



The SNAP-tag technology revised: an effective *chemo-enzymatic approach* by using a universal azide-based substrate

Rosa Merlo , Diego Caprioglio , Michele Cillo , Anna Valenti , Rosanna Mattosovich , Castrese Morrone , Alberto Massarotti , Franca Rossi , Riccardo Miggiano , Antonio Leonardi , Alberto Minassi & Giuseppe Perugino

To cite this article: Rosa Merlo , Diego Caprioglio , Michele Cillo , Anna Valenti , Rosanna Mattosovich , Castrese Morrone , Alberto Massarotti , Franca Rossi , Riccardo Miggiano , Antonio Leonardi , Alberto Minassi & Giuseppe Perugino (2021) The SNAP-tag technology revised: an effective *chemo-enzymatic approach* by using a universal azide-based substrate, Journal of Enzyme Inhibition and Medicinal Chemistry, 36:1, 85-97, DOI: [10.1080/14756366.2020.1841182](https://doi.org/10.1080/14756366.2020.1841182)

To link to this article: <https://doi.org/10.1080/14756366.2020.1841182>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 29 Oct 2020.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

SHORT COMMUNICATION



The SNAP-tag technology revised: an effective *chemo-enzymatic* approach by using a universal azide-based substrate

Rosa Merlo^{a*}, Diego Caprioglio^{b*}, Michele Cillo^{c*}, Anna Valenti^a, Rosanna Mattosovich^a, Castrese Morrone^b, Alberto Massarotti^{b,d}, Franca Rossi^b, Riccardo Miggiano^{b,d}, Antonio Leonardi^c, Alberto Minassi^b and Giuseppe Perugino^a

^aInstitute of Biosciences and BioResources, National Research Council of Italy, Naples, Italy; ^bDepartment of Pharmaceutical Sciences, University of Piemonte Orientale, Novara, Italy; ^cDepartment of Molecular Medicine and Medical Biotechnology, University of Naples "Federico II", Naples, Italy; ^dIXTAL srl, Novara, Italy

ABSTRACT

SNAP-tag[®] is a powerful technology for the labelling of protein/enzymes by using benzyl-guanine (BG) derivatives as substrates. Although commercially available or ad hoc produced, their synthesis and purification are necessary, increasing time and costs. To address this limitation, here we suggest a revision of this methodology, by performing a *chemo-enzymatic approach*, by using a BG-substrate containing an azide group appropriately distanced by a spacer from the benzyl ring. The SNAP-tag[®] and its relative thermostable version (SsOGT-H⁵) proved to be very active on this substrate. The stability of these tags upon enzymatic reaction makes possible the exposition to the solvent of the azide-moiety linked to the catalytic cysteine, compatible for the subsequent conjugation with DBCO-derivatives by azide-alkyne Huisgen cycloaddition. Our studies propose a strengthening and an improvement in terms of biotechnological applications for this self-labelling *protein-tag*.

ARTICLE HISTORY

Received 16 September 2020
Revised 5 October 2020
Accepted 14 October 2020

KEYWORDS

Protein-tag; protein labelling; enzymatic reaction; click chemistry; biotechnology

1. Introduction

The advent of the self-labelling *protein-tags* (SLPs) has led to a huge push in modern biotechnology, especially in the field of cell biology, where auto-fluorescent proteins (AFPs) for a long time dominated for their versatility in the localisation experiments of proteins, organelles, and membranes¹. But the use of SLPs clearly goes beyond: they catalyse the covalent, highly specific and irreversible attachment of a part of their synthetic ligands upon reaction. This offers the opportunity to label them by conjugation to those ligands of an infinite number of chemical groups, such as fluorescent dyes, affinity molecules, or solid surfaces, expanding the application fields². Among SLPs, of particular note are the *Halotag*[®], the *SpyTag*³, the SNAP- and the CLIP-tag[®]. The Promega *Halotag*[®] is a halo-alkane dehalogenase with a genetically modified active site, which reacts irreversibly with primary alkyl-halides^{4,5}.

SNAP-tag[®] from New England Biolabs (NEB) is the engineered variant of the natural suicide human O⁶-methylguanine DNA-methyltransferase protein (hMGMT). Alkylated DNA-alkyltransferases (AGTs, MGMTs or OGTs, E.C. 2.1.1.63) are ubiquitous and conserved proteins involved in the repair of the DNA alkylation damage, in particular, they remove alkyl adducts at the level of O⁶-position on guanine base^{6,7}. The peculiar single-step mechanism are called "*suicide enzymes*," in which the alkylated base is directly repaired by the irreversible transfer of the alkyl group

from the damaged guanine to the catalytic cysteine in the protein active site⁸. The protein is permanently inactivated upon the trans-alkylation reaction and susceptible to *in vivo* degradation via the proteasome.

In 2003, the group of Kai Johnsson developed a new strategy to exploit the hMGMT suicidal reaction in biotechnology, adopting a directed-evolution approach to engineer a variant to be used as an innovative *protein-tag*, that is, the SNAP-tag[®]. The rationale behind the SNAP-tag technology is the low substrate specificity of some AGT proteins, being able to efficiently recognise also the O⁶-benzyl-guanine (BG) nucleobase⁹. Likely, the reaction of these enzymes with BG-derivatives could happen: upon the irreversible transfer to the catalytic cysteine, they indeed demonstrated the specific labelling of the hMGMT with molecules, as fluorophores, previously conjugated to the 4-position of the BG benzyl ring. Because of the small dimension of this protein, it was mutagenized to abolish any DNA binding activity and utilised as *protein-tag* for the indirect labelling of proteins of interest genetically fused to it (Figure 1)^{9–13}. Later, the same group further engineered the SNAP-tag[®] to obtain the CLIP-tag[®], which specifically reacts with O²-benzyl-cytosine derivatives, expanding that technology for *in vivo* and *in vitro* multi-protein labelling¹⁴.

Apart from cell biology and fluorescence imaging, hundreds of papers are present in the literature showing many applications of SNAP-tag[®] in several fields, among which RNA-editing¹⁵, the

CONTACT Giuseppe Perugino ✉ giuseppe.perugino@ibbr.cnr.it 📧 Institute of Biosciences and BioResources, National Research Council of Italy, Via Pietro Castellino 111, Naples 80131, Italy; Alberto Minassi ✉ alberto.minassi@uniupo.it 📧 Department of Pharmaceutical Sciences, University of Piemonte Orientale, Largo Donegani 2/3, Novara 28100, Italy

*These Authors contributed equally and joined as First Author.

📄 Supplemental data for this article can be accessed [here](#).

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

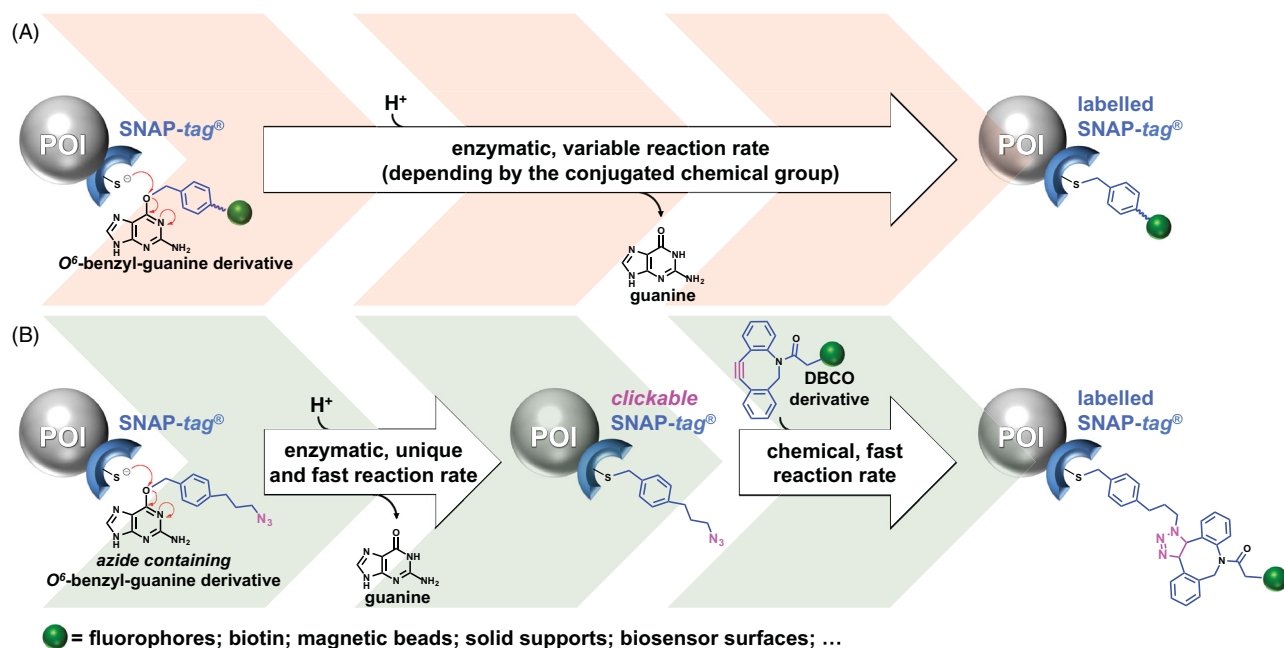


Figure 1. Single-step reaction vs chemo-enzymatic approach. (A) The SNAP-tag[®] technology is based on BG-derivatives singularly synthesised and purified, and not excluding that the conjugated chemical group (green sphere) could affect the enzymatic reaction rate. (B) The SNAP-tag[®] technology revised uses a unique and universal azide BG-derivative, converting SNAP-tag[®] in a clickable form, prone to perform a fast and efficient cycloaddition with DBCO-based chemical groups. POI, protein of interest genetically fused to the SNAP-tag[®].

development of SNAP-based sensors for small molecules^{16–18} and ions^{19,20}, and protein-DNA complexes in “DNA Origami” structures²¹.

Following the same approach, Perugino and co-workers expanded this technology to extremophilic organisms and to all the applications which require harsh reaction conditions, not fully suitable for the employing of the mesophilic SNAP-tag[®]. To this aim, they developed a “thermo-SNAP-tag” by the production of a variant of the OGT from *Saccharolobus solfataricus* (previously *Sulfolobus solfataricus*, *SsOGT-H⁵*, hereinafter *H⁵*), an enzyme which revealed extremely resistant to high temperature, high ionic strength, proteases attack, and, in general, to common physical and chemical denaturants^{22,23}. The intrinsic stability of *H⁵* made it compatible with expression and utilisation *in vivo* as protein-tag in thermophilic organisms, as *Thermus thermophilus*²⁴ and *Sulfolobus islandicus*²⁵ as well as in an *in vitro* expression system using *Sulfolobus* lysates²⁶. Recently, *H⁵* became a part of the new ASL^{tag} system²⁷, which was particularly useful for the *in vivo* immobilisation and contemporary labelling of proteins and enzymes of interest, stabilising them without any purification procedures needed²⁸.

SNAP-tag[®] technology is essentially based on BG-substrates: although many of them are commercially available, the possibility of conjugation of infinite desired molecules to the 4-position on BG leads to the synthesis of *ad hoc* substrates. This is generally possible through the crosslinking reaction of the so-called “BG-building block” (such as the amine-reactive BG-NH₂) with NHS-ester derivative compounds. The main disadvantage is the need to purify the final compounds before the reaction with the enzyme, increasing the times and costs of the experiments (Figure 1(A)). Furthermore, the presence of chemical groups conjugated to the benzyl moiety of the BG could affect the reaction efficiency of the SNAP-tag[®]^{29–33}, sometimes making this enzyme not fully applicable to particular requests.

In this work, we analysed and confirmed the catalytic dependence of SNAP-tag[®] and *H⁵* by several substrates having different

chemical groups conjugated to the O⁶-position of the guanine. To overcome these limitations, in the current study we suggest a further improvement of this technology with the application of a chemo-enzymatic approach, by using a unique and universal azide decorated BG-derivative, to obtain the specific labelling of the tag (clickable-SNAP), that can be easily coupled with a potentially infinite number of commercially available di-benzo-cyclo-octyl (DBCO)-based molecules, through the copper-free azide-alkyne Huisgen cycloaddition (Figure 1(B)). This approach could mainly offer the advantage to take into account of a unique reaction rate for the enzyme (with the azide-based BG), saving costs and times for the linking to the tag of an infinite number of commercially available DBCO-molecules. Here, we successfully proved the labelling of the SNAP-tag[®] with several DBCO-based fluorophores and the covalent immobilisation of this protein on alkyne-coated surface sensors.

2. Materials and methods

2.1. Reagents

BG was from Activate Scientific GmbH (UK), whereas **MGPA** was a gift of Prof D. Prosperi (University of Bicocca, Milan, Italy). SNAP-Vista[®] Green (**SVG**), SNAP Cell[®] Block (**SCB**), SNAP Cell[®] 430 (**SC430**), BG-PEG-NH₂ (**BGPA**), pSNAP-tag(m) plasmid, DNA restriction endonucleases and DNA modification enzymes were purchased from New England Biolabs (USA). Molecular biology kits for plasmid preparations were from Macherey-Nagel GmbH (Germany). Oligonucleotides synthesis and DNA sequencing service were performed by Eurofins Genomics (Germany). **BDP FL alkyne**, **BDP FL DBCO**, **Cy5 DBCO** were purchased from Lumiprobe GmbH (Germany). **DBCO-PEG₄-Fluor 545**, Tris(2-carboxyethyl)phosphine (TCEP), Tris [(1-benzyl-1H-1,2,3-triazol-4-yl)-methyl]amine (TBTA) were from Sigma-Aldrich (St. Louis, MO). Pierce[™] Premium Grade Sulfo-NHS (N-hydroxy-sulfo-succinimide) and Pierce[™] Premium Grade 1-ethyl-3-(3-dimethyl-amino-propyl)-

carbodiimide hydrochloride (EDC) were from Thermo Fisher Scientific (Carlsbad, CA).

2.2. Compounds synthesis: general procedures

^1H (400 MHz) and ^{13}C (100 MHz) NMR spectra were measured on Bruker Advance Neo 400 MHz spectrometer. Chemical shifts were referenced to the residual solvent signal (CDCl_3 : $\delta_{\text{H}} = 7.26$, $\delta_{\text{C}} = 77.0$; DMSO: $\delta_{\text{H}} = 2.50$, $\delta_{\text{C}} = 39.5$). Low-resolution ESI-MS were obtained on an LTQ OrbitrapXL (Thermo Scientific) mass spectrometer. IR spectra were registered on Shimadzu DR 8001 spectrophotometer. Silica gel 60 (70–230 mesh) used for gravity column chromatography (CC) was purchased from Macherey–Nagel. Reactions were monitored by TLC on Merck 60 F254 (0.25 mm) plates, visualised by staining with 5% H_2SO_4 in ethanol or KMnO_4 and heating. Organic phases were dried with Na_2SO_4 before evaporation. Chemical reagents and solvents were from Aldrich, Alfa Aesar, and TCI and were used without any further purification unless stated otherwise.

2.3. Synthesis of BGN3

BGN3 was synthesised according to the method of Zhang et al.³⁴, whose experimental spectra were comparable. White solid. ^1H NMR (400 MHz, DMSO- d_6) δ 7.84 (s, 1H), 7.53 (d, $J = 7.8$ Hz, 2H), 7.39 (d, $J = 7.8$ Hz, 2H), 6.28 (bs, 2H), 5.49 (s, 2H), 4.45 (s, 2H) (Figure S1(A)). ^{13}C NMR (100 MHz, DMSO- d_6) δ 159.69, 136.80, 135.43, 128.82, 128.58, 66.44, 53.37. IR (KBr) cm^{-1} : 3638, 3462, 3322, 2799, 2132, 1424, 1257, 1163, 912, 790, 656, 514. ESI/MS: m/z [$\text{M} + \text{H}^+$] 297 (Figure S1(B)).

2.4. Synthesis of BGSN3

BGSN3 was synthesised by following the scheme in Figure S2.

2.4.1 Synthesis of 4-azido-N-(4-(hydroxymethyl) benzyl) butanamide (compound 3)

A stirred solution of compound **1** (see Figure S2; 1.176 g, 9.115 mmol, 1 eq/mol) was prepared according to the method by Huang et al.³⁵ in DCM (30 ml), compound **2** (1.500 g, 10.939 mmol, 1.2 eq/mol; prepared according to the method by Leng et al.³⁶ and TEA (5.08 ml, 36.460 mmol, 4 eq/mol) were added. The mixture was stirred for 10 min at room temperature, then T3P (50% solution in EtOAc, 10.85 ml, 18.230 mmol, 2 eq/mol) was slowly added dropwise, and the stirred reaction was left overnight at room temperature until the complete conversion of the starting material (TLC: PE-EtOAc 4:6; Rf 1 = 0.47; Rf 3 = 0.16). The reaction was quenched by the addition of BRINE and extraction with DCM. After drying (Na_2SO_4) and evaporation, the residue was purified by gravity column chromatography on silica gel (gradient PE-EtOAc from 6:4 to 3:7) to afford compound **3** as a white solid (660 mg, 30%). ^1H NMR (400 MHz, CDCl_3) δ 7.28 (d, $J = 8.0$ Hz, 2H), 7.20 (d, $J = 7.9$ Hz, 2H), 6.36 (t, $J = 5.6$ Hz, 1H), 4.63 (s, 2H), 4.35 (d, $J = 5.7$ Hz, 2H), 3.31 (t, $J = 6.6$ Hz, 2H), 2.27 (t, $J = 7.3$ Hz, 2H), 1.89 (p, $J = 6.9$ Hz, 2H) (Figure S3(A)). ^{13}C NMR (100 MHz, CDCl_3) δ 171.93, 140.30, 137.28, 127.69, 127.17, 64.53, 50.65, 43.19, 32.96, 24.69. IR (KBr) cm^{-1} : 3276, 3055, 2921, 2880, 2103, 1635, 1540, 1418, 1257, 1015, 827, 747, 677, 553. ESI/MS: m/z [$\text{M} + \text{H}^+$] 249 (Figure S3(B)).

2.4.2. Synthesis of (N-(4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)-4-azidobutanamide) (BGSN3)

To a cooled solution (0 °C) of compound **3** (400 mg, 1.611 mmol, 1 eq/mol) in dry DMF (10 ml) in a dry flask under N_2 atmosphere, NaH (60% dispersion in mineral oil, 202 mg, 5.059 mmol, 3.14 eq/mol) was slowly added. The mixture was stirred at 0 °C for 10 min, then DMAP (16 mg, 0.129 mmol, 0.08 eq/mol) and compound **4** (451 mg, 1.772 mmol, 1.1 eq/mol; prepared according to the method by Kindermann et al.³⁷ were sequentially added. The reaction was then heated at room temperature and stirred for 4 h until the complete conversion of the starting material (TLC: DCM-MeOH 9:1; Rf 4 = 0.70; Rf BGSN = 0.55), then quenched by slow addition of BRINE and extraction with EtOAc. After drying (Na_2SO_4) and evaporation, the residue was purified by gravity column chromatography on silica gel (gradient DCM-MeOH from pure DCM to 20:1) to afford **BGSN3** as a white solid (413 mg, 67%). ^1H NMR (400 MHz, DMSO- d_6) δ 12.48 (bs, NH purine, 1H), 8.43 (t, $J = 5.9$ Hz, 1H), 7.85 (s, 1H), 7.49 (d, $J = 7.7$ Hz, 2H), 7.30 (d, $J = 7.8$ Hz, 2H), 6.31 (bs, NH2 purine 2H), 5.50 (s, 2H), 4.30 (d, $J = 5.9$ Hz, 2H), 3.36 (t, $J = 6.8$ Hz, 2H), 2.26 (t, $J = 7.4$ Hz, 2H), 1.81 (p, $J = 7.1$ Hz, 2H) (Figure S4(A)). ^{13}C NMR (100 MHz, DMSO- d_6) δ 171.38, 159.91, 159.69, 155.23, 139.48, 137.90, 135.31, 128.60, 127.34, 113.57, 66.59, 50.36, 41.95, 32.23, 24.58. IR (KBr) cm^{-1} : 3647, 3484, 3379, 3282, 2794, 2100, 1580, 1403, 1282, 1163, 938, 835, 789, 635, 553. ESI/MS: m/z [$\text{M} + \text{H}^+$] 382 (Figure S4(B)).

2.5 Plasmids and protein purification

The cloning procedures in the pQE31 expression vector (Qiagen, Germany) were similar for both proteins. In particular, the pSNAP-tag(m) Vector was used as a template to amplify the DNA fragment relative to the SNAP-tag[®] gene, by using QE_SNAP-Fwd/QE_SNAP-Rev oligonucleotides pairs (5'-ATGGCAGGATCCAA TGGACAAAGACTGCGAAATG-3'/5'-CTATCAAAGCTTAACCCAGCCCAG GCTTGCCCA G-3'; BamH I and Hind III sites, respectively, are underlined). Afterwards, the resulting fragment and the pQE31 vector were digested with BamH I and Hind III restriction enzymes and ligated, leading to the final pQE-SNAP plasmid. The final SNAP-tag[®] protein was expressed with an extra N-terminal amino-acidic sequence, comprising a His₆-tag (MRGSHHHHHHTDP-). The ligation mixture was used to transform *E. coli* KRX competent cells and positive colonies were confirmed by PCR and DNA sequence analyses.

H^5 was cloned as previously described²². SNAP-tag[®] and H^5 proteins were expressed in *E. coli* ABLE C cells, grown at 37 °C in Luria–Bertani (LB) medium supplemented with 50 mg/l kanamycin and 100 mg/l ampicillin. The protein expression was induced with 1 mM isopropyl-thio- β -D-galactoside (IPTG) at an absorbance value of 0.5–0.6 A_{600} nm. After overnight growth, cells were collected and resuspended 1:3 (w/v) in purification buffer (50 mM phosphate, 300 mM NaCl; pH 8.0) supplemented with 1% Triton X-100 and stored overnight at –20 °C. Subsequently, the biomass was treated in ice with lysozyme and DNase for 60 min and then sonicated as described (Perugino et al., 2012). After centrifugation of 30 min at 60,000 $\times g$, the cell extract was recovered and applied to a Protino Ni-NTA Column 1 ml (Macherey–Nagel) for His₆-tag affinity chromatography. The eluted fractions containing the protein were collected and dialysed against phosphate-buffered saline (PBS 1 \times , 20 mM phosphate buffer, NaCl 150 mM, pH 7.3). Pooled fractions were concentrated and protein purification was confirmed by SDS-PAGE analysis. Aliquots were finally stored at –20 °C.

Table 1. Substrate specificity of SNAP-tag[®] and H⁵ by competitive inhibition method (IC₅₀) by using SVG as substrate, and second order rate constant of the enzymatic reaction of these protein-tags only on the BGSN3 substrate.

Structure	Name	SNAP-tag [®]		SsOGT-H ⁵		Note
		IC ₅₀ (μM)	k ^a (s ⁻¹ M ⁻¹)	IC ₅₀ (μM)	k (s ⁻¹ M ⁻¹)	
	SVG	–	2.8 × 10 ⁴ ^b	–	1.6 × 10 ⁴	[14,24]
	BG	36.8 ± 5.6	–	10.1 ± 1.0	–	This work
	SCB	2.1 ± 0.5	–	4.4 ± 0.8	–	This work
	BGN3	15.6 ± 0.3	–	23.5 ± 1.0	–	This work
	BG430	ND ^c	–	ND	–	This work
	BGPA	86.0 ± 6.7	–	14.3 ± 1.9	–	This work
	MGPA ^d	–	–	268.9 ± 19.1 ^e	–	This work
	BGSN3	17.8 ± 1.1	4.64 ± 1.04 × 10 ⁵	10.0 ± 0.7	1.40 ± 0.47 × 10 ⁴	This work

For each compound, the guanine moiety is drawn *in black* and the chemical group conjugated to the benzyl ring *in blue*. The fluorescein moiety of the SVG is *in green*, whereas SCB differs from the other derivatives by the presence of a benzylic ring (*in red*). Azide group is conventionally coloured *in magenta*.

^aReaction rates at 25 °C; ^bthis value was obtained by using a BG-fluorescein substrate (BG-FL) very similar to SVG; ^cnot determined; ^dthis molecule is a O⁶-methyl-guanine derivative; ^ecompetitive assay for H⁵ was performed at 65 °C.

2.6 AGTs' substrate assay by competitive inhibition method

Competitive inhibition assay was performed as described^{23,38}. Briefly, by using a fixed concentration of the fluorescent SVG (5 μM) and enzymes (5 μM), an increasing amount of guanine-derivatives (0–2 mM) was added to the mixtures. The reactions were incubated for 30 min at 25 °C and 50 °C for SNAP-tag[®] and H⁵ respectively, and loaded on SDS-PAGE. Subsequently, fluorescent bands were measured by *gel-imaging* on a VersaDoc 4000[™] system (Bio-Rad), by applying a blue LED/530 bandpass filter. Obtained data were finally plotted by Equation (1),

$$RF = \frac{100\%}{1 + \left(\frac{[I]}{IC_{50}}\right)^n [S]} \quad (1)$$

where RF is the obtained Relative Fluorescence, [I] and [S] are the concentration of the inhibitor and the substrate, respectively, and

finally the IC₅₀ is the concentration needed to reduce by 50% the fluorescence intensity of the protein band.

We evaluated the activity of SNAP-tag[®] and H⁵ enzymes on BGN3 and BGSN3 by the afore-mentioned IC₅₀ method (Figure S5(A,B)) and other O⁶-guanine-derivatives (Table 1).

2.7. In vitro Huisgen Cu(I)-catalysed cycloaddition reaction

The Huisgen chemical reaction was evaluated on SNAP-tag[®] and H⁵ previously incubated with BGN3 and BGSN3. An opportune amount of purified proteins was incubated within an equimolar ratio of these substrates for 60–120 min at 25 °C and 37 °C respectively, to ensure the complete enzymatic labelling reaction. Later, we performed the subsequent cycloaddition using 5 μM of an alkyne-derivative of the fluorescein (BDP FL alkyne), in the

presence of copper (1 mM), TCEP (1 mM), TBTA (0.1 mM) and, where indicated, of SDS (0.05%). Finally, mixtures were loaded on SDS-PAGE and analysed as described in Section 4 (Figure S5(C,D)).

2.8. Molecular modelling

All molecular modelling studies were performed on a Tesla workstation equipped with two Intel Xeon X5650 2.67 GHz processors and Ubuntu 14.04 (<http://www.ubuntu.com>). The protein structures and 3D chemical structures were generated in PyMOL (The PyMOL Molecular Graphics System, version 2.2.3, Schrödinger LLC, 2019).

2.9. Molecular dynamics (MD) simulation

The MD simulations were carried out using the Desmond simulation package of Schrödinger LLC (Schrödinger Release 2019–1: Desmond Molecular Dynamics System; D. E. Shaw Research: New York, NY, 2019; Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2019). The X-ray structure of the H^5 covalently bound to **SVG** was used in this study, entry code 6GAO³⁹, water molecules were removed, and all hydrogen atoms and charges were added. The NPT ensemble with the temperature of 300 K and a pressure 1 bar was applied in all runs. The simulation length was 100 ns with relaxation time 1 ps. The OPLS_2005 force field parameters were used in all simulations⁴⁰. The long-range electrostatic interactions were calculated using the particle mesh Ewald method⁴¹. The cut-off radius in Coulomb interactions was 9.0 Å. The water molecules were explicitly described using the simple point charge model⁴². The Martyna–Tuckerman–Klein chain coupling scheme⁴³ with a coupling constant of 2.0 ps was used for the pressure control and the Nosé–Hoover chain coupling scheme⁴⁴ for the temperature control. Non-bonded forces were calculated using an r-RESPA integrator where the short-range forces were updated every step and the long-range forces were updated every three steps. The trajectory sampling was done at an interval of 1.0 ps. The behaviour and interactions between the ligands and protein were analysed using the Simulation Interaction Diagram tool implemented in the Desmond MD package. The stability of MD simulations was monitored by looking at the RMSD of the ligand and protein atom positions in time.

2.10. Determination of the rate constants of the chemo-enzymatic labelling approach

Rate constants of the enzymatic reactions with the only **BGSN3** were determined by the method of Gautier et al.¹⁴. In this case, purified proteins (5 μ M) were incubated with the substrate (5 μ M) in PBS 1 \times buffer at 25 °C. Aliquots were taken at different times, the reactions were immediately stopped in Leamli Buffer 1 \times in addition with 10 μ M of **Cy5 DBCO** fluorophore and placing tubes on ice.

Rate constants for the chemical reaction needed for the preliminary achievement of the *clickable*-SNAP and *clickable*- H^5 with **BGSN3**, which was obtained by the afore-described protocol, in order to get the complete labelling. Then, to each aliquot of 5 μ M of *clickable* proteins, 20 μ M of **DBCO-PEG₄-Fluor 545** fluorophore was added. At different times, an excess of sodium azide (NaN₃, 300 mM) was immediately added to each aliquot and then placing tubes on ice, in order to stop the click reaction between the azide group on the **BGSN3** and the **DBCO-PEG₄-Fluor 545** molecule.

Finally, for both the experiments, all aliquots were boiled in an SDS buffer for 5 min, and immediately loaded on a SDS-PAGE, for the *gel-imaging* and *coomassie staining* analyses, as previously described. Data were fitted to a pseudo-first-order reaction model using the GraFit 5.0 software package (Erithacus Software Ltd.). Second-order rate constants k (in s⁻¹ M⁻¹) were then obtained by dividing the pseudo-first-order constant by the concentration of the substrate (Figure 2 and Table 1). Values given are an average of at least three independent measurements.

2.11. In vitro Huisgen copper-free cycloaddition reaction with different DBCO-fluorophores

For the copper-free click reaction, aliquots of 5 μ M of each *clickable*-protein were incubated for 60 min at room temperature in the dark with 5 μ M of fluorescent DBCO-derivative substrates (**BDP FL DBCO**, **Cy5 DBCO**, and **DBCO-PEG₄-Fluor 545**) in a total volume of 10 μ l of PBS 1 \times buffer (Figure 4 and Figure S6). The reactions were finally stopped in Leamli Buffer 3 \times , loaded on SDS-PAGE, and analysed as described in Section 4, by applying a blue LED/530 bandpass filter, red LED/695 bandpass filter and green LED/605 bandpass filter as excitation/emission parameters for each DBCO-fluorophores, respectively. The click reaction was also performed on 5 μ M of both the enzymes, but in the presence of an *EcCFE* diluted in PBS 1 \times buffer.

2.12. Procedure for protein immobilisation on bio layer interferometry (BLI), by following the chemo-enzymatic approach

OctetRED96TM (ForteBio, Fremont, CA) was used to immobilise specifically SNAP-*tag*[®] and H^5 with the *chemo-enzymatic approach* (Figure 5(A,B)). Samples and reaction buffers were located in black 96-well plates (OptiPlate-96 Black, Black Opaque 96-well Microplate, PerkinElmer, Billerica, MA) in a maximum reaction volume of 300 μ l per well with 800 rpm shaking for each step. For the immobilisation procedure, AR2G sensors were first wetted in 200 μ l of pure water for at least 15 min, followed by an equilibration step (3 min) in acetate buffer 0.1 M, pH 5.0. Afterwards, they were activated with 20 mM 1-ethyl-3-(3-dimethyl-amino-propyl) carbodiimide hydrochloride (EDC)/20 mM N-hydroxy-sulfo-succinimide (sulfo-NHS) mixture in acetate buffer (60 min) and covered with 2 mM **propargyl-PEG₃-amine** bifunctional linker (BroadPharm, San Diego, CA) in Loading step (20 min). To avoid the presence of any free amine groups on the biosensors, a Blocking step with Ethanolamine 1 M (30 min) was performed. Subsequently, a Washing step (15 min) with water and an Equilibration step in click-reaction buffer (15 min) are followed.

During the afore-described procedure, proteins were labelled with **BGSN3**. Finally, the immobilisation step for each sample via Huisgen reaction was carried out at 30 °C for 80 min, followed by a Washing step (20 min), in order to remove all the unbound molecules. This procedure was the same in the presence of the *EcCFE*. All measurements were performed in triplicates.

2.13. Permeability of eukaryotic and prokaryotic cells to BGSN3

HEK293T cells were maintained at 37 °C with 5% CO₂ in Dulbecco's Modified Essential Medium (Invitrogen, Carlsbad, CA) supplemented with 10% Foetal Bovine Serum (FBS) (Invitrogen) and 100 U/ml Penicillin/Streptomycin (Roche, Switzerland). HEK293T cells were transfected with SNAP-*tag*[®] plasmid by using

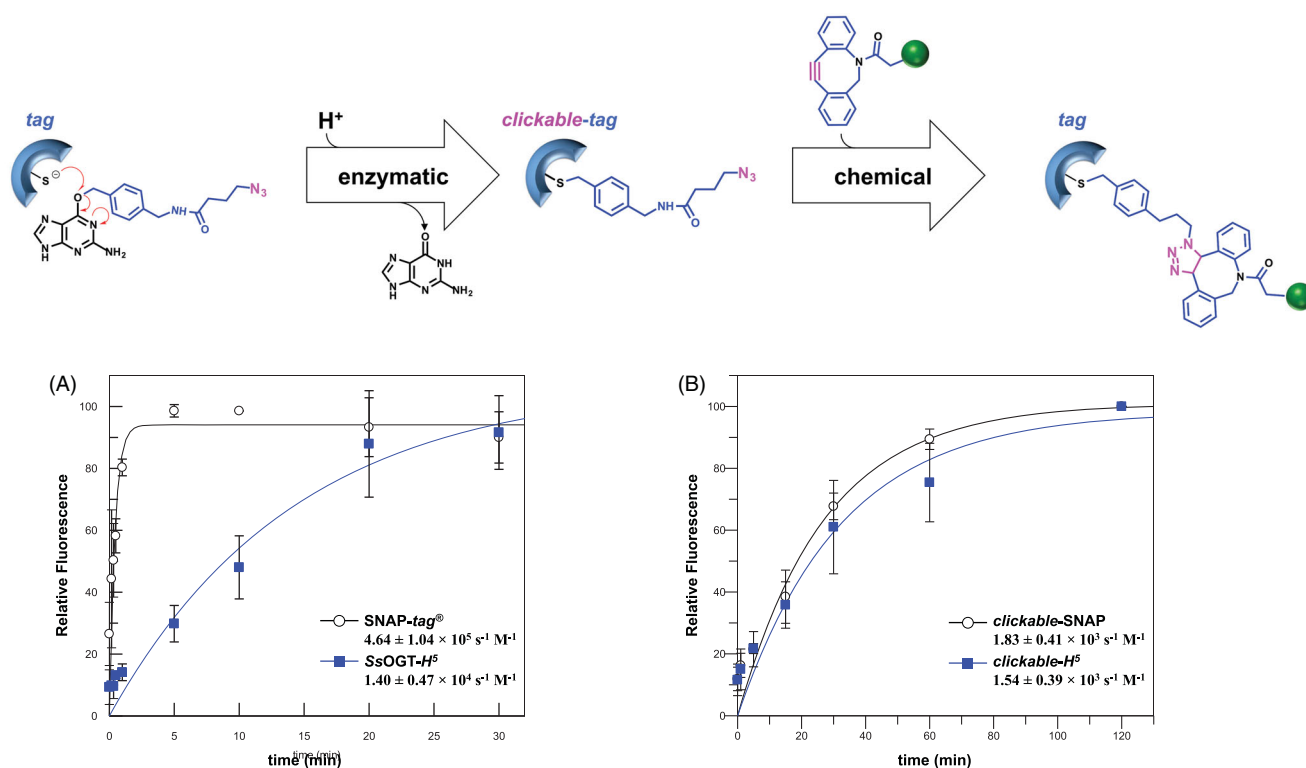


Figure 2. Reaction rates of the chemo-enzymatic approach. Pseudo-first-order reaction of *protein-tags* for (A) the enzymatic reaction with BGSN3 (see *k* values also in Table 1), and of *clickable-tags* for (B) Huisgen reaction with DBCO-PEG₄-Fluor 545 (see values in the main text). Values given are an average of three independent measurements. The reaction scheme was an exemplification of Figure 1(B) in the main text. Data are represented as mean \pm SEM.

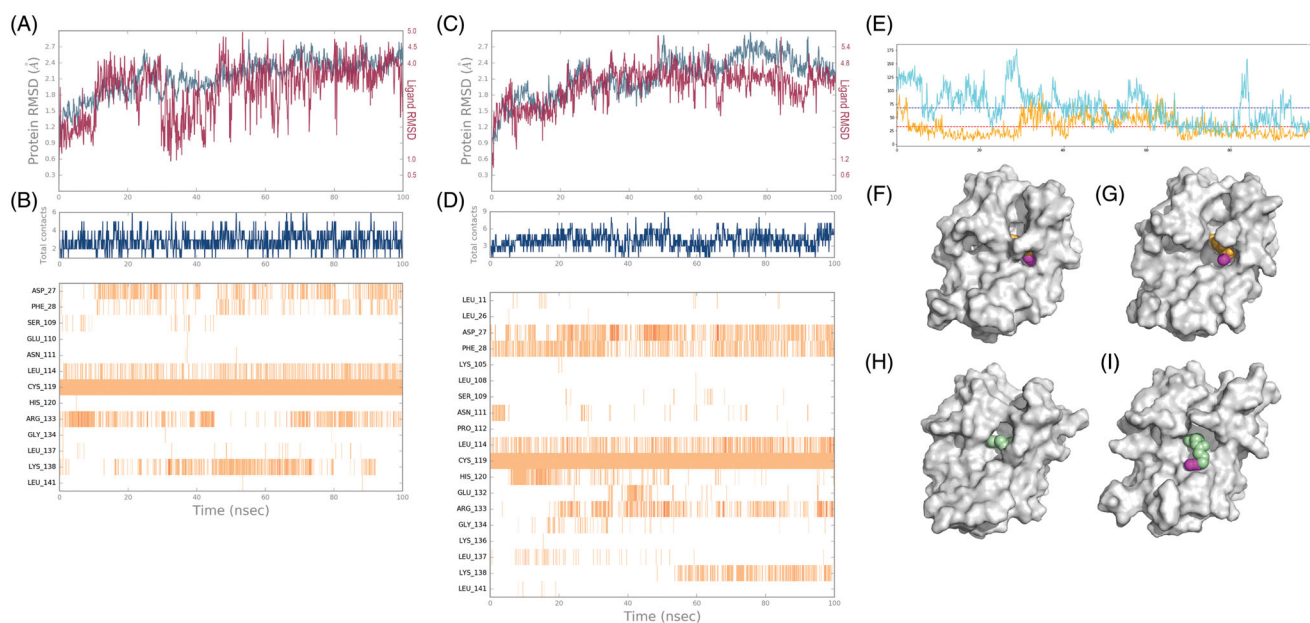


Figure 3. Molecular modelling on H⁵ with BG-azides. (A) RMSD of the atomic positions for the compound BGN3 (Lig fit Prot, *in red*) and the protein H⁵ (C α positions, *in blue*) of the 100 ns molecular dynamics simulations using Desmond package. (B) A timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). (C) RMSD of the atomic positions for the compound BGSN3 (Lig fit Prot, *in red*) and the protein H⁵ (C α positions, *in blue*) of the 100 ns molecular dynamics simulations using Desmond package. (D) A timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). (E) Solvent Accessible Surface Area (SASA) of BGN3/H⁵ (*in orange*) and BGSN3/H⁵ (*in cyan*) complexes over the MD simulation time (mean values are depicted as dot lines). Frames of H⁵-probe complexes with lower (F, H) and higher (G, I) SASA value for BGN3 (F, G) and BGSN3 (H, I), respectively.

Lipofectamine 2000 (Invitrogen) following manufacturer's protocol. The treatment with BGSN3 were performed, at the concentration and time indicated for each experiment. Twenty-four hours after transfection, we treated cells with BGSN3 for 2 h at different

concentrations ranging (from 1 to 25 μM), directly dissolving the compound in complete culture medium. Then cells were harvested, washed with PBS 1 \times buffer and lysed with 50mM Tris-HCl pH 7.4, 150mM NaCl, 0.5mM EDTA, 0.1% Triton X-100

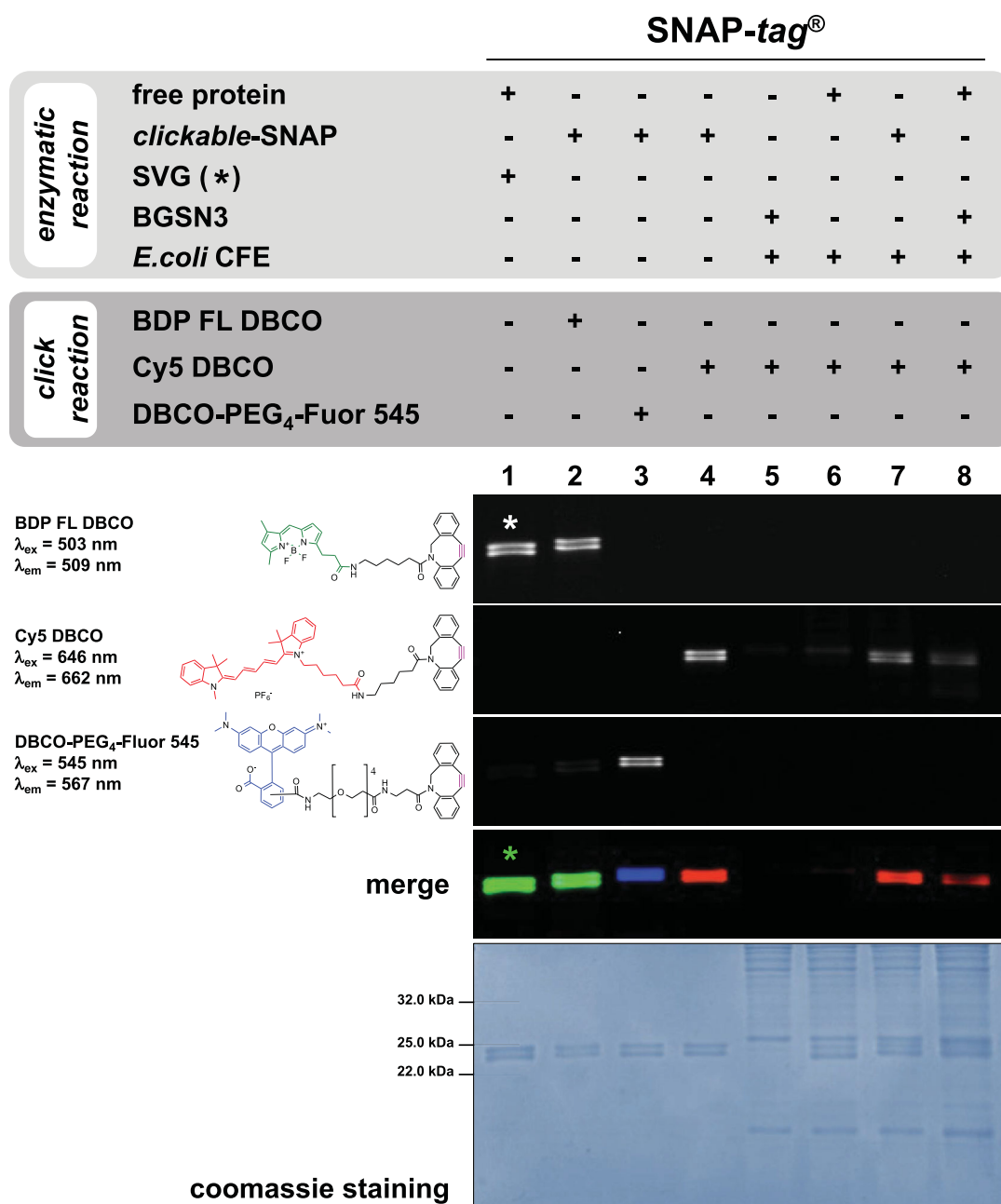


Figure 4. Specificity of the Huisgen reaction. *Gel-imaging* analysis of SNAP-tag® labelling by a chemo-enzymatic approach with BGSN3 and three different DBCO-derivative fluorophores. Protein (5 μM) was incubated with 5 μM of the azide-based BG for 60 min at 25 °C; then, an equimolar amount of DBCO-based substrate was added for the chemical click reaction, keeping the same time and temperature conditions. As control, SNAP-tag® was incubated only with SVG (lane 1, signal marked with an asterisk).

supplemented with complete protease (Roche, Switzerland) and phosphatase (SERVA Electrophoresis, Germany) inhibitors. Afterwards, transfected cells were treated with a fixed concentration of BGSN3 (10 μM) at different time points (from 30 to 120 min). Again, HEK293T cells were washed and lysed as described before. To confirm the reaction with BGSN3, the same amount of protein extract (0.91 μg/μL for each sample) was incubated for 30 min at 25 °C with SVG. Subsequently, proteins were loaded on SDS-PAGE and analysed by *gel-imaging* on a VersaDoc 4000™ system (Bio-Rad), by applying a blue LED/530 bandpass filter (Figure 6).

For flow cytometry analysis, HeLa cells were seeded in 24-well plates and transfected with SNAP-tag® plasmid by using

Lipofectamine 2000 (Invitrogen, USA) following manufacturer's protocol. Twenty-four hours after the transfection, cells were treated with 25 μM BGSN3 for 1 h, and the excess of the substrate was washed out by 2 × 15 min, followed by 1 × 30 min washes. Cells were then treated with 2.5 μM BDP FL DBCO for 30 min and unbound fluorophore was removed by following the same procedure performed for the BGSN3. All treatments and washes were performed at 37 °C in a complete culture medium. Lastly, cells were harvested by trypsinization, and fluorescence was measured using FACS CANTO II instrument. The analysis was performed on live singlet cells using FlowJo software (Figure S7(A)).

E. coli ABLE C strain was transformed with SNAP-tag® plasmid and protein expressed as previously described. After overnight

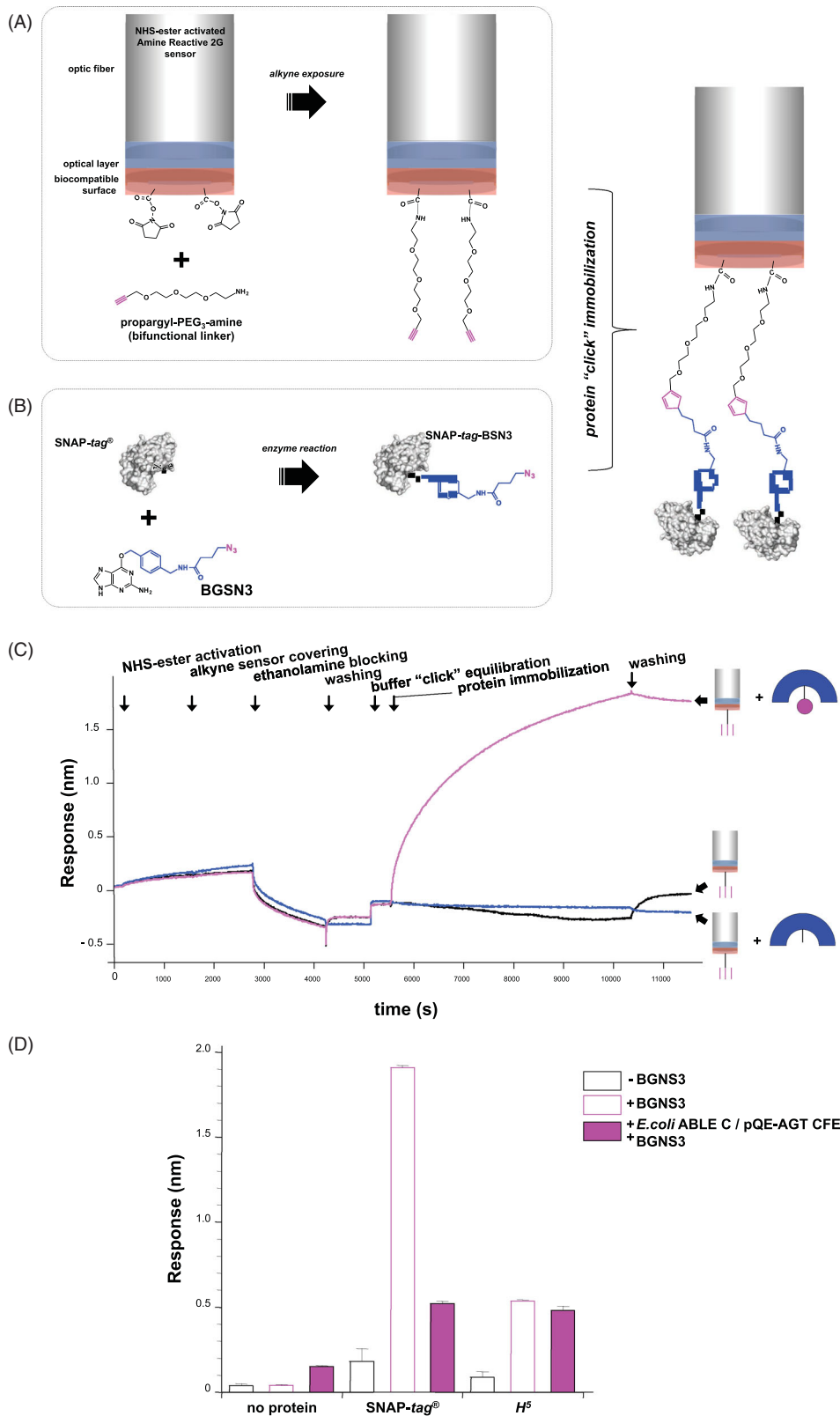


Figure 5. Covalent immobilisation of clickable-tags on the BLI sensor. (A) Covering of the BLI sensor with a bi-functional linker, exposing alkyne groups for the Huisgen cycloaddition reaction; (B) reaction of the SNAP-tag[®] with BGSN3; (C) chemo-enzymatic SNAP-tag[®] immobilisation on BLI. The alkyne-covered sensor (*silver cylinder*) was immersed in wells containing the buffer (*in black*), the free SNAP-tag[®] (*in blue*) and the *clickable*-SNAP (*in magenta*); (D) column chart relative to the BLI immobilisation of purified *protein-tags* alone (*black-bordered bars*) or in the presence of BGSN3 (*magenta-bordered bars*). Filled magenta bars represent the BLI immobilisation using the *EcCFE* upon heterologous expression of *protein-tags*. Standard deviations were obtained from three independent experiments. Data are represented as mean \pm SEM.

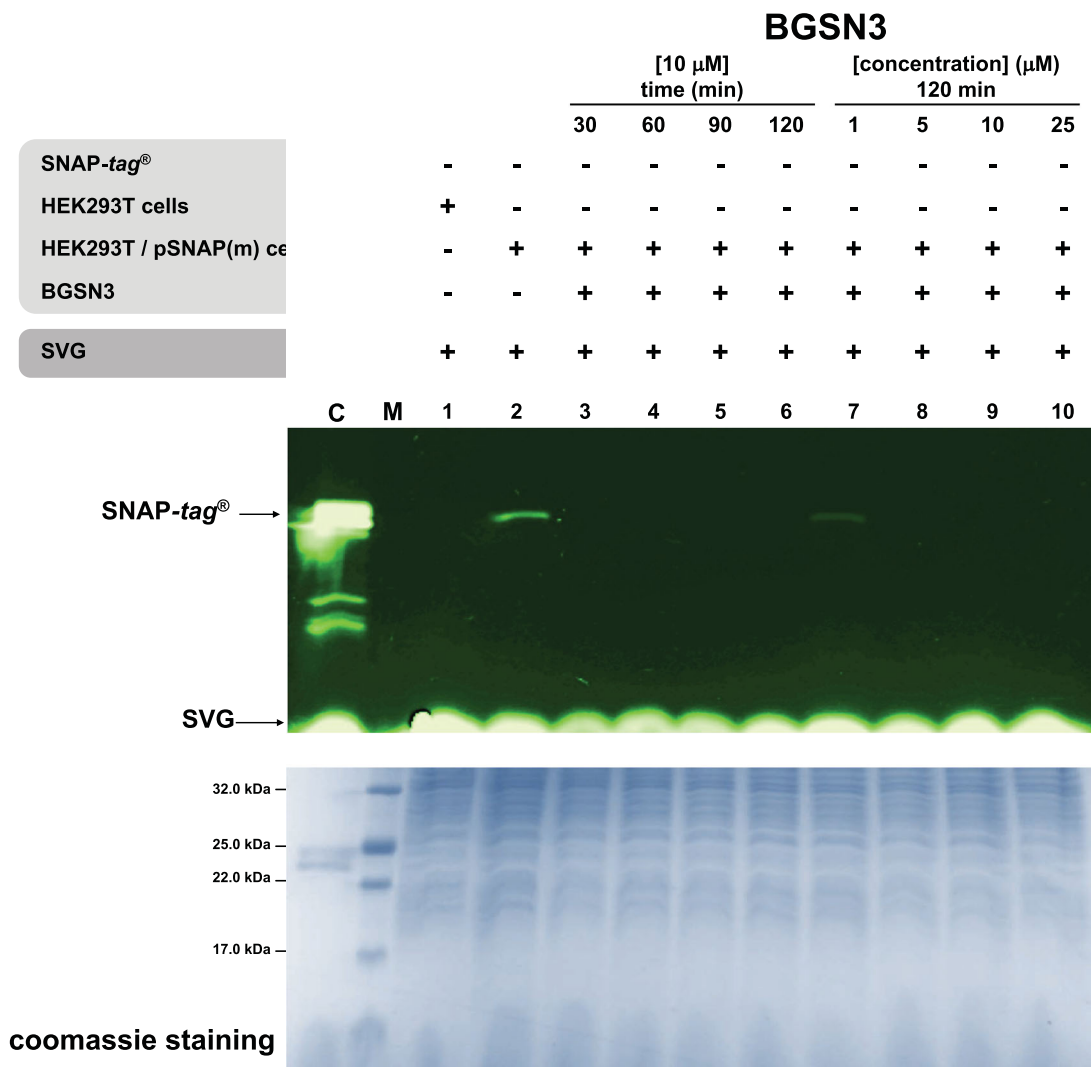


Figure 6. Eukaryotic permeability to BGSN3. SDS-PAGE analysis by *gel-imaging* and *coomassie staining* of HEK293T cell lysates. After BGSN3 in medium treatment, lysates were incubated with SVG.

growth, samples of 2 ml were treated with 100 μ M of **BGSN3** for 2 h at 25 °C and then collected by centrifugation at 2000 $\times g$. Cell pellets of 0.05 g were resuspended 1:3 (w/v) in PBS 1 \times supplemented with 1% Triton X-100 and subjected to cell lysis, by applying 5 cycles of freeze-thawing. After a centrifugation at 13,000 $\times g$, the supernatants containing the protein extract were incubated 30 min at 25 °C with **SVG**, and proteins were loaded on SDS-PAGE. Finally, fluorescent bands were analysed by *gel-imaging* techniques (Figure S7(B)).

3. Result and discussion

3.1. Substrate specificity of AGTs on BG-based substrates

Following the irreversible reaction shown in Figure 1, we evaluated the activity of two enzymes in our possession on several *O*⁶-guanine-derivatives (Table 1). Because most of them are non-fluorescent compounds, we performed an AGTs' competitive inhibition assay by using the fluorescein-derivative SNAP-Vista[®] Green as substrate (**SVG**), as previously described^{22–24,45}. Briefly, the reaction of an AGT with **SVG** led to a fluoresceinated protein, which can be visualised as a fluorescent band in *gel-imaging* analysis after SDS-PAGE. The presence of increasing amounts of a

non-fluorescent competitor in the reaction causes a decrease of the fluorescent signals, which can be measured and plotted for the IC₅₀ values determination^{23,46}. As shown in Table 1, SNAP-tag[®] and *H*⁵ displayed different behaviours versus these competitors, without any rationale for the dimension and/or polarity of the conjugated chemical groups. While SNAP-Cell[®] 430 (**SC430**) completely lost the competition with **SVG**, both the enzymes are extremely active on the SNAP Cell[®] Block (**SCB**), displaying the lowest IC₅₀ value measured. This result was expected, because **SCB** has a structure very similar to the Lomeguatrib, one of the most efficient inhibitors of the hMGMT protein, employed in the cancer treatment in combination with alkylating agents-based chemotherapeutics⁴⁷.

In general, all commercially available products used (**SVG**, **SCB**, **BG430**, and **BG-PEG-NH2**, **BGPA**) are good substrates for the SNAP-tag[®] and *H*⁵ enzymes, completing their labelling reaction in few hours (data not shown). However, based on our results, the choice of the chemical group to be conjugated to the *O*⁶-guanine for customized substrates may present risks, with consequent decreases in the reaction rate for these *protein-tags*. This was the case of methyl-guanine-PEG-NH2 (**MGPA**), which is an *O*⁶-methyl-guanine derivative, used for the immobilisation of SNAP-tag[®] on nanoparticles⁴⁸. The latter is not a preferred substrate, probably

because of the absence of the benzyl ring, which leads to complete labelling of the SNAP-tag[®] and H⁵ after over-night incubation at 4 °C⁴⁸ and 65 °C (data not shown), respectively.

3.2. In vitro enzymatic reaction of engineered AGTs with BG-azide substrates

Recent studies were focussed on the synthesis of alternative “BG-building blocks,” which offer the opportunity to produce SNAP-substrates by following easier and faster protocols: an alkyne substituted O⁶-BG was employed in the synthesis of compounds by the Huisgen cycloaddition with azide-based fluorescent probes⁴⁹ or, inversely, by using the O⁶-BG-N₃ (**BGN3**, Figure S1) for the conjugation with alkyne-based chemical groups³⁴. We evaluated the enzymatic reaction of the H⁵ and the SNAP-tag[®] directly on **BGN3** and a synthesised BG-derivative containing a benzyl ring opportunely spaced from the azide group (**BGSN3**, Figure S4): after the reaction, no fluorescent signal was obtained on SDS-PAGE *gel-imaging* upon the addition of **SVG** (Figure S5(A,B)). This indicates that the catalytic cysteine was completely blocked by the benzyl-azide moiety, impeding the access of the fluorescent substrate to the active site. Compared to the classical BG-derivatives, these *protein-tags* showed a reasonable activity on both these BG-azides, as resulted by the calculated IC₅₀ (Table 1 and Figure S5(A,B)).

After the enzymatic reaction of H⁵ with **BGN3** and **BGSN3**, we performed the subsequent cycloaddition using an alkyne-derivative of the fluorescein (**BDP FL alkyne**): however, the chemical reaction was less efficient using the former substrate (Figure S5(C), lane 2). In this case, the complete fluorescein labelling of the protein was achieved only in the presence of a small amount of SDS during the cycloaddition step (lane 3), suggesting that the protein is still folded after the enzymatic reaction and the azide is hidden in the active site core. The addition of the denaturant could have slightly opened the protein structure, favouring a better exposure of the azide group to the solvent, and allowing the click reaction to occur.

On the contrary, using **BGSN3** as substrate, the labelling of both the enzymes was comparable to the classical reaction with **SVG** without any denaturing agent, likely the longer spacer of **BGSN3** could sufficiently move away from the azide group from the protein surface for the Huisgen reaction (Figure S5(D), lanes 2 and 4). From now on, experiments were only performed by using the longer BG-azide. We first calculated the rate of the enzymatic reaction, demonstrating that both *protein-tags* show a high catalytic activity comparable to the commercial BG-derivatives currently used (Figure 2(A) and Table 1), also indicating that the complete protein labelling in less of an hour can be performed^{13,14,24}.

3.3. Molecular modelling on the H⁵ with BG-azides

BGN3 and **BGSN3** differ in length since the chemical spacer between the benzyl ring and the active azide makes the latter potentially more prone to the labelling reaction. It could be assumed that this aspect alone influences the availability of the azide moiety to react. However, proteins are not a static system, the amino acids side-chain movements could mask the azide and prevent the “click” chemistry reaction. The covalent complexes of these compounds with H⁵ were analysed with Molecular Dynamics (MD) simulations using the Desmond package (see Experimental Section). The complexes were simulated for 100 ns at 300 K using a standard protocol. The protein structure has been stabilised, as shown in the RMSDs for both the IDO1 C α and the ligand (Figure

3(A,C)). The MD results were analysed in terms of Solvent Accessible Surface Area (SASA) of the compounds: more time the compounds are exposed to the solvent, the higher is the possibility to react⁵⁰. In Figure 3 is reported the fluctuation of the SASA values over the simulation time together with the structure model of the H⁵ protein in complex with **BGN3** and **BGSN3**, respectively. The former is less exposed to the solvent with a SASA value of $32.967 \pm 18.573 \text{ \AA}^2$ compared to **BGSN3**, which shows a higher SASA value $68.302 \pm 32.455 \text{ \AA}^2$. This simulation confirmed our biochemical data, proposing the BG-derivative with the spacer as a better substrate for our *chemo-enzymatic approach*.

3.4. Specificity and versatility of the chemo-enzymatic reaction

The O⁶-BG-based **BGSN3** is a good substrate for the two *protein-tags* used (Table 1 and Figure 2(A)) and offering the advantage to sufficiently expose the azide group for the Huisgen reaction. This was the starting point to examine: (i) the labelling efficiency of the *clickable*-SNAP and *clickable*-H⁵ by using different DBCO-based fluorophores; (ii) the specificity of the “click” reaction.

Upon the reaction with **BGSN3**, all cycloaddition reactions with three different DBCO-based fluorophores were complete in ca. 30–45 min in PBS 1 \times buffer (Figure 4, lanes 2–4), with a protein-labelling as efficient as the enzymatic reaction using the sole **SVG** (lane 1). We quantitatively evaluated the rate (*k*) of the click reaction by using the **DBCO-PEG4-Fluor 545** fluorophore: as expected, both the *clickable-tags* were labelled with the same efficiency ($1.83 \pm 0.41 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$ for SNAP-tag[®]; $1.54 \pm 0.39 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$ for H⁵), demonstrating that the chemical reaction is sufficiently fast and independent from the *tags* (Figure 2(B)).

Concerning the specificity, we added a crude protein extract from *Escherichia coli* ABLE C (ECCFE), without any AGT activity at the *gel-imaging* analysis (Figure 4, lane 5). In this context, the only presence of the free *protein-tag* and the DBCO-fluorophore also did not result in any fluorescent signal (lanes 6), whereas the previously purified *clickable*-SNAP (lane 7), as well as its free form in the presence of **BGSN3** (lane 8), was specifically able to complete the chemo-enzymatic reaction, giving an evident fluorescent signal. The high specificity of our approach was also confirmed by using the H⁵ enzyme, which displays a better labelling reaction than the mesophilic SNAP-tag[®] (Figure S6). Probably, something in the extract might impede SNAP-tag[®] activity. These results clearly demonstrated the high efficiency of our *chemo-enzymatic approach* for the labelling of both the *protein-tags* used.

3.5. Application to the bio layer interferometry

The possibility to apply the SNAP-tag[®] technology to the Surface Plasmon Resonance (SPR) for the covalent immobilisation of a protein of interest was first explored by the group of Kai Johnsson³⁷, followed by other groups with the same substrate⁵¹ or a biotin BG-derivative⁵². Their approaches, again, required preliminarily the synthesis and the purification of a compatible substrate to cover the sensor chip surface. We used, instead, the **BGSN3** substrate for the immobilisation of the SNAP-tag[®] directly on an alkyne-derived sensor chip of the bio layer interferometry (BLI) equipment, as shown in Figure 5. This technique is more advantageous with respect to the SPR because: (i) it needs a smaller amount of sample, making it more compatible to higher throughput (the capacity of running up to 96 samples in a parallel); (ii) the possibility to reuse samples, and (iii) of the total independency from any microfluidic issues.

Given the lack of any available BLI alkyne-derived sensors, we first activated the AR2G type by a bi-functional linker (**propargyl-PEG 3-amine**) in order to expose an alkyne group on the surface (Figure 5(A)). This modified protocol provides first the coating of the sensor tips with alkyne groups (approx. 80 min), during that the reaction between the *protein-tag* and **BGSN3** inside the 96-wells rack takes place (Figure 5(B)). Only the contemporary presence of the *clickable*-SNAP and the alkyne-coated sensor led to a measurable response (Figure 5(C)). After washing procedures, the signal did not significantly drop-down, given the covalent reaction between the protein and the sensor. We successfully achieved results with both the enzymes, although temperature and times of the enzymatic reaction on BLI (30 °C) favoured the SNAP-*tag*[®] respect to the thermophilic *H*^{513,24}. Furthermore, in *EcCFEs* where both the enzymes were expressed, a specific and efficient immobilisation on BLI sensor tips occurred (Figure 5(D)), although the SNAP-*tag*[®] displayed a lower labelling efficiency in the *EcCFE*, as expected (compare lane 8 in Figure 4 and Figure S6). As for other techniques, this specific surface immobilisation of SNAP-*tag*[®] gives the opportunity to perform a directly *on-chip purification* of a tagged-POI from a crude lysate. without any purification step, in an indirect manner, which favours a better orientation of the POI for its biological activities.

3.6. Permeability of eukaryotic and prokaryotic cells to BG-azides

One of the major applications of the SNAP-*tag*[®] technology concerns the field of cell biology, where detecting fluorescent-tagged-POIs in living cells represents an important tool to study protein functions and locations⁵³. To test our *chemo-enzymatic approach*, we first investigated the permeability of **BGSN3**. Lysates of HEK293T cells pre-treated with **BGSN3** were then incubated with the **SVG** substrate: the absence of any fluorescent signal by *gel-imaging* only in BG-azide treated lysates demonstrated that the internalisation of **BGSN3** was fast (ca. 30 min; Figure 6, lane 3) and at concentrations comparable with commercial cell biology BG-substrates (in the range of <5 μM; Figure 6, lane 8). Preliminary experiments by FACS analysis confirmed that the *in vivo* cycloaddition between **BGSN3** and the **BDP-FL DBCO** occurred (Figure S7(A)). This was also confirmed for *E. coli* bacterial cells (Figure S7(B), lane 2).

4. Conclusions and perspectives

We developed an innovative modification of the SNAP-*tag*[®] technology, in order to overcome times and costs relative to the production and the utilisation of commercial or purified customised BG-derivatives. Although they are compatible in terms of catalytic activity as for the SNAP-*tag*[®], as well as for the others AGTs^{22–24,37,46,54} the risk of lowering the catalytic activity of these *tags* with customised BG-derivatives should not be underestimated (Table 1). We started by the knowledge that: (i) self-labelling *protein-tags* are still folded and enough stability in their benzylated form after the enzymatic reaction^{13,24}; (ii) the Huisgen cycloaddition is extremely versatile, fast and specific. Recently, the latter was used for the entrapment of catalytic activities by azide-based pseudo-substrates in a well-known powerful method, the *in vivo* activity-based protein profiling (ABPP)⁵⁵. For these reasons, a *chemo-enzymatic approach* (Figure 1(B)) with an opportunely selected azide-based BG-substrate (**BGSN3**) was set up: the efficient exposition of the azide outside the protein surface allows the reaction with a huge number of commercially DBCO-based

molecules, more than those BG-derivatives, keeping high the specificity in the presence of *in vitro* “perturbing” proteins (like in cell lysates) and the *in vivo* labelling of expressed SNAP-*tag*[®] in eukaryotic cells. Finally, **BGSN3** proved to be a good substrate for the direct immobilisation of these *tags* on solid surfaces. We demonstrated that splitting the SNAP-*tag*[®] reaction into two fast steps, as experimentally measured (Figure 2(A,B)), does not affect the overall rate and efficiency of the protein labelling^{13,24}, thus opening new perspectives and widening the applications of this powerful biotechnology.

Authors’ contributions

Conceptualisation, G.P. and A.Mi.; Methodology, R.Me., G.P.; Investigation, R.Me., D.C. and M.C.; Formal Analysis, C.M., R.Mi. and A.Ma.; Validation, R.Me., R.Ma., A.Mi. and G.P.; Writing – Original Draft, R.Me. and G.P.; Writing – Review & Editing, C.M., F.R. and A.V. and G.P.; Funding Acquisition, A.V. and G.P.; Resources, R.Me., R.Ma. and D.C.; Supervision, A.L., A.Mi. and G.P.

Acknowledgement

We are grateful to prof. Davide Prospero (University of Milan Bicocca) for the kind gift of the **MGPA** substrate.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by Ministero dell’Istruzione Università e Ricerca (MIUR) National Operational Program (PON) Research and Innovation 2014–2020 [CCI 2014IT16M2OP005], European Social Fund, Action 10.1 “Innovative Doctorates with Industrial characterization.” This research was also funded by Università del Piemonte Orientale (project: RIVmiggianorFAR_2019).

References

1. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* 2010;90:1103–63.
2. Liss V, Barlag B, Nietschke M, Hensel M. Self-labelling enzymes as universal tags for fluorescence microscopy, super-resolution microscopy and electron microscopy. *Sci Rep* 2015;5:17740–52.
3. Hatlem D, Trunk T, Linke D, Leo JC. Catching a SPY: using the SpyCatcher-SpyTag and related systems for labeling and localizing bacterial proteins. *Int J Mol Sci* 2019;20:2129–39.
4. Los GV, Encell LP, McDougall MG, et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* 2008;3:373–82.
5. England CG, Luo H, Cai W. HaloTag technology: a versatile platform for biomedical applications. *Bioconjug Chem* 2015; 26:975–86.
6. Pegg AE. Repair of O(6)-alkylguanine by alkyltransferases. *Mutat Res* 2000;462:83–100.
7. Daniels DS, Mol CD, Arvai AS, et al. Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *Embo J* 2000;19:1719–30.

8. Daniels DS, Woo TT, Luu KX, et al. DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nat Struct Mol Biol* 2004;11:714–20.
9. Keppler A, Gendreizig S, Gronemeyer T, et al. A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat Biotechnol* 2003;21:86–9.
10. Keppler A, Pick H, Arrivoli C, et al. Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc Natl Acad Sci USA* 2004;10:9955–9.
11. Juillerat A, Gronemeyer T, Keppler A, et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules *in vivo*. *Chem Biol* 2003;10:313–7.
12. Gronemeyer T, Chidley C, Juillerat A, et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng Des Sel* 2006;19:309–16.
13. Mollwitz B, Brunk E, Schmitt S, et al. Directed evolution of the suicide protein O6-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* 2012;51:986–94.
14. Gautier A, Juillerat A, Heinis C, et al. An engineered protein tag for multiprotein labeling in living cells. *Chem Biol* 2008;15:128–36.
15. Vogel P, Moschref Merkle MM, Li Q, et al. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat Methods* 2018;15:535–8.
16. Srikun D, Albers AE, Nam CI, et al. Organelle-targetable fluorescent probes for imaging hydrogen peroxide in living cells via SNAP-Tag protein labeling. *J Am Chem Soc* 2010;132:4455–65.
17. Abo M, Minakami R, Miyano K, et al. Visualization of phagosomal hydrogen peroxide production by a novel fluorescent probe that is localized via SNAP-tag labeling. *Anal Chem* 2014;86:5983–90.
18. Griss R, Schena A, Reymond L, et al. Bioluminescent sensor proteins for point-of-care therapeutic drug monitoring. *Nat Chem Biol* 2014;10:598–603.
19. Tomat E, Nolan EM, Jaworski J, Lippard SJ. Organelle-specific zinc detection using zinpyr-labeled fusion proteins in live cells. *J Am Chem Soc* 2008;130:15776–7.
20. Kamiya M, Johnsson K. Localizable and highly sensitive calcium indicator based on a BODIPY fluorophore. *Anal Chem* 2010;82:6472–9.
21. Saccà B, Meyer R, Erkelenz M, et al. Orthogonal protein decoration of DNA origami. *Angew Chem Int Ed Engl* 2010;49:9378–83.
22. Perugino G, Vettone A, Illiano G, et al. Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. *J Biol Chem* 2012;287:4222–31.
23. Perugino G, Miggiano R, Serpe M, et al. Structure-function relationships governing activity and stability of a DNA alkylation damage repair thermostable protein. *Nucleic Acids Res* 2015;43:8801–16.
24. Vettone A, Serpe M, Hidalgo A, et al. A novel thermostable protein-tag: optimization of the *Sulfolobus solfataricus* DNA-alkyl-transferase by protein engineering. *Extremophiles* 2016;20:1–13.
25. Visone V, Han W, Perugino G, et al. *In vivo* and *in vitro* protein imaging in thermophilic archaea by exploiting a novel protein tag. *PLoS One* 2017;12:e0185791.
26. Lo Gullo G, Mattossovich R, Perugino G, et al. Optimization of an *in vitro* transcription/translation system based on *Sulfolobus solfataricus* cell lysate. *Archaea* 2019;2019:9848253.
27. Merlo R, Del Prete S, Valenti A, et al. An AGT-based protein-tag system for the labelling and surface immobilization of enzymes on *E. coli* outer membrane. *J Enzyme Inhib Med Chem* 2019;34:490–9.
28. Del Prete S, Merlo R, Valenti A, et al. Thermostability enhancement of the α -carbonic anhydrase from *Sulfurihydrogenibium yellowstonense* by using the anchoring-and-self-labelling-protein-tag system (ASL^{tag}). *J Enzyme Inhib Med Chem* 2019;34:946–54.
29. Moschel RC, McDougall MG, Dolan ME, et al. Structural features of substituted purine derivatives compatible with depletion of human O6-alkylguanine-DNA alkyltransferase. *J Med Chem* 1992;35:4486–91.
30. Chae MY, McDougall MG, Dolan ME, et al. Substituted O6-benzylguanine derivatives and their inactivation of human O6-alkylguanine-DNA alkyltransferase. *J Med Chem* 1994;37:342–7.
31. Terashima I, Kawate H, Sakumi K, et al. Substrate specificity of human O6-methylguanine-DNA methyltransferase for O6-benzylguanine derivatives in oligodeoxynucleotides. *Chem Res Toxicol* 1997;10:1234–9.
32. McElhinney RS, Donnelly DJ, McCormick JE, et al. Inactivation of O6-alkylguanine-DNA alkyltransferase. 1. Novel O6-(hetaryl)methylguanines having basic rings in the side chain. *J Med Chem* 1998;41:5265–71.
33. Sun G, Fan T, Zhang N, et al. Identification of the structural features of guanine derivatives as MGMT inhibitors using 3D-QSAR modeling combined with molecular docking. *Molecules* 2016;21:823–44.
34. Zhang CJ, Li L, Chen GYJ, et al. One- and two-photon live cell imaging using a mutant SNAP-Tag protein and its FRET substrate pairs. *Org Lett* 2011;13:4160–3.
35. Huang M, Deng Z, Tian J, Liu T. Synthesis and biological evaluation of salinomycin triazole analogues as anticancer agents. *Eur J Med Chem* 2017;127:900–8.
36. Leng S, Qiao Q, Miao L, et al. A wash-free SNAP-tag fluorogenic probe based on the additive effects of quencher release and environmental sensitivity. *Chem Commun* 2017;53:6448–51.
37. Kindermann M, George N, Johnsson N, Johnsson K. Covalent and selective immobilization of fusion proteins. *J Am Chem Soc* 2003;125:7810–1.
38. Mattossovich R, Merlo R, Fontana A, et al. A journey down to hell: new thermostable protein-tags for biotechnology at high temperatures. *Extremophiles* 2020;24:81–91.
39. Rossi F, Morrone C, Massarotti A, et al. Crystal structure of a thermophilic O6-alkylguanine-DNA alkyltransferase-derived self-labeling protein-tag in covalent complex with a fluorescent probe. *Biochem Biophys Res Commun* 2018;500:698–703.
40. Banks JL, Beard HS, Cao Y, et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem* 2005;26:1752–80.
41. Toukmaji AY, Board JA. Jr., Ewald summation techniques in perspective: a survey. *Comput Phys Commun* 1996;95:73–92.
42. Zielkiewicz J. Structural properties of water: comparison of the SPC, SPCE, TIP4P, and TIP5P models of water. *J Chem Phys* 2005;123:104501.
43. Martyna GJ, Klein ML, Tuckerman M. Reversible multiple time scale molecular dynamics. *J Chem Phys* 1992;97:2635–43.

44. Evans DJ, Holian BL. The Nose–Hoover thermostat. *J Chem Phys* 1985;3:4069.
45. Miggiano R, Perugino G, Ciaramella M, et al. Crystal structure of *Mycobacterium tuberculosis* O6-methylguanine-DNA methyltransferase protein clusters assembled on to damaged DNA. *Biochem J* 2016;473:123–33.
46. Morrone C, Miggiano R, Serpe M, et al. Interdomain interactions rearrangements control the reaction steps of a thermostable DNA alkyltransferase. *Biochim Biophys Acta Gen Subj* 2017;1861:86–96.
47. Ranson M, Middleton MR, Bridgewater J, et al. Lomeguatrib, a potent inhibitor of O6-alkylguanine-DNA-alkyltransferase: phase I safety, pharmacodynamic, and pharmacokinetic trial and evaluation in combination with temozolomide in patients with advanced solid tumors. *Clin Cancer Res* 2006;12:1577–84.
48. Colombo M, Mazzucchelli S, Montenegro JM, et al. Protein oriented ligation on nanoparticles exploiting O6-alkylguanine-DNA transferase (SNAP) genetically encoded fusion. *Small* 2012;8:1492–7.
49. Song X, Wang C, Han Z, et al. Terminal alkyne substituted O6-benzylguanine for versatile and effective syntheses of fluorescent labels to genetically encoded SNAP-tags. *RSC Adv* 2015;5:23646–9.
50. Ferraris DM, Spallek R, Oehlmann W, et al. Structure of *Thermococcus litoralis* Δ^1 -pyrroline-2-carboxylate reductase in complex with NADH and L-proline. *Proteins* 2014;82:2268–74.
51. Huber W, Perspicace S, Kohler J, et al. SPR-based interaction studies with small molecular weight ligands using hAGT fusion proteins. *Anal Biochem* 2004;333:280–8.
52. Niesen J, Sack M, Seidel M, et al. SNAP-Tag technology: a useful tool to determine affinity constants and other functional parameters of novel antibody fragments. *Bioconjug Chem* 2016;27:1931–41.
53. Testa I, Wurm CA, Medda R, et al. Multicolor fluorescence nanoscopy in fixed and living cells by exciting conventional fluorophores with a single wavelength. *Biophys J* 2010;99:2686–94.
54. Miggiano R, Valenti A, Rossi F, et al. Every OGT is illuminated ... by fluorescent and synchrotron lights. *Int J Mol Sci* 2017;18:2613–31.
55. Zweerink S, Kallnik V, Ninck S, et al. Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea. *Nat Commun* 2017;8:15352.