO conjugates were present under non-reducing conditions (lanes 6, 7 and 9), whereas addition of DTT in the sample buffer completely disrupted the thioester bond of the conjugates (lanes 1, 2 and 4). Moreover, the SUMO activating enzyme SAE1/SAE2 was absolutely required for formation of the complex between Ubc9 and SUMO as shown in lanes 8 and 10. Like with the ubiquitin conjugation enzyme Rad6B fused to the SNAP-tagTM, we were able to prove that Ubc9 fully retained its functional activity with the biotinylated BioEaseTM and AviTagTM extensions fused to its C-terminus.

3.4. In vivo biotinylation large-scale by co-expression of birA under control of inducible araBAD promoter

Experiments were performed with a second birA construct containing an araBAD promoter instead of the usual T7lac promoter where expression is induced and regulated with arabinose instead of IPTG. This construct was used to potentially increase the soluble form of the biotin protein ligase birA in the cells by the more stringent control of expression levels due to the tunable activity of the araBAD promoter. More soluble birA should also improve the in vivo biotinylation efficiency of the BioEaseTM tag. To test this hypothesis, we performed large-scale experiments with GB1-His-S-Ubc9-BioEase and MBP-His-S-Ubc9-Avi expressed in parallel in BL21(DE3) Tuner cells co-transformed with pACYC184 birA expression vector pACYC184 containing either the T7lac or the araBAD promoter. Cells transformed with the plasmid pACYC184 araBAD were induced for the expression of AviTagTM and BioEaseTM tagged Ubc9 with 0.1 mM IPTG as previously described, whereas birA expression was initiated from the beginning by addition of 2 mg/ml L-arabinose directly to LB medium. Biotinylated GB1-His-S-Ubc9-BioEase and MBP-His-S-Ubc9-Avi were purified with a single Ni-chelate affinity chromatography step. The non-cleaved fusion proteins were characterized by LC–MS in order to determine the degree of biotinylation for each tag and co-expression approach. Table 1 shows that the efficiency of the in vivo biotinylation was significantly increased when tagged proteins were co-expressed with the new birA construct containing the araBAD promoter compared to the T7 promoter. We especially improved the biotin incorporation for the BioEaseTM tag. Only a small fraction (<5%) of Ubc9-Avi and Ubc9-BioEase could be attributed to the expected biotinylated GB1-His-S-Ubc9-BioEase (37 641.4 Da) when the target protein was co-expressed with birA under control of the T7 promoter. In contrast, the co-expression under control of araBAD gave a biotin incorporation of almost 100 % with only one detected mass of 37 641.4 Da.

3.5. In vitro biotinylation assays

In order to verify if in vitro biotinylation of BioEaseTM and AviTagTM by the BPL birA would be as efficient as the in vivo reaction, birA was expressed in E. coli and purified. New batches of non-biotinylated BioEaseTM and Avi-tagged Ubc9 were produced as substrates for in vitro biotinylation experiments. Prior to use, Ubc9-BioEase and Ubc9-Avi were characterized by LC–MS to determine their degree of biotinylation. As expected, only a small fraction (<5%) of Ubc9-Avi and Ubc9-BioEase was found to be biotinylated. The reaction mixtures were prepared and a time course of the in vivo biotinylation

<table>
<thead>
<tr>
<th>Table 1</th>
<th>LC–MS characterization of GB1-His-S-Ubc9-BioEase and MBP-His-S-Ubc9-Avi</th>
</tr>
</thead>
<tbody>
<tr>
<td>GB1-His-S-Ubc9-BioEase expressed in BL21(DE3) Tuner cells co-transformed with T7 lac promoter</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>Measured by LC–MS</td>
</tr>
<tr>
<td>37 641.6 Da</td>
<td>37 641.6 Da</td>
</tr>
<tr>
<td>Non-biotinylated protein</td>
<td>Expected non-biotinylated protein (first main peak)</td>
</tr>
<tr>
<td>37 641.6 Da</td>
<td>37 641.6 Da</td>
</tr>
<tr>
<td>Biotinylated protein</td>
<td>Expected biotinylated protein (second main peak)</td>
</tr>
<tr>
<td>37 641.6 Da</td>
<td>37 641.6 Da</td>
</tr>
<tr>
<td>GB1-His-S-Ubc9-BioEase expressed in BL21(DE3) Tuner cells co-transformed with araBAD promoter</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>Measured by LC–MS</td>
</tr>
<tr>
<td>37 641.6 Da</td>
<td>37 641.6 Da</td>
</tr>
<tr>
<td>Non-biotinylated protein</td>
<td>Expected non-biotinylated protein (first main peak)</td>
</tr>
<tr>
<td>37 641.6 Da</td>
<td>37 641.6 Da</td>
</tr>
<tr>
<td>Biotinylated protein</td>
<td>Expected biotinylated protein (second main peak)</td>
</tr>
<tr>
<td>37 641.6 Da</td>
<td>37 641.6 Da</td>
</tr>
<tr>
<td>MBP-His-S-Ubc9-Avi expressed in BL21(DE3) Tuner cells co-transformed with T7 lac promoter</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>Measured by LC–MS</td>
</tr>
<tr>
<td>66 750.4 Da</td>
<td>66 750.4 Da</td>
</tr>
<tr>
<td>Non-biotinylated protein</td>
<td>Expected non-biotinylated protein (second main peak)</td>
</tr>
<tr>
<td>66 750.4 Da</td>
<td>66 750.4 Da</td>
</tr>
<tr>
<td>Biotinylated protein</td>
<td>Expected biotinylated protein (second main peak)</td>
</tr>
<tr>
<td>66 750.4 Da</td>
<td>66 750.4 Da</td>
</tr>
<tr>
<td>MBP-His-S-Ubc9-Avi expressed in BL21(DE3) Tuner cells co-transformed with araBAD promoter</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>Measured by LC–MS</td>
</tr>
<tr>
<td>66 750.4 Da</td>
<td>66 750.4 Da</td>
</tr>
<tr>
<td>Non-biotinylated protein</td>
<td>Expected non-biotinylated protein (second main peak)</td>
</tr>
<tr>
<td>66 750.4 Da</td>
<td>66 750.4 Da</td>
</tr>
<tr>
<td>Biotinylated protein</td>
<td>Expected biotinylated protein (second main peak)</td>
</tr>
<tr>
<td>66 750.4 Da</td>
<td>66 750.4 Da</td>
</tr>
</tbody>
</table>
reaction was conducted. Two different analysis formats were used to determine the in vitro biotinylation efficiency of birA. First we checked the results of the kinetics by Western blot analysis (results not shown). No biotinylation of Ub9-BioEase and Ub9-Avi could be detected in the samples at the beginning of the reaction because birA had not yet been added to the mixtures. Biotinylated proteins were then observed with bands of similar intensities at 30, 60, 120 and 240 min incubation time. Apparently, an incubation time of 30 min was sufficient to reach the maximal yield of modification. For more quantitative information, the reaction mixtures were subjected to LC–MS in order to quantify birA incorporation. The analysis was performed after two incubation times (60 and 240 min) and only the LC–MS results obtained after 60 min are shown in Table 2 because they were quite similar for both incubation times. LC–MS analysis confirmed that biotin incorporation was almost identical for the two reaction times. Both proteins were biotinylated nearly to 100%, demonstrating the high in vitro activity of recombinant birA. Therefore, purified birA could be used to attach biotin molecules to target proteins containing either the AviTag™ or BioEase™ tags if in vivo biotinylation is not feasible or yields a low degree of modification.

4. Conclusion

With the growing necessity to develop a large panel of tools for the study of proteins and their interactions, we decided to investigate the new SNAP-tag™, BioEase™ and AviTag™ systems as a novel approach for in vivo or in vitro site-specific modification of proteins. The current paper describes that we could efficiently express and specifically label the ubiquitin-conjugating enzyme Ubc9 with the BioEase™ or the AviTag™ fused to its C-terminus. We proved by a SDS-PAGE band shift assay that the integrity of our protein of interest was preserved and labelled Rad6B or biotinylated Ub9 kept their enzymatic activity. As already stated above, these technologies combine several advantageous features such as the ease of application, speed of the procedure, as well as a high efficiency and specificity of the protein modification without any noticeable negative effect on the tagged protein of interest. Concerning the SNAP-tag™ technology, our feasibility study comprised expression of AGT-tagged proteins in E. coli, and the in vitro labelling aspect of the technology. However, the method has general applicability and versatility as in vivo modifications in different host cells using a broad variety of labels or immobilization supports were already successfully performed and reported [1,3]. We evaluated the expression and in vivo biotinylation of BioEase™ and Avi-tagged proteins in E. coli, nevertheless in vivo modifications in different host cells were already tested [6,19]. As we demonstrated, the BioEase™ and AviTag™ systems are also suitable for in vitro biotinylation catalyzed by birA which can be obtained from recombinant expression in E. coli [12]. Eventually, the AviTag™ is our preferred choice for site-specific biotinylation of proteins because of its smaller size and its somewhat higher efficiency for in vivo biotinylation. Taken together, we expect that these technologies for protein modification will have a major impact on modern protein chemistry and will enable many novel applications with labelled proteins in the field of basic and applied biomedical research.

Acknowledgement

We would like to thank Dr. Jutta Blank and Dr. Fraser Glickman for carefully reading the manuscript and bringing up highly valuable input to these studies.

References