A method for site-specific labeling of multiple protein thiols

Johanna M. Kuiper, Radek Pluta, Wim H. C. Huibers, Fabrizia Fusetti, Eric R. Geertsma, and Bert Poolman*

Department of Biochemistry, Groningen Biomolecular Science and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

Received 8 August 2008; Revised 23 December 2008; Accepted 5 February 2009

DOI: 10.1002/pro.113

Published online 16 March 2009 proteinscience.org

Abstract: We present a generic method for the site-specific and differential labeling of multiple cysteine residues in one protein. Phenyl arsenic oxide has been employed as a protecting group of two closely spaced thiols, allowing first labeling of a single thiol. Subsequently, the protecting group is removed, making available a reactive dithiol site for labeling with a second probe. For proof-of-principle, single and triple Cys mutants of the sulphate binding protein of an ABC transporter were constructed. The closely spaced thiols were engineered on the basis of the crystal structure of the protein and placed in different types of secondary structure elements and at different spacing. We show that phenyl arsenic oxide is a good protecting group for thiols spaced 6.3–7.3 Å. Proteins were labeled with two different fluorescent labels and the labeling ratios were determined with UV-Vis spectroscopy and MALDI-ToF mass spectrometry. The average labeling efficiency was ~80% for the single thiol and 65–90% for the dithiol site.

Keywords: covalent modification; thiol chemistry; sulphate-binding protein; arsine oxide derivatives; fluorescent labeling

Introduction

Functionalization of proteins is used for the introduction of fluorescent and other spectroscopic tags,¹ for the coupling of cofactors for catalysis,² for electrochemical detection of ligand binding,³ and for the immobilization of proteins on solid surfaces.⁴ Several methods for chemical modification are available, whereby the functionality is introduced to the protein in either a covalent or noncovalent manner. Methods have been reported in which a fluorescent tag is com-

plexed to the protein via a metal-ion chelating moiety, which binds reversibly to a polyhistidine sequence engineered at either the amino- or carboxyl-terminus of the protein.^{5,6} A covalent bond can be formed via a reaction between an activated ester of the tag and a primary amine of the protein. However, proteins generally contain several amine groups and labeling at multiple sites will occur. A more specific method involves the reaction between a unique cysteine residue of the protein and a maleimide functionalized tag. Under physiological conditions, the thiol moiety of cysteine is generally the most reactive site in a protein and, importantly, cysteines are relatively rare and often proteins can be engineered to contain only one or a few cysteine residues.

However, in several occasions, it is desirable to create multiple, different covalent modifications in a protein. For instance, one functionality may serve as a detection aid whereas a second one may be needed for immobilization of the protein. Furthermore, labeling with two different fluorescent tags can yield a couple with fluorescence resonance energy transfer (FRET),

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: The European Membrane Protein Consortium; Grant number: E-MeP, 504601; Grant sponsors: Zernike Institute for Advanced Materials, Nanoned, and The Netherlands Proteomics Centre.

*Correspondence to: Bert Poolman, Department of Biochemistry, Groningen Biomolecular Science and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands. E-mail: b.poolman@rug.nl

Figure 1. Principle of labeling with two different probes, using phenylarsine oxide (PhAsO) as a protecting agent. Procedure: (A) Protection of dithiol by PhAsO; (B) labeling of single thiol; (C) removal of protecting group by DTT; (D) labeling of dithiol site, either with thiol (dark green) or dithiol-directed (light green) functionalities.

allowing protein dynamics to be studied either in ensemble or at the single-molecule level. Most of the published methods for the preparation of a protein molecule with a donor and an acceptor fluorophore require additional purification steps to isolate the correctly labeled protein from species with undesired (non) labeling (Refs. 7,8, and references cited therein). Recently, Smith *et al.*⁸ published a method to label independently three different sites, each with a unique fluorophore. The same method was used to first couple a fluorescent tag and then immobilize the labeled protein on a surface containing maleimide residues.⁹ In both studies, a zinc-finger domain was fused to a single cysteine-containing protein. The zinc-finger domain is a 32 residue polypeptide with a Cys₂His₂ primary coordination sphere that binds Zn²⁺ reversibly and with high affinity. Zn²⁺ binding protects the zinc-finger domain from labeling with thiol-specific reagents, which react with Cys residue(s) in the protein of interest. Subsequently, the Zn²⁺ is removed by treatment with ethylenediamine tetraacetic acid (EDTA), allowing labeling of the zinc-finger domain. A third position for labeling was created using the truncated version of the zinc-finger domain, in which the α -helix containing the two histidines was deleted. This truncated version was fused to the other terminus of the protein of interest and still contained the dithiol that readily oxidized to a disulfide. This dithiol was protected against labeling via formation of the disulfide bond.

Here, we present an alternative method to specifically modify a protein with two different labels (Fig. 1). The protein is engineered such that it contains one isolated and two vicinal thiols. In the procedure, the closely spaced thiols are protected by a reaction with phenylarsine oxide (PhAsO). After labeling the single, free thiol, the protecting group is removed by treatment with a reducing agent and the resulting reduced thiols are now accessible for labeling with a second fluorophore. PhAsO complexes specifically with closely spaced dithiols and forms a high affinity ring structure.¹⁰ The cyclic complexes formed with dithiols are markedly more stable than the noncyclic compounds

formed with monothiol due to entropic considerations.¹¹ The dissociation constant for the arsenite complex of dithrothreitol (DTT) was found to be 0.33 μ M.¹² An advantage of this method is that no extra (flexible) domains are introduced into the protein and, second, the positions of the thiols can be chosen strategically, especially if a crystal structure of the protein is available.

For proof-of-principle of the method, we addressed the sulphate-binding protein (SBP) from *Salmonella typhimurium*. SBP is a periplasmic receptor protein, which is an essential component of ABC transporters involved in solute uptake.¹³ SBP and other solute-binding proteins are known for their high stability and consist of two domains linked by a hinge region where the ligand is bound. A high-resolution crystal structure for SBP is available.¹⁴ The wild-type SBP is devoid of cysteine residues, and a set of mutants was constructed containing an isolated cysteine and a site with two vicinal cysteines.

Results and Discussion

Overexpression, purification, and ligand binding activity of SBP

The gene encoding SBP from *Salmonella typhimurium* was expressed in *Escherichia coli* under the control of the *P_{BAD}* promoter. A sequence encoding a tobacco etch virus (TEV) protease cleavage site, followed by a deca-His tag, was added to the 3' end of the gene. As amplified expression of SBP with its signal sequence resulted in incomplete processing of the signal sequence (Supporting Information Figs. A and B), we removed the signal sequence coding region and produced mature SBP in the cytoplasm of *E. coli* with an approximate yield of 6 mg SBP per liter of cell culture; the expression conditions are presented in Supporting Information Fig. C). Using one-step metal affinity chromatography, the protein was obtained at >99% purity.

The mass of the mature protein was determined by MALDI-*Tof* mass spectrometry (Supporting Information Figs. D and E). The mature protein had an

Table I. Engineered Sulphate Binding Proteins, Their Protein Yields from 1 L of Cell Culture, and the K_D Values for Sulphate (the Standard Deviations are Given)

Protein	Amino acid substitution	Yield (mg of protein per liter)	K_D (μM)
WT	none	7.0	0.29 ± 0.06
ST	G289C	5.7	0.18 ± 0.01
RP1	Q20C, K23C, G289C	4.3	0.12 ± 0.02
RP2	E19C, K23C, G289C	4.8	0.14 ± 0.01
RP3	Q20C, I255C, G289C	5.7	0.17 ± 0.03
RP5	V38C, R40C, G289C	8.6	0.39 ± 0.01

observed mass of 37,246 Da, which is 133 Da less than the predicted mass, most probably resulting from the removal of the N-terminal methionine (mass of 131 Da). It is known that many bacterial proteins are subject to post-translational modification in which a methionine, in fact *formyl*-methionine, is excised.¹⁵ Moreover, we generally observed a second minor peak, which for wild-type SBP corresponded to 36,972 Da (and m/z of 18,488 for the doubly charged (M^{2+}) species, see Supporting Information Fig. D, upper panel). On cleavage of the proteins with TEV protease, the doublet band disappeared and the observed mass was in excellent agreement with the calculated mass of the protein lacking the N-terminal methionine (Supporting Information Fig. D, lower panel). Apparently, there was heterogeneity at the carboxyl-terminus of the expressed protein, presumably due to aberrant translation termination or post-translational processing. MS/MS analysis of tryptic digests of SBP confirmed that a fraction of the protein was synthesized with eight rather than 10 C-terminal histidines (data not shown). The mass of the cleaved protein was 1638 a.m.u. lower than full length SBP with the deca-His tag, which corresponds to the loss of the C-terminal 13 residues. For the most critical experiments, we used the TEV protease-treated monodisperse SBP as described later.

The ligand binding activity of SBP was determined as described by Pardee.¹⁶ Radiation-less energy transfer from excited tryptophan residues to CrO_4^{2-} decreased the fluorescence intensity of SBP by a maximum of 85% (Supporting Information Fig. F, subpanel I). By titrating the protein with increasing amounts of CrO_4^{2-} , the dissociation constant (K_D) of SBP for this substrate analogue could be determined; the K_D for CrO_4^{2-} was 1.6 μM (Supporting Information Fig. F, subpanels I and II). Using saturating amounts of CrO_4^{2-} (150 μM), we titrated the SBP- CrO_4^{2-} complex with sulphate. From the increase in fluorescence (chasing of CrO_4^{2-} ; Supporting Information Fig. F, subpanel III) and the K_D for CrO_4^{2-} , we determined the K_D for the natural substrate sulphate. The K_D for sulphate was about 0.3 μM ; the data are summarized in Table I. Because CrO_4^{2-} was able to lower the protein fluorescence by ~85%, most if not all of the SBP must have been functional in ligand binding.

Engineering of surface-accessible thiol residues

Using the crystal structure of the cysteine-less SBP,¹⁴ we designed and constructed five mutants, each containing a single isolated thiol (G289C) without additional thiols (ST) or combined with two closely spaced thiols (RP1, RP2, RP3, and RP5, Table I; Supporting Information Fig. G). The distance between the closely spaced thiols, as measured from the crystal structure, ranged from 5.3 to 7.3 Å (distance between C_β atoms). Each of the proteins was purified to homogeneity and characterized biochemically (e.g., Supporting Information Fig. H). Importantly, the amino acid substitutions did not negatively affect the binding of sulphate (Table I).

Labeling strategy and properties of dithiol sites

Wild-type and mutant SBP derivatives were first labeled with Oregon green-maleimide in the presence and the absence of the protecting group PhAsO. After binding to the Ni^{2+} -Sephacrose resin, the protein was incubated with 1 mM of DTT to reduce the cysteines. To one sample, five column volumes (CV) of 1 mM of PhAsO were applied; the other sample received five CV of solvent. After incubation, Oregon green was added to both samples to a final concentration of 83 μM . The Oregon green solution was added stepwise to the column (five CV in total). In preliminary experiments, we observed that the highest degrees of labeling were obtained by treating Ni^{2+} -Sephacrose-bound SBP with the Oregon green solution without prior removal of the excess of DTT. After incubation for 60 min at room temperature, the nonreacted label was removed by extensive washing of the resin with 20 mM KPi, pH 7.0., and the proteins were recovered as described in the Materials and Method section. The labeling efficiencies of ST, RP1, RP2, and RP5 are presented in Figure 2. For RP3, virtually complete labeling of the single thiol site was achieved but significant labeling of the two vicinal thiols was not possible. Most likely, the two neighboring thiols in RP3 rapidly oxidized to form a disulfide bond, that is, before significant reaction with the label could take place. Strong indication for disulfide formation was obtained from the differential migration of RP3 on SDS-polyacrylamide gels in the presence and absence of reducing agents; oxidization prevented RP3 from complete unfolding and this protein species migrated faster than the reduced form (Supporting Information Fig. H). In RP3, the vicinal thiols are the closest spaced (distance $C_\beta-C_\beta$ is 5.3 Å). Because the vicinal thiols in RP3 readily oxidized, this protein was not used in further studies.

PhAsO protects dithiols from labeling with maleimides

To demonstrate that PhAsO protects the two vicinal thiols, the small “non-absorbing” thiol-specific reagent *N*-ethylmaleimide (NEM) was used in combination

Figure 2. Labeling of SBP, wild-type and mutant derivatives with Oregon green-maleimide. The degrees of labeling with (white bars) and without (black bars) PhAsO pretreatment are presented. The extent of labeling was estimated from the ratio of the Oregon green to protein as described in the materials and method section. To 0.25 mL bed volume of Ni^{2+} -Sepharose resin was added 50 μL of protein (25–40 μM). For protection of the dithiol, five CV of 1 mM PhAsO in 20 mM KPi, pH 7.0, were subsequently added. The protein was then treated with four CV of 83 μM Oregon green maleimide in 20 mM KPi, pH 7.0. Another CV was added and the labeling reaction was continued for 60 min as described in the Methods section.

with PhAsO. Protein RP1 was either or not pretreated with PhAsO as indicated in Table II. Next, the proteins were either or not incubated with NEM, after which excess maleimide was inactivated by DTT and the reaction of PhAsO was reversed (samples A and C). B and D represent the corresponding control samples. Finally, the proteins were labeled with Oregon green. The two scenarios for labeling of the vicinal thiols by Oregon green are indicated. For condition D, no labeling with Oregon green was observed as expected, indicating that the (irreversible) reaction of the thiols with NEM was complete. For conditions B and C labeling of the single

thiol site (Cys289) and the vicinal thiols was observed. The apparent number of Oregon green labels per SBP approached 2. Importantly, on PhAsO treatment, subsequent labeling with NEM and removal of PhAsO (condition A), the number of Oregon green labels per SBP was approximately one, and half that of the samples B and C. The residual labeling was presumably at the vicinal thiol sites, and Cys289 was blocked by NEM.

MS analysis of singly labeled proteins

To confirm the identity of the labels at the single thiol and dithiol sites, SBP was digested by trypsin and the peptides *w/wo* thiol-specific labels were identified by tandem mass spectrometry (MALDI-ToF MS-MS). Relevant MS and MS/MS spectra demonstrating the identity of the labels at the thiol-289 site are shown in Supporting Information Figure I. When PhAsO-pretreated RP1 was labeled with NEM, a 125.05 Da mass increase was observed for LFTIDEVFC₂₈₉GWAK; labeling with Oregon green resulted in a mass increase of 463.05. After removal of PhAsO and subsequent labeling of SBP with Oregon green, a 926.1 Da mass increase was observed for the tryptic peptide ELYE-C₂₀YNC₂₃AFSAHWK, corresponding to labeling of both Cys20 and Cys23. The sites of labeling were confirmed by MS/MS of the tryptic peptides.

To estimate the degree of labeling, MALDI-ToF mass spectra were recorded of the full length proteins treated according to conditions A, B, C, and D (Table II). The doubly charged proteins (M^{2+}) gave the best resolution in linear mode MS and therefore these m/z values were used for the determination of the degree of labeling. In Table III, the increases of theoretical mass on labeling of RP1 are presented. Because MALDI mass spectra of large proteins lack the high resolution of the MS and MS/MS spectra of the peptides, the quantification of the degrees of labeling are presented as approximate numbers (i.e., rounded off to 5% increments). From the MS analysis, it was clear

Table II. Overview of the Procedure and Observed Labeling Efficiencies for RP1.

Experimental conditions	A	B	C	D
Reduction of thiols with DTT	x	x	x	x
Pretreatment with PhAsO (blocking of dithiols)	x		x	
Incubation with NEM	x			x
Inactivation of NEM and removal of PhAsO by DTT	x		x	
Incubation with Oregon green	x	x	x	x
Labelling efficiency				
Theoretical (1 Oregon green/thiol)	2	3	3	0
Theoretical (1 Oregon green/dithiol)	1	2	2	0
Experimental (Oregon green/SBP)	0.9	1.8	1.8	0

To 0.25 mL bed volume of Ni^{2+} -Sepharose was added 50 μL of RP1 (56 μM). For protection of the dithiol, five CV of 1 mM PhAsO in 20 mM KPi, pH 7.0 were added. For labeling with NEM, five CV of 1 mM NEM in 20 mM KPi, pH 7.0 were subsequently added. Another CV of NEM in KPi was added and the labeling reaction was continued for 60 min. The proteins were then treated with 4 CV of 112 μM Oregon green maleimide in 20 mM KPi, pH 7.0. Another CV Oregon green in KPi was added and the labeling reaction was continued for 60 min as described in the Methods section. x, step carried out to prepare the samples.

Table III. The Expected and Observed Mass Increases Determined by MALDI-ToF Mass Spectrometry

Condition	Observed mass (Da)		Expected labels	Expected mass increase	Observed mass increase
	M ²⁺	M ⁺			
A	n.a.		1 NEM plus 2 OG	1051	n.a.
B	19,327	38,655	3 OG	1389	1431
C	19,320	38,641	3 OG	1389	1417
D	18,800	37,601	3 NEM	375	377

M²⁺ and M⁺ are doubly and singly charged protein species. OG, Oregon green.

that both thiols of the dithiol site of RP1 were labeled with Oregon green.

Dye-dye interactions at the dithiol site

For the wild-type cysteine-less SBP, no significant labeling was observed as expected. In three independent experiments, we determined by absorbance measurements a labeling efficiency of 0.8 ± 0.2 for the ST protein, close to the theoretical maximum of 1 Oregon green per SBP. Pretreatment with PhAsO had no significant effect on the labeling efficiency (see Fig. 2). The RP1, RP2, and RP5 proteins had approximately 0.8 Oregon green per SBP when pretreated with PhAsO, which is in good agreement with the labeling efficiency of the ST protein. Without PhAsO treatment, the labeling degree increased to 1.7–1.8. These numbers are consistent with one label on Cys288 and a second label on either one of the two vicinal thiols. However, the MS analyses indicated that both thiols of the dithiol site could be labeled with Oregon green. It thus seems likely that due to dye-dye interactions of closely spaced labels,¹⁷ the Oregon green concentration was underestimated in the absorbance measurements. In fact, the increased intensity of the shoulder of Oregon green at 474 nm points to dye-dye interactions and thus a lower than expected absorbance at 494 nm (λ_{max} of Oregon green; see Fig. 3). Most importantly, the spectroscopy experiments together with the MS analyses indicate that PhAsO protects the vicinal thiols from labeling without having an effect on the derivatization of Cys288.

Labeling of SBP with two different fluorophores: Oregon green and Alexa 350

In this series of experiments, the WT protein and the mutants were labeled as outlined in Figure 1 with Oregon green as label 1 and Alexa 350 as label 2. First, the closely spaced thiols were protected by treatment with PhAsO. Subsequently, the Cys288 was labeled with Oregon green. After removal of the protecting group by DTT, the resulting free thiols were labeled with Alexa 350. As a control, the wild-type SBP was treated in the same way. In Figure 4, the resulting UV-Vis absorption spectrum of Oregon green-Alexa 350-labeled RP5 protein is presented. Clearly, the absorbance peaks of the protein (~ 280 nm), Alexa 350 (~ 345 nm), and Oregon green (~ 495 nm) can be

observed. In Figure 5, the degrees of labeling, estimated from the UV-vis absorption spectra, are presented. As expected, the wild-type protein was not labeled. The mutants RP1, RP2, and RP3 were labeled with Oregon green for 80–85% at Cys288. For Alexa 350, we estimated by UV-Vis absorbance measurements more than one label per dithiol, indicating that closely spaced Alexa dyes influenced each other less than Oregon green dyes did.

MS analysis of doubly labeled proteins

To obtain an estimate of the number of labels at the closely spaced dithiol site, mass spectra of the proteins were recorded. The overall results are presented in Table IV. The results for the control wild-type protein was as expected, that is, no labeling. For ST, the majority of protein was labeled with one fluorophore. For RP1 and RP5, we could observe peaks corresponding to protein without label, protein with one label, corresponding to Oregon green at Cys289, and protein with two or three labels, corresponding to Oregon green at

Figure 3. UV-Vis Spectrum of RP1 after treatment as described in Table II, procedure A (continuous line); the concentration of Oregon green-labeled protein was 2.7 μM . For comparison (the shoulder at 474 nm in the absorption spectrum is much less pronounced when not bound to the sulphate binding protein), a spectrum of Oregon green incubated with excess mercaptoethanol is also shown (dashed line). For the control sample, Oregon green was first solubilized in DMF (4.2 mM) and then diluted with 20 mM KPi, pH 7.0, and 1 mM mercaptoethanol to a final concentration of 2.5 μM .

Figure 4. UV-Vis Spectrum of RP5 (3.2 μ M of protein in 100 μ M EDTA, 2 mM Tris-HCl, pH 8.0) after labeling with Oregon green and Alexa 350. To 0.25 mL bed volume of Ni^{2+} -Sepharose was added 50 μ L of RP5 (40 μ M). For protection of the dithiol, five CV of 1 mM PhAsO in 20 mM KPi, pH 7.0, were subsequently added. The protein was labeled with 4 CV of 83 μ M Oregon green maleimide in 20 mM KPi, pH 7.0. Another CV of Oregon green in KPi was added and the labeling reaction was continued for 60 min. Subsequently, after deprotection, the protein was treated with four CV of 80 μ M of Alexa Fluor 350 maleimide in 20 mM KPi, pH 7.0. The column was incubated with one extra CV of the Alexa Fluor 350 solution and the labeling reaction was continued for 60 min.

Cys289 and one or two Alexa 350 dyes at the dithiol site (Fig. 6; data not shown). In different independent experiments, we observed $\sim$ 80% labeling of Cys288 (Oregon green) and 65–90% labeling of the dithiol site (Alexa 350).

Resonance energy transfer between Trp and Alexa 350

The Alexa 350 and/or Oregon green-labeled proteins were also assessed by fluorescence spectroscopy. The

Figure 5. Labeling efficiency of wild-type and mutant SBP. Oregon green was the first label and Alexa 350 the second. Values were determined by UV-Vis spectroscopy.

Table IV. The Expected and Experimental Mass Increases and Labeling Efficiencies as Determined by MALDI-Tof mass Spectrometry

Mutant	Expected binding	Expected mass increase	Observed mass increase (% of protein)
WT	No label	0	0 (100%)
ST	No label	0	(15%)
	1 OG	463	472 (85%)
RP1	No label	0	(5%)
	1 OG	463	470 (20%)
	1 OG plus 1 Alexa 350	939	944 (35%)
	1 OG plus 2 Alexa 350	1415	1408 (40%)
RP5	No label	0	(10%)
	1 OG	463	454 (25%)
	1 OG plus 1 Alexa 350	939	952 (35%)
	1 OG plus 2 Alexa 350	1415	1422 (30%)

The fractions of different protein species were estimated from the peak heights within the mass spectra (an example is given in Fig. 6). Given the uncertainty of quantification from MS spectra, measurements from different experiments were used to estimate and summarize the labeling efficiencies and the data were rounded to 5%.

chromophores were excited at their respective wavelengths maxima (283 nm for SBP; 348 nm for Alexa Fluor 350; and 489 nm for Oregon green) and the emission spectra were recorded. Calibration curves were constructed from solutions with known concentrations of free dye. Intrinsic protein fluorescence measurements with excitation at 283 nm and emission at 334 nm suggested concentrations much lower ($\sim$ 50-fold) than estimated from UV/Vis absorbance measurements. In contrast, fluorescence measurements at the emission maximum of Alexa 350 (i.e., 441 nm) and excitation at 348 nm indicated a three-fold higher signal than expected. These observations point toward resonance energy transfer (RET) between excited tryptophans in the protein and Alexa 350. We did not observe significant changes in the RET signal on binding of sulphate to RP1, RP2, and RP5. For the wild-type protein (control, no labels), the fluorescence

Figure 6. MALDI-Tof spectrum of RP1 after TEV protease cleavage and labeling with Oregon green at Cys289 and Alexa 350 at the dithiol site.

intensities in the UV were in accordance with the UV/vis measurements.

Conclusions

We have demonstrated that it is possible to label a protein via thiol chemistry with two different dyes, that is, by using PhAsO as protecting group. Two closely spaced thiols can be protected reversibly by PhAsO. In comparison with the method of De Lorimier *et al.*,⁹ the modifications made to the protein (three single amino acid substitutions) are small and the engineered triple-thiol mutants were fully functional. The SH groups of the dithiol sites, which were approximately 6.3–7.3 Å apart, did not readily form disulfide bridges. We noted that pairs of cysteines that readily form disulfides (RP3) are well protected from labeling with maleimide-derivatives. However, it is difficult to keep such pairs reduced under conditions of labeling with the second fluorophore. On the contrary, two cysteines separated by two (RP1) and three (RP2) residues present in an α -helix did not rapidly reoxidize on removal of DTT. The same holds true for two cysteines separated by one residue in a β -sheet (RP5). The strategy for dual labeling of proteins outlined here and the conditions for spacing of the vicinal thiols should be generally applicable. In case a protein structure is not available, suitable surface-exposed sites might be obtained from secondary structure prediction algorithms.

Methods

Chemicals and reagents

Ni²⁺-Sephacrose was purchased from GE Healthcare, Imidazole from Carl Roth Chemicals (Kahlsruhe, Germany) and Phenylarsine oxide from Sigma-Aldrich (Zwijndrecht, The Netherlands). Oregon green and Alexa Fluor 350, both with a maleimide reactive group, were purchased from Invitrogen (Breda, The Netherlands). NAPTM five columns were purchased from GE Healthcare (Buckinghamshire, UK).

Construction of the mutants

Plasmid pBADMyHisB (Invitrogen) was cut with *Pst*I and *Sal*I, which yielded a vector fragment of 4020 bp and an exit fragment of 72 bp. A linker fragment, containing a *Pst*I and *Xba*I site, followed by a TEV-encoding sequence, a *Spe*I site, a 10-histidine tag-encoding sequence (in frame with the TEV sequence) and finally a *Sal*I site, was ligated into the vector fragment. The linker fragment was obtained by hybridizing two complementary oligonucleotides with the following sequences: 5'-gt cta gag aaa acc tgt att ttc agg gca cta gtc atc atc atc acc atc atc atc acc atc att aag 3' and 3'-acg tca gat ctc ttt tgg aca taa aag tcc cgt gat cag tag tag tag tgg tag tag tag tgg tag taa ttc agc t 5'. The resulting vector was named pERG1.

Table V. Engineered Sulphate Binding Proteins and Plasmid Designations

Mutant	Plasmid	Mutations	Predicted secondary structure of dithiol site
ST1	pJMK288	G289C	
RP1	pRP1	Q20C, K23C, G289C	α -helix
RP2	pRP2	E19C, K23C, G289C	α -helix
RP3	PRP3	Q20C, I255C, G289C	—
RP5	PRP5	V38C, R40C, G289C	β -sheet

The gene encoding the sulphate binding protein (SBP) from *Salmonella typhimurium* LT2 was amplified by PCR, using genomic DNA as template; forward and reverse primers introduced a *Nco*I site (at the 5' end) and a *Xba*I site (at the 3' end), respectively. Following digestion of pERG1 with *Nco*I and *Xba*I, the PCR fragment containing the *sbp* gene was ligated into vector, resulting in pJMK1. To express the *sbp* gene without signal sequence (MKKWGVGFTLLLASTSILA), the first 57 base pairs of *sbp* were removed; the mature SBP without signal sequences started with the following sequence: MAKDI. For convenience, the numbering of the amino acids is the same as used in the crystal structure,¹⁴ that is, *K* is amino acid number 1, *D* is two, *I* is three, etc. Site directed mutagenesis (Stratagene Quik-Change mutagenesis kit) was employed to mutate amino acids to cysteine residues. Point mutations that have been made are as follows: E19C: GAG → TGC; Q20C: CAG → TGC; K23C: AAA → TGT; V38C: GTG → TGC; R40C: CGC → TGC; G289C: GGC → TGC; and I255C: ATA → TGC. Through multiple rounds of mutagenesis, several mutations were combined as summarized in Table V.

Cell growth and protein purification

The final constructs were transformed to *E. coli* MC1061. The bacteria were cultivated at 37°C in Luria Broth under vigorous aeration (shake flasks). The medium was supplemented with 50 µg/mL ampicillin to maintain the plasmids. Protein expression was induced by adding 0.01% (wt/vol) L-(+)-arabinose to the cultures when OD₆₆₀ reached 0.5. Ninety minutes after induction, the cultures were cooled on ice and centrifuged to pellet the cells at 4°C (12,250g, 12 min). All steps during the purification were performed at 4°C, unless indicated otherwise. The cells were washed with 20 mM Tris-HCl, pH 8.0. After centrifugation at 12,250g for 12 min, the pellet was concentrated to an OD₆₆₀ of 40 in 20 mM Tris-HCl, pH 8.0. To this suspension was added 1 mM MgCl₂, RNAase, and DNAase (both ~0.1 mg/mL). Subsequently, the cells were disrupted by a threefold passage through a French pressure cell at 10,000 psi. The cell lysate was centrifuged for 10 min at 10,000g to remove cell debris and unbroken cells, and the supernatant was centrifuged at 184,000g. The cytosolic lysate fraction was aliquoted

(2 mL volumes), rapidly frozen in liquid nitrogen, and stored at -80°C until further use.

For the purification of SBP, 8 mL of cytosolic lysate material was incubated with 1 mL of washed (Millipore water) Ni^{2+} -Sephacrose resin (2 mL slurry) for 1 to 2 h at 4°C with rotation. The resin was washed with six CV of 50 mM Tris-HCl, pH 8.0, 20 mM imidazole and then with 15 CV of 50 mM Tris-HCl, pH 8.0, 50 mM imidazole. The protein was eluted with 50 mM Tris-HCl, pH 8.0, 500 mM imidazole and the purity was checked by SDS-PAGE. To the purified protein sample, Na-EDTA was added to a final concentration of 1 mM. The excess of imidazole was removed on a NAP size-exclusion column. The elution buffer contained 1 mM EDTA, 20 mM Tris-HCl, pH 8.0. Next, the protein solution was dialyzed against a solution of 2 mM Tris-HCl, pH 8.0 plus 1 mM EDTA and supplemented with Dowex X1/2, a strongly basic anion resin, to remove contaminating sulphate. In the second dialysis step, EDTA was omitted from the buffer. Aliquots of the purified protein were flash-frozen in liquid nitrogen and stored at -80°C and proved stable for several months.

When appropriate the protein was cleaved with TEV protease to remove the C-terminal histidine tag. This digest was performed in 50 mM Tris-HCl, pH 8.0, 0.5 mM EDTA, and 1 mM DTT overnight at 4°C . The TEV protease was added in weight ratio of protein to protease of 40:1.

Protein labeling

Labeling in the presence and absence of PhAsO. To the purified SBP, bound to the Ni^{2+} -Sephacrose resin, was added a solution of 1 mM DTT in 20 mM KPi, pH 7.0. The mixture was incubated for 60 min at room temperature with gentle mixing. Next, the resin was washed with four CV of 1 mM DTT in 20 mM KPi, pH 7.0. For protection of the dithiol, five CV of 1 mM phenylarsine oxide (PhAsO) in 20 mM KPi, pH 7.0, were subsequently added. After incubation at room temperature for 15 min, four CV of Oregon green in 20 mM KPi, pH 7.0 were subsequently added. The concentration of label (e.g., Oregon green) was at least 10 times higher than the concentration of SBP. Another CV of Oregon green solution was added and the mixture was incubated for 60 min at room temperature with gentle mixing. Then, unreacted Oregon green was removed by excessive washing (20 CV) with 20 mM KPi, pH 7.0. Arsenite was removed by subsequent addition of 10 CV of 1 mM DTT in 20 mM KPi, pH 7.0. DTT was then removed by excessive washing (20 CV) with 20 mM KPi, pH 7.0. Labeled SBP was eluted as described in the purification section. After elution, EDTA was added to a final concentration of 1 mM. The excess of imidazole and possible unbound dye were removed by chromatography over a NAP column. The elution buffer contained 1 mM EDTA, 20 mM Tris-HCl, pH 8.0.

Labeling with two fluorophores. The initial steps were the same as described in the previous section, that are, the reduction with DTT, protection with PhAsO, labeling with Oregon green and removal of the protecting group. The protein labeled with Oregon green was then treated with 4 CV of Alexa Fluor 350 in 20 mM KPi, pH 7.0. Next, the column resin with bound SBP was incubated with 1 CV of Alexa Fluor 350 solution for 60 min at room temperature and gentle rotation. Unreacted Alexa Fluor 350 was removed by excessive washing with 20 mM KPi, pH 7.0. Labeled SBP was eluted as described in the purification section. After elution, EDTA was added to a final concentration of 1 mM. The excess of imidazole and possible unbound dye were removed by chromatography over a NAP column. The elution buffer contained 1 mM EDTA, 20 mM Tris-HCl, pH 8.0.

Determination of ligand dissociation constants

Purified SBP (stock concentrations of 20–50 μM) was diluted to 0.2–0.5 μM into 2 mM Tris-HCl, pH 8.0 at 20°C ; the buffer solution was pretreated with Dowex X1/2 to remove contaminating SO_4^{2-} . Increasing amounts of chromate were added under continuous stirring and the decrease in fluorescence was recorded on a Fluorolog 3 fluorescence spectrophotometer; the excitation and emission wavelengths were 285 and 325 nm, respectively.¹⁸ To determine the dissociation constant for SO_4^{2-} , chromate was used at a final concentration of 150 μM and the fluorescence increases were recorded on stepwise addition of sulphate. The dissociation constants were determined using the equation of Kragh-Hansen¹⁹: $[\text{SBP-SO}_4^{2-}]/[\text{SBP-CrO}_4^{2-}] = K_D(\text{CrO}_4^{2-})/K_D(\text{SO}_4^{2-}) \times [\text{SO}_4^{2-}]_{\text{tot}}/[\text{CrO}_4^{2-}]_{\text{tot}}$. $[\text{SBP-SO}_4^{2-}]$ refers to the concentration of SBP with bound SO_4^{2-} , $[\text{SBP-CrO}_4^{2-}]$ to the concentration of SBP with bound CrO_4^{2-} , $[\text{SO}_4^{2-}]_{\text{tot}}$ to the total concentration of SO_4^{2-} , and $[\text{CrO}_4^{2-}]_{\text{tot}}$ to the total concentration of CrO_4^{2-} .

Determination of the fluorophore labeling stoichiometry

For labeling efficiencies, UV-Vis absorbance spectra were recorded on a Cary 100 UV/Vis spectrophotometer. The percentage of labeling was determined, using the molecular absorption coefficients shown in Table VI and making the appropriate corrections. The molecular absorption coefficient for SBP was calculated according to Pace *et al.*²⁰ The molecular absorption

Table VI. Molecular Absorption Coefficients

	ϵ (280 nm) $\text{M}^{-1} \text{cm}^{-1}$	ϵ (345 nm) $\text{M}^{-1} \text{cm}^{-1}$	ϵ (496 nm) $\text{M}^{-1} \text{cm}^{-1}$
Oregon green	11,550	5780	82,500
Alexa Fluor 350	4180	19,000	0
SBP	57,000	0	0

coefficients for Alexa Fluor 350 and Oregon green were obtained experimentally.

Mass spectrometry

The protein samples were further purified on C8 Top Tip 10 μ L (Glygen Corp., Columbia, USA). Mass spectra were recorded on an Applied Biosystems 4700 proteomics analyzer and standard protein samples (bovine serum albumin; yielding multiply charged species at m/z values of 66,431, 33,216, 22,144, and 16,608) were used for (internal) calibration of the spectra. Protein samples were spotted on a stainless steel MALDI target with α -cyano-4-hydroxy cinnamic acid at a final concentration of 2.5 mg/mL.²¹ For the full length proteins, the MALDI-*Tof* was operated in linear positive ionization mode.

To confirm the nature of the labels on the protein, purified SBP was subjected to digestion for 6 h at 37°C with trypsin (Trypsin Gold, MS-grade; Promega, Leiden, The Netherlands) in 0.04M ammonium bicarbonate, pH 8.0, containing 10% acetonitrile; the trypsin to SBP ratio was 1:50. The reaction was stopped by addition of trifluoroacetic acid (TFA) to a final concentration of 0.8% (v/v). Subsequently, the peptide digests were mixed 1:1 (v/v) with a solution of α -cyano-4-hydroxycinnamic acid matrix (5 mg/mL, LaserBio Labs, Sophia-Antipolis, France), spotted directly on a stainless steel MALDI target and analyzed using a 4800 Proteomics Analyzer MALDI-*Tof/Tof* mass spectrometer (Applied Biosystems, Foster City, CA). The laser has a constant output of 20 μ J and on average 4000 laser shots were used to acquire the mass spectra. For the digested peptide samples, the MALDI-*Tof/Tof* was operated in reflectron positive ionization mode in the m/z range 800–4000.

References

1. Hermanson GT (1996) Bioconjugate techniques, 1st ed. San Diego, CA: Academic Press.
2. Kaiser ET (1988) Catalytic activity of enzymes altered at their active sites. *Angew Chem Int Ed Engl* 27:913–922.
3. Benson DE, Conrad DW, de Lorimier RM, Trammell SA, Hellinga HW (2001) Design of bioelectronic interfaces by exploiting hinge-bending motions in proteins. *Science* 293:1641–1644.
4. Cao L (2005) Carrier-bound immobilized enzymes: principles, applications, and design. 1st ed. Weinheim: Wiley VCH.
5. Guignet EG, Hovius R, Vogel H (2004) Reversible site-selective labeling of membrane proteins in live cells. *Nature Biotechnol* 22:440–444.

6. Krishnan B, Szymanska A, Gierasch LM (2007) Site-specific fluorescent labeling of poly-histidine sequences using a metal-chelating cysteine. *Chem Biol Drug Des* 69: 31–40.
7. Allen MW, Urbauer RJB, Zaidi A, Williams TD, Urbauer JL, Johnson CK (2004) Fluorescence labeling, purification, and immobilization of a double cysteine mutant calmodulin fusion protein for single-molecule experiments. *Anal Biochem* 325:273–284.
8. Smith JJ, Conrad DW, Cuneo MJ, Hellinga HW (2005) Orthogonal site-specific protein modification by engineering reversible thiol protection mechanisms. *Prot Sci* 14:64–73.
9. De Lorimier RM, Tian Y, Hellinga HW (2006) Binding and signaling of surface-immobilized reagentless fluorescent biosensors derived from periplasmic binding proteins. *Prot Sci* 15:1936–1944.
10. Adams E, Jeter D, Cordes AW, Kolis JW (1990) Chemistry of organometalloid complexes with potential antidotes: structure of an organoarsenic(III) dithiolate ring. *Inorg Chem* 29:1500–1503.
11. Stocken LA, Thompson RHS (1946) British Anti-Lewisite. II. Dithiol compounds as antidotes for arsenic. *Biochem J* 40:535–548.
12. Zahler WL, Cleland WW (1968) A specific and sensitive assay for disulfides. *J Biol Chem* 243:716–719.
13. Biemans-Oldehinkel H, Doeven MK, Poolman B (2006) ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett* 580:1023–1035.
14. Pflugrath JW, Quiocho FA (1988) The 2 Å resolution structure of the sulfate-binding protein involved in active transport in *Salmonella typhimurium*. *J Mol Biol* 200: 163–180.
15. Castanha ER, Fox A, Fox KF (2006) Rapid discrimination of *Bacillus anthracis* from other members of the *B. cereus* group by mass and sequence of “intact” small acid soluble proteins (SASPs) using mass spectrometry. *J Microbiol Methods* 67:230–240.
16. Pardee AB (1966) Purification and properties of a sulfate-binding protein from *Salmonella typhimurium*. *J Biol Chem* 241:5886–5892.
17. Kasha M, Rawls HR, El-Bayoumi MA (1965) The exciton model in molecular spectroscopy. *Pure Appl Chem* 11: 371–392.
18. Kragh-Hansen U (1983) Graphical analysis of competitive binding of comparable concentrations of ligand, inhibitor and protein. *Biochem Pharmacol* 32:2679–2681.
19. Pace CN, Vajdos F, Fee L, Grimsley G, Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* 4:2411–2423.
20. Cadene M, Chait BT (2000) A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins. *Anal Chem* 72:5655–5658.
21. Lanfermeijer F, Picon A, Konings WN, Poolman B (1999) Kinetics and consequences of binding of nona- and dodecapeptides to the oligopeptide binding protein (OppA) of *Lactococcus lactis*. *Biochemistry* 38:14440–14450.