3. Conclusions

Early inhibitors of the 20S proteasome have been invaluable tools for improving our understanding of the proteasome degradation pathway *in vitro*. However, they are relatively non-specific compounds and exert their inhibitory activity by forming a covalent bond with the enzyme. We have identified a new class of non-covalent inhibitors able to modulate the chymotrypsin-like activity of the 20S proteasome in ways not previously possible. The increased potency and selectivity of the compounds described in this paper should allow us to investigate and eventually confirm the potential therapeutic application of this approach.

[1] a) K. Tanaka, *J. Biochem.* **1998**, *123*, 195; b) K. Tanaka, M. Kasahara, *Immunol. Rev.* **1998**, *163*, 161.
[2] M. Orlowski, C. Cardozo, C. Michaud, *Biochemistry* **1993**, *32*, 1563.
[3] For a recent review on 20S proteasome inhibitors, see C. García-Echeverría, *Mini-Reviews in Medicinal Chemistry* **2002**, *2*, 247.
[4] A dipeptide boric acid analogue (PS-341, Velcade™) is currently in Phase III clinical trials for the treatment of multiple myeloma. For preclinical data on this compound, see: J. Adams, V.J. Palombella, E.A. Sausville, J. Johnson, A. Destree, D.D. Lazarus, J. Maas, C.S. Pien, S. Prakash, P.J. Elliott, *Cancer Res.* **1999**, *59*, 2615.
[5] Several classes of 20S proteasome inhibitors have been described: peptide aldehydes, peptide α-keto aldehydes (glyoxals), peptide α',β'-epoxyketones, peptide vinyl sulfones, peptide indanones, peptide boric acids (*e.g.* PS-341, see [4]), bifunctional inhibitors or natural products (*e.g.* lactacystin, epoxomicin, TMC-95A-D or UCK 14A2); see also [3].
[6] C. García-Echeverría, P. Imbach, D. France, P. Fuerst, M. Lang, M. Noorani, D. Scholz, J. Zimmermann, P. Furet, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1317.
[7] Other inhibitors of the HIV-1 protease have also been described as inhibitors of the 20S proteasome: a) P. Andre, V. Lotteau, R.M. Zinkernagel, P. Klenerman, M. Groettrup, WO 9963998, **1999**; b) P. Andre, M. Groettrup, P. Klenerman, R. de Giuli, B.L. Booth, Jr., V. Cerundolo, M. Bonneville, F. Jotereau, R.M. Zinkernagel, V. Lotteau, *Proc. Nat. Acad. Sci. U.S.A.* **1998**, *95*, 13120.
[8] M. Groll, L. Ditzel, J. Löwe, D. Stock, M. Bochtler, H.D. Bartunik, R. Huber, *Nature* **1997**, *386*, 463.
[9] P. Furet, P. Imbach, P. Fuerst, M. Lang, M. Noorani, J. Zimmermann, C. García-Echeverría, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1321.
[10] P. Furet, P. Imbach, P. Fuerst, M. Lang, M. Noorani, J. Zimmermann, C. García-Echeverría, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1331.

Covalent Labeling of Fusion Proteins with Chemical Probes in Living Cells

Susanne Gendreizig[a,b]*, Antje Keppler[a,b], Alexandre Juillerat[a,b], Thomas Gronemeyer[a,b], Horst Pick[b], Horst Vogel[b], and Kai Johnsson[a,b]

Abstract: A general method for the specific and covalent labeling of fusion proteins *in vivo* is described. The approach is based on the unusual mechanism of the human DNA repair protein O^6-alkylguanine-DNA alkyltransferase, which irreversibly transfers an alkyl group from O^6-alkylguanine-DNA to one of its cysteine residues. Labeling is possible in bacterial as well as eukaryotic cells and is independent of the nature of the label, thereby opening up new ways to study proteins *in vivo*.

Keywords: O^6-Alkylguanine-DNA alkyltransferase · Fluorescent probes · Protein labeling

Correspondence: S. Gendreizig[a,b]
Tel.: +41 21 693 94 42
Fax: +41 21 693 93 65
E-Mail: susanne.gendreizig@epfl.ch
[a]Swiss Federal Institute of Technology Lausanne (EPFL)
Institute of Molecular and Biological Chemistry
CH–1015 Lausanne
[b]Swiss Federal Institute of Technology Lausanne (EPFL)
Institute of Biomolecular Sciences
CH–1015 Lausanne

Understanding the function of proteins on a molecular level often requires characterizing their location and interaction partners within the living cell. Despite the importance of this task, there are few methods that allow proteins to be studied *in vivo*. Most existing approaches rely on genetically fusing the protein of interest to a reporter protein that gives the observer a direct or indirect measure of the location and microenvironment of the tagged protein. The most prominent examples for such reporter proteins are fluorescent proteins [1]. However, all currently existing approaches possess various disadvantages and we have recently introduced a general method for the covalent labeling of fusion proteins *in vivo* that complements the existing strate-

3. Conclusions

Early inhibitors of the 20S proteasome have been invaluable tools for improving our understanding of the proteasome degradation pathway *in vitro*. However, they are relatively non-specific compounds and exert their inhibitory activity by forming a covalent bond with the enzyme. We have identified a new class of non-covalent inhibitors able to modulate the chymotrypsin-like activity of the 20S proteasome in ways not previously possible. The increased potency and selectivity of the compounds described in this paper should allow us to investigate and eventually confirm the potential therapeutic application of this approach.

[1] a) K. Tanaka, *J. Biochem.* **1998**, *123*, 195; b) K. Tanaka, M. Kasahara, *Immunol. Rev.* **1998**, *163*, 161.
[2] M. Orlowski, C. Cardozo, C. Michaud, *Biochemistry* **1993**, *32*, 1563.
[3] For a recent review on 20S proteasome inhibitors, see C. García-Echeverría, *Mini-Reviews in Medicinal Chemistry* **2002**, *2*, 247.
[4] A dipeptide boric acid analogue (PS-341, Velcade™) is currently in Phase III clinical trials for the treatment of multiple myeloma. For preclinical data on this compound, see: J. Adams, V.J. Palombella, E.A. Sausville, J. Johnson, A. Destree, D.D. Lazarus, J. Maas, C.S. Pien, S. Prakash, P.J. Elliott, *Cancer Res.* **1999**, *59*, 2615.
[5] Several classes of 20S proteasome inhibitors have been described: peptide aldehydes, peptide α-keto aldehydes (glyoxals), peptide α',β'-epoxyketones, peptide vinyl sulfones, peptide indanones, peptide boric acids (*e.g.* PS-341, see [4]), bifunctional inhibitors or natural products (*e.g.* lactacystin, epoxomicin, TMC-95A-D or UCK 14A2); see also [3].
[6] C. García-Echeverría, P. Imbach, D. France, P. Fuerst, M. Lang, M. Noorani, D. Scholz, J. Zimmermann, P. Furet, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1317.
[7] Other inhibitors of the HIV-1 protease have also been described as inhibitors of the 20S proteasome: a) P. Andre, V. Lotteau, R.M. Zinkernagel, P. Klenerman, M. Groettrup, WO 9963998, **1999**; b) P. Andre, M. Groettrup, P. Klenerman, R. de Giuli, B.L. Booth, Jr., V. Cerundolo, M. Bonneville, F. Jotereau, R.M. Zinkernagel, V. Lotteau, *Proc. Nat. Acad. Sci. U.S.A.* **1998**, *95*, 13120.
[8] M. Groll, L. Ditzel, J. Löwe, D. Stock, M. Bochtler, H.D. Bartunik, R. Huber, *Nature* **1997**, *386*, 463.
[9] P. Furet, P. Imbach, P. Fuerst, M. Lang, M. Noorani, J. Zimmermann, C. García-Echeverría, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1321.
[10] P. Furet, P. Imbach, P. Fuerst, M. Lang, M. Noorani, J. Zimmermann, C. García-Echeverría, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1331.

Covalent Labeling of Fusion Proteins with Chemical Probes in Living Cells

Susanne Gendreizig[a,b]*, Antje Keppler[a,b], Alexandre Juillerat[a,b], Thomas Gronemeyer[a,b], Horst Pick[b], Horst Vogel[b], and Kai Johnsson[a,b]

Abstract: A general method for the specific and covalent labeling of fusion proteins *in vivo* is described. The approach is based on the unusual mechanism of the human DNA repair protein O[6]-alkylguanine-DNA alkyltransferase, which irreversibly transfers an alkyl group from O[6]-alkylguanine-DNA to one of its cysteine residues. Labeling is possible in bacterial as well as eukaryotic cells and is independent of the nature of the label, thereby opening up new ways to study proteins *in vivo*.

Keywords: O[6]-Alkylguanine-DNA alkyltransferase · Fluorescent probes · Protein labeling

Correspondence: S. Gendreizig[a,b]
Tel.: +41 21 693 94 42
Fax: +41 21 693 93 65
E-Mail: susanne.gendreizig@epfl.ch
[a]Swiss Federal Institute of Technology Lausanne (EPFL)
Institute of Molecular and Biological Chemistry
CH–1015 Lausanne
[b]Swiss Federal Institute of Technology Lausanne (EPFL)
Institute of Biomolecular Sciences
CH–1015 Lausanne

Understanding the function of proteins on a molecular level often requires characterizing their location and interaction partners within the living cell. Despite the importance of this task, there are few methods that allow proteins to be studied *in vivo*. Most existing approaches rely on genetically fusing the protein of interest to a reporter protein that gives the observer a direct or indirect measure of the location and microenvironment of the tagged protein. The most prominent examples for such reporter proteins are fluorescent proteins [1]. However, all currently existing approaches possess various disadvantages and we have recently introduced a general method for the covalent labeling of fusion proteins *in vivo* that complements the existing strate-

gies and should open up new ways of studying and manipulating proteins in living cells [2]. The covalent attachment of a small reporter molecule to a given protein is achieved by genetically fusing the protein of interest to the human DNA repair protein O[6]-alkylguanine-DNA alkyltransferase (hAGT). During DNA repair, hAGT irreversibly transfers the alkyl group from its substrate, O[6]-alkylguanine-DNA, to one of its cysteine residues (Fig. 1A) [3]. The substrate specificity of hAGT is relatively low, as it reacts also readily with O[6]-benzylguanine (BG), either as the nucleobase or when incorporated in oligonucleotides [4]. Based on these observations we reasoned that hAGT fusion proteins could be labeled by cell-permeable BG derivatives of the type **1** both *in vivo* and *in vitro* (Fig. 1B). To test the feasibility of the approach we first investigated the labeling of hAGT with different reporter groups in bacterial and yeast cells. In these experiments we demonstrated the specific labeling of hAGT fusion proteins with the affinity tags biotin and digoxigenin by simply adding either BG derivatives BGBT or BGDG to bacterial or

yeast cell cultures expressing the corresponding fusion protein (Fig. 1C) [2]. It is important to note that BG derivatives of the type **1** can be readily synthesized starting from 6-chloroguanine (Fig. 1D).

Clearly, the most interesting applications for such a specific protein labeling can be foreseen for mammalian cells. As a first example for the extension of our approach to mammalian cells, we investigated the labeling of a hAGT fusion protein with fluorescein in Chinese hamster ovarian (CHO) cells using BGFL (Fig. 1C). To facilitate the evaluation of the specificity of the fluorescence labeling, a hAGT fusion protein was constructed which was targeted to the nucleus of the CHO cell by fusing hAGT to three consecutive nuclear localization sequences (NLS), yielding hAGT-NLS$_3$ [5]. Furthermore, we used the non-fluorescent diacetylated derivative BGAF as labeling reagent, as it possesses improved cell permeability compared to fluorescent BGFL. Inside the cell BGAF is rapidly hydrolyzed by esterases, yielding BGFL (Fig. 1C) [6]. CHO cells transiently expressing hAGT-NLS$_3$ were incubated with BGAF (5 μM)

for 5 min and then washed to remove excess substrate. The labeling was monitored with confocal fluorescence microscopy (Fig. 2A–C). During and after the washing procedure unbound fluorophores leaked out of the cell and after 25 min only the nucleus of the cell displayed a strong fluorescence signal (Fig. 2A–C). Control experiments in CHO cells not expressing the hAGT fusion protein showed no fluorescence labeling of the nucleus (data not shown) [2]. The data thus demonstrate a covalent and specific labeling of the hAGT fusion protein with fluorescein inside the mammalian cell.

The DNA repair protein hAGT possesses a variety of properties that make it a good choice for the specific and covalent labeling of fusion proteins *in vivo*. Most importantly, it has high reactivity on a substrate which is otherwise chemically inert and which can be derivatized by a wide variety of labels such as dyes, crosslinkers, or affinity tags. None of the tested substrates so far showed chemical toxicity during incubation with bacterial or eukaryotic cells and due to the irreversibility of the labeling, excess substrate can be easily washed away.
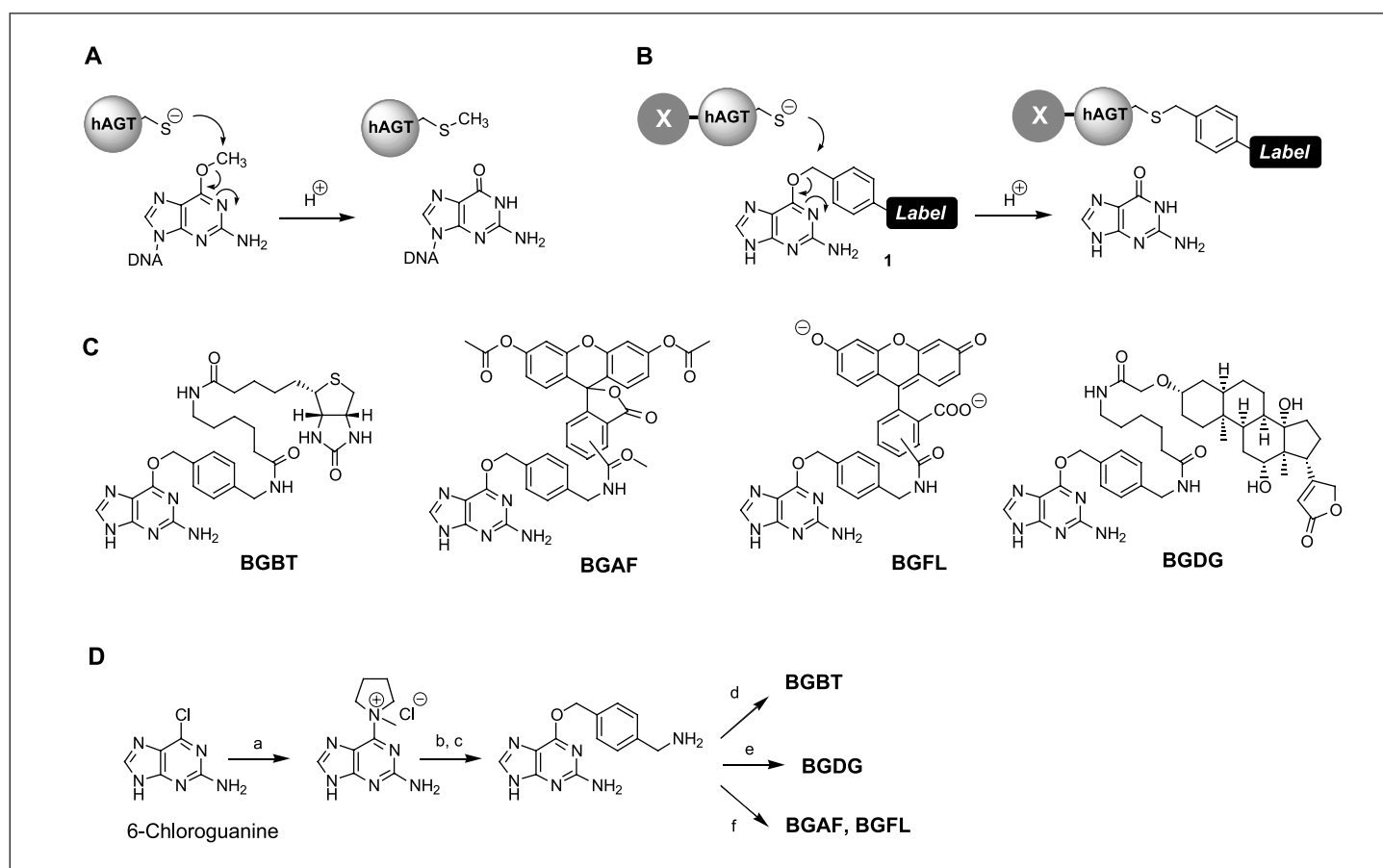

Fig. 1. (A) Mechanism of DNA repair by hAGT. (B) Covalent labeling of an X-hAGT fusion protein using O[6]-benzylguanine (BG) derivatives of the type **1**. (C) Structures of the BG derivatives BGBT, BGAF, BGFL and BGDG. (D) Scheme of synthesis of BGBT, BGAF, BGFL and BGDG; a) 1-methylpyrrolidin, DMF, 66%; b) 2,2,2-trifluoro-N-(4-hydroxymethyl-benzyl)-acetamide, potassium tert-butoxide, DMF, 88%; c) K$_2$CO$_3$, methanol, 85%; d) N-(+)-biotinyl-6-aminocaproic acid N-hydroxy-succinimidyl ester, triethylamine, DMF, 69%; e) Digoxigenin-3-O-methylcarbonyl-6-aminocaproic acid N-hydroxy-succinimidyl ester, DMF, triethylamine, 87%; f) 5(6)-carboxyfluorescein diacetate N-hydroxy-succinimidyl ester (mixture of isomers), triethylamine, DMF, 8% (BGFL), 2% (BGAF).
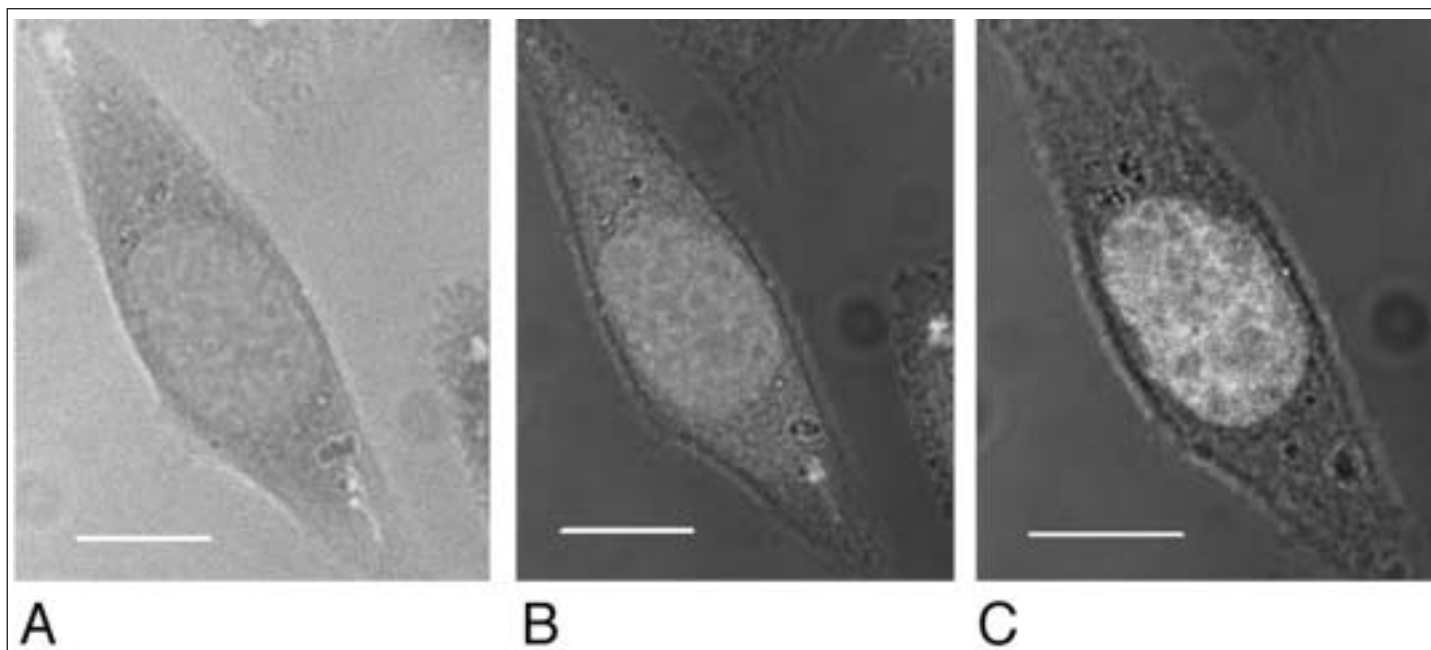
Fig. 2. Covalent labeling of nuclear targeted ^{W160}hAGT-NLS$_3$ (NLS$_3$: three consecutive nuclear localization sequences) in a cell line deficient of endogenous AGT [7]. The hAGT mutant G160W (^{W160}hAGT) is used because it has been reported to possess an increased activity against BG [2]. In these experiments, confocal micrographs A–C show overlays of transmission and fluorescence channels (exc 488 nm). Size bar in A–C corresponds to 10 µm. Confocal micrographs (A–C) illustrate the time course of the labeling of transiently expressed ^{W160}hAGT-NLS$_3$ with BGAF in AGT-deficient CHO cells. (A) CHO cell transiently expressing ^{W160}hAGT-NLS$_3$ during incubation with BGAF (5 µM). (B) Same cell as in (A) after 5 min of incubation with BGAF (5 µM) and 3 washes with PBS. (C) Same cell as in (B) after additional 25 min of incubation in PBS, illustrating the specific labeling of hAGT-NLS$_3$ in the nucleus of a transiently transfected CHO cell.

Furthermore, hAGT is a monomer of 207 residues, thereby reducing the likelihood that its fusion to other proteins will influence their oligomeric state. The protein of interest can be fused either to the N or the C terminus of hAGT without affecting its reactivity. Experiments in mammalian cells should be performed in AGT-deficient cell lines, preventing the labeling of endogenous AGT [7]. To circumvent the necessity to use AGT-deficient cell lines, the generation of hAGT mutants with substrate specificities orthogonal to wild-type hAGT or of artificial alkyltransferases is a possible solution [8][9]. Using directed evolution, we have already generated hAGT mutants with 20-fold increased activity against BG derivatives of the type **1** over wild type [8].

In conclusion, we have developed a general method for the covalent and specific labeling of fusion proteins *in vivo*. The method can be applied in different organisms and allows the attachment of a wide variety of labels, making it a valuable tool for functional studies of proteins in living cells.

[1] R.Y. Tsien, *Annu. Rev. Biochem.* **1998**, *67*, 509.

[2] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.*, **2003**, *21*, 86.

[3] A.E. Pegg, *Mutat. Res.* **2000**, *462*, 83.

[4] R. Damoiseaux, A. Keppler, K. Johnsson, *ChemBioChem* **2001**, *2*, 285.

[5] D. Kalderon, B.L. Roberts, W.D. Richardson, A.E. Smith, *Cell* **1984**, *39*, 499.

[6] M.M. Wu, J. Llopis, S. Adams, J.M. McCaffery, M.S. Kulomaa, T.E. Machen, H.P. Moore, R.Y. Tsien, *Chem. Biol.* **2000**, *7*, 197.

[7] B. Kaina, G. Fritz, S. Mitra, T. Coquerelle, *Carcinogenesis* **1991**, *12*, 1857.

[8] A. Juillerat, T. Gronemeyer, A. Keppler, H. Pick. H. Vogel, K. Johnsson, *Chem. Biol.*, in press.

[9] R. Damoiseaux, P.G. Schultz, K. Johnsson, *ChemBioChem* **2002**, *3*, 573