DOI: 10.1002/cbic.200900029

A Targeted Releasable Affinity Probe (TRAP) for In Vivo Photocrosslinking

Ping Yan,[a] Ting Wang,[b] Gregory J. Newton,[c] Tatyana V. Knyushko,[d] Yijia Xiong,[b] Diana J. Bigelow,[b] Thomas C. Squier,[b] and M. Uljana Mayer*[b]

Protein crosslinking, especially coupled to mass-spectrometric identification, is increasingly used to determine protein binding partners and protein–protein interfaces for isolated protein complexes. The modification of crosslinkers to permit their targeted use in living cells is of considerable importance for studying protein-interaction networks, which are commonly modulated through weak interactions that are formed transiently to permit rapid cellular response to environmental changes. We have therefore synthesized a targeted and releasable affinity probe (TRAP) consisting of a biarsenical fluorescein linked to benzophenone that binds to a tetracysteine sequence in a protein engineered for specific labeling. Here, the utility of TRAP for capturing protein binding partners upon photoactivation of the benzophenone moiety has been demonstrated in living bacteria and mammalian cells. In addition, ligand exchange of the arsenic–sulfur bonds between TRAP and the tetracysteine sequence to added dithiols results in fluorophore transfer to the crosslinked binding partner. In isolated protein complexes, this release from the original binding site permits the identification of the proximal binding interface through mass spectrometric fragmentation and computational sequence identification.

Introduction

The essential role of both stable and transient protein–protein interactions in cell function is becoming increasingly apparent since dynamic changes in protein complex composition regulate cellular processes fundamental to life.[1,2] Thus, the development of new methods that permit the identification of protein interactions in living cells will be crucial to understanding mechanisms that regulate biological processes. Currently, crosslinking after cell lysis provides one of the most reliable methods for extrapolation of in vivo protein binding interactions.[3] However, cell lysis has the potential to disrupt intracellular protein complexes. For this reason the identification of functional interactions between proteins in supramolecular complexes, as well as of critical transient protein–protein interactions involved in signaling pathways, requires in vivo crosslinking methods that stabilize functional associations prior to cell disruption.

Current methods that permit the facile identification of interfacial interaction sites in known protein complexes commonly rely on the formation of disulfide bonds following the introduction of cysteines at the putative binding interface.[4,5] However, in the absence of reliable information regarding the binding interface, this technology is not practical even for simple isolated protein complexes. Instead, covalent crosslinking strategies involving either conventional chemistries (for example, formaldehyde or nonspecific amino- or thiol-reactive crosslinkers),[6] or nontargeted photoreactive amino acids are useful in stabilizing in vivo protein complexes prior to cell lysis, which can then be analyzed by using mass spectrometry following proteolytic digestion.[7] However, the identification of macromolecular binding interfaces by using these global approaches is complicated by the plethora of peptide masses that arise from widespread intra- and intermolecular crosslinking interactions. Therefore, the identification of crosslinking sites currently requires multiple purification steps and high-resolution mass spectrometry coupled with the development of new computational algorithms to deconvolute the resulting crosslinked dipeptides.[6,8–10] An alternative approach has been demonstrated that involves reengineering aminoacyl-tRNA synthetases in *E. coli* and *S. cerevisiae* to allow the selective incorporation of *p*-benzoyl-L-phenylalanine at defined sites.[11–13] This latter method takes advantage of the extensive (>50%) crosslinking efficiencies of benzophenones upon their incorporation within protein binding interfaces to identify binding partners.[14,15]

To simplify the identification of interactions within protein complexes in living cells we introduce here a new crosslinker,

[a] *Dr. P. Yan*
Novozymes, Inc., 1445 Drew Ave
Davis, CA 95618 (USA)

[b] *Dr. T. Wang, Dr. Y. Xiong, Dr. D. J. Bigelow, Dr. T. C. Squier, Dr. M. U. Mayer*
Cell Biology and Biochemistry Group,
Pacific Northwest National Laboratory
999 Battelle Boulevard,Stop P7-54, Richland, WA 99354 (USA)
Fax: (+1) 509-376-6767
E-mail: uljana.mayer@pnl.gov

[c] *G. J. Newton*
The University of Tokyo, Komaba Open Laboratory
Room 305, 4-6-1, Komaba, Meguro-ku, Tokyo, 153-8904 (Japan)

[d] *Dr. T. V. Knyushko*
Harris IT Services, NIH/NCI
9030 Old Georgetown Road, 82, Bethesda, MD 20814 (USA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200900029.

5-(4-benzoylbenzamido)-4′,5′-di(1,3,2-dithiarsolan-2-yl) fluorescein. This targeted reversible affinity probe (TRAP) extends the functionality of biarsenical multiuse affinity probes, which have previously been used to fluorescently label proteins,[16–19] by adding crosslinking as a means to retain native protein complexes during cell disruption and analysis. These approaches build upon the ability of the fluorescent biarsenical affinity probe FlAsH-EDT_2 and its derivatives to selectively label a range of soluble and membrane proteins in both eukaryotic and bacterial organisms following the introduction of an engineered tetracysteine binding sequence, CysCysXXCysCys, in which X is any amino acid.[20–23] Addition of this short tagging sequence has been shown to not change the structure or function of the protein under study.[24–26] However, as with all tags, the tetracysteine tag has to be placed judiciously for accessibility by the probe and to avoid steric effects. Therefore the tag is on existing loops or at a terminus that does not have a cellular localization tag. Site-specific labeling is simplified compared to unnatural amino acid engineering because no changes need to be made to the structure of the organism's aminoacyl-tRNA synthetases or codon usage. Fluorescence that is specifically associated with a protein of interest provides a means to monitor protein localization and protein turnover in living cells.[21] The large (>20-fold) increase in fluorescence upon binding and stability under electrophoretic conditions provides a rapid means to assess the abundance and labeling of tagged proteins.[16,27] In addition, it has previously been shown in the identification of binding partners through complex isolation[26,28] and through label transfer[29] that biarsenical fluorophores can be released from the binding motif through competition with a high concentration of dithiols.

In TRAP, the arsenic-derivatized carboxyfluorescein, CrAsH, which binds to tetracysteine motifs of the sequence CysCysXXCysCys,[30] is coupled through an ethylenediamine linkage to carboxybenzophenone. Benzophenone was chosen as the photocrosslinking group because the active radical regenerates after reaction with water, thus leading to high photocrosslinking yields when the crosslinker is proximal to a protein binding partner.[14,15] We have named the new crosslinker Targeted Releasable Affinity Probe (TRAP), because it is an *affinity probe* that combines 1) the ability to selectively label (*target*) an engineered tagging site on a target protein in a complex mixture (Figure 1 B), 2) the ability of the benzophenone-photoactivated crosslinker to capture (*trap*) binding partners (Figure 1 C), and 3) an ability to *release* the bound fluorophore and transfer it to the binding partner for efficient mass spectrometric identification of the crosslinking site (Figure 1 D). Here, we have demonstrated the utility of TRAP as a general method for stabilizing transient and low-affinity binding interactions between bacterial and eukaryotic proteins in vivo prior to cell lysis. The utility of this approach in the identification of binding interfaces in isolated protein complexes is demonstrated following ligand exchange and mass spectrometric identification of the crosslinking site following proteolytic digestion.

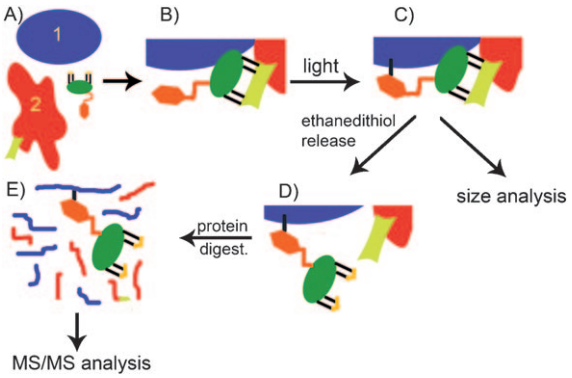
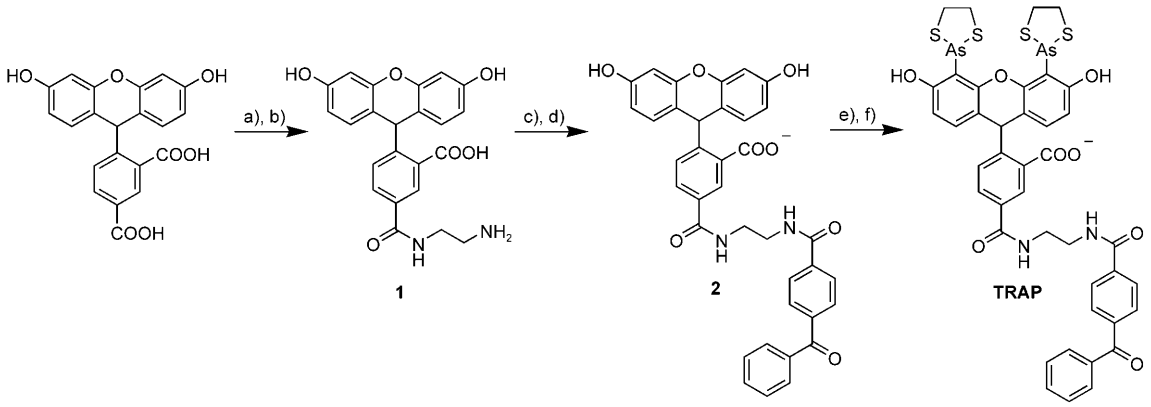


Figure 1. Summary cartoon of TRAP function. A) A mixture of TRAP (green with orange benzophenone and yellow dithiols caps), protein 1 and protein 2 (additional proteins and cellular material not shown); B) TRAP selectively binds the tetracysteine tag engineered onto protein 2, depicted in yellow; C) The TRAP benzophenone subunit is photocrosslinked to protein 1. The crosslinked complex can then be visualized, for example by SDS-PAGE, to determine the size of the crosslinked species, or D) the crosslinker can be released from protein 2 by adding excess dithiol; E) After protein digestion the peptide attached to the released crosslinker can be determined through tandem mass spectrometry.

Results and Discussion

Synthesis

We report the synthesis of TRAP, a biarsenical crosslinker with a fluorescein scaffold linked through ethylenediamine to a photoreactive benzophenone moiety. As shown in Scheme 1, TRAP synthesis proceeded in several multistep reactions through *N*-hydroxysuccinimide (NHS) ester-mediated amide couplings, followed by metalation. EDC was found to be the most effective reagent for activation of the carboxylates to form the NHS esters. The first synthetic step involved the transient creation of the carboxyfluorescein-NHS ester, followed by coupling of *N*-Boc-ethylenediamine; this compound was deprotected and added to the newly generated carboxybenzophenone-NHS ester. Then the fluorescein moiety was mercurated, followed by a transmetalation to arsenic and capping with ethanedithiol, as described previously.[16,17,26,30] No benzophenone-metalated isomers were identified after elution from silica. An analogue, 5-(4-benzoylbenzamido)-4′,5′-di(1,3,2-dithiarsolan-2-yl) fluorescein, which has a shorter linker length, was synthesized by generating the acetyl chloride of carboxybenzophenone, followed by coupling to aminofluorescein. In this case, the reaction using NHS esters did not proceed due to the low reactivity of aminofluorescein compared to the aliphatic amine. This short crosslinker was tested in all of the protein complexes described below, but is a lower-yielding crosslinker than TRAP in all cases. A third crosslinker, which contained a hexanediamine linker showed no reactivity in most cases and will not be described further (data not shown). We hypothesize that TRAP is the most efficient crosslinker of the three, because it has sufficient flexibility to reach binding partners, but is not so long that it samples mostly protein-free space.



Scheme 1. Synthesis of TRAP. A) EDC, Et_3N, NHS, dry DMF, 30 min; B) *N*-Boc-ethylenediamine, 16 h; C) 20% TFA/CH_2Cl_2, 2 h, D) 4-Benzoylbenzoic acid, EDC, NHS, iPr_2EtN, DMF, 16 h; E) HgO, TFA, 70 °C; F) 1) $AsCl_3$, PdOAc, iPr_2EtN, NMP, 4 h; 2) EDT, 20% acetone/H_2O (overall yield: 2%).

Because reduction of benzophenone crosslinking efficiency through coupling to a chromophore was a concern, the crosslinking efficiencies of the TRAP reagent and of nonderivatized benzophenone were tested through photoreaction at 30 °C in isopropanol solvent by using the Rayonet photoreactor at 350 nm. As measured by the integration of substrate and product peaks on HPLC, TRAP retained 60% of the efficiency of unmodified benzophenone (data not shown). Because benzophenone is considered one of the more efficient photocrosslinkers,[14,15] this result indicates that TRAP should be a good crosslinker in the study of protein complexes.

In vitro crosslinking of calmodulin to MLCK peptide

The utility of TRAP for effecting the crosslinking of protein complexes was validated by using a model system involving the high-affinity association between calcium-activated calmodulin and a synthetic peptide corresponding to the calmodulin-binding sequence of smooth-muscle myosin light-chain kinase, also known as RS20.[31,32] This calmodulin-binding peptide, henceforth abbreviated MLCK, contains an appended C-terminal amino acid sequence, CCKACCA, which serves as the site for TRAP crosslinker binding and has arbitrarily been depicted as a helix connected to the calmodulin-binding helix (cartoon in Figure 2A). This tagging sequence was previously shown to facilitate arsenic exchange to dithiols,[19,33] which is expected to aid in the transfer of the TRAP fluorophore to the crosslinked protein to permit facile identification.

The tagged MLCK peptide was labeled with TRAP, as described in the Experimental Section. Essentially complete crosslinking was achieved as evidenced by the disappearance of the band associated with free (unbound) MLCK; this is apparent from a comparison of either the Coomassie-stained or fluorescence-imaged SDS-PAGE gel (Figure 2B). There is a corresponding appearance of two high-molecular mass bands that are consistent with the formation of a crosslinked complex between MLCK and calmodulin (near 22 kDa) and a TRAP-bound calmodulin (near 17 kDa) that results from thiol exchange under the reducing conditions associated with SDS-PAGE. Consistent with this latter suggestion, addition of dithiothreitol or

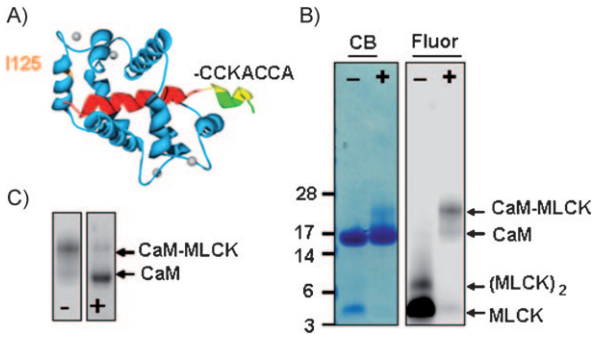


Figure 2. Photodependent crosslinking of protein complex and fluorophore transfer. A) Depiction of the cartoon structure of binding complex between calcium (white spheres) bound calmodulin (CaM; blue, with I123 shown in orange) and tetracysteine-tagged calmodulin-binding sequence of smooth-muscle myosin light-chain kinase (MLCK; red; PDB ID: 1cdl) that illustrates the relative size of the tagging sequence (CCKACCA; displayed in green with yellow cysteine residues). B) Extent of crosslinking of the isolated complex (150 μM) is visualized before (–) and after (+) photoactivation of TRAP following SDS-PAGE by Coomassie blue protein staining (CB) or fluorescence (shown as an inverted image) from the TRAP crosslinker (Fluor), as described in the Experimental Section. C) Crosslinked complex before disruption (–) and after TRAP fluorescence is transferred from MLCK to calmodulin upon addition of dithiothreitol (5 mM; +).

other dithiols to the crosslinked calmodulin-MLCK peptide complex induces nearly complete transfer of TRAP from the MLCK peptide to calmodulin (Figure 2C). No crosslinking products were observed when the reaction was performed in the presence of the calcium scavenger EGTA (data not shown); this result indicates proper functioning of the tagged peptide and also signifies a requirement for close proximity between benzophenone and reactant because the calcium activation of calmodulin is necessary for binding to MLCK.

We have thus shown here that in purified complexes, TRAP fluorescence can be used to evaluate labeling and crosslinking. We showed that at the relatively high protein complex concentration of 150 μM (compared to 2 μM complex concentration for in vitro crosslinking of larger complexes)[34] no crosslinking takes place if the protein complex is prevented from forming by calcium chelation (data not shown). Labeling, crosslinking and label transfer efficiencies are similar to those for a similar

transfer agent with oxidative crosslinking chemistry that has been applied in simple systems.[29]

Mass spectrometric identification of a crosslinking site

The ability of added dithiols to efficiently promote ligand exchange of the arsenic ligands and release the TRAP fluorophore to the crosslinked binding partner facilitates the use of mass spectrometry to identify the site of crosslinking. Calmodulin–peptide complexes have previously been studied by classical crosslinkers combined with high-resolution tandem mass spectrometry.[35–37] Herein, we used the complex between calmodulin and the MLCK peptide to validate the ability to use TRAP with low-resolution tandem mass spectrometry for the identification of crosslinking sites. Accordingly, the TRAP fluorophore was transferred to calmodulin within the MLCK protein complex as described above and in the Experimental Section, followed by protein digestion with trypsin. After LC–MS/MS, the fragment patterns of the mass spectra were searched against a database containing the mouse proteome with the modified MLCK peptide added by using the search algorithm OMSSA.[38] A crosslinking site was identified at Ile125 within the peptide I^{125}READIDGDGQVNYEEFVQM_{ox}MTAK, and all possible fragment ions were identified within the spectrum (Figure 3). Other spectra that had fewer fragment ions showed different crosslinking sites. As shown in Figure 2, this binding site is particularly interesting because it suggests that the MLCK peptide binds calmodulin in the direction opposite to that in the crystal structure shown, in support of previous findings.[32,39] This crosslink cannot be achieved in the structure that was found after crystallization based on crosslinker length. These findings seem to indicate a problem inherent in structure determination, namely that crystallization only captures stable, rigid binding modes, but crosslinking and single-molecule fluorescence spectroscopy can be used to characterize dynamic and transient interactions.

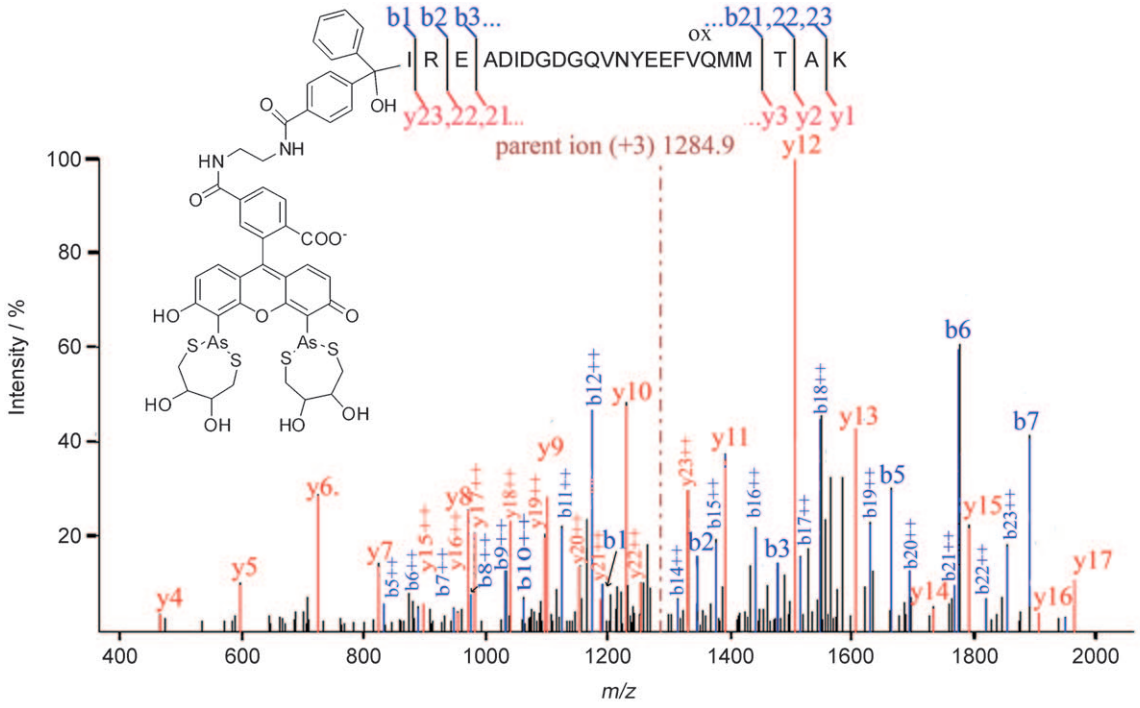
In a control sample corresponding to the uncrosslinked MLCK peptide bound to calmodulin, no spectrum was identified whose fragmentation pattern was consistent with a peptide with an appended TRAP moiety. Equivalent results were obtained by using the X!tandem algorithm;[40] this indicates the broad utility of the TRAP reagent in the identification of crosslinking sites located at binding interfaces within protein complexes.

This simple in vitro protein complex of calmodulin and TRAP-labeled MLCK peptide shows the utility of TRAP for the identification of protein interaction sites in purified systems. Moreover, the bright fluorescence of the fluorescein moiety and its strong binding permit the facile assay of protein specificity by using electrophoresis gels. At the same time, the ability to transfer TRAP from the original tagged protein to its crosslinked partner, followed by bottom-up proteomics with the transferred TRAP as variable modification, permits pinpointing the exact sites of protein–protein interactions by using existing and publically available search software.[38,40] This is neither possible with crosslinkers that are not targeted to one specific tagging site prior to crosslinking, nor with those that are not releasable. In those cases, mass spectra must be searched by hand in simple systems or through new computationally intensive approaches that have the ability to match fragmentation patterns made by linked peptides.[8–10,36] Searching for crosslinked species from protein complexes in complex samples, for example, whole proteomes, becomes problematic because every peptide has to be searched as possibly crosslinked to every other peptide; this results in a combinatorial expansion. We have shown here that TRAP overcomes this problem through its combination of unique properties: 1) Targeting of TRAP affords knowledge of the initial binding site. 2) TRAP binds tightly to the targeted tag sequence until it is released through dithiols. 3) Tandem-MS analysis by using released TRAP as post-translational modification is made possible because TRAP is relatively small (an added weight of 1077 u compared to 1709 u for DOPA-biotin-FlAsH,[29] assuming the same release conditions), which permits the identification of TRAP-containing singly and doubly charged fragments by analysis through low-resolution electron spray ionization tandem mass spectrometry. In addition, the biarsenical unit of FlAsH could be used to reduce sample complexity, if necessary, by removing the dithiol that was used for release from the bait protein, and capturing the proteins to which TRAP has been transferred with a tetracysteine-peptide bound to a resin.

Ex vivo crosslinking of SlyD identifies protein interactions

To test the utility of TRAP as a crosslinking reagent within a complex proteome, TRAP was incubated with lysates from *Shewanella oneidensis* MR-1 that contain the peptidylprolyl isomerase, SlyD. SlyD is an important chaperone and metallase that has been implicated in the maturation of hydrogenases and other proteins.[41–44] SlyD contains a high-affinity binding sequence that has been demonstrated to efficiently and specifically bind the tetracysteine-binding dye FlAsH in cell lysates.[33] Moreover, a synthetic peptide corresponding to its naturally occurring tetracysteine-containing motif, GCCGGSGNDAGGCCGG, has been shown to exhibit a similar binding affinity, but brighter fluorescence and less propensity for ethanedithiol release than the tag sequence CCKACCA used with MLCK; this suggests that SlyD is suitable for the sensitive detection of TRAP-binding specificity.[19,33]

For the crosslinking experiment, *S. oneidensis* was grown under suboxic conditions to increase expression of SlyD for optimal labeling because it is not an abundant protein in *S. oneidensis* (< 0.2% of soluble proteins). The cells were then harvested, lysed, and the soluble fraction was labeled with TRAP. We find that under these conditions, incubation of TRAP with the lysate before photoactivation results in selective labeling of SlyD, which is apparent upon visualization of the fluorescence signal of a single band with an apparent molecular weight near 28 kDa (Figure 4). As previously determined through mass spectrometry, this band corresponds to the 23.5 kDa SlyD protein.[33] After photoactivation, a family of fluorescently labeled crosslinked protein complexes is observed concomitant with loss of the 28 kDa fluorescent band. A parallel protein complex isolation performed for MS analysis provid-



ion	*M*+1 (u)	*M*+2 (u)	fragment sequences	ion	*M*+1 (u)	*M*+2 (u)
b1	1191.48(0.41)		TRAP-I | READIDGDGQVNYEEFVQmMTAK	y23		1331.95 (0.36)
b2	1347.61(0.43)		TRAP-IR | EADIDGDGQVNYEEFVQmMTAK	y22		1253.76 (0.22)
b3	1476.6(0.38)		TRAP-IRE | ADIDGDGQVNYEEFVQmMTAK	y21		1188.93 (-0.09)
b4	1547.5 (0.25)		TRAP-IREA | DIDGDGQVNYEEFVQmMTAK	y20		1153.70 (0.19)
b5	1662.74(0.45)	831.91 (0.26)	TRAP-IREAD | IDGDGQVNYEEFVQmMTAK	y19		1095.83 (0.22)
b6	1775.72(0.35)	888.26 (0.07)	TRAP-IREADI | DGDGQVNYEEFVQmMTAK	y18		1039.62 (0.18)
b7	1890.86(0.47)	945.94 (0.24)	TRAP-IREADID | GDGQVNYEEFVQmMTAK	y17	1962.82 (-0.04)	981.95 (0.01)
b8	1947.56(0.14)	974.35 (0.14)	TRAP-IREADIDG | DGQVNYEEFVQmMTAK	y16	1906.10 (0.26)	953.27 (-0.15)
b9		1032.03 (0.31)	TRAP-IREADIDGD | GQVNYEEFVQmMTAK	y15	1790.68 (-0.13)	895.80 (-0.11)
b10		1060.52 (0.28)	TRAP-IREADIDGDG | QVNYEEFVQmMTAK	y14	1733.69 (-0.1)	
b11		1124.48 (0.21)	TRAP-IREADIDGDGQ | VNYEEFVQmMTAK	y13	1605.68 (-0.05)	
b12		1174.05 (0.25)	TRAP-IREADIDGDGQV | NYEEFVQmMTAK	y12	1506.59 (-0.08)	
b13		1231.28 (0.46)	TRAP-IREADIDGDGQVN | YEEFVQmMTAK	y11	1392.53 (-0.09)	
b14		1312.59 (0.24)	TRAP-IREADIDGDGQVNY | EEFVQmMTAK	y10	1229.51 (-0.05)	
b15		1376.95 (0.07)	TRAP-IREADIDGDGQVNYE | EFVQmMTAK	y9	1100.44 (-0.07)	
b16		1441.74 (0.35)	TRAP-IREADIDGDGQVNYEE | FVQmMTAK	y8	971.50 (0.03)	
b17		1515.28 (0.35)	TRAP-IREADIDGDGQVNYEEF | VQmMTAK	y7	824.41 (0.01)	
b18		1564.44 (-0.02)	TRAP-IREADIDGDGQVNYEEFV | QmMTAK	y6	725.41 (0.08)	
b19		1628.25 (-0.24)	TRAP-IREADIDGDGQVNYEEFVQ | mMTAK	y5	597.31 (0.03)	
b20		1693.96 (-0.06)	TRAP-IREADIDGDGQVNYEEFVQm | MTAK	y4	466.30 (0.06)	
b21		1767.52 (-0.01)	TRAP-IREADIDGDGQVNYEEFVQmM | TAK	y3		
b22		1818.92 (0.42)	TRAP-IREADIDGDGQVNYEEFVQmMT | AK	y2		
b23		1853.92(0.35)	TRAP-IREADIDGDGQVNYEEFVQmMTA | K	y1		

Figure 3. Identification of interfacial crosslinking site in the protein complex after LC-MS/MS in an LTQ Orbitrap. MS/MS spectrum of TRAP-bound C-terminal peptide, $I^{125}READIDGDGQVNYEEFVQM_{ox}MTAK^{148}$, with the triply-charged precursor mass indicated as a hatched line (top) and identified masses (bottom) from calmodulin following release of complex and tryptic digestion. All *b* (blue) and *y* (red) fragment ions are identified with a difference of less than 0.5 u from the theoretical mass of the fragment ion (indicated in brackets in the table). The monoisotopic mass of TRAP (1076.89 u) is associated with N-terminal fragment ions (*b* ions; blue) containing Ile125. In comparison, no C-terminal fragment ion (*y* ions; red) contains TRAP.

ed a tentative identification of these multiple binding partners; this is consistent with the role of SlyD as a chaperone in promoting the maturation of a range of different proteins and constitutes a first attempt to identify these previously un-

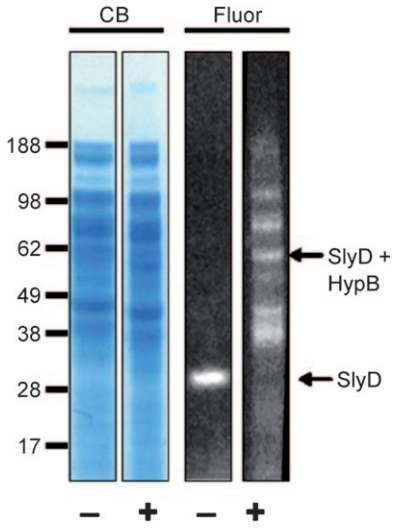


Figure 4. Selective labeling of bacterial peptidyl–prolyl isomerase SlyD and photo-dependent crosslinking of binding partners. SDS-PAGE (4–12%) visualized by using Coomassie blue protein stain (CB) or TRAP fluorescence (Fluor) following addition of TRAP (20 μM) to lysates prepared from *Shewanella oneidensis* either prior to (–) or following (+) photoactivation. Positions of molecular mass markers are shown on the left. Positions of SlyD (SO3417) and SlyD+HypB (SO2093) previously shown to bind SlyD[43] are shown on the right. Proteins identified as putative SlyD interaction partners following affinity isolation, identified from MS/MS as described in the Experimental Section, are: ornithine decarboxylase speF, (SO0314); heat shock protein HtpG (SO2016); glucosamine-fructose-6-phosphate aminotransferase, glmS (SO4741); molybdopterin biosynthesis MoeA (SO4723); Fe–S cluster-binding protein (SO1521); phosphoribosylaminoimidazolecarboxamide formyltransferase, purH (SO0442); polysaccharide biosynthetic protein (SO3190); 2-oxoglutarate dehydrogenase E2 component, dihydrolipoamide succinyl transferase, sucB (SO1931); conserved hypothetical protein (SO1190); hydrogenase accessory protein, HypB (SO2093); catabolite gene activator (SO0624); peptidyl-prolyl cis-trans isomerase Fkbp family (SO1995); molybdenum cofactor biosynthesis protein C, MoaC (SO4451); and ferric uptake regulatory protein, fur (SO1937).

known substrate proteins. On the other hand, the identification of the hydrogenase accessory protein, HypB (29.3 kDa), in complex with SlyD, apparent as a 54 kDa fluorescent gel band, supports the validity of these identifications based on previous work that demonstrates the role of the SlyD-HypB complex in the metalation of the [Ni–Fe] hydrogenases.[43,45] Thus, this work demonstrates the specific binding of TRAP to a single tetracysteine-containing protein expressed at native levels within the complex protein mixture of a bacterial proteome and the utility of TRAP for photoactivation-dependent capture of multiple protein interactions, that can be visualize through SDS-PAGE with fluorescent readout.

In vivo crosslinking of the bacterial RNA polymerase α subunit identifies new interactions

The utility of biarsenical crosslinkers for use in living Gram-negative bacteria was demonstrated in living *S. oneidensis* MR-1 bacterial cells by using the tetracysteine-tagged α subunit of RNA polymerase. The bacterial RNA polymerase core complex is localized to the cytoplasm and is responsible for all DNA transcription.[46] In addition, RNA polymerase represents a supramolecular complex that contains more than 70 different proteins in actively growing *S. oneidensis*,[26] in which complex protein–protein interactions regulate RNA polymerase transcription in this facultative anaerobe.[47] Further, the core $\alpha_2\beta\beta'$ complex represents approximately 0.6% of the cellular protein in *S. oneidensis*; this permits the efficient in vivo labeling of the tagged α subunit (containing the CCPGCC tetracysteine tag, a V5 epitope and a His_6 sequence at the C terminus) in this protein complex by using biarsenical probes.[26] Consistent with these measurements, high levels of specific labeling are also apparent upon incubation of TRAP with living bacteria (Figure S1 in the Supporting Information); this permits an investigation of the ability of TRAP to identify binding partners within large supramolecular complexes in living *S. oneidensis* MR-1 bacterial cells.

RNA polymerase was labeled by incubation of bacteria with either FlAsH-EDT_2, as a control sample incapable of crosslinking, or the TRAP reagent as described in the Experimental Section. It was previously shown that at the conditions used, the tetracysteine-tagged α subunit represents about half of the total α subunit and is incorporated in the complex without a change in function.[26] As indicated, a single fluorescent band associated with the 42 kDa tagged α subunit of RNA polymerase is observed on SDS-PAGE in the absence of photoactivation (Figure S1). Likewise, a single major band is apparent by using immunoblots against the V5 epitope of the α subunit (Figure 5); this permits a targeted evaluation of the crosslinking of the RNA polymerase core complex. After photoactivation, proteins with a higher molecular mass are apparent by using antibodies against either the tagged α subunit or the β subunit of the RNA polymerase (Figure 5). These immunoblots reveal facile crosslinking of α subunits even with low illumination as evidenced by the major band migrating with an apparent molecular mass of 90 kDa, which corresponds to the

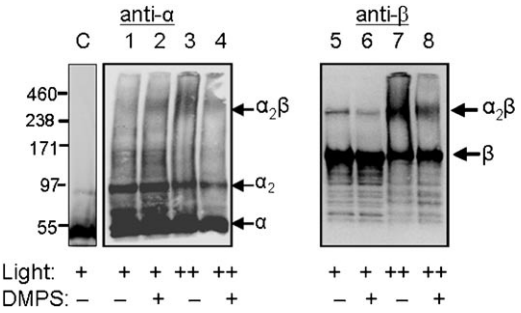


Figure 5. In vivo labeling and crosslinking of tagged α subunit of bacterial RNA polymerase. Immunoblots using primary antibodies directed against i) the V5 epitope of the tag sequence of the α subunit (lanes C, 1–4) or ii) the β subunit (lanes 5–8) of the RNA polymerase in lysates from *Shewanella oneidensis* MR-1 expressing tagged α subunits of RNA polymerase with engineered binding sequence for FlAsH or TRAP at the C terminus, as previously described.[26] Live cells were labeled with either the parent compound, FlAsH-EDT_2 (lane C) or the TRAP crosslinking reagent (lanes 2–8) prior to photoreaction and lysis. After light exposure for 1 h (+) or 3 h (++), cells were incubated in the absence (–) or presence (+) of 1 mM dimercaptopropanesulfonate (as described in the Experimental Section) prior to protein separation by SDS-PAGE (3–8% acrylamide) and immunoblotting. Arrows on the right of panels indicate the expected positions of RNA polymerase crosslinked species based on positions relative to molecular mass markers (left).

α homodimer. Longer light exposures result in some loss of the α_2 species and a concomitant increase in the density of a band migrating with a mass of 260 kDa; this is consistent with the formation of the $\alpha_2\beta$ heterotrimer. This latter assignment is confirmed by the use of antibodies against the β subunit. The complex banding pattern apparent by using an antibody against the V5 epitope of the α subunit suggests additional crosslinking interactions between the α subunit and multiple associating proteins that are consistent with the large supramolecular complex structure of RNA polymerase.[26] The lack of higher bands in the FlAsH control and the relatively weak anti-V5 antibody binding to the species with the higher molecular weight compared to the anti-β subunit antibody is further evidence that the new bands are due to specific crosslinks in which the proximity of the crosslinking site to the antibody-binding site appears to partially occlude antibody binding.

The specificity of the crosslinking reaction was further confirmed by dimercaptopropanesulfonate-induced release of the biarsenical probe from the CCPGCC tetracysteine tag (Figure 5, lanes 2, 4, 6, 8), which results in partial loss of both α_2 and $\alpha_2\beta$ crosslinked products as detected by antibodies to both α and β subunits of RNA polymerase. As expected, the CCPGCC motif, which was chosen to ensure good TRAP binding under in vivo conditions shows only partial ligand exchange compared to the facile exchange from the CCKACC tag that was used in the calmodulin experiment (Figure 2). The gel region where an αβ-heterodimer would migrate, that is, 192 kDa, is notably devoid of bands, which are, however, seen if a purified RNA-polymerase complex is crosslinked with TRAP in vitro (data not shown).

We chose to place the TRAP-binding site on the C terminus of the RNA polymerase α subunit; the spatial arrangement of these domains represented an unknown because they are not resolved in the crystal structure.[48,49] The C termini of the two α subunits were known to interact with DNA and transcription factors.[50] As opposed to the N termini, these domains were not believed to interact with one another or other RNA polymerase subunits, at least in the transition initiation complex,[48] despite findings of dimer formation for the truncated C-terminal α subunits.[51] Our finding of an α_2 dimer indicates that this previously seen interaction of the isolated C-terminal domains is physiologically functional. One can hypothesize that while the C termini of the RNA polymerase subunit interact with DNA in the transcription initiation complex, they swing back to interact with the β subunit in the course of transcription. These latter results emphasize the utility of in vivo crosslinking methods as a means to reliably capture the physiological conformation of RNA polymerase in complex with DNA and other protein associations. The ability of FlAsH and TRAP to enter cells and label proteins in the cytoplasm as they are formed under reducing conditions has an additional advantage; they can immediately bind to the tetracysteine tag before misfolding due to disulfide formation with native cysteines can occur. In addition, prior to release, TRAP (like specifically directed unnatural photoreactive amino acids)[13] thus permits the identification of in vivo protein interactions following protein separation by SDS-PAGE without prior knowledge of interaction sites.

Identification of unknown binding interactions in living myocytes

Prior measurements have demonstrated an unexpected trafficking of the regulatory protein phospholamban in myocytes that involves the directed transport of small vesicles of unknown origin to the plasma membrane.[21] To identify possible binding partners, and explore the utility of TRAP in capturing transient binding interactions in eukaryotic cells, we used the abundant muscle membrane protein phospholamban, in which amino acids 23, 24, 27, and 28 were mutated to cysteines to create an internal PQCCRQCCQN motif.[52] This tag location was chosen based on the work of Robia et al.,[52] who showed that the tagged and FlAsH-labeled phospholamban behaves like the wild-type in forming native oligomeric complexes. The engineered protein was expressed at nearly native levels in the C2C12 myocyte cell line.[21] Previous experiments have demonstrated specific labeling of tagged phospholamban with FlAsH in C2C12 myocytes.[21] The following experiments focus on the use of Western immunoblotting against phospholamban, because immunoblotting has greater sensitivity than fluorescence signals after SDS-PAGE.

Prior to photoactivation, a phospholamban pentamer is apparent as a 28 kDa species that predominates on the immunoblot (Figure 6).[53] After photoactivation, a major new phospholamban-containing species is observed that migrates with an apparent molecular mass above 250 kDa. Mass spectrometric analyses after a parallel pull-down of phospholamban identified this as a complex with fibronectin, which was the only protein to be confidently identified in this mass range. As part of the actin cytoskeleton, the appearance of fibronectin–phospholamban interactions is consistent with the observed trafficking of phospholamban to the plasma membrane.[21,54]

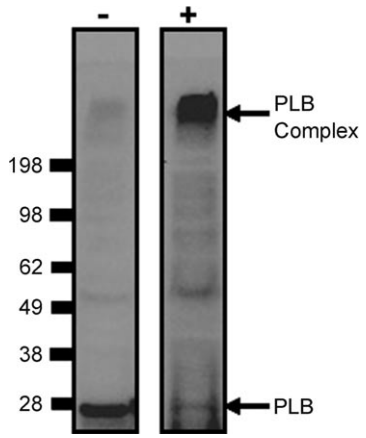


Figure 6. In vivo labeling and crosslinking of phospholamban in C2C12 myocytes. Western immunoblot against phospholamban (PLB) in lysates prepared from living C2C12 myocytes labeled with TRAP crosslinker (2 μM) either before (−) or after (+) photoactivation by using a Rayonet photoreactor. In all cases, myocytes were incubated overnight at 37 °C with TRAP crosslinker (2 μM) and, following buffer exchange to remove unbound TRAP, cells were subjected to photolysis (where indicated) and following lysis were probed using an antibody against phospholamban. Myocytes were engineered to express a tagged phospholamban that contains a binding sequence for TRAP, as described in the Experimental Section.

These measurements demonstrate that the cell-permeable TRAP reagent is capable of crosslinking of interacting proteins after labeling of a tetracysteine-tagged intracellular membrane protein in a living eukaryotic cell. Combined with the results in Gram-negative bacteria, it is reasonable to believe that TRAP would also cross cell walls and have utility in Gram-positive bacteria and archaea.

Conclusions

We have synthesized a new class of photocrosslinkers that can be targeted to a tetracysteine tag to capture (that is, trap) protein binding partners upon light activation. TRAP, which has the photocrosslinking benzophenone connected through an ethylenediamine linker to the fluorescent targeting unit is the focus of this report. Targeting crosslinkers through cysteine mutations and crosslinking in vitro has permitted the identification of protein–protein interaction sites.[55,56] TRAP constitutes the first step to extending this targeting capability to protein complexes in vivo with the help of standard mutagenesis; this avoids more demanding methods, such as changing the translation machinery to accept unnatural amino acids.[12,13] The simplicity of this method will facilitate the high-throughput identification of protein complexes in a range of different organisms. We first showed that precise targeting and photocrosslinking is easily achieved in a pure protein complex. By choosing a tag of appropriate relative binding affinity, that is, CCKACC, as opposed to CCPGCC or CCGGSGNDAGGCC, release of the crosslinker from the initial tag to the newly bound protein is easily achieved. This capability and the small size of TRAP compared to other targeted and multifunctional crosslinkers enable identification of the site of crosslinking by searching for the added molecular weight of the transferred crosslinker after low-resolution tandem mass spectrometry, as shown in the case of the calmodulin–MLCK complex. The fact that a search for the added molecular mass of released TRAP through publically available software permits determination of the site of protein binding interfaces has the potential to open up protein crosslinking followed by mass spectrometry to high-throughput approaches and the general research community, rather than specialized mass spectrometry groups with access to high-resolution mass spectrometers and the resources to develop new computational algorithms.[8,10,35]

In addition, in purified protein complexes and cell lysates the tight binding and fluorescence of the crosslinker enables evaluation of crosslinked protein complex formation through visualization of TRAP fluorescence on SDS-PAGE gels. Further, TRAP is cell-permeable and capable of crosslinking cytoplasmic proteins in Gram-negative bacteria, and internal membrane proteins in eukaryotic cells. Lower labeling yields and longer photoactivation times in vivo permit the use of the fluorescence feature prior to, but not after photoactivation.

The studies detailed herein, which were designed to showcase the ability of TRAP to cross cellular membranes and selectively label the tagged target proteins expressed at near-native concentrations, have further shown that TRAP will find uses in the identification of new binding partners. In the cases of SlyD and phospholamban, using a combination of TRAP technology with the protein complex purification capability of biarsenical multiuse affinity probes, we were able to tentatively identify new phospholamban and SlyD binding partners. In the case of the RNA polymerase tagged on the C terminus of the α subunit, a large number of protein complexes were seen, as expected from prior results.[26] By using immunoblots, we were able to identify α_2 dimers, and a crosslink of the α_2 dimer to the β subunit. Because the C-terminal domains of the α subunits are not visible in the crystal structures, this is the first time the C termini of the α subunits have been confidently localized to the RNA polymerase core, rather than on DNA, as previously hypothesized.[48]

In future work we will use TRAP in combination with targeted in vivo photocrosslinking and mass spectral identification of interfacial binding sites to identify in high-throughput how environmental conditions affect protein–protein interaction networks. It is expected that useful results can then be achieved with the existing TRAP crosslinker after reaction scale-up, retagging of the proteins, and expression of RNA polymerase in SlyD deletion mutants. Parallel efforts will aim to enhance crosslinking efficiencies through the synthesis of new targeted reversible crosslinkers based on the multiuse affinity probes AsCy3,[57] ReAsH,[16] as well as perfluorinated,[58] and spirolactam biarsenicals,[59] which due to their shifted fluorescence will be less-efficient quenchers for benzophenone. Azide and nitrobenzene-based crosslinking units will be substituted for benzophenone in an effort to reduce the in vivo photoactivation times.

Experimental Section

Materials and equipment: All reagents and solvents were obtained from Sigma–Aldrich at the highest purity available unless stated otherwise. The synthetic MLCK peptide corresponding to the smooth-muscle myosin light-chain kinase calmodulin-binding sequence with a C-terminal extension of 7 amino acids containing a tetracysteine motif, ARRKWQKTGHAVRAIGRLSSA**CCKACCA**, was custom-synthesized by CPC Scientific (San Jose, CA, USA). The antibodies against α and β subunits of RNA polymerase were products of Neoclone (Madison, WI, USA). Ni-NTA Superflow affinity resin was purchased from Qiagen (Valencia, CA, USA). Gels were Invitrogen 4–12% Bis–Tris gels run with MES buffer according to manufacturer's protocol, unless mentioned otherwise. NMR spectra were obtained with a Varian 500 MHz instrument (Palo Alto, CA, USA). LC–ESI-MS was performed on the Agilent 1100 series with the Agilent MSD SL detector (Santa Clara, CA, USA), LC–MS/MS data were taken on an LTQ Orbitrap (Thermo Scientific, Waltham, MA, USA), and MALDI-MS was performed on the Bruker Autoflex II (Billerica, MA, USA). Photolysis was carried out in the Rayonet RPR-200 photoreactor (Southern New England Ultra Violet Co., Branford, CT, USA) equipped with 16 UV lamps (350 nm).

Synthesis of TRAP

1) *TRAP precursor* ***2***: 5-(2-(4-Benzoylbenzamido)ethylcarbamoyl)-fluorescein: 5-carboxyfluorescein (50 mg, 0.13 mmol) and EDC·HCl (50 mg, 0.26 mmol) were dissolved in dry DMF (2 mL), and Et_3N (2 drops) and *N*-hydroxysuccinimide (15 mg, 0.13 mmol) were added. The mixture was stirred at room temperature for 30 min before the addition of *N*-Boc-ethylenediamine (20 μL, 0.13 mmol)

and allowed to react overnight. The desired product was eluted from a silica gel column by EtOAc and concentrated to give a yellow liquid, which was dissolved in TFA in CH_2Cl_2 (20%, 5 mL total volume) and stirred at room temperature for 2 h to remove the Boc protecting group. The deprotected product (intermediate **1**) was obtained after evaporation of the reaction mixture to dryness and used directly in the next step without further purification.

4-Benzoylbenzoic acid (35 mg, 0.15 mmol) was dissolved in DMF (1.5 mL), and EDC·HCl (50 mg, 0.13 mmol), *N*-hydroxysuccinimide (15 mg, 0.13 mmol), and iPr_2EtN (40 μL, 0.23 mmol) were added. The mixture was stirred at room temperature for 30 min before the addition of the solution of intermediate **1** in DMF. After overnight reaction, the mixture was partially evaporated and H_2O (5 mL) was added to produce an orange-red precipitate, which was filtered and purified by flash chromatography (1:9 to 1:4 methanol/CH_2Cl_2) to give the desired product (28 mg, 34% total yield). ^{1}H NMR (500 MHz, CD_3OD): $\delta = 4.28$ (t, $J = 6.7$ Hz, 4H), 6.56 (dd, $J = 9.0$, 2.2 Hz, 2H), 6.68 (d, $J = 2.2$ Hz, 2H), 6.77 (d, $J = 9.0$ Hz, 2H), 7.31 (d, $J = 7.6$ Hz, 1H), 7.54 (t, $J = 7.7$ Hz, 2H), 7.62 (dd, $J = 3.3$, 5.6 Hz, 2H), 7.66 (t, $J = 7.5$ Hz, 1H), 7.72 (dd, J = Hz, 2H), 7.78 (d, $J = 7.9$ Hz, 2H), 7.84 (d, $J = 8.3$ Hz, 2H), 7.98 (d, $J = 8.3$ Hz, 2H), 8.14 (dd, $J = 7.9$ Hz, 1H), 8.47 ppm (br s, 1H); ^{13}C NMR: ($CDCl_3$): $\delta = 41.09$, 41.11, 103.92, 128.65, 129.80, 130.02, 131.14, 131.21, 132.51, 133.72, 134.31, 138.51, 139.32, 141.52, 169.44, 169.82, 192.99, 197.83, 200.99 ppm; ESI-MS: *m/z* calcd for $C_{37}H_{27}N_2O_8$: 627.18 $[M+H]^+$ (monoisotopic); found: 626.8.

2) Final TRAP synthesis: *5-(2-(4-Benzoylbenzamido)ethylcarbamoyl)-4′,5′-di(1,3,2-dithiarsolan-2-yl)fluorescein:* TRAP precursor **2** (70 mg, 0.11 mmol) and mercury oxide (55 mg, 0.22 mmol) were dissolved in TFA (6 mL) and reacted at 70 °C for 1 h, and then at room temperature for 5 h. The mixture was evaporated under vacuum and then H_2O was added to precipitate the product, which was collected by filtration and washed with H_2O and MeOH. The obtained red solid was thoroughly dried under vacuum and subsequently suspended in *N*-methylpyrrolidinonone (1.5 mL), into which arsenic trichloride (120 μL, 2.8 mmol) was added slowly; this resulted in dissolution of the mercurates and an exothermic reaction. Palladium acetate (a few mg) and iPr_2EtN (100 μL, 1 mmol) were added to the mixture, and the reaction was allowed to proceed at room temperature for 4 h. The mixture was then poured into phosphate buffer (20 mM, pH 7)/acetone (40 mL total volume, 1:1, *v/v*) and 1,2-ethanedithiol (0.4 mL, excess) was added; this produced a rapid precipitation. The slurry was stirred at room temperature for 30 min before extraction with $CHCl_3$ (3×25 mL). The combined organic layers were dried over anhydrous Na_2SO_4, concentrated, and purified by column chromatography (5% methanol in toluene) to yield final product (2 mg). Yield: 2%; ^{1}H NMR (500 MHz, $CDCl_3$): $\delta = 3.60$ (m, 8H), 4.12 (m, 2H), 4.20 (m, 2H), 6.50 (d, 2H), 6.58 (d, 2H), 7.50 (m, 1H), 7.62 (m, 2H), 7.79 (d, 2H), 7.83 (s, 1H), 7.87 (d, 2H), 7.94 (d, 2H), 8.17 (d, 1H), 8.40 (s, 1H), 9.92 ppm (s, 1H); ^{13}C NMR was not obtained in good quality due to lack of sample; MALDI MS: *m/z* calcd for $C_{41}H_{33}As_2N_2O_8S_4$: 958.96 $[M+H]^+$ (monoisotopic); found: 959.09.

Synthesis of 5-(2-(4-benzoylbenzamido)ethyl carbamoyl): 4′,5′-di(1,3,2-dithiarsolan-2-yl)fluorescein

1) 5-(4-Benzoylbenzamido)fluorescein: 4-Benzoylbenzoic acid (0.2 g, 0.88 mmol) was suspended in dry benzene (4 mL), thionyl chloride (0.3 mL, 4.1 mmol) was added dropwise, and the mixture was reacted at reflux for 5 h. The solution was then cooled and evaporated to yield a white solid residue. A solution of 5-aminofluorescein (0.1 g, 0.29 mmol) dissolved in acetone (8 mL) was added to the while residue. The mixture was stirred at room temperature for 2 h and subsequently the formed yellow precipitates were collected by filtration, and washed with acetone and EtOAc. Further purification was carried out by using column chromatography to give the desired compound (0.14 g). Yield: 88%; ^{1}H NMR (500 MHz, CD_3OD): $\delta = 6.56$ (dd , $J = 8.6$ Hz, 2.5 Hz, 2H), 6.67 (d, $J = 8.6$ Hz, 2H), 6.68 (d, $J = 2.5$ Hz, 2H), 7.22 (d, $J = 8.4$ Hz, 1H), 7.57 (t, $J = 7.7$ Hz , 2H), 7.68 (d, $J = 7.7$ Hz, 1H), 7.83 (d, $J = 7.6$ Hz, 2H), 7.92 (d, $J = 8.2$ Hz, 2H), 8.07 (dd, $J = 8.2$ Hz, 1.8 Hz, 1H), 8.14 (d, $J = 8.5$ Hz, 2H), 8.49 ppm (d, $J = 1.8$ Hz, 1H); ^{13}C NMR ($CDCl_3$): $\delta = 103.69$, 129.06, 129.84, 130.02, 130.43, 131.20, 131.25, 132.50, 134.38, 138.48, 139.49, 141.91, 141.93, 168.19, 193.04, 197.76, 200.96, 210.35 ppm; ESI-MS: *m/z* calcd for $C_{34}H_{22}NO_7$: 556.14 $[M+H]^+$ (monoisotopic); found: 556.0.

2) Synthesis of 5-(4-benzoylbenzamido)-4′,5′-di(1,3,2-dithiarsolan-2-yl)-fluorescein: 5-(4-Benzoylbenzamido)fluorescein (60 mg, 0.11 mmol) and mercuric oxide (55 mg, 0.22 mmol) were dissolved in HOAc (2 mL) and TFA (0.5 mL) and reacted at room temperature for 5 h. The mercurated product was precipitated by water and collected by filtration and washed with water and methanol. The obtained red solid was thoroughly dried under vacuum and subsequently suspended in *N*-methylpyrrolidinone (1.5 mL), into which arsenic trichloride (120 μL, 2.8 mmol) was added slowly; this resulted in dissolution of the mercurates and an exothermic reaction. Palladium acetate (a few mg) and iPr_2EtN (100 μL, 1 mmol) were added to the mixture, and the reaction was allowed to proceed at room temperature for 4 h. The mixture was then poured into phosphate buffer (20 mM pH 7)/acetone (total volume 40 mL 1:1, *v/v*), and 1,2-ethanedithiol (0.4 mL, excess) was added; this produced a rapid precipitation. The slurry was stirred at room temperature for 30 min before extraction with $CHCl_3$ (3×25 mL). The combined organic layers were dried over anhydrous Na_2SO_4, concentrated, and purified by column chromatography (30% methanol in toluene) to yield the final product (33 mg). Yield: 34%; ^{1}H NMR (500 MHz, $CDCl_3$): $\delta = 2.34$ (s, OH), 3.56 (m , 8H), 6.49 (d, $J = 8.9$ Hz, 2H, H-2′, 7′), 6.63 (d, $J = 8.9$ Hz, 2H, H-1′, 8′), 7.16 (d, $J = 7.0$ Hz, 1H, H-4), 7.49 (t, $J = 7.4$ Hz, 2H), 7.62 (t, $J = 7.4$ Hz, 1H), 7.79 (d, $J = 8.3$ Hz, 2H), 7.85 (d, $J = 7.8$ Hz, 2H), 8.04 (d, $J = 7.8$ Hz, 2H), 8.26 (br s, 2H, H-1, 5), 8.75 (s, 1H, NH), 9.94 ppm (s, 1H, COOH); ^{13}C NMR ($CDCl_3$): $\delta =$ 43.65, 43.72, 110.43, 112.55, 115.02, 115.20, 124.97, 127.45, 127.68, 128.04, 128.41, 128.72, 129.21, 130.44 (m), 130.84, 131.03, 133.35, 136.88, 137.46, 140.10, 140.84, 147.91, 152.71, 163.13, 165.55, 167.58, 195.99, 196.90 ppm; ESI-MS: *m/z* calcd for $C_{38}H_{28}As_2NO_7S_4^-$: 887.92 $[M+H]^+$ (monoisotopic); found: 887.6.

In vitro crosslinking of MLCK peptide to calmodulin

1) TRAP labeling of peptide: The synthetic peptide corresponding to the calmodulin-binding sequence of smooth-muscle myosin light-chain kinase containing CCKACC was incubated with equimolar amounts of TRAP in DMSO (200 μL) in the presence of tris[2-carboxyethyl] phosphine (TCEP, 2 mM) at room temperature for 2 h. Completion of labeling was confirmed by reversed-phase HPLC (Phenomenex C18 column), by the disappearance of the unlabeled peptide. The labeled peptide was then purified by HPLC.

2) Crosslinking: TRAP-labeled MLCK peptide (150 μM) was incubated with calmodulin (expressed in *E. coli* and purified as previously described)[60,61] at a molar ratio of 1:1 in phosphate buffer (20 mM, pH 7.6), NaCl (150 mM) and either EGTA (0.5 mM) or $CaCl_2$ (0.5 mM) for 30 min at room temperature prior to transfer to a quartz test tube. The solution was subjected to illumination with 350 nm light for 2 h in a Rayonet photoreactor after three cycles of air evacuation and argon purge. Gel electrophoresis and HPLC analysis of the

sample after photoreaction identified the crosslinked product and determined its yield to be around 30%. The crosslinked product was then purified by HPLC on an Agilent Zorbax SBC-18 column (3.5 μm, 4.6×150 mm), on a gradient from MeCN (10%) in TFA (0.1%)/H_2O at 0 min to MeCN (60%) at 25 min The eluted fraction (top of peak at 18.2 min) was then evaporated to nearly dryness, and redissolved in 100 mM ammonium bicarbonate (pH 8.0).

3) Cleavage and transfer of TRAP to calmodulin: Both the purified crosslinked complex and a control sample containing the same amount of calmodulin and MLCK peptide were incubated with dithiothreitol (5 mM) at 60 °C for 30 min prior to trypsin digestion at 37 °C for 3 h. C18 cleanup was carried out according to a procedure described earlier to remove dithiothreitol and undesired contaminants.[26]

Mass spectrometric identification of crosslinking sites on calmodulin: The trypsinized samples were submitted for MS analysis and analyzed as follows: The HPLC system consisted of a custom configuration of 100 mL Isco Model 100DM syringe pumps (Isco, Inc., Lincoln, NE, USA), 2-position Valco valves (Valco Instruments Co., Houston, TX, USA), and a PAL autosampler (Leap Technologies, Carrboro, NC, USA); this permitted fully automated sample analysis across four separate HPLC columns as described previously.[62] Reversed-phase capillary HPLC columns were manufactured in-house by slurry packing 3 μm Jupiter C18 stationary phase (Phenomenex, Torrence, CA, USA) into a 60 cm length of 360 μm o.d.×75 μm i.d. fused silica capillary tubing (Polymicro Technologies Inc., Phoenix, AZ, USA) that incorporated a 0.5 μm retaining screen in a 1/16" custom laser-bored 75 μm i.d. union (screen and union, Valco Instruments Co., Houston, TX, USA; laser bore, Lenox Laser, Glen Arm, MD, USA). The mobile phase consisted of A) HOAc (0.2%) and TFA (0.05%) in H_2O and B) TFA (0.1%) and MeCN (90%) in H_2O (B). The mobile phase was degassed by using an in-line Alltech vacuum degasser (Alltech Associates, Inc., Deerfield, IL, USA). The HPLC system was equilibrated at 10 kpsi with mobile phase A (100%), and then a mobile phase selection valve was switched 50 min after injection, which created a near-exponential gradient as mobile phase B displaced A in a 2.5 mL active mixer. A 30 cm length of 360 μm o.d.×15 μm i.d. fused silica tubing was used to split ~20 μL min^{-1} of flow before it reached the injection valve (5 μL sample loop). The split flow controlled the gradient speed under conditions of constant pressure operation (10 kpsi). Flow through the capillary HPLC column when equilibrated to mobile phase A (100%) was ~400 nL min^{-1}. MS analysis was performed by using a ThermoFinnigan LTQ-Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA) with electrospray ionization (ESI). The HPLC column was coupled to the mass spectrometer by using an in-house-manufactured interface. Homemade 150 μm o.d. × 20 μm i.d. chemically etched electrospray emitters were used.[63] The heated capillary temperature and spray voltage were 200 °C and 2.2 kV, respectively. Data was acquired for 100 min, beginning 65 min after sample injection (15 min into gradient). Orbitrap spectra (AGC 1×106) were collected from *m/z* 400–2000 at a resolution of 100 k followed by data-dependent ion trap MS/MS spectra (AGC 1×104) of the six most abundant ions using a collision energy of 35%. A dynamic exclusion time of 60 s was used to discriminate against previously analyzed ions.

Mascot files were generated from the Thermoelectron RAW output by using the mascot distiller. The resulting MS/MS data was searched 20 times, once for each amino acid, by using OMSSA[38] against FASTA databases containing the complete mouse proteome with +1076.98 u for the mass of the crosslinker and +16 u for methionine oxidation as variable modifications. The data was evaluated by using OMSSAGUI[64] with an E-value cut-off of 0.01 (that is, spectra with an E-value of less than 0.01 were taken into consideration) and individual spectra containing the cleaved crosslinker were further evaluated manually by using OMSSAGUI and theoretical values that were generated by mMass.[65] The data was also searched by using X!tandem (http://ppp.thegpm.org/tandem/ppp.html), and the mouse database choice is given in the online version.

Crosslinking of SlyD: Wild-type *Shewanella oneidensis* MR-1 containing a vector conferring kanamycin resistance was grown in Luria–Bertani media at 30 °C overnight under exclusion of air to a cell density of at least 1 as measured by absorbance at 600 nm. Subsequently, cells were lysed in phosphate-buffered saline (pH 7.5) with β-mercaptoethanol (BME, 5 mM) and tris(2-carboxyethyl)phosphine (TCEP, 1 mM) by using 3×1 min sonication cycles and the resulting lysate (200 μL) was incubated with TRAP (20 μM) at room temperature for 3 h prior to purging with argon. Photoactivation was performed by 1 h of illumination in the Rayonet photoreactor. Mass-spectrometric identification of SlyD binding partners was performed by a parallel pulldown experiment as previously described.[26]

In vivo phospholamban crosslinking: Tetracysteine mutants of phospholamban (1 A and 1 B) were kindly provided by Seth Robia (University of Minnesota) in plasmid pVL1393.[52] The genes were excised by restriction digestions with BamHI and NotI (New England Biolabs) and then gel-purified by using the UltraClean™ 15 DNA purification system (MoBio Laboratories, Inc., Carlsbad, CA, USA). The DNA was ligated into the linearized retroviral vector, pBM-IRESpuro[66] by using T4 DNA ligase, followed by transformation of the ligation mixture into subcloning efficiency *E. coli* DH5α (Invitrogen) and selection of clones on LB agar plates containing ampicillin (100 μg mL^{-1}). After isolating the plasmid DNA of several clones by using the Qiaprep spin miniprep kit (Qiagen), positive clones were identified by the appearance of insert bands after restriction digestion with BamHI and NotI. Those clones identified as having inserts were sequenced by ACGT, Inc. (Northbrook, IL, USA) and were designated pBM-PLB1A (6356 bp) and pBM-PLB1B (6356 bp). Phoenix cells (ATCC SD3443)[66] were grown in Dulbecco's modified Eagle's medium (DMEM) that was supplemented with fetal bovine serum (5%, Gibco), defined/supplemented bovine calf serum (5%, HyClone, Logan, UT, USA), L-glutamine (2 mM, Gibco) and antibiotic (100 units mL^{-1} penicillin and 100 μg mL^{-1} streptomycin) at 37 °C in an atmosphere of CO_2 (5%). The day prior to transfection, Phoenix cells (2.0×10^6) were used to seed a 60 mm plate, which was incubated overnight at 37 °C. After 16–20 h, the medium was removed and replaced with fresh medium (3 mL), to which chloroquine (25 mM, 10 μL) was added. To transfect the cells, a solution of Hank's Balanced Salt solution (1 mL) containing plasmid DNA (8 μg, pBM-PLB1 A), and $CaCl_2$ (125 mM) was added to one plate of cells, and the plate was incubated for 7 h at 37 °C, after which the medium was replaced with supplemented DMEM. The transfected Phoenix cells were incubated for 24 h, then transferred to a 100 mm dish and grown until 70–80% confluent. The medium was replaced fresh medium (5 mL) and the plate incubated overnight at 32 °C. After approximately 16 h, the medium that contained the virus was collected, spun at 1200 rpm in a swinging-bucket centrifuge for 5 min, and the viral supernatant (4 mL) was added to the C2C12 myocyte target cells that had been washed once in PBS. To improve the transduction efficiency, polybrene (4 μg mL^{-1}) was added to the viral supernatant. After a 3 h incubation at 32 °C, the viral supernatant that contained polybrene was removed, replaced with C2C12 medium, and

incubated at 37 °C until the cells were approximately 80% confluent (typically 1 d). The cells were then transferred to a 100 mm dish, incubated for 24 h, and put under selection by using C2C12 medium that contained puromycin (2 μg mL^{-1}). Those cells surviving the selection were maintained in complete growth medium with puromycin at 37 °C.

For the crosslinking experiment, C2C12 myocytes expressing tetracysteine-tagged phospholamban 1B, in which amino acids 23, 24, 27, and 28 are mutated to cysteines, were cultured for 5 d at 37 °C and CO_2 (5%) in Dulbecco's Modified Eagle Medium (DMEM) that was supplemented with fetal bovine serum (10% Gibco), penicillin/streptomycin, sodium pyruvate (1 mM), and L-glutamine (4 mM) prior to the replacement with differentiation medium (DMEM, 2% horse serum, penicillin/streptomycin, 1 mM sodium pyruvate, and 4 mM L-glutamine) as described previously.[21] Cells were grown in standard growth or differentiation media to a density of about 90%. TRAP (2 μM) was added in sterile DPBS (5 mL, 20 mM phosphate-buffered saline, pH 7.5) that had been prewarmed to 37 °C (for this amount of cells 5–10 mL growth media are normal) and replaced into the incubator for 2 h. The crosslinking solution was replaced with DPBS (50 mL), followed by 350 nm illumination in the Rayonet photoreactor for four cycles of 30 min each in phosphate-buffered saline (20 mM, pH 7.5). Cells were lysed for immunoblotting. Alternatively, in separate affinity pull-down experiments, FlAsH-tagged phospholamban and associated proteins in myocytes were enriched by using FlAsH immobilized on glass beads for mass spectrometric identification, as previously described for RNA polymerase.[26]

Crosslinking of RNA polymerase alpha in vivo: A vector that contained a tetracysteine tag (CCPGCC) as well as V5 and His_6-tags at the C terminus of the RNA polymerase α subunit (RpoA-C4) was constructed and transformed into *Shewanella oneidensis* MR-1 as described previously.[26,28] *S. oneidensis* was grown in LB media at 30 °C to a density of 1.2 at 600 nm under conditions ensuring good aeration. Cells were incubated for 1 h with either FlAsH-EDT_2 or TRAP (10 μM), followed by two sequential low-speed centrifugation steps in PBS buffer (20 mM phosphate-buffered saline, pH 7.8) in the presence of 1) DMPS (1 mM) and then 2) DMSO (10%, *v/v*) to remove excess TRAP. Photocrosslinking was performed in PBS at 350 nm by using the Rayonet photoreactor for 1, 2, and 4 h. A set volume was then replaced into LB media, and the growth rate determined. The cells were lysed in SDS-PAGE running buffer and separated on SDS-PAGE.

Abbreviations: BME: β-mercaptoethanol; Boc: *t*-butoxycarbonyl; CaM: calmodulin; DMEM: Dulbecco's modified Eagle media; DMF: *N,N*-dimethylformamide; DMPS: 2,3-dimercaptopropylsulfonate; DMSO: dimethylsulfoxide; EDC: *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide; EDT: ethanedithiol; FlAsH-EDT_2: 4′,5′-bis(1,3,2-dithoarsolan-2-yl)fluorescein; EGTA: ethylene glycol-bis(2-aminoethylether)-*N,N,N*′,*N*′-tetraacetic acid; HEPES: 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; HPLC: high pressure liquid chromatography; MES: 4-morpholinethanesulfonic acid; MLCK: CaM-binding sequence of myosin light-chain kinase containing a CCKACCA tag at the C terminus; MS: mass spectrometry; NHS: *N*-hydroxysuccinimide; NMP: *N*-methyl-2-pyrrolidinone; OMSSA: open mass spectrometry search algorithm; *S. oneidensis*: *Shewanella oneidensis* MR-1; SDS-PAGE: sodium dodecylsufate-polyacrylamide gel electrophoresis; SlyD: protein sensitive to lysis by bacteriophage D, a bacterial peptidyl–prolyl isomerase and metallochaperone; TCEP: tris(carboxyethyl)phosphine; TFA: trifluoroacetic acid; TRAP: targeted releasable affinity probe, that is, 5-(2-(4-benzoylbenzamido)ethyl carbamoyl)-4′,5′-di(1,3,2-dithiarsolan-2-yl)fluorescein.

Acknowledgements

This work was supported by the Genomics: GTL Program of the US Department of Energy Office of Biological and Environmental Research. Pacific Northwest National Laboratory is operated by Battelle for the DOE under Contract DE-AC05–76 L0 1830. We thank Seth Robia for providing phospholamban clones and helpful discussions. We thank Janani Shutthanandan for her help with the myocyte experiments, Karla Thrall and David Wuenschel for mass spectrometer access, and Saiful M. Chowdhury, Joshua N. Adkins and Colette A. Sacksteder for helpful discussions. We thank Ronald J. Moore and the EMSL mass spectrometry team for running the calmodulin mass spectrometric analysis. A portion of the research was performed by using EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research located at Pacific Northwest National Laboratory.

Keywords: arsenic · bioorganic chemistry · fluorescent probes · mass spectrometry · photoaffinity labeling

[1] I. Remy, F. X. Campbell-Valois, S. W. Michnick, *Nat. Protoc.* **2007**, *2*, 2120–2125.
[2] G. Butland, J. M. Peregrin-Alvarez, J. Li, W. Yang, X. Yang, V. Canadien, A. Starostine, D. Richards, B. Beattie, N. Krogan, M. Davey, J. Parkinson, J. Greenblatt, A. Emili, *Nature* **2005**, *433*, 531–537.
[3] A. M. Edwards, B. Kus, R. Jansen, D. Greenbaum, J. Greenblatt, M. Gerstein, *Trends Genet.* **2002**, *18*, 529–536.
[4] L. R. Jones, R. L. Cornea, Z. Chen, *J. Biol. Chem.* **2002**, *277*, 28319–28329.
[5] C. A. Studdert, J. S. Parkinson, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15623–15628.
[6] C. Tagwerker, K. Flick, M. Cui, C. Guerrero, Y. Dou, B. Auer, P. Baldi, L. Huang, P. Kaiser, *Mol. Cell. Proteomics* **2006**, *5*, 737–748.
[7] M. Suchanek, A. Radzikowska, C. Thiele, *Nat. Methods* **2005**, *2*, 261–267.
[8] G. A. Anderson, N. Tolic, X. T. Tang, C. X. Zheng, J. E. Bruce, *J. Proteome Res.* **2007**, *6*, 3412–3421.
[9] O. W. Nadeau, G. J. Wyckoff, J. E. Paschall, A. Artigues, J. Sage, M. T. Villar, G. M. Carlson, *Mol. Cell. Proteomics* **2008**, *7*, 739–749.
[10] O. Rinner, J. Seebacher, T. Walzthoeni, L. Mueller, M. Beck, A. Schmidt, M. Mueller, R. Aebersold, *Nat. Methods* **2008**, *5*, 315–318.
[11] J. W. Chin, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11020–11024.
[12] J. W. Chin, P. G. Schultz, *ChemBioChem* **2002**, *3*, 1135–1137.
[13] N. Mohibullah, S. Hahn, *Genes Dev.* **2008**, *22*, 2994–3006.
[14] G. Dorman, G. D. Prestwich, *Biochemistry* **1994**, *33*, 5661–5673.
[15] Y. Yao, J. Gao, T. C. Squier, *Biochemistry* **1996**, *35*, 12015–12028.
[16] S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis, R. Y. Tsien, *J. Am. Chem. Soc.* **2002**, *124*, 6063–6076.
[17] B. A. Griffin, S. R. Adams, J. Jones, R. Y. Tsien, *Methods Enzymol.* **2000**, *327*, 565–578.
[18] B. A. Griffin, S. R. Adams, R. Y. Tsien, *Science* **1998**, *281*, 269–272.
[19] B. Chen, H. Cao, P. Yan, M. U. Mayer, T. C. Squier, *Bioconjugate Chem.* **2007**, *18*, 1259–1265.
[20] G. Gaietta, T. J. Deerinck, S. R. Adams, J. Bouwer, O. Tour, D. W. Laird, G. E. Sosinsky, R. Y. Tsien, M. H. Ellisman, *Science* **2002**, *296*, 503–507.
[21] D. L. Stenoien, T. V. Knyushko, M. P. Londono, L. K. Opresko, M. U. Mayer, S. T. Brady, T. C. Squier, D. J. Bigelow, *Am. J. Physiol. Cell. Physiol.* **2007**, *292*, C2084–2094.
[22] J. Enninga, J. Mounier, P. Sansonetti, G. T. Van Nhieu, *Nat. Methods* **2005**, *2*, 959–965.
[23] L. Rudner, S. Nydegger, L. V. Coren, K. Nagashima, M. Thali, D. E. Ott, *J. Virol.* **2005**, *79*, 4055–4065.
[24] M. Andresen, R. Schmitz-Salue, S. Jakobs, *Mol. Biol. Cell* **2004**, *15*, 5616–5622.

[25] B. Chen, D. F. Lowry, M. U. Mayer, T. C. Squier, *Biochemistry* **2008**, *47*, 9220–9226.
[26] S. Verma, Y. Xiong, M. U. Mayer, T. C. Squier, *Biochemistry* **2007**, *46*, 3023–3035.
[27] G. Feldman, R. Bogoev, J. Shevirov, A. Sartiel, L. Margalit, *Electrophoresis* **2004**, *25*, 2447–2451.
[28] M. U. Mayer, L. Shi, T. C. Squier, *Mol. BioSyst.* **2005**, *1*, 53–56.
[29] B. Liu, C. T. Archer, L. Burdine, T. G. Gillette, T. Kodadek, *J. Am. Chem. Soc.* **2007**, *129*, 12348–12349.
[30] H. Cao, B. Chen, T. C. Squier, M. U. Mayer, *Chem. Commun.* **2006**, 2601–2603.
[31] W. E. Meador, A. R. Means, F. A. Quiocho, *Science* **1992**, *257*, 1251–1255.
[32] M. Afshar, L. S. D. Caves, L. Guimard, R. E. Hubbard, B. Calas, G. Grassy, J. Haiech, *J. Mol. Biol.* **1994**, *244*, 554–571.
[33] T. Wang, P. Yan, T. C. Squier, M. U. Mayer, *ChemBioChem* **2007**, *8*, 1937–1940.
[34] S. C. Alley, M. A. Trakselis, M. U. Mayer, F. T. Ishmael, A. D. Jones, S. J. Benkovic, *J. Biol. Chem.* **2001**, *276*, 39340–39349.
[35] D. M. Schulz, C. Ihling, G. M. Clore, A. Sinz, *Biochemistry* **2004**, *43*, 4703–4715.
[36] A. Sinz, S. Kalkhof, C. Ihling, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1921–1931.
[37] S. Kalkhof, C. Ihling, K. Mechtler, A. Sinz, *Anal. Chem.* **2005**, *77*, 495–503.
[38] L. Y. Geer, S. P. Markey, J. A. Kowalak, L. Wagner, M. Xu, D. M. Maynard, X. Yang, W. Shi, S. H. Bryant, *J. Proteome Res.* **2004**, *3*, 958–964.
[39] A. Barth, S. R. Martin, P. M. Bayley, *J. Biol. Chem.* **1998**, *273*, 2174–2183.
[40] R. Craig, R. C. Beavis, *Bioinformatics* **2004**, *20*, 1466–1467.
[41] T. G. Bernhardt, W. D. Roof, R. Young, *Mol. Microbiol.* **2002**, *45*, 99–108.
[42] S. Hottenrott, T. Schumann, A. Pluckthun, G. Fischer, J. U. Rahfeld, *J. Biol. Chem.* **1997**, *272*, 15697–15701.
[43] J. W. Zhang, G. Butland, J. F. Greenblatt, A. Emili, D. B. Zamble, *J. Biol. Chem.* **2005**, *280*, 4360–4366.
[44] J. W. Zhang, M. R. Leach, D. B. Zamble, *J. Bacteriol.* **2007**, *189*, 7942–7944.
[45] M. R. Leach, J. W. Zhang, D. B. Zamble, *J. Biol. Chem.* **2007**, *282*, 16177–16186.
[46] D. Voet, J. G. Voet, *Biochemistry*, John Wiley & Sons, New York, **1995**.
[47] K. H. Nealson, D. Saffarini, *Annu. Rev. Microbiol.* **1994**, *48*, 311–343.
[48] K. S. Murakami, S. Masuda, E. A. Campbell, O. Muzzin, S. A. Darst, *Science* **2002**, *296*, 1285–1290.
[49] K. S. Murakami, S. Masuda, S. A. Darst, *Science* **2002**, *296*, 1280–1284.
[50] K. Murakami, M. Kimura, J. T. Owens, C. F. Meares, A. Ishihama, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 1709–1714.
[51] Y. H. Jeon, T. Yamazaki, T. Otomo, A. Ishihama, Y. Kyogoku, *J. Mol. Biol.* **1997**, *267*, 953–962.
[52] S. L. Robia, N. C. Flohr, D. D. Thomas, *Biochemistry* **2005**, *44*, 4302–4311.
[53] Q. Yao, J. L. Bevan, R. F. Weaver, D. J. Bigelow, *Protein Expression Purif.* **1996**, *8*, 463–468.
[54] Y. Mao, J. E. Schwarzbauer, *Matrix Biol.* **2005**, *24*, 389–399.
[55] A. Sinz, *Mass Spectrom. Rev.* **2006**, *25*, 663–682.
[56] M. A. Trakselis, S. C. Alley, F. T. Ishmael, *Bioconjugate Chem.* **2005**, *16*, 741–750.
[57] H. Cao, Y. Xiong, T. Wang, B. Chen, T. C. Squier, M. U. Mayer, *J. Am. Chem. Soc.* **2007**, *129*, 8672–8673.
[58] C. C. Spagnuolo, R. J. Vermeij, E. A. Jares-Erijman, *J. Am. Chem. Soc.* **2006**, *128*, 12040–12041.
[59] A. K. Bhunia, S. C. Miller, *ChemBioChem* **2007**, *8*, 1642–1645.
[60] G. M. Strasburg, M. Hogan, W. Birmachu, D. D. Thomas, C. F. Louis, *J. Biol. Chem.* **1988**, *263*, 542–548.
[61] C. B. Boschek, H. Sun, D. J. Bigelow, T. C. Squier, *Biochemistry* **2008**, *47*, 1640–1651.
[62] E. A. Livesay, K. Tang, B. K. Taylor, M. A. Buschbach, D. F. Hopkins, B. L. LaMarche, R. Zhao, Y. Shen, D. J. Orton, R. J. Moore, R. T. Kelly, H. R. Udseth, R. D. Smith, *Anal. Chem.* **2008**, *80*, 294–302.
[63] R. T. Kelly, J. S. Page, Q. Luo, R. J. Moore, D. J. Orton, K. Tang, R. D. Smith, *Anal. Chem.* **2006**, *78*, 7796–7801.
[64] R. Tharakan, L. Martens, J. E. Van Eyk, D. R. Graham, *Proteomics* **2008**, *8*, 2376–2378.
[65] M. Strohalm, M. Hassman, B. Kosata, M. Kodicek, *Rapid Commun. Mass Spectrom.* **2008**, *22*, 905–908.
[66] Y. Hitoshi, J. Lorens, S. I. Kitada, J. Fisher, M. LaBarge, H. Z. Ring, U. Francke, J. C. Reed, S. Kinoshita, G. P. Nolan, *Immunity* **1998**, *8*, 461–471.

Received: January 19, 2009
Published online on May 13, 2009

