Labeling of fusion proteins with synthetic fluorophores in live cells

Antje Kepper*, Horst Pick*, Claudio Arrivoli, Horst Vogel, and Kai Johansson†

Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Edited by Peter G. Schultz, The Scripps Research Institute, La Jolla, CA, and approved May 24, 2004 (received for review March 18, 2004)

A general approach for the sequential labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase (AGT) with different fluorophores in mammalian cells is presented. AGT fusion proteins with different localizations in the cell can be labeled specifically with different fluorophores, and the fluorescence labeling can be used for applications such as multicolor analysis of dynamic processes and fluorescence resonance energy transfer measurements. The facile access to a variety of different AGT substrates as well as the specificity of the labeling reaction should make the approach an important tool to study protein function in live cells.

Advances in fluorescence microscopy and the isolation and engineering of different autofluorescent proteins have revolutionized the way protein function is studied in the living cell (1–3). Beyond localization studies, fusion proteins of autofluorescent proteins can be used to study dynamic processes, follow conformational changes, and detect protein–protein interactions (1, 2, 4). Despite this enormous range of applications for autofluorescent fusion proteins, there are two reoccurring limitations of the approach. The first limitation is the restriction to the fluorophores of the existing autofluorescent proteins. Properties such as emission and excitation wavelength, extinction coefficient, and photostability are characteristic properties of each autofluorescent protein that do not necessarily match the requirements of the envisioned application. The second limitation is caused by the fusion of the autofluorescent protein to the protein of interest. In general, autofluorescent proteins possess a molecular mass of >25 kDa, and certain autofluorescent proteins furthermore form oligomers, raising the question of to what extent the creation of a fusion protein affects the function of the protein of interest (1, 3).

Alternative approaches to make proteins amenable to fluorescence studies in live cells include introduction of chemically labeled proteins through invasive techniques such as microinjection, site-specific incorporation of unnatural fluorescent amino acids, and the selective labeling of fusion proteins (5–10). An example of the latter approach is the so-called tetracysteine tag, which reacts with biarsenical fluorophores to form stable and highly fluorescent complexes (7, 8). We recently developed an approach that allows for the labeling of fusion proteins of human O6-alkylguanine-DNA alkyltransferase (AGT) with synthetic molecules (9). The labeling is based on the irreversible and specific reaction of AGT with O6-benzylguanine (BG) derivatives, leading to the transfer of the synthetic probe to a reactive cysteine residue (Fig. 1). Wild-type human AGT is a monomeric protein of 207 aa, the 30 C-terminal residues of which can be deleted without affecting the activity against BG, making it slightly smaller than autofluorescent proteins (11). Importantly, the rate of the reaction of AGT fusion proteins with BG derivatives is independent of the nature of the label, opening up the possibility to label a single AGT fusion protein with a variety of different probes (9, 12–14). Specific labeling of AGT fusion proteins in mammalian cells can be achieved by using AGT-deficient cell lines, and we have shown previously that nuclear-localized AGT can be labeled specifically with fluorescein in such cells (9). Here we demonstrate that synthetic BG derivatives can be used for the sequential labeling of AGT fusion proteins with a variety of different fluorophores in live mammalian cells and how the approach can be used for multicolor analysis of cellular processes and fluorescence resonance energy transfer (FRET) measurements.

Materials and Methods

Syntheses of AGT Substrates. All substrates for the labeling of AGT fusion proteins were prepared by reacting a previously described BG derivative with commercially available N-hydroxysuccinimide esters of the corresponding fluorophores (9), SNARF-1 (SF), fluorescein, Oregon green (OG), and tetramethylrhodamine (TMR) were used as mixtures of isomers. Reactions were performed at ambient temperature in N,N-dimethylformamide in the presence of 1 eq of triethylamine. AGT substrates were purified by reversed-phase HPLC on a C18-column by using linear gradients from 0% trifluoroacetic acid to acetonitrile. After solvent evaporation substrates were dissolved in dimethyl sulfoxide and stored at −20°C. Each AGT substrate was analyzed by high-resolution mass spectrometry and purity-checked by HPLC. Electrospray ionization (ESI)-MS (m/z) calculated for BGTMR [C38H35N8O5]⁺: 683.2730; found: 683.2747. ESI-MS (m/z) calculated for BGCy3 [C44H50N8O8S2]⁺: 749.1807; found: 749.1807. ESI-MS (m/z) calculated for BGSF [C42H34N7O7]⁺: 748.2519; found: 748.2524. ESI-MS (m/z) calculated for BGCy3 [C44H50N8O8S2]⁺: 882.3193; found: 882.3203. For the determination of the fluorescence spectra of the fluorophores bound to AGT, purified glutathione S-transferase (GST)-AGT fusion protein (10 μM) was incubated with an excess of the corresponding AGT substrate (20 μM) in 50 mM Hepes/1 mM DTT, pH 7.2. After 1 h, excess AGT substrate was removed through gel filtration by using a NAP-10 column (Amersham Pharmacia Biosciences), and fluorescence spectra were recorded on a Spectramax Gemini plate reader (Molecular Devices). Absorption and fluorescence spectra for BG- and AGT-bound fluorophores as well as their relative quantum yields are shown in the supporting information, which is published on the PNAS web site.

Construction of AGT Fusion Proteins. The AGT mutant used throughout this study is based on human AGT, which was modified by the following amino acid changes: Lys-125 → Ala, Ala-127 → Thr, Arg-128 → Ala, Asn-157 → Gln, and Ser-159 → Glu (14). The mutations at positions 157 and 159 have been shown to increase the activity of AGT against BG, and the mutations at positions 125, 127, and 128 were introduced to disrupt the interaction of the protein with DNA (14, 15). For

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: AGT, O6-alkylguanine-DNA alkyltransferase; BG, O6-benzylguanine; FRET, fluorescence resonance energy transfer; OG, Oregon green; TMR, tetramethylrhodamine; SF, SNARF-1; β-Gal, β-galactosidase; tsVSVG, thermosensitive glycoprotein of vesicular stomatitis virus; NK6, neurokinin-1; EGF, enhanced GFP; NLS3, three consecutive simian virus 40 nuclear localization sequences; CHO, Chinese hamster ovary; AF, diaциolated fluorescein; SP-h.o, NK-specific ligand substance P.

* A.K. and H.P. contributed equally to this work.
† To whom correspondence should be addressed. E-mail: kai.johansson@epfl.ch.

© 2004 by The National Academy of Sciences of the USA
targeting AGT to different subcellular localizations, the following constructs were made. For expression of farnesylated AGT (AGT-CaaX), the peptide RSKLNPDESGFGCMSCCKVLS containing a CaaX motif was fused to the C terminus of AGT (16). The corresponding gene was inserted into the NheI and EcoRI sites of the vector pEFGFP-F (Clontech). For expression of AGT-β-galactosidase (β-Gal), AGT (residues 1–177) was fused via an RSGSG linker to the N terminus of β-Gal. The corresponding gene was inserted into the NheI and XbaI sites of the vector pECPF-Nuc (Clontech). For expression of the thermosensitive glycoprotein of vesicular stomatitis virus (tsVSVG)-AGT, AGT was fused to the C terminus of tsVSVG via an SGLRS linker (17). The corresponding gene was inserted into the NheI and BamHI sites of the vector pECPF-Nuc (Clontech). For expression of AGT-α-tubulin, AGT was fused to the N terminus of human α-tubulin via an RSRR linker by ligating the AGT gene into the vector pEFGFP-Tub (Clontech) using the NheI and BglII restriction sites. For the expression of AGT-neurokinin-1 (NK1), the signal sequence of the 5-HT1 receptor (Sig5HT1) was fused to the N terminus of AGT via a DYS linker, and NK1 was fused to the C terminus of AGT via a TS linker (18, 19). In the resulting construct, a FLAG tag and a 6×HIS tag were also attached to the C terminus of NK1. The corresponding gene of the fusion protein was inserted into the NheI and BamHI sites of the vector pCEP4 (Invitrogen). For the expression of AGT-EGFP-NLS3, AGT was fused to the N terminus of enhanced GFP (EGFP) via a GSG linker, and three consecutive simian virus 40 nuclear localization sequences (NLSs) were attached to the C terminus of EGFP (20). The corresponding gene was inserted into the NheI and BamHI sites of the vector pECEFP-Nuc (Clontech).

Fluorescence Labeling of AGT Fusion Proteins. Chinese hamster ovary (CHO)-9-neo-C5, deficient of AGT, and HEK293 were grown in DMEM/F12 supplemented with 2.2% FCS (GIBCO/BRL) (21). Transient transfection was performed as described (9). For the labeling experiments, all AGT substrates were added to a final concentration of 5 μM in PBS. BG linked to diacylated fluorescein (BGAf), BGOG, and BGCy3 were added at room temperature for 5 min before washing the cells three times with PBS. Cells were imaged in PBS 30 min after washing using appropriate filter sets for each fluorophore. The labeling with BGSF was followed online for 1–2 h at room temperature. If necessary, excess BGSF was removed through washing steps as described above. For labeling with BGTRM, cells were incubated for at least 1.5 h before excess fluorophore was removed through washing steps.

For AGT-NLS3, the fluorescence intensity of the nucleus of at least 20 randomly picked cells was determined 80, 160, 285, 430, and 510 min after labeling (supporting information). In CHO cells expressing AGT-NLS3 in medium containing 10 μM BG, the fluorescence intensity of at least 12 randomly picked cells was determined 70, 180, and 400 min after labeling (supporting information).

Results and Discussion

Fluorophores for the Labeling of AGT Fusion Proteins. AGT fusion proteins offer the possibility to label individual fusion proteins with a variety of different fluorophores. Although we have shown previously that AGT fusion proteins can be labeled specifically with fluorescein, fluorophores exhibiting higher photostability and higher emission wavelengths would be advantageous for functional studies in living cells (9). We therefore synthesized AGT substrates in which BG is linked to the corresponding acetates to increase cell permeability. Inside the cell, the nonfluorescent acetates are hydrolyzed by esterases, liberating the fluorophore. Except for BGCy3, all of the fluororescent AGT substrates proved to be cell-permeable, allowing for the labeling of nuclear-localized AGT (AGT-NLS3) in AGT-deficient CHO cells (Fig. 2A–C). In general, labeling comprises incubation of the cell with the AGT substrate for a period characteristic for each substrate and subsequent washing steps to remove excess fluorophores. An exception is BGSF, for which fluorescence labeling of AGT-NLS3 with SF can be followed online without the need to remove excess dye through washing steps (Fig. 2C). This can be rationalized assuming that intracellular BGSF is scavenged by AGT-NLS3 before it is hydrolyzed to liberate the fluorophore. A subsequent washing step can then be used to further decrease background fluorescence.

Fluorescence Labeling of Different AGT Fusion Proteins. To test the scope of our labeling approach for live-cell applications, we first...
investigated the fluorescence labeling of AGT fusion proteins at different subcellular localizations. The AGT fusion proteins investigated included β-Gal, α-tubulin, a farnesylation motif, the tsVSVG, and the G protein-coupled receptor NK1.

When β-Gal was fused to the C terminus of AGT, fluorescence labeling of the corresponding AGT-β-Gal fusion in CHO cells revealed a diffuse cytosolic localization of the fusion protein (Fig. 2D). These data indicate that fusion of AGT to cytosolic proteins does not lead to active transport of these fusions into the nucleus of a cell, an important prerequisite for functional studies of cytosolic proteins. It should be noted that the predominant nuclear localization of wild-type human AGT is due to its affinity for DNA and that the AGT mutant used throughout this study possesses mutations that interfere with its interaction with DNA and enhance its activity against BG (12, 14). Both EGFP and fluorescein-labeled AGT display an almost even distribution between cytosol and nucleus when expressed without any fusion information. To investigate whether AGT fusion proteins can be transiently expressed in CHO cells at the permissive temperature of 34°C, we examined whether AGT fusion proteins, despite possessing a reactive cysteine, can also be functionally expressed on the extracellular side of the plasma membrane. Toward this end, AGT was fused to the exoplasmic N terminus of the G protein-coupled NK1 receptor (AGT-NK1) (19). HEK293 cells transiently expressing AGT-NK1 could be labeled with fluorescein as well as with the nonpermeable dye Cy3, indicating a preserved activity of extracellular AGT toward its substrate (Fig. 3A and B). The use of a nonpermeable AGT substrate for extracellular applications has the advantage that both endogenous AGT as well as not properly translocated or internalized AGT-NK1 will not be labeled. This differentiation between different receptor populations is not possible when using protein fusions with autofluorescent proteins. To investigate whether AGT-NK1 is functional, the binding of the TMR-labeled, NK1-specific ligand substance P (SP-rho) to cells expressing AGT-NK1 was studied. Fluorescence signals resulting from labeling of AGT-NK1 with fluorescein and binding of its cognate ligand SP-rho overlapped (Fig. 3A and C). In addition, fluorescence labeling resulting from binding of SP-rho could be readily reversed by adding excess of unlabeled SP, indicating preserved functionality of the AGT-NK1 receptor in terms of ligand binding (Fig. 3D). Ligand-induced activation of the AGT-NK1 fusion protein also couples to the intracellular Ca2+ signaling pathway. Binding of SP (1 μM) to HEK293 cells coexpressing the AGT-NK1 receptor and the promiscuous Goα protein elicited a transient Ca2+ response, with a rapid increase in intracellular Ca2+ concentration and slower decay (10–20 s) back to baseline as detected by confocal microscopy using a calcium-sensitive fluorescence indicator (unpublished data) (24). These experiments demonstrate that the localization of the investigated AGT fusion proteins was determined exclusively by either specific targeting sequences or the natural subcellular localization of the fusion partner to which the AGT was genetically attached.

A characteristic property of AGT is that it is degraded after transfer of the alkyl group from O6-alkylated guanine-DNA to the reactive cysteine. It is believed that a conformational change of the alkylated protein triggers its degradation, presumably through ubiquitinylation of the alkylated protein and subsequent proteolysis (25). To investigate both the rate of degradation of AGT fusion proteins and to what extent the degradation after labeling interferes with fluorescence measurements, the decrease in fluorescence intensity for two AGT fusions proteins (AGT-NLS3 and AGT-CaaX) was determined over a period of 7 h. During these measurements, the fluorescence intensity for both proteins decreased, but not more than 12% per hour (supporting information). After 7 h of incubation, the intensity of the fluorescence signal for AGT-NLS3 and AGT-CaaX was 30% and 50%, respectively, of the value measured directly after the labeling. The observed rate of degradation is comparable with that measured in earlier experiments in which the fate of alkylated wild-type HAGT was followed by Western blotting using anti-hAGT antibodies (26). These data demonstrate that labeled AGT fusion proteins can be followed inside living cells for several hours after the labeling.
FRET and Multicolor Analysis. FRET measurements between pairs of (auto)fluorescent proteins are an important tool to investigate the spatial distance between proteins in a time-resolved manner (27). Furthermore, various reporter systems to monitor cellular parameters are based on intramolecular FRET between two autofluorescent proteins fused in a single polypeptide chain (1). To demonstrate the use of AGT fusion proteins in FRET applications, we constructed a fusion protein between EGFP and AGT-NLS3 (AGT-EGFP-NLS3). Here, EGFP would be the donor (488 nm) and SF-labeled AGT would be the acceptor for intramolecular FRET. The broad absorbance of SF-labeled AGT in the region from 500 to 580 nm at pH 7.4 makes it an ideal acceptor in such experiments (supporting information). Labeling of AGT-EGFP-NLS3 with $^{5}$H9262 M BGSF for 1 h led to a 98% decrease in EGFP emission at 505–530 nm, indicating both efficient FRET and an efficient labeling of AGT-EGFP-NLS3 with SF (Fig. 3 E–I). Because SF can also be excited to some extent at 488 nm, we assume that the observed emission at 650 nm is caused by both direct laser excitation and FRET (Fig. 3 H).

Preincubation of cells with BG (20 $^{9262}$ M) completely prevented the quenching of the EGFP fluorescence of AGT-EGFP-NLS3 after incubation with BGSF (Fig. 3 I). Furthermore, simultaneous expression of AGT-NLS3 and EGFP-NLS3 as two separate proteins and subsequent labeling with SF did not lead to any significant quenching of the fluorescence of EGFP-NLS3 (unpublished data). Potential applications of intramolecular FRET after labeling with SF (excitation 488 nm/emission 505–530 nm). (I) Overlay of transmission and fluorescence channel for SF after labeling with SF (excitation 488 nm/emission > 650 nm). (J) Overlay of transmission and fluorescence channel for SF after labeling with SF (excitation 488 nm/emission > 650 nm). (K) Overlay of transmission and fluorescence channel for SF after labeling with SF (excitation 488 nm/emission > 650 nm). (L) Relative fluorescence intensities (RFI) of the EGFP channel before and after labeling with SF. Columns E and G correspond to the fluorescence intensities (RFI) of the EGFP channel of the cell shown in E and G. Columns J and K correspond to the fluorescence intensities (RFI) of the EGFP channel of a cell before and after incubation with SF but previously incubated with BG (20 $^{9262}$ M). (Scale bars 10 $^{9262}$.)
between an autofluorescent protein and a fluorescence-labeled AGT are as reporter systems for measuring biochemical events in live cells (1). Here, such a FRET pair might prove superior to the currently used donor–acceptor pair of enhanced cyan fluorescent protein and yellow fluorescent protein (1). In addition, these experiments demonstrate that AGT fusion proteins in general are well suited for FRET applications.

Sequential labeling of proteins with different fluorophores offers the possibility to distinguish older and newer copies of the same protein within one cell through multicolor analysis. Multicolor analysis using the tetracysteine tag and biarsenical dyes has been applied to study the dynamics of gap-junction channels (28). To demonstrate the feasibility of multicolor analysis of AGT fusion proteins, we attempted to differentiate tsVSVG-AGT populations expressed before and after a temperature shift through sequential labeling with different fluorophores (Fig. 4).

In these experiments, CHO cells transiently expressing tsVSVG-AGT were incubated for 20 h at 34 °C and subsequently labeled with fluorescein. To quench residual AGT fusion protein not labeled with fluorescein, cells were incubated for 15 min with the AGT inhibitor Oβ-(4-bromothenyl)-guanine (29). The temperature of the medium then was shifted to 40 °C, which should lead to misfolding of newly synthesized tsVSVG-AGT and its retention in the endoplasmic reticulum. After incubation of the cells for 75 min at 40 °C, tsVSVG-AGT synthesized at this temperature was labeled with SF. Subsequent fluorescence imaging of cells demonstrated that fluorescein-labeled tsVSVG-AGT was located predominantly in the plasma membrane (77%), whereas SF-labeled tsVSVG-AGT was predominantly located in internal membrane structures (71%), most likely perinuclear endoplasmic reticulum or Golgi (Fig. 4 A–C) (17, 30). The exact localization of retained tsVSVG-AGT was not characterized further. The labeling and washing at 40 °C unavoidably lead to temperature fluctuations, and because tsVSVG rapidly folds at permissive temperatures (78 °C), one would expect a heterogeneous distribution of retained tsVSVG-AGT (31, 32). Despite the ambiguities with respect to the localization of retained tsVSVG-AGT, the data clearly demonstrate that within live cells older and newer copies of AGT fusion proteins can be discriminated by sequential labeling with different fluorophores. In addition to sequential labeling, AGT fusion proteins can also be simultaneously with two different fluorophores. If fluorophores with different photostability would be used, such double-labeling experiments could be exploited to study dynamic processes by using fluorescence localization after photobleaching (33). The experiments suggest that multicolor analysis of AGT fusion proteins has the potential to become an important tool to study dynamic processes in live cells.

Labeling in Cell Lines Expressing Endogenous AGT. The scope of applications for the fluorescence labeling of AGT fusion proteins would be significantly broadened if AGT fusion proteins could be labeled specifically in cells expressing endogenous wild-type AGT. Because of the 20-fold higher activity of the here-used AGT mutants relative to wild-type AGT, a specific labeling of (overexpressed) AGT fusion proteins in cells possessing wild-type AGT is already possible in certain cases. An example is the fluorescein labeling of transiently expressed AGT-CaaX in HEK293 cells; Transfected cells can be identified by a fluorescent plasma membrane, whereas nontransfected cells display a fluorescent nucleus (Fig. 4 D and E). We assume that the specificity of the labeling observed in the transfected cells, i.e., the absence of a fluorescent nucleus, relies on the increased activity of AGT-CaaX relative to wild-type AGT and a substrate-specific labeling.

Conclusions

AGT fusion proteins can be labeled specifically in a sequential manner with synthetic fluorophores in live cells. The localization of the investigated fusion proteins was determined exclusively by either specific targeting sequences or the natural subcellular localization of the fusion partner to which the AGT was genetically attached. The stability of fluorescence-labeled AGT fusion proteins allows following their fate over periods of hours. Importantly, a single AGT fusion protein can be (sequentially) labeled with a variety of different fluorophores for applications such as multicolor analysis or FRET applications. These features should make AGT fusion proteins an important tool for fluorescence-based studies of protein function in live cells.

We thank Jan Ellenberg for advice and for providing the gene of tsVSVG. This work was supported by Swiss National Science Foundation Grant 3152A0-100103 (to K.J.) and by the TopNano program (H.V.).
Absorbance and fluorescence spectra of BG- and AGT-coupled fluorophores

For the determination of the fluorescence spectra of the fluorophores bound to AGT, purified glutathione S-transferase-AGT (GST-AGT) fusion protein (10 \( \mu \)M) was incubated with an excess of the corresponding AGT substrate (20 \( \mu \)M) in 50 mM Hepes, 1 mM DTT, pH 7.4. After 1 h, excess AGT substrate was removed through gel filtration using a NAP-10 column (Amersham Biosciences) and fluorescence spectra were recorded on a Spectramax Gemini plate-reader (Molecular Devices). In the following figures, absorption and fluorescence spectra for BG- and AGT-bound fluorophores are shown. The relative quantum yields for BG- and AGT-bound fluorophores with respect to the underivatized fluorophore are listed in Supporting Information Table 1. The data for BGAF, BGSF and BGOG are those of the corresponding deacetylated molecules.

**Supporting Information Figure 1:** Absorption and fluorescence spectra for BG- and AGT-bound fluorescein.
Supporting Information Figure 2: Absorption and fluorescence spectra for BG- and AGT-bound Oregon Green.

Supporting Information Figure 3: Absorption and fluorescence spectra for BG- and AGT-bound Cy3.
Supporting Information Figure 2: Absorption and fluorescence spectra for BG- and AGT-bound Oregon Green.

Supporting Information Figure 3: Absorption and fluorescence spectra for BG- and AGT-bound Cy3.
Supporting Information Figure 4: Absorption and fluorescence spectra for BG- and AGT-bound TMR.

Supporting Information Figure 5: Absorption and fluorescence spectra for BG- and AGT-bound SNARF-1.
Supporting Information Figure 4: Absorption and fluorescence spectra for BG- and AGT-bound TMR.

Supporting Information Figure 5: Absorption and fluorescence spectra for BG- and AGT-bound SNARF-1.
Supporting Information Figure 6: A) Intensity of fluorescence signal (RFI) resulting from labeling of AGT-CaaX as a function of time. Intensity of signal obtained directly after labeling was set to 1. B) Intensity of fluorescence signal (RFI) resulting from labeling of AGT-NLS₃ as a function of time. Intensity of signal obtained directly after labeling was set to 1.
**Distribution of transiently expressed EGFP-α-tubulin in CHO cells.**

For expression of EGFP-α-tubulin, the vector pEGFP-Tub (Clontech) was used. CHO-9-neo-C5 were grown in DMEM/F12 (Dulbecco’s modified Eagle medium) supplemented with 2.2 % fetal calf serum (Gibco, BRL). Transient transfection was performed as described in the manuscript for AGT fusion proteins. Laser-scanning confocal micrographs were recorded using a 488 nm argon/krypton laser line on a Zeiss LSM 510 microscope (Carl Zeiss AG) with a 63 x water (1.2 numerical aperture) objective. Scanning speed and laser intensity were adjusted to avoid photobleaching of the fluorescent probes and damage or morphological changes of the cells. The fluorescence micrograph (Supporting Information Figure 7A) indicate that, as observed for AGT-α-tubulin (Supporting Information Figure 7B), EGFP-α-tubulin appears is located both in the cytosol and microtubules.

**Supporting Information Figure 7:** A) Fluorescence micrograph of EGFP-α-tubulin transiently expressed in CHO-9-neo-C5. B) Fluorescence micrograph of AGT-α-tubulin transiently expressed in CHO-9-neo-C5 and after labeling with fluorescein using BGAF.