Bioorthogonal Reactions

DOI: 10.1002/anie.201400681

Protein-Templated Peptide Ligation**

*Nicolas Brauckhoff, Gernot Hahne, Johannes T.-H. Yeh, and Tom N. Grossmann**

Abstract: *Molecular templates bind particular reactants, thereby increasing their effective concentrations and accelerating the corresponding reaction. This concept has been successfully applied to a number of chemical problems with a strong focus on nucleic acid templated reactions. We present the first protein-templated reaction that allows N-terminal linkage of two peptides. In the presence of a protein template, ligation reactions were accelerated by more than three orders of magnitude. The templated reaction is highly selective and proved its robustness in a protein-labeling reaction that was performed in crude cell lysate.*

Controlling the selectivity and the rate of chemical reactions is a central hallmark of the natural sciences. In classic chemical synthesis, reactants of interest are used at high concentrations to ensure effective product formation. A fundamentally different approach is applied in biological systems that co-host a multitude of reactive compounds at relatively low concentrations. In this case, reactants are recognized by biomacromolecules that mediate the desired chemical transformation. Such macromolecules act as templates thereby increasing the effective concentration of reaction partners. Given a correct alignment of functional groups, the template mediates specific and efficient product formation at concentrations where the occurrence of non-templated reactions is inefficient.[1] The concept of template-induced reactions was applied to various challenging chemical problems. Examples of designed templated reactions involve the self-replication of nucleic acids[2] and peptides,[3] the generation of dynamic combinatorial[4] or DNA-encoded libraries,[5] and the detection of nucleic acids[6] as well as proteins.[7] The biocompatible modification of proteins is another appealing field for the use of templated reactions. So far, this field has been dominated by bioorthogonal reactions,[8] enzymatic transformations,[8a] and ligand-directed chemistry.[9] Nonetheless, some examples of template-mediated peptide modifications have been reported. They involve the de novo design of peptides and foldamers capable of templating peptide ligations,[10] and the nucleic acid templated ligation of peptide fragments.[11] However, these approaches have not yet proven their applicability in complex biological environments and have not been used for the labeling of entire proteins.

Our goal was the design of a template-mediated peptide ligation (Figure 1 A) allowing selective reactions within a complex biological matrix. Taking the exceptional surface-recognition properties of proteins into consideration, we decided to search for protein domains capable of binding two different peptide sequences simultaneously thereby serving as a template. The KIX domain (**T**, gray) of CBP (CREB binding protein)[12] was considered an appropriate candidate (Figure 1 B). This domain binds several short peptide sequences of corresponding interaction partners. Among those, we chose peptide **A** (blue) representing a fragment of CREB (cAMP response element-binding protein, aa 120–146),[13]

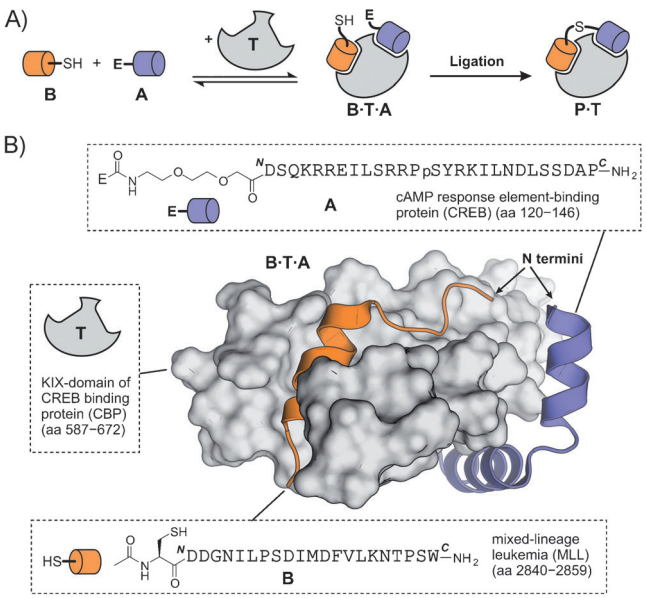

Figure 1. A) Concept of protein (**T**)-templated linkage of electrophile (**E**)-modified peptide **A** and thiol-bearing peptide **B** yielding product **P**. B) NMR structure (PDB 2LXT)[13] of protein **T** (KIX domain of CBP) bound to the two peptides **A** and **B** (KIX-binding domains of CREB and MLL, respectively) including their sequences. The structure shows amino acids used for the design of protein and peptide ligands (Table S1 and Figure S4). Sequence **B** bears an additional C-terminal tryptophan (W) to allow UV-spectroscopic quantification.

[*] N. Brauckhoff, Dr. G. Hahne, Dr. T. N. Grossmann
Chemical Genomics Centre of the Max Planck Society
Otto-Hahn-Strasse 15, 44227 Dortmund (Germany)
E-mail: tom.grossmann@cgc.mpg.de
Dr. J. T. H. Yeh
Department of Genetics, Harvard Medical School
77 Avenue Louis Pasteur, Boston MA 02115 (USA)
Dr. G. Hahne, Dr. T. N. Grossmann
Technical University Dortmund
Department of Chemistry and Chemical Biology
Otto-Hahn-Strasse 6, 44227 Dortmund (Germany)

[**] T.N.G. thanks the Fonds der Chemischen Industrie and the German Research Foundation (DFG, Emmy Noether program GR3592/2-1) for financial support. This work was supported by AstraZeneca, Bayer CropScience, Bayer HealthCare, Boehringer Ingelheim, Merck KGaA, and the Max Planck Society. We thank Sascha Gentz for support with automated solid-phase peptide synthesis and Dr. Christian Hedberg for useful discussions.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201400681.

and peptide **B** (orange) constituting the KIX-binding domain of MLL (mixed-lineage leukemia, aa 2840–2859)[13] as ligation partners. All of these proteins are involved in the transcriptional machinery of eukaryotic cells which circumvents interference with the prokaryotic expression system that we anticipated to use in this study. An NMR structure of the trimeric complex **B**·**T**·**A** has been reported (Figure 1B, PDB 2LXT)[13] and served as the starting point for our design process.

Initially, the affinity of protein domain **T** to fluorescein-labeled versions of peptides **A** and **B** was determined using fluorescence polarization (FP). The required protein was expressed heterologously from *Escherichia coli* and size-exclusion chromatography was applied as the last purification step (Figure S5). Peptides were synthesized on solid support following Fmoc-based protocols. Peptide identity and purity were verified by HPLC in combination with electrospray ionization mass spectrometry (ESI-MS, Table S1). The FP measurements revealed rapid complex formation (< 5 min) and dissociation constants around 0.1 μM (Figure S6), which we considered sufficient for a templated reaction.[14] To ensure a fast reaction on the template, the correct orientation of functional groups is crucial. An inspection of the NMR structure reveals the close proximity of the two peptide N termini (Figure 1B). Consequently, we decided to design a reaction that would facilitate an N-terminal linkage of the two peptides. Peptide **B** was equipped with an N-terminally acetylated cysteine introducing a thiol function. For peptide **A**, we considered the installation of a thiol-selective electrophilic group (**E**) attached by means of a polyethylene glycol (PEG) spacer (Figure 1B).

Various electrophiles were selected for initial tests (Figure 2A) involving a maleimide (**A1**) known to rapidly react with thiols at neutral pH, two different acrylamides (**A2** and **A3**), and a chloroacetamide (**A4**). Fluoroacetamide **A5** served as a negative control, as we did not expect it to react with thiols at neutral pH. The peptide sequences were assembled on solid support, coupled with an N-terminal PEG spacer, and then capped with the appropriate electrophilic group (Figure S1). At first, we investigated the general reactivity of these electrophile-modified peptides (**A1**–**A5**). High concentrations of the water-soluble tripeptide glutathione (**G-SH**, 10 mM) were used to facilitate nontemplated ligations in a buffered solution at pH 7.4 and 30 °C. Product formation was monitored for 24 h using HPLC as the readout, and initial rates (v) were determined (Figure S8). Chloroacetamide derivative **A4** showed relatively high reactivity and was used as reference for the calculation of relative initial rates ($v_{rel.}$, Figure 2B). As expected, the maleimide-modified peptide **A1** proved to be the most reactive derivative showing quantitative product formation after only 10 s ($v_{rel.} > 7 \times 10^2$). Compared to **A4**, the two acrylamide-modified peptides **A2** and **A3** showed 6.7- and 11-fold reduced reactivity, respectively. Negative control **A5** did not undergo reaction under these conditions ($v_{rel.} < 4 \times 10^{-4}$).

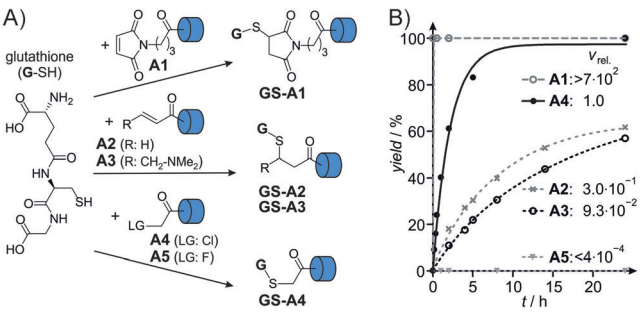

Figure 2. A) Nontemplated reactions of glutathione (**G-SH**) with electrophile-modified peptides **A1**–**A5**. B) Time courses of the reactions determined by HPLC including relative initial rates ($v_{rel.}$). Relative initial rates were calculated using the initial rate of **A4** as reference (Figure S8). Peptides **A1**–**A5** (150 μM) were reacted with glutathione (**G-SH**, 10 mM) in buffer (20 mM Na_2PO_4, 2 mM tris(2-carboxyethyl)phosphine, pH 7.4) at 30 °C.

Having observed rapid complex formation in the FP assay, we expected the peptide ligation to be rate limiting in the templated reaction. Consequently, a fast reaction would be desired to ensure efficient templated ligation. As maleimides have been reported to react in nontemplated reactions even at low concentration,[15] chloroacetamide **A4**, which proved to be the second most reactive electrophile in our panel, was the preferred choice for the following templated reactions. The ligation of peptide **A4** and cysteine-bearing peptide **B** (each 20 μM) at 30 °C and pH 7.4 was monitored using HPLC. After 3 h in the absence of template **T**, we did not observe product formation (Figure 3A, top). However, the disappearance of **A4** and **B** and the formation of a new peak were observed after the same time in the presence of **T** (Figure 3A, bottom). The new peak was assigned to ligation product **P4** using MALDI-TOF mass spectrometry (Figure S12). To obtain more detailed information, the kinetics of templated and nontemplated reactions were recorded at three different pH values (7.4, 8.4, and 9.4; Figure 3B). In accordance with the increased nucleophilicity of the thiol group, reactions were accelerated with increasing pH (v/pM s^{-1} × 10^3: 3.4 < 14 < 30). Notably, we were not able to detect product formation at any pH in the absence of **T** after 3 h (Figure 3B).

In order to determine the initial rates, we extended the reaction times for the nontemplated reactions to 24 h. At pH 8.4 and 9.4 low product yields (13% and 19%, respectively) were observed, allowing the determination of the initial rates (v/pM s^{-1}: 30 and 44, respectively). At pH 7.4 even after 24 h, we were unable to detect product formation. Taking the HPLC detection limit into consideration, we estimated a maximum initial rate of 14 pM s^{-1} (Figure S11). Figure 3C summarizes all v values for the ligation reactions of peptides **A4** and **B** and it shows the resulting rate acceleration (a) which refers to the ratio of the ligation rates of the templated versus nontemplated reaction. Under these conditions, the templated reactions appear to be accelerated by a factor of 250 to 670 (depending on the pH) relative to the corresponding nontemplated reaction (Figure 3C).

The fact that increasing pH values result in higher templated reaction rates (Figure 3C) indicates that the ligation reaction, and not the complex formation, is rate limiting. Therefore, we felt confident to use the previously presented highly reactive maleimide-modified peptide **A1** in protein-templated reactions. Knowing that maleimides rap-

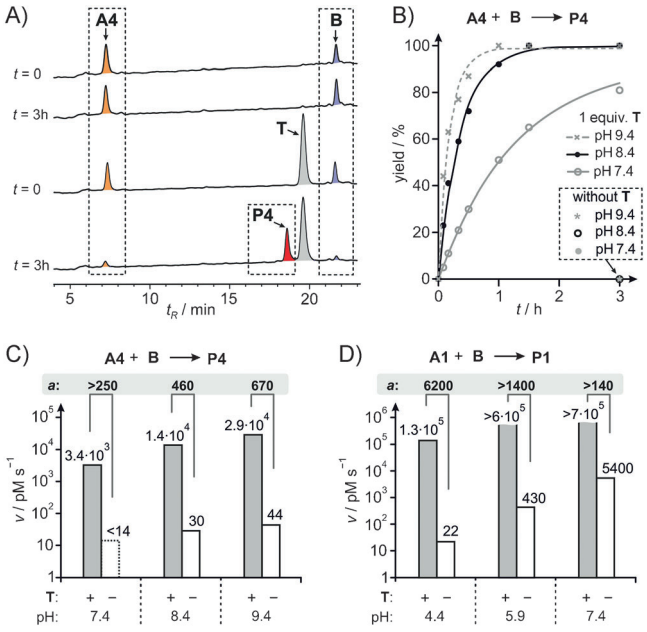

Figure 3. A) HPLC traces of ligation reactions at pH 7.4 between **A4** (orange) and **B** (blue) in the absence (top) and presence of **T** (bottom, gray). Product formation (**P4**, red) was observed after 3 h in the presence of **T**. B) Time courses of reactions performed at different pH values (7.4, 8.4, and 9.4). C) Logarithmic plot of observed rate constants (v/pM s^{-1}) for the ligation of **A4** and **B**. Rate acceleration ($a = v(+\mathbf{T})/v(-\mathbf{T})$) is given for each pH value (reactions were performed in buffer (20 mM Na_2PO_4, 20 μM tris(2-carboxyethyl)phosphine) at 30 °C with 20 μM **A4**, **B**, and with or without **T**). For details see Figures S10 and S11. D) Same plot as in (C) for ligation of **A1** and **B** (reactions were performed in buffer (20 mM Na_2PO_4, 5 μM tris(2-carboxyethyl)phosphine at 0 °C with 5 μM **A1**, 5.75 μM **B**, and with or without 5 μM **T**). For details see Figures S13 and S14.

idly react with thiols at room temperature and neutral pH, we lowered the reaction temperature to 0 °C. In addition, three different pH values (4.4, 5.9, and 7.4) were tested, and the concentration of peptide **A1** was reduced to 5 μM (previously 20 μM). Otherwise the described protocol was followed and initial rates for the ligation reaction between peptide **A1** and cysteine-bearing peptide **B** were determined (Figure 3D). In all cases, we observed considerable reaction rates in the absence of template **T** (v/pM s^{-1}: 22–5400). However, the rates of the templated reaction increased disproportionately. After 5 s at pH 5.9 and 7.4, already more than 60 % product was formed which exceeds the linear range. Therefore, only a lower limit for the initial rates could be determined (v/pM s^{-1}: $> 6 \times 10^5$, Figure S14) providing rate accelerations of > 140 (pH 7.4) and > 1400 (pH 5.9). At pH 4.4, the templated reaction was slow enough that the rate constant could be determined (v/pM s^{-1}: 1.3×10^5) resulting in a rate acceleration of 6200. This value exceeds accelerations of previously reported templated peptide ligations ($a = 81$–1800),[10–11] which is remarkable taking into consideration that we did not optimize spacer lengths and peptide affinities.

Having investigated these templated reactions under very defined conditions, we were interested in their performance in a more complex setup such as a protein-labeling reaction within cell lysate. Aiming for labeling under physiological pH (7.4), we decided to use a fluorescein-modified version of the chloroacetamide-modified peptide **A4*** (Figure 4A). As the target protein we chose human HSP70 (heat shock protein 70), which consists of 641 amino acids. To mark HSP70 for the ligation reaction, we fused the sequence of peptide **B** to its N and C termini, respectively (**B-HSP** and **HSP-B**, Figure 4A).

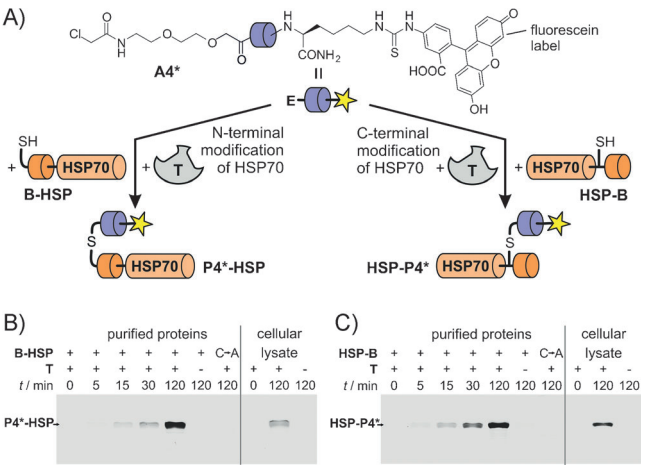

Figure 4. A) Protein-templated chemical modification of HSP70 using a fluorescein-labeled electrophilic peptide (**A4***). B) Labeling reactions with N-terminally modified HSP70 (**B-HSP**) resolved by SDS gel electrophoresis and fluorescence imaging. C) Analogous reactions with C-terminally modified HSP70 (**HSP-B**). Reactions were performed at 30 °C in buffer (20 mM Na_2PO_4, 5 μM tris(2-carboxyethyl)phosphine, pH 7.4) with 20 μM **A4*** and 20 μM **T**. As control, mutant fusion proteins were used which lack the reactive cysteine (C→A mutation).

These two fusion proteins were heterologously expressed in *E. coli*. For initial studies, we used purified proteins **B-HSP** and **HSP-B** (Figure S7), which we incubated with template **T** and labeled peptide **A4*** (each 20 μM). Thereafter, the proteins were analyzed by gel electrophoresis, and the covalently bound fluorescein label was visualized using fluorescence imaging (Figure 4B and C, Coomassie-stained gels in Figure S16). For both proteins, successful labeling was shown reflected by increasing product signals (**P4*-HSP** and **HSP-P4***) over incubation time ($t_{max} = 120$ min). As expected, we did not observe any labeling in the absence of template **T**. However, HSP70 exposes numerous solvent-accessible nucleophilic groups including one cysteine residue[16] that might participate in the templated reaction. To test for this possibility, we designed two additional fusion proteins in which the essential cysteine residue in sequence **B** was replaced by alanine (**B-HSP(C→A)** and **HSP-B(C→A)**). After 120 min in the presence of template **T** and labeled peptide **A4***, we did not observe any ligation product for the two mutated proteins (Figure 4B and C). This confirms the cysteine in sequence **B** to be the covalently modified residue in the templated reaction. Finally, we tested the efficiency of the labeling reaction with crude lysate of *E. coli* cells overexpressing fusion proteins **B-HSP** and **HSP-B**, respectively (Figure 4B and C). After 120 min of reaction only in the presence of template **T** and **A4*** we were able to detect

a single band corresponding to the size of the labeled HSP70 fusion proteins (**P4*-HSP** and **HSP-P4***).

In summary, we present the first example of a designed protein-templated ligation reaction involving two peptide fragments. Using an enzymatically inactive protein domain (KIX domain of CBP) as template **T**, we designed two substrates (**A** and **B**) based on known peptide ligands. The peptides were equipped with an electrophilic group and a thiol function, respectively. For the templated reactions, chloroacetamide (**A4**) and maleimide (**A1**) were selected as electrophiles. All tested reactions exhibited excellent rate accelerations ($a > 140$) with the maleimide-modified peptide **A1** ($a = 6200$) significantly exceeding accelerations of reported peptide ligations that use peptides, foldamers, or nucleic acids as templates ($a = 81$–1800).[10,11] To confirm the efficiency and selectivity of the protein-templated peptide ligation, we applied this reaction successfully to the selective labeling of a protein in cell lysate. With this proof-of-concept study we extend the portfolio of templated reactions, so far dominated by nucleic acid templates.[1] Given the elaborated biomolecular tools for protein engineering and peptide ligand discovery, protein-templated peptide ligations hold the potential to contribute improved techniques for the manipulation and detection of proteins in complex biological systems.

Received: January 21, 2014
Published online: March 18, 2014

Keywords: bioorthogonal reactions · peptide ligation · protein modification · solid-phase peptide synthesis · templated reactions

[1] K. Gorska, N. Winssinger, *Angew. Chem.* **2013**, *125*, 6956 – 6980; *Angew. Chem. Int. Ed.* **2013**, *52*, 6820 – 6843.
[2] a) G. von Kiedrowski, *Angew. Chem.* **1986**, *98*, 932 – 934; *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 932 – 935; b) W. S. Zielinski, L. E. Orgel, *Nature* **1987**, *327*, 346 – 347.
[3] a) D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525 – 528; b) K. Severin, D. H. Lee, J. A. Martinez, M. R. Ghadiri, *Chem. Eur. J.* **1997**, *3*, 1017 – 1024; c) S. Yao, I. Ghosh, R. Zutshi, J. Chmielewski, *J. Am. Chem. Soc.* **1997**, *119*, 10559 – 10560; d) B. Rubinov, N. Wagner, H. Rapaport, G. Ashkenasy, *Angew. Chem.* **2009**, *121*, 6811 – 6814; *Angew. Chem. Int. Ed.* **2009**, *48*, 6683 – 6686.
[4] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, et al., *Chem. Rev.* **2006**, *106*, 3652 – 3711.
[5] R. E. Kleiner, C. E. Dumelin, D. R. Liu, *Chem. Soc. Rev.* **2011**, *40*, 5707 – 5717.
[6] a) A. P. Silverman, E. T. Kool, *Chem. Rev.* **2006**, *106*, 3775 – 3789; b) T. N. Grossmann, A. Strohbach, O. Seitz, *ChemBioChem* **2008**, *9*, 2185 – 2192.
[7] K. K. Sadhu, T. Eierhoff, W. Romer, N. Winssinger, *J. Am. Chem. Soc.* **2012**, *134*, 20013 – 20016.
[8] a) C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem.* **2008**, *120*, 10182 – 10228; *Angew. Chem. Int. Ed.* **2008**, *47*, 10030 – 10074; b) E. M. Sletten, C. R. Bertozzi, *Angew. Chem.* **2009**, *121*, 7108 – 7133; *Angew. Chem. Int. Ed.* **2009**, *48*, 6974 – 6998; c) M. Grammel, H. C. Hang, *Nat. Chem. Biol.* **2013**, *9*, 475 – 484.
[9] a) M. Volgraf, P. Gorostiza, R. Numano, R. H. Kramer, et al., *Nat. Chem. Biol.* **2006**, *2*, 47 – 52; b) C. J. Krusemark, P. J. Belshaw, *Org. Biomol. Chem.* **2007**, *5*, 2201 – 2204; c) S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura, I. Hamachi, *Nat. Chem. Biol.* **2009**, *5*, 341 – 343; d) S. S. Gallagher, J. E. Sable, M. P. Sheetz, V. W. Cornish, *ACS Chem. Biol.* **2009**, *4*, 547 – 556; e) Y. Takaoka, A. Ojida, I. Hamachi, *Angew. Chem.* **2013**, *125*, 4182 – 4200; *Angew. Chem. Int. Ed.* **2013**, *52*, 4088 – 4106.
[10] a) K. Severin, D. H. Lee, A. J. Kennan, M. R. Ghadiri, *Nature* **1997**, *389*, 706 – 709; b) R. R. Araghi, B. Koksch, *Chem. Commun.* **2011**, *47*, 3544 – 3546.
[11] A. Erben, T. N. Grossmann, O. Seitz, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4993 – 4997.
[12] N. Vo, R. H. Goodman, *J. Biol. Chem.* **2001**, *276*, 13505 – 13508.
[13] S. Brüchweiler, R. Konrat, M. Tollinger, *ACS Chem. Biol.* **2013**, *8*, 1600 – 1610.
[14] Nucleic acid templated reactions indicate that ligand concentrations should be higher than the dissociation constants of involved complexes (e.g. T. N. Grossmann, O. Seitz, *Chem. Eur. J.* **2009**, *15*, 6723 – 6730). As we planned to use peptide ligands at micromolar concentrations, the observed dissociation constants (≈ 0.1 μM) should ensure efficient complex formation.
[15] D. S. Wilbur, *Bioconjugate Chem.* **1992**, *3*, 433 – 470.
[16] a) R. Kityk, J. Kopp, I. Sinning, M. P. Mayer, *Mol. Cell* **2012**, *48*, 863 – 874; b) X. C. Gao, C. J. Zhou, Z. R. Zhou, M. Wu, et al., *J. Biol. Chem.* **2012**, *287*, 6044 – 6052.

