Protein Engineering, Design & Selection vol. 19 no. 7 pp. 309–316, 2006
Published online April 25, 2006 doi:10.1093/protein/gzl014

Directed evolution of O^6-alkylguanine-DNA alkyltransferase for applications in protein labeling

Thomas Gronemeyer, Christopher Chidley, Alexandre Juillerat, Christian Heinis and Kai Johnsson[1]

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

To whom correspondence should be addressed.
E-mail: kai.johnsson@epfl.ch

The specific reaction of O^6-alkylguanine-DNA alkyltransferase (AGT) with O^6-benzylguanine (BG) derivatives allows for a specific labeling of AGT fusion proteins with chemically diverse compounds in living cells and *in vitro*. The efficiency of the labeling depends on a number of factors, most importantly on the reactivity, selectivity and stability of AGT. Here, we report the use of directed evolution and two different selection systems to further increase the activity of AGT towards BG derivatives by a factor of 17 and demonstrate the advantages of this mutant for the specific labeling of AGT fusion proteins displayed on the surface of mammalian cells. The results furthermore identify two regions of the protein outside the active site that influence the activity of the protein towards BG derivatives.
Keywords: directed evolution/O^6-alkylguanine-DNA alkyltransferases/phage display/protein labeling

Introduction

The specific labeling of proteins in living cells with synthetic probes is an important approach to study protein function. One way to achieve such a labeling is through expression of the protein of interest as a fusion protein with an additional polypeptide (a so-called tag) that mediates the labeling (Johnsson and Johnsson, 2003). Factors that affect the feasibility and attractiveness of such an approach are (i) the size of the tag, (ii) the specificity and speed of the labeling, (iii) the availability of a broad range of different probes and (iv) properties of the tag that affect the function of the fusion protein such as tag-dependent localization or stability. We have previously used human O^6-alkylguanine-DNA alkyltransferase (AGT) as a tag for the specific labeling of AGT fusion proteins in living cells as well as *in vitro* (Keppler *et al*., 2003). The labeling is achieved through the reaction of O^6-benzylguanine (BG) derivatives with AGT, leading to an irreversible transfer of the derivatized benzyl group to the active site cysteine of AGT. In earlier experiments we optimized the properties of AGT with respect to a number of important parameters (Table I) (Juillerat *et al*., 2003; Juillerat *et al*., 2005). Firstly, we used saturation mutagenesis of active site residues to increase the activity of AGT versus BG derivatives ~20-fold relative to wild-type human AGT (Juillerat *et al*., 2003). Importantly, such an increased activity directly translates into a more efficient labeling of AGT fusion proteins in living cells and *in vitro*. Secondly, we truncated the size of AGT to 182 residues, removed non-essential cysteines to facilitate folding under oxidizing conditions and introduced mutations to prevent the binding of AGT fusion proteins to DNA (Juillerat *et al*., 2005). Thirdly, we introduced mutations into AGT to render it resistant against an inhibitor of wild-type AGT (Juillerat *et al*., 2005). The latter point is important in case endogenous AGT interferes with the specific labeling of AGT fusion proteins. We then assembled all 15 different mutations in a single protein which was named ^{M}AGT (Table I). Although ^{M}AGT conserved beneficial properties of its predecessors such as lack of affinity for DNA and resistance against an inhibitor of wild-type AGT, its activity decreased by a factor of ~8 relative to its most active predecessor (Juillerat *et al*., 2005). The observed decrease in activity diminishes the practical value of the mutant for applications in living cells as well as *in vitro*.

Here, we describe our efforts to restore the decreased activity of this AGT mutant. Using directed evolution, we were able to generate an AGT mutant with a 17-fold increased activity, generating the most active AGT mutant characterized so far.

Materials and methods

Phage selections

The error-prone library insert was prepared by a modified PCR as described elsewhere (Zaccolo and Gherardi, 1999). Saturation mutagenesis of residues 31–35 and 150–154 was performed by overlap extension PCR using the synthetic oligonucleotides with the corresponding codons randomized as NNK. The PCR products were ligated into the phagemid pAK100 and electroporated into *Escherichia coli* JM109 (Krebber *et al*., 1997). The colonies were scraped off the plates, resuspended in 2YT medium containing 20% glycerol and stocked at −80°C prior to further use.

Phages were prepared as described previously (Juillerat *et al*., 2003). For the phage selections, BGDG or BGFL were added to 500 µl of phage preparation to a final concentration of either 5 or 50 nM and incubated for 10 min. Phages were subjected to two cycles of precipitation with 4% w/v polyethylene glycol 8000 and 3% w/v NaCl. After resuspension of the phages in 500 µl of phosphate-buffered saline (PBS), 500 µl of PBSMM (PBS + 4% w/v skimmed milk powder) was added and the solution rotated for 60 min at room temperature. The sample was added to 25 µl of anti-digoxigenin antibody-coated magnetic beads (Roche) or to 100 µl of anti-fluorescein antibody-coated magnetic beads (Polyscience) and rotated for 30 min. The beads were washed twice with PBS and blocked for 60 min with PBSMM before addition of the phage. After addition of the phages the beads were washed three times with PBSM (PBS + 2% w/v skimmed milk powder), five times with PBST [PBS/0.05% (v/v) Tween-20] and twice with PBS. Phages were eluted by incubating the beads with 100 µl of 0.1 M glycine (pH 2.5) for 5 min. The supernatant was neutralized with 50 µl of 1 M

Table I. Engineering of AGT for labeling applications

Mutations	Phenotype	Reference
N157G, S159E	23-Fold increase in activity relative to wild-type	Juillerat *et al.* (2003)
C62A, C150N	Removal of non-essential cysteines	Juillerat *et al.* (2005)
G131K, G132T, M134L, R135S	Resistance against N^9-substituted BG derivatives used for inhibition of wild-type	Juillerat *et al.* (2005)
Q115S, Q116H, K125A, A127T, R128A, S151I, S152N,	Suppress affinity towards DNA	Keppler *et al.* (2004b); Juillerat *et al.* (2005)
Delete residues 183–207	Reduce size of protein	Keppler *et al.* (2004b)
Combination of mutations listed above (^{M}AGT)	8-Fold decrease in activity versus BG relative to most active mutant described previously	Juillerat *et al.* (2005)

Tris–HCl (pH 8.0). *E.coli* JM109 (OD_{600} = 0.5–0.6) were infected with the eluted phage, plated on 2YT plates containing 25 μg/ml Cam and incubated overnight at 37°C. The colonies were scraped off the dishes, aliquoted and stored at −80°C prior to the next round of selection. After the selections, the mutants were characterized as described below.

Yeast three-hybrid selections

The error-prone PCR library was generated by two sequential PCR reactions. First, the gene of ^{M}AGT was amplified using *Taq* polymerase under standard conditions used for error-prone PCR and the resulting PCR product reamplified using the GENEMORPH kit (Stratagene). The final PCR product was ligated into the pHybLex/Zeo plasmid (Invitrogen) and electroporated into *E.coli* XL1-Blue, yielding a library of at least 3.2×10^5 independent clones. To check the success of this strategy, the sequences of 10 randomly picked clones were determined by DNA sequencing. Colonies were scraped off the plates and the plasmid DNA was isolated for further transformation into L40 yeast cells.

The large-scale transformation of the yeast strain L40 containing the plasmid for expression of B42–DHFR (dihydrofolate reductase) was performed using standard procedures (Gietz and Woods, 2002). For the selections, the transformed yeast were grown on a plate lacking uracil, tryptophane and histidine but containing zeocin (200 μg/ml) and BGMtx (0.25 μM) at 30°C for 5 days. The plasmid DNA (pHybLex/Zeo–AGT) of 24 of the fastest growing colonies was isolated by standard phenol:chloroform precipitation using acid-washed glass beads, electroporated in XL1-Blue and analyzed by DNA sequencing. The genes of the selected AGT mutants were subcloned into the pGEX-2T plasmid (Amersham Biosciences) as described elsewhere (Juillerat *et al.*, 2003), allowing for the expression of the AGT mutants as GST fusion proteins. Activity assays and competition assays were performed by a fluorescence assay as described elsewhere (Juillerat *et al.*, 2005). The observed rate constants were obtained by measuring the rate of labeling under pseudo first-order conditions and dividing the resulting pseudo first-order rate constant by the concentration of BGCy3. If not specified otherwise, given values are average values of at least three measurements.

Expression and labeling of GPI-anchored AGT mutants on mammalian cells

The gene of AGT with a C-terminal HA epitope tag was inserted downstream of the endoplasmic reticulum signal sequence of huCD59 (amino acids 1–25: MGIQGGSVLFGLLLVLAVFCHSGHSRS) and upstream of the glycosylphosphatidylinositol (GPI) signal sequence of huCD59 (amino acids 100–128: EFLENGGTSLSEKTVLLLVTPFLAAAWSLHP Stop), which contains the GPI attachment site at the underlined asparagine 102 (Chen *et al.*, 2001; Hiscox *et al.*, 2002). The propeptide 103–128 is removed in the mature form and the numbering corresponds to the huCD59 sequence. This construct was inserted into the mammalian expression vector pEGFP-F (Clontech) using the *Nhe*I and *Sal*I restriction sites. Labeling experiments were carried out in HEK293T cells. Cells were cultured in DMEM/F12 medium (Cambrex) supplemented with 5% FCS in humidified atmosphere under 5% CO_2. Twenty-four hours before transfection, cells were seeded in a 12-well cell culture plate to a density of 100 000 cells per well. One hour before transfection, the medium was exchanged for DMEM/F12, 3% FCS. Transfection was performed with calcium phosphate as described elsewhere (Keppler *et al.*, 2004a). An aliquot of 2.5 μg DNA per well (1.5 μg AGT containing plasmid and 1.0 μg pEGFP-N1) were used for transfection. For *in vivo* labeling, the medium was aspirated 24 h after transfection and the cells were incubated for 5 min with 2.5 μM BGCy3 in Dulbecco's PBS (DPBS). Afterwards, the cells were washed three times for each 10 min with DPBS.

Imaging was carried out with a Zeiss Axiovert200 microscope, equipped with LD Plan Neofluar ×40 (0.6 corr. phase 2) and ×63 (0.75 corr. phase 2) objectives and an AxioCam MR digital camera (Zeiss). Zeiss filter sets 10 (excitation 450–490 nm; emission 515–565 nm) and 43 (excitation 545 ± 25 nm; emission 605 ± 70 nm) were used for fluorescence microscopy. Image analysis was performed with the AxioVision 4.0 (Zeiss) software.

For western blot analysis, HEK293T cells were grown in suspension culture in ExCell-293 medium (JRH Biosciences). Transfection and protein expression was carried out in a 12-well cell culture plate. For transfection, 7.5 μg linear PEI (polyethylenimine in water, pH 7.1) and 2.5 μg DNA where mixed and diluted to a final volume of 100 μl (in 150 mM NaCl) and incubated for 10 min at room temperature. This transfection cocktail was added to 1 ml of cell suspension [2×10^6 cells/ml in RPMI-1640 medium (Cambrex) containing 25 mM HEPES pH 7.3]. After 4 h of incubation (humidified atmosphere, 5% CO_2) the transfection mixture was diluted with 1 ml of Pro293s-CD medium (Cambrex) and incubation was continued for 24 h (humidified atmosphere, 5% CO_2). Suspension cultures were agitated (150–200 r.p.m.) during all incubations. The cells were harvested by centrifugation, washed once with 500 μl of PBS and resuspended in lysis buffer (150 mM KH_2PO_4 pH 7.0; 100 mM NaCl; 0.5 mM EDTA; 0.1% TX100) and subjected to three freeze-thaw cycles (liquid nitrogen/37°C water bath). Cell debris and the nuclei were separated by centrifugation and the resulting

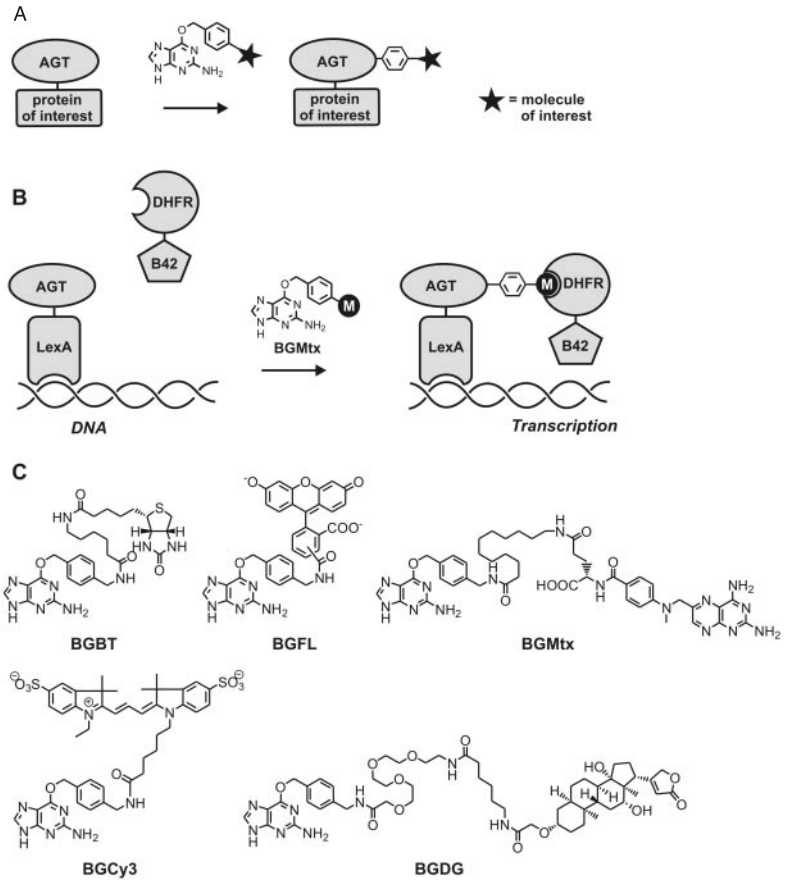

Fig. 1. (**A**) General scheme for labeling of AGT fusion proteins. (**B**) AGT-based three-hybrid system. AGT is expressed as fusion with the DNA-binding domain LexA and DHFR as fusion protein with the transcriptional activation domain B42. Labeling of the AGT fusion protein in yeast with methotrexate using BGMtx (C) induces dimerization of the AGT and DHFR fusion proteins and transcription of reporter genes (here *HIS3* and *lacZ*) under LexA. (**C**) BG derivatives used in this work.

extract was directly subjected to western blot analysis using a mouse anti-HA antibody as primary antibody and a goat anti-mouse IgG antibody–peroxidase conjugate as secondary antibody.

Results

Random mutagenesis of AGT

As a strategy to isolate AGT mutants with increased activity against BG derivatives we focused on random mutagenesis of ^{M}AGT. To select AGT mutants out of libraries generated through random mutagenesis we used phage display (Juillerat *et al*., 2003). Here, phages displaying AGT on the tip of the phage as a pIII fusion protein were incubated first with BG-digoxigenin (BGDG, Fig. 1) and subsequently with beads covered with anti-digoxigenin antibodies, thereby allowing for an enrichment of phages displaying reactive AGT. For the selections using phage display, the gene of ^{M}AGT was subjected to error-prone PCR and subsequently cloned into phagemid pAK100, yielding a library of 4×10^7 independent clones. ^{M}AGT contained in average 2–3 base mutations per gene as determined by the sequencing of five randomly picked clones. Two selections were run in parallel: in one selection, phages were incubated with 5 nM BGDG for 10 min before recovering labeled phage; in the other the BGDG concentration was increased to 50 nM. Both selections were repeated for six rounds after which individual clones were assayed in cell extracts for activity. A total of 49 clones from both selections were analyzed and the sequences of 21 clones with significantly higher activity determined through DNA sequencing (Table II). A clone with the single mutation A154T, which lead to an ~4-fold increase in activity, was found seven times and 15 clones in total contained the mutation A154T alone or in combination with other mutations. The most active mutant possessed a 6-fold higher activity than ^{M}AGT and contained, in addition to the mutation A154T, the mutations L33F, V44A and V52A.

To complement the phage selections with another selection system which might be better suited for proteins used in intracellular applications we focused on a previously established AGT-based yeast three-hybrid system in which the reaction of AGT with BG-methotrexate (BGMtx, Fig. 1) leads to transcription of the reporter genes *HIS3* and *lacZ* (Gendreizig *et al*., 2003). As the growth of yeast on plates lacking histidine and containing BGMtx critically depends on the activity of AGT, we envisioned isolating AGT mutants out of libraries generated through error-prone PCR in a simple growth assay. For the yeast three-hybrid selections, the gene of ^{M}AGT was also subjected to random mutagenesis using error-prone PCR and subsequently cloned into the plasmid pHybLex/Zeo, yielding a library of 3×10^5 independent clones after transformation of *E.coli*. In this construct, ^{M}AGT is expressed as a LexA-^{M}AGT fusion protein and ^{M}AGT contained in average three base mutations per gene as

Table II. AGT mutants selected using phage display[a]

Clone	Mutations	k_{rel}
A1	A154T	4.2
A2	K104E, T127A, A154T	4.5
A3	A154T	4.4
A4	A154T	4.1
A5	K107R, A154T	4.7
A6	K8T, T127A, A154T, H174R	5.8
A7	T127A	2.2
A8	N123S	1.3
A9	A154T	4.1
A10	A154T	3.0
A11	L33F, V44A, V52A, A154T	6.1
A12	V46A, A50V, P58V, A154T	3.4
A13	V149I, A154T	4.5
A14	A154T	4.2
A15	H71Y, A154T	3.3
A16	K8T, A51T, I112V, A154T	5.5
A17	N150D	1.8
A18	A154T	4.2
A19	M1V, V164M	1.9
A20	K8R, K104E, I151T	1.4
A21	V52A, I151S, K178E	2.3

[a]Activities of mutants were measured as GST–AGT-His_6 fusion proteins in *E.coli* cell extracts.

Table III. AGT mutants selected using yeast three-hybrid system[a]

Clone	Mutations	k_{rel}
B1	K36M, Q97L, R175L	nd
B2	E92D, I151V, R175W	1.0
B3	L66M, K131R	1.5
B4	R175L	0.6
B5	D42E, P47L, V155L, K178M	1.9
B6	E110D, L120M	0.7
B7	T11I, N67K, Q72L	1.0
B8	G37W, L66V, V81M	nd
B9	M60K, A68V, K104M, A121T, E166K	nd
B10	V52I, V164M	2.1
B11	Q90R, K101N, F108I, V164L	1.2
B12	L33F, A68T	2.1
B13	D42E, P47L, V155L, K178M	1.9
B14	A41D, S115T, I151N	1.6
B15	T38M, A41D, A64T, G173C	0.8
B16	G122C	0.9
B17	D42E, A51T, A64V, K104M	0.5
B18	D42E, P47L, V155L, K178M	1.9
B19	F79I, V88I, F89L	0.7
B20	V44G, V106A, I151T, A170T	1.4
B21	E25K	0.7
B22	L33F, N123Y	2.2
B23	No mutations	nd
B24	No mutations	nd

[a]Activities of mutants were measured as GST–AGT fusion proteins after purification; nd: not determined; standard deviation of all values below 20%.

determined by the sequencing of 10 randomly picked clones. Plasmid DNA of the library was transformed into the *Saccharomyces cerevisiae* yeast strain L40, which also contained a plasmid encoding a B42–DHFR fusion protein. The transformed yeast cells were plated on selective plates containing 0.25 μM BGMtx and after 5 days of incubation at 30°C the plasmids from 24 of the faster growing colonies were isolated. The ^{M}AGT-based genes were subcloned into the vector pGEX-2T, allowing for the expression of a GST–AGT fusion protein in *E.coli*, the DNA sequence was determined and the activities of purified proteins measured (Table III). The mutation L33F was found in two different clones which displayed activities ~2-fold higher than ^{M}AGT. The mutation L33F was also one of the mutations found in the most active clone isolated from the phage selections. Another mutation that was found in both selections was the mutation V164M, which appears to increase the activity of AGT versus BG by a factor of ~2. Furthermore, a clone with 1.9-fold higher activity and the mutations D42E, P47L, V155L and K178M was found three times and ^{M}AGT was recovered twice. The clone with the mutation A154T, which was the most active mutant in the phage display selections, was not isolated in the yeast three-hybrid experiments.

Saturation mutagenesis of residues 150–154 and 31–35

The selections based on random mutagenesis point to two regions of the protein outside the active site that affect the activity against BG derivatives: the loop around Thr154 and the residues around Leu33. Point mutations accessible through error-prone PCR are generally limited to those arising from one or two base-per-codon mutations (Kuchner and Arnold, 1997). Based on the results from the previous selections and this reasoning we decided to create two new libraries through saturation mutagenesis of residues 31–35 and 150–154. The corresponding regions of the gene were randomized using standard procedures, yielding libraries of 1.2×10^7 and 3.8×10^6 independent clones, respectively. The library with randomized residues 31–35 was subjected to four rounds of phage panning at a substrate concentration of 50 nM BGDG. Twenty-two clones were screened for activity in cell extracts and the four most active clones were analyzed through DNA sequencing (Table IV). Except at residue 31, where hydrophobic side chains were conserved, no obvious consensus at any of the other positions was found, indicating that these residues are neither critical for folding nor activity of the protein. The two most active clones C1 and C2 displayed a 4.3- and 3.5-fold increase in activity, respectively. However, expression of clone C1 in *E.coli* yielded significantly lower amounts of soluble protein compared to C2 or ^{M}AGT. Concerning the library with the randomized residues 150–154, four rounds of selections were performed at a concentration of 5 nM BGFL. Forty-four clones were screened for activity in cell extracts. Of these 44, 7 clones displayed significantly higher activity than ^{M}AGT and were analyzed through DNA sequencing and their activity

Table IV. AGT mutants selected after saturation mutagenesis of residues 31–35 and 150–154[a]

Clone	Position					k_{rel}	Clone	Position					k_{rel}
	31	32	33	34	35			150	151	152	153	154	
^{M}AGT	I	K	L	L	G	1	^{M}AGT	N	I	N	G	A	1
C1	L	G	S	S	R	4.3	D1	Q	G	D	L	D	7.3
C2[b]	I	I	F	L	G	3.5	D2	E	G	E	G	R	6.8
C3	A	Y	M	T	F	2.1	D3	E	G	R	E	R	4.1
C4	L	S	P	M	G	1.9	D4	V	S	R	D	D	7.8
							D5	V	N	D	L	D	5.8
							D6	E	G	W	N	E	4.4
							D7	E	S	H	G	K	5.7

[a]Activities of mutants C1–C4 were measured as GST–AGT-His_6 fusion proteins in *E.coli* cell extracts and of mutants D1–D7 as GST–AGT-His_6 fusion proteins after purification.
[b]C2 also possesses the point mutation G182V. Standard deviation of all values below 13%.

measured after purification of the protein as GST fusion (Table IV). Clone D4 showed an 8-fold increased reactivity relative to ^{M}AGT and thus was the fastest mutant isolated up to this point in these selections. We then combined the mutations of clone C2 resulting from saturation mutagenesis of residues 31–35 (Table IV) and D4 in a single clone and measured the activity of the purified mutant. The resulting mutant, abbreviated in the following as ^{N}AGT, possesses a 17-fold higher activity than ^{M}AGT and a 52-fold higher activity than wild-type AGT, and is the most active AGT mutant against BG derivatives described so far (Table V).

Characterization of the selected mutant ^{N}AGT and its application for protein labeling

One of the advantages of AGT fusion proteins for covalent labeling is that a single fusion protein can be derivatized with a large variety of different labels. This is possible as previously used AGT mutants did not discriminate between BG derivatives carrying different probes attached to the benzyl group. To verify that the mutations K32I and L33F or any of the other mutations in ^{N}AGT did not result in discrimination between different substrates, we performed a competition assay between BGCy3 and BGFL. In these experiments, the reaction of ^{N}AGT with BGCy3 (1 μM) was almost completely quenched through the addition of BGFL (1 μM) (Fig. 2). Furthermore, ^{N}AGT can be efficiently biotinylated using BGBT (data not shown). These experiments demonstrate that ^{N}AGT can react with BG derivatives with different probes and linkers.

Another important property is the affinity of the engineered AGT towards (alkylated) DNA. The activity of wild-type AGT towards alkylated guanine incorporated in DNA compared to the corresponding alkylated nucleobase is at least four orders of higher magnitude (Pegg, 2000). Consequently, the reaction of wild-type AGT with BGCy3 is efficiently quenched in the presence of oligonucleotides containing BG (Fig. 2) (Juillerat *et al*., 2005). In contrast, the reaction of ^{N}AGT is not significantly quenched in the presence of oligonucleotides containing BG (Fig. 2), indicating that ^{N}AGT does not possess significant affinity towards DNA.

To investigate the use of ^{N}AGT for applications as a tag in protein labeling we compared the efficiency of labeling of different AGT mutants expressed as cell surface proteins. To allow for a labeling of an AGT fusion displayed on the cell surface of a mammalian cell it must fold and remain active under the oxidizing conditions of the secretory pathway and the cell surface. We have previously expressed AGT as an N-terminal fusion with the G protein-coupled NK_1 receptor, leading to the display of AGT (carrying the mutations N157G, S159E, K125A, A127T and R128A; abbreviated as ^{K}AGT) on the extracellular side of the plasma membrane (Keppler *et al*., 2004b). While this protein could be specifically labeled with fluorescent BG derivatives, the signal proved to be relatively weak compared to the fluorescence signal observed with NK_1 constructs with EGFP (data not shown) or a labeling strategy based on expression of NK_1 as fusion with acyl carrier protein and subsequent chemical labeling (George *et al*., 2004). We decided to compare the efficiency of the labeling of this mutant ^{K}AGT on the surfaces of living cells with that of ^{N}AGT. Compared to ^{K}AGT, ^{N}AGT has an ~3-fold higher activity towards BG derivatives and two non-essential cysteines have been removed through site-directed

Table V. Kinetic properties of AGT mutants[a]

Mutant	k_{obs} (M^{-1} s^{-1})	k_{rel}
wt AGT (human)	100	1
wt AGT (mouse)	20	0.2
^{M}AGT	300	3
^{N}AGT	5200	52

[a]Activities of mutants were measured as GST–AGT-His_6 fusion proteins after purification. Standard deviation of all values below 20%.

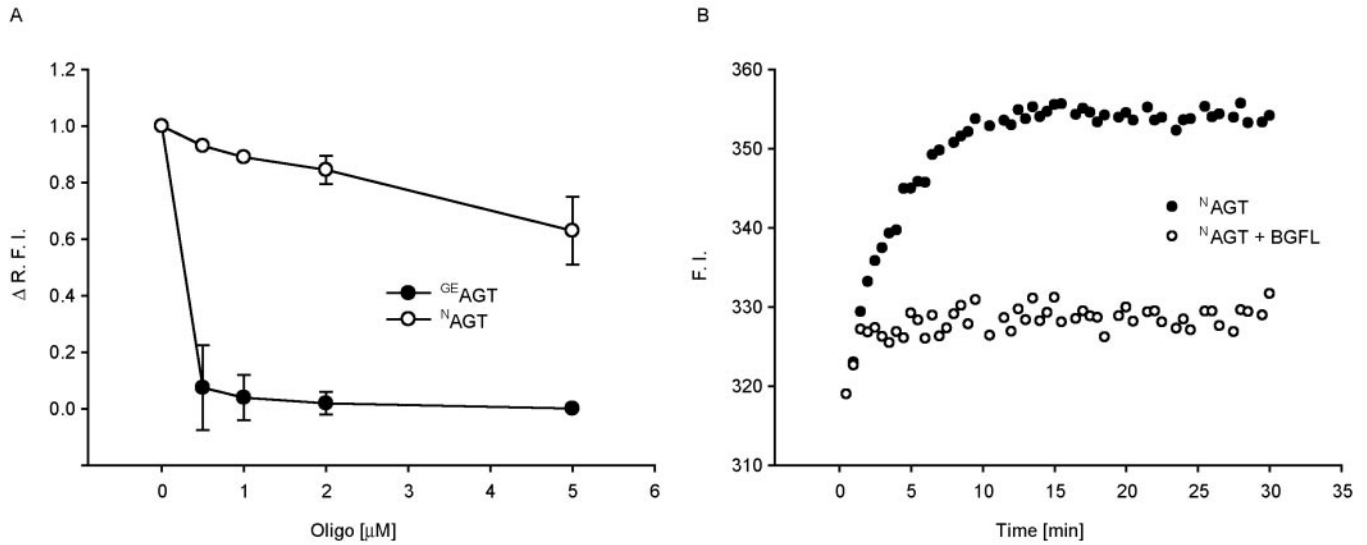

Fig. 2. Reaction of ^{N}AGT with BGCy3 in the presence of competing substrates. The assay is based on a 3-fold increase in the intensity of fluorescence upon reaction of ^{N}AGT with BGCy3 and the quenching of this fluorescence increase upon reaction of AGT with a competing substrate. (**A**) Reaction of ^{N}AGT or wild-type-like AGT containing the mutations N157G, S159E with BGCy3 (1 μM) in the presence of varying concentrations of an oligonucleotide (22mer) containing BG. Reactions were analyzed by measuring differences in relative fluorescence intensity at t_∞ and t_0 (ΔRFI) and plotting these values as a function of concentration of oligonucleotide containing BG. Values measured in the absence of oligonucleotide were arbitrarily set to 1. (**B**) Reaction of ^{N}AGT with BGCy3 (1 μM) in presence or absence of BGFL (1 μM). Reactions were analyzed by following the increase in fluorescence intensity (FI) as a function of time.

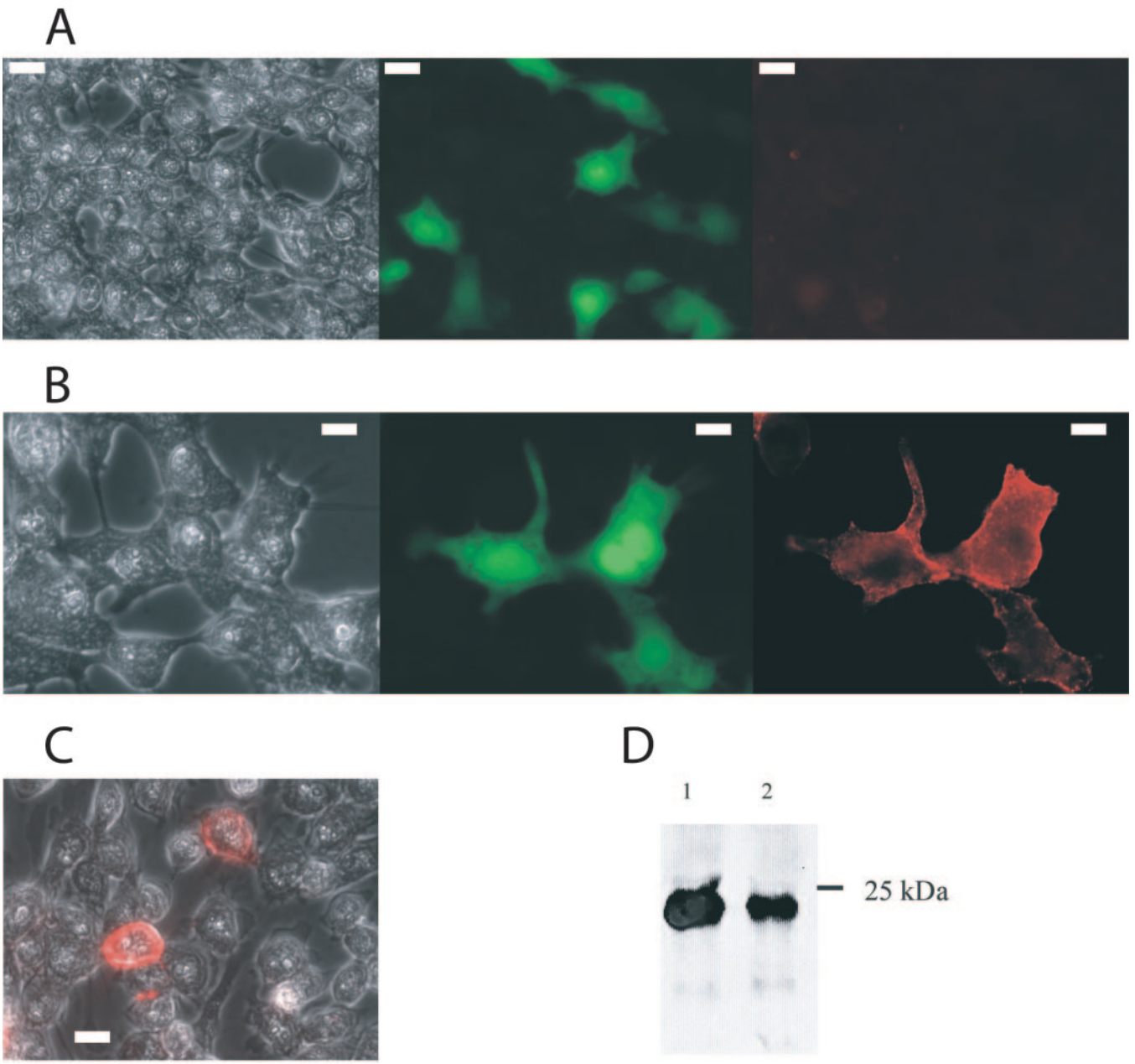

Fig. 3. Cell surface labeling of AGT mutants expressed as GPI fusion proteins. (**A**) Labeling of HEK293T cells transiently expressing ^{K}AGT–GPI and EGFP. Images from left to right: phase contrast; EGFP (green), Cy3 (red). (**B**) Labeling of HEK293 cells transiently expressing ^{N}AGT–GPI and EGFP. Images from left to right: phase contrast; EGFP (green), Cy3 (red). (**C**) Labeling of CHO cells transiently expressing ^{N}AGT–GPI. Image shown is overlay of phase contrast and fluorescence (Cy3, red). (**D**) Western blot of HEK293T cell extracts transiently expressing ^{K}AGT–GPI (lane 1) or ^{N}AGT–GPI (lane 2) probed with an anti-HA antibody.

mutagenesis. To investigate if ^{N}AGT allows for a more efficient labeling of AGT fusion proteins on cell surfaces, we transiently expressed both ^{N}AGT and ^{K}AGT with a C-terminal signal sequence that leads to display of the proteins on the cell surface via a GPI anchor. HEK293T cells were co-transfected with mixtures of plasmids leading to co-expression of EGFP and either ^{K}AGT–GPI or ^{N}AGT–GPI. The expression of the GPI fusions was verified by western blotting and both proteins were expressed to about the same level (Fig. 3D). When cells expressing ^{K}AGT–GPI were incubated with BGCy3, no significant labeling of cell surface proteins was observed, indicating that most of the protein is not active or properly translocated (Fig. 3A). However, when these experiments were repeated with cells expressing ^{N}AGT–GPI, a strong fluorescence labeling of cell surface proteins upon addition of BGCy3 was observed (Fig. 3B). Similar results were also observed when these experiments were repeated with CHO cells as expression host (Fig. 3C).

These experiments clearly demonstrate the value of ^{N}AGT over previously described AGT mutants for labeling applications on the cell surfaces of live cells.

Discussion

The value of AGT as a tag in protein labeling depends on a number of key properties. These include (i) the speed and selectivity of the labeling, (ii) its size, (iii) expression yields in different hosts, (iv) absence of AGT-dependent localization,

(v) absence of activity against any other cellular substrates and (vi) the availability of a wide range of different substrates for protein labeling in living cells and *in vitro*. Using a combination of rational design and directed evolution we have previously generated AGT mutants that were optimized with respect to activity, substrate specificity, size, absence of DNA binding and expression properties (Juillerat *et al*., 2003, 2005). However, when the corresponding mutations were assembled in a single protein, the activity of this AGT mutant, ^{M}AGT, dropped by a factor of ~8 while retaining the other beneficial properties of its predecessors (Juillerat *et al*., 2005). To restore the activity of ^{M}AGT we decided to use directed evolution based on random mutagenesis. Previous directed evolution experiments were based on saturation mutagenesis of mainly active site residues. Random mutagenesis might allow the identification of other regions of the protein, more distant from the active site, that affect activity. Furthermore, the earlier directed evolution experiments were exclusively based on phage display of AGT and the labeling and subsequent isolation of phage displaying active AGT. Phage selections in general are biased towards proteins that can be secreted to and folded in the periplasm of *E.coli*, properties not necessarily important for AGT fusion proteins. To complement the selections based on phage display we established a yeast three-hybrid system in which the alkylation of an AGT fusion protein leads to formation of a functional transcription factor (Fig. 1A). When error-prone libraries based on ^{M}AGT were subjected to selections using either phage display or the yeast three-hybrid system, a number of mutants with significantly increased activity were isolated. These two independent selections pointed towards two regions of AGT that significantly affect its activity towards BG derivatives. Firstly, mutations of Thr154 or amino acids located in the loop around Thr154 (i.e. N150, I151) appear to affect the activity of ^{M}AGT (Fig. 4). The mutations C150N, S151I and S152N were incorporated into ^{M}AGT to remove the non-essential cysteine C150 and to suppress the interaction of AGT with DNA as S151 and S152 contribute to the protein–DNA interaction (Daniels *et al*., 2004). The present study results of the performed directed evolution experiments thus suggest that at least one of the mutations is responsible for the decreased activity of ^{M}AGT towards BG derivatives. It should also be noted that the region 150–154 and in particular residues 150, 152 and 154 have been previously identified as hot spots for resistance against BG (Encell *et al*., 1998; Davis *et al*., 2001), indicating that this loop not only plays a role in the interaction of AGT with DNA but that its conformation also modulates the activity of the protein versus BG. Secondly, mutations at position L33 also appear to affect the activity of ^{M}AGT versus BG derivatives. When BG is docked into the active site of AGT, $C_{\delta 1}$ of L33 is within 4.5 Å of the C-4 of the benzyl ring of

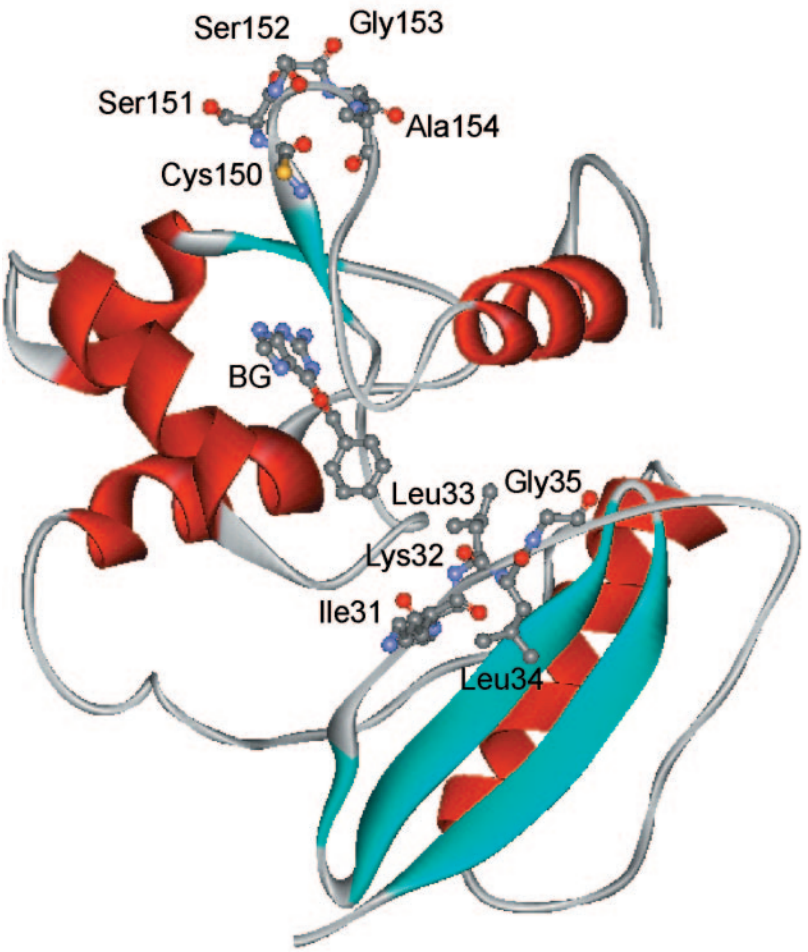

Fig. 4. Structure of wild-type AGT with BG docked into the active site. Highlighted are residues 31–35 and 150–154.

BG. Steric repulsion between the linker attached to the C-4 of BG and L33 might be attenuated through mutations at this position or favorable interactions introduced.

It is instructive to compare the efficiency of phage display and the yeast three-hybrid system for the selection of AGT mutants with increased activity. The selections based on phage display required six rounds of panning to identify the best mutant out of a library of randomly mutated AGT; analysis of clones after four rounds did not identify any mutants with significantly increased activity (data not shown). In contrast, after a single round of growth selection with the yeast three-hybrid system 40% of the analyzed mutants possessed an increased activity. Mutants isolated after six rounds of phage selections possessed in average slightly higher activity than those isolated after one round of growth selections using the yeast three-hybrid system, but these differences should be interpreted with caution, as only 24 mutants have been analyzed after the selections based on the yeast three-hybrid system. The higher activity of the mutants isolated from the phage selections might also be due to the 100-fold larger size of the library used in phage display compared to that used in the yeast three-hybrid system. In general, the transformation efficiency of *E.coli* is larger than that of yeast. Another potential advantage of phage display is the possibility to perform the selections *in vitro*. This yields access to a much larger range of reaction conditions, such as variations in pH and temperature, and should also allow covering a larger dynamic range of the selections with respect to rate enhancements. Nevertheless, the here established yeast three-hybrid system is a valuable alternative to phage display for the directed evolution of AGT and another example of a growing list of yeast three-hybrid systems successfully used for the directed evolution of enzymatic activity (Firestine *et al*., 2000; Baker *et al*., 2002; Lin *et al*., 2004; Carter *et al*., 2005).

Based on the results from these two selections, we decided to subject the regions of residues 31–35 and 150–154 to saturation mutagenesis. Libraries based on error-prone PCR often allow the identification of so-called hot spots of activity but not necessarily the best mutation at these hot spots (Kuchner and Arnold, 1997). We decided to subject the resulting libraries to selections using phage display and not the yeast three-hybrid system mainly for two practical reasons. Firstly, transformation of libraries into an expression host is generally more efficient with *E.coli* than with yeast. Secondly, isolation of plasmids for subcloning is also easier from *E.coli* than from yeast. These selections then yielded mutants with activities increased up to 8-fold (saturation mutagenesis of residues 150–154) and up to 4-fold (saturation mutagenesis of residues 31–35). When recombining mutants from both selections, a mutant with a 17-fold increased activity relative to the starting mutant ^{M}AGT was identified, indicating that the effects of the different mutations on the activity are at least partially additive. In addition to its increased activity the mutant, which is referred to as ^{N}AGT, reacts with structurally diverse BG derivatives, displays no significant affinity towards DNA and retains its activity even under a variety of different conditions, including the oxidizing conditions of the secretory pathway and the cell surface. Previously used AGT mutants often showed diminished activities under such conditions. One important application we foresee for this mutant should be in labeling experiments in living animals such as mice. Here, the 260-fold higher activity of ^{N}AGT compared to mouse wild-type AGT (Table V), as well as the possibility to express ^{N}AGT functionally on cell surfaces makes the protein an attractive candidate for *in vivo* applications.

In summary, we report here the generation of an AGT mutant with optimized conditions for the labeling of AGT fusion proteins in living cells and *in vitro* using directed evolution. The high activity of the mutant towards BG derivatives, its lack of affinity towards DNA and its stability under oxidizing conditions should make this protein a valuable tool for a variety of applications.

Acknowledgements

This work was supported by the EPFL, the Swiss National Science Foundation and the Bundesamt fuer Berufsbildung und Technologie (CTI Nr. 7042.1). We thank Prof. F. Wurm, Prof. H. Vogel, Prof. A. Plückthun and Covalys Biosciences for generous support and Nathalie George for technical assistance.

References

Baker,K., Bleczinski,C., Lin,H., Salazar-Jimenez,G., Sengupta,D., Krane,S. and Cornish,V.W. (2002) *Proc Natl Acad. Sci. USA*, **99**, 16537–16542.

Carter,B.T., Lin,H., Goldberg,S.D., Althoff,E.A., Raushel,J. and Cornish,V.W. (2005) *Chembiochem*, **6**, 2055–2067.

Chen,R., Knez,J.J., Merrick,W.C. and Medof,M.E. (2001) *J. Cell Biochem*., **84**, 68–83.

Daniels,D.S., Woo,T.T., Luu,K.X., Noll,D.M., Clarke,N.D., Pegg,A.E. and Tainer,J.A. (2004) *Nat. Struct. Mol. Biol*., **11**, 714–720.

Davis,B.M., Encell,L.P., Zielske,S.P., Christians,F.C., Liu,L., Friebert,S.E., Loeb,L.A. and Gerson,S.L. (2001) *Proc. Natl Acad. Sci. USA*, **98**, 4950–4954.

Encell,L.P., Coates,M.M. and Loeb,L.A. (1998) *Cancer Res*., **58**, 1013–1020.

Firestine,S.M., Salinas,F., Nixon,A.E., Baker,S.J. and Benkovic,S.J. (2000) *Nat. Biotechnol*, **18**, 544–547.

Gendreizig,S., Kindermann,M. and Johnsson,K (2003) *J. Am. Chem. Soc*., **125**, 14970–14971.

George,N., Pick,H. , Vogel,H. , Johnsson,N. and Johnsson,K. (2004) *J. Am. Chem. Soc*., **126**, 8896–8897.

Gietz,R.D. and Woods,R.A. (2002) *Methods Enzymol*., **350**, 87–96.

Hiscox,S., Hallett,M.B., Morgan,B.P. and van den Berg,C.W. (2002) *Biochem. Biophys. Res. Commun*., **293**, 714–721.

Johnsson,N. and Johnsson,K. (2003) *Chembiochem*, **4**, 803–810.

Juillerat,A., Gronemeyer,T., Keppler,A., Gendreizig,S., Pick,H., Vogel,H. and Johnsson,K. (2003) *Chem. Biol*., **10**, 313–317.

Juillerat,A., Heinis,C., Sielaff,I., Barnikow,J., Jaccard,H., Kunz,B., Terskikh,A. and Johnsson,K. (2005) *Chembiochem*, **6**, 1263–1269.

Keppler,A., Gendreizig,S., Gronemeyer,T., Pick,H., Vogel,H. and Johnsson,K. (2003) *Nat. Biotechnol*., **21**, 86–89.

Keppler,A., Kindermann,M., Gendreizig,S., Pick,H., Vogel,H. and Johnsson,K. (2004a) *Methods*, **32**, 437–444.

Keppler,A., Pick,H., Arrivoli,C., Vogel,H. and Johnsson,K. (2004b) *Proc. Natl Acad. Sci. USA*, **101**, 9955–9959.

Krebber,A., Bornhauser,S., Burmester,J., Honegger,A., Willuda,J., Bosshard,H.R. and Pluckthun,A. (1997) *J. Immunol. Methods*, **201**, 35–55.

Kuchner,O. and Arnold,F.H. (1997) *Trends Biotechnol*., **15**, 523–530.

Lin,H., Tao,H. and Cornish,V.W. (2004) *J. Am. Chem. Soc*., **126**, 15051–15059.

Pegg,A.E. (2000) *Mutat Res*., **462**, 83–100.

Zaccolo,M. and Gherardi,E. (1999) *J. Mol. Biol*., **285**, 775–783.

Received December 13, 2005; revised February 13, 2006; accepted March 14, 2006

Edited by Ashutosh Chilkoti