Saturation fluorescence labeling of proteins for proteomic analyses

Elizabeth Pretzer a, John E. Wiktorowicz b,*

a Bayer HealthCare Pharmaceuticals, Berkeley, CA 94710, USA
b Department of Biochemistry and Molecular Biology, Biomolecular Resource Facility, and Sealy Center for Molecular Medicine, University of Texas Medical Branch, Galveston, TX 77550, USA

Received 8 August 2007
Available online 23 December 2007

Abstract

We present here an optimized and cost-effective approach to saturation fluorescence labeling of protein thiols for proteomic analysis. We investigated a number of conditions and reagent concentrations, including the disulfide reducing agent tris(2-carboxyethyl)phosphine (TCEP), pH, incubation time, linearity of labeling, and saturating dye/protein thiol ratio with protein standards to gauge specific and nonspecific labeling. Efficacy of labeling under these conditions was quantified using specific fluorescence estimation, defined as the ratio of fluorescence pixel intensities and Coomassie-stained pixel intensities of bands after digital imaging. Factors leading to specific versus nonspecific labeling in the presence of thiourea are also discussed. We found that reproducible saturation of available Cys residues of the proteins used as labeling standards (human carbonic anhydrase I, enolase, and α-lactalbumin) is achieved at 50- to 100-fold excess of the uncharged maleimide-functionalized BODIPY dyes over Cys. We confirmed our previous findings, and those of others, that the maleimide dyes are not affected by the presence of 2 M thiourea. Moreover, we established that 2 mM TCEP used as reductant is optimal. We also established that labeling is optimal at pH 7.5 and complete after 30 min. Low nonspecific labeling was gauged by the inclusion of non-Cys-containing proteins (horse myoglobin and bovine carbonic anhydrase) to the labeling mixture. We also showed that the dye exhibits little to no effect on the two-dimensional mobilities of labeled proteins derived from cells.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Saturation fluorescence labeling; Cys labeling; Proteomics; Two-dimensional electrophoresis

One of the many criticisms leveled against two-dimensional gel electrophoresis (2DGE)1 is that the staining techniques used to visualize separated proteins are semiquantitative, exhibit a small dynamic range and low sensitivity, and are not very reproducible (for recent reviews, see Refs. [1–3]). This may be due in part to the variation in protein reactivity with noncovalent dyes, protein precipitation at their isoelectric points (pI), incomplete solubilization by sodium dodecyl sulfate (SDS), variability of gel composition and quality, and other factors. Quantitative standards are of little use because of the uniqueness of individual proteins with respect to dye reactivity and other difficult-to-control factors. With the introduction of fluorescent dyes and high-quality imagers, postseparation staining has improved in terms of sensitivity and dynamic range, but high coefficients of variation in abundance estimates remain a problem even with these dyes [4]. Nevertheless, 2DGE remains the “gold standard” for differential proteomic analyses, due primarily to its exquisite resolving power and its ability to efficiently compare

---

1 Abbreviations used: 2DGE, two-dimensional gel electrophoresis; SDS, sodium dodecyl sulfate; SILAC, stable isotope labeling of amino acids in culture; iTRAQ, isotope tag for relative and absolute quantitation; MS, mass spectrometer; DIGE, differential in-gel electrophoresis; enol, yeast enolase I; bCA, bovine carbonic anhydrase II; hCA, human carbonic anhydrase I; α-lac, bovine α-lactalbumin A; myo, horse myoglobin; Tris, trishydroxymethyl aminomethane; 2-ME, 2-mercaptoethanol; DMSO, dimethyl sulfoxide; colloidal Coomassie Blue, Brilliant Blue G Colloidal Concentrate; BODIPY FL-Mal, BODIPY FL-5(N-(2-aminoethyl) maleimide; BODIPY FL-Iaa, BODIPY FL Cl-iodoacetamide; Chaps, 3-(3-cholamidopropyl)dimethylammonio)-1-propanesulfonate; TCEP, tris(2-carboxyethyl)phosphine; PAGE, polyacrylamide gel electrophoresis; IEF, isoelectric focusing; IPG, immobilized pH gradient; ROI, region of interest; MALDI, matrix-assisted laser desorption/ionization; AAA, amino acid analysis.
protein abundances from multiple sources. It has been our long-term goal to address some of the difficulties in 2DGE, and accurate quantification is, in our view, one of the most important.

One promising approach applies fundamental biochemical principles, namely, preseparation covalent labeling of proteins for proteomic analysis. This approach has garnered attention from the proteomics community because it promises accurate quantification, multiplexed separation and detection, and high sensitivity. Covalent labeling techniques generally fall into two categories: isotope (stable and radio) and fluorescence labeling. Stable isotope labeling schemes, such as stable isotope labeling of amino acids in culture (SILAC) [5,6] and isotope tag for relative and absolute quantitation (iTRAQ) [7], exploit the exquisite resolution obtainable by even comparatively low-cost mass spectrometers (MSs) to the degree that labels differing by 1 amu are easily discriminated. Thus, to compare two protein populations, each mixture can be covalently labeled with a reagent that differs by 1 amu from the label on the other mixture. The two populations can then be mixed and their detection can be multiplexed to obtain relative quantification of parental proteins. The advantage to this approach is that the stable isotopes permit ratiometric normalization of protein (peptide) recoveries to minimize the impact of artificial loss. The difficulty in these approaches is the requirement of most MSs for low mass/charge ratios of analytes, hence the need to digest proteins to their constituent peptides. Moreover, the increased complexity that this imparts to the protein extracts dictates pre-MS separation strategies, leading to potential loss of sample. This is troublesome with respect to proteins that are in low abundance; artifactual loss may decrease their signals below the detection limits, and the stable isotopes do not, in and of themselves, increase MS sensitivity. In addition, relative quantification is performed at the analytical stage (MS), rather than at the separation stage, thereby dictating that every peptide that is generated from the extract must be analyzed regardless of its constitutive or inducible status. Thus, there is no opportunity for a differential cut, and even uninformative proteins must be analyzed. Finally, for quantification to work, the cognate peptides from the different sources must coelute from the fractionation step and appear in the same MS scan; otherwise, the opportunity for quantification is lost. Failure of cognate peptides to coelute might occur, for example, when proteins undergo posttranslational modification(s) as a result of the differential treatment of the cell that alters chromatographic or MS behavior. In this case, cognate peptides will not appear in the same MS scan and, therefore, will not be quantified.

The alternative strategy, covalent derivatization of protein samples with fluorescent dyes prior to separation by 2DGE, is a technique that is finding increasing utility in proteomics, such as differential in-gel electrophoresis (DIGE) [8–10] and saturation labeling [8,11], due to its potential for quantitative accuracy and sensitivity. The advantages of this approach include the opportunity for multiplexing samples and inclusion of internal standards, the potential for detecting low-abundance proteins, the ability to realize relative and absolute linear quantification (after the protein is identified), and the potential for real-time monitoring of electrophoresis. In contrast to the stable isotope strategies, quantification is performed after separation but before MS identification. This means that the sample complexity is reduced before the identification phase because only those proteins whose abundance levels are of interest need to be analyzed by the MS instruments for the purpose of identification.

These advantages come with a price, however. Quantification requires reproducible saturation of targeted sites on any protein as well as absence of nonspecific modifications. Detection of low-abundance proteins requires instrumenta tion that exhibits low noise and low background as well as the ability to prolong exposure times. In our view, the advantages outweigh the disadvantages, and with attention given to both the chemical and instrument demands, preseparation labeling with fluorescence can contribute greatly to the needs of the proteomics community and, most important, can permit accurate and reproducible detection of low-abundance proteins.

We and others [11–13] have previously examined conditions for labeling proteins with thiol-reactive dyes and defined the basic properties that these dyes must exhibit to function as viable candidates; these include high extinction coefficients, good quantum yields, pH insensitivity of fluorescence, low photobleaching, minimal effects on protein pI values, a high degree of chemical specificity, and the ability to achieve reproducible saturation of protein residues. In our previous study [11], we noted that thiourea, an important component of many protein solubilization cocktails, severely inhibited labeling of protein thiols by iodoacetamide dyes, but not by a maleimide dye, in agreement with others who observed similar reactivities [14]. For preseparation fluorescence labeling to find widespread use in proteomics as a quantitative tool, its reactivity in varied chemical environments must be studied systematically. For this reason, and in view of the widespread use of thiourea in proteomic sample preparation, we continued our investigations into the use of fluorescence dyes for proteomic applications and present here optimal conditions for labeling protein thiol groups with a maleimide dye in the presence of thiourea, such that saturation labeling of thiols and minimal nonspecific labeling occur.

**Materials and methods**

**Materials**

Yeast enolase I (enol), bovine carbonic anhydrase II (bCA, pI 5.4), human carbonic anhydrase I (hCA), bovine α-lactalbumin A (α-lac), horse myoglobin (myo), trishydr oxyethyl aminomethane (Tris), sodium bicarbonate, SDS, 2-mercaptoethanol (2-ME), dimethyl sulfoxide
(DMSO), thiourea, individual fluorescent molecular weight markers, Triton X-100, and Brilliant Blue G Colloidal Concentrate (colloidal Coomassie Blue) were obtained from Sigma–Aldrich (St. Louis, MO, USA). Urea was added to the mix to prevent undesired alkylation. All incubations were carried out overnight with acetone and pelleted at 25,000 g for 1 h. Then 5 μg was removed and analyzed for protein and Cys content by amino acid analysis (Hitachi model L-8800, Hitachi High Technologies America, San Jose, CA, USA). The pellet was resuspended in 8 M urea and 1% Chaps.

Protein labeling

In general, proteins (12.5–50.0 μM) were denatured with 7 M urea/2 M thiourea or with 8 M urea alone along with 0.1% Triton X-100 or reduced Triton X-100 with 2% Chaps in 50 mM Tris–HCl and reduced with TCEP for 20 to 30 min. In some cases, 0.5 mM EDTA was added during the final 10 min of reduction. Dye (50 mM in DMSO) was added to the desired concentration, and the mixtures were incubated for up to 1.5 h. Unlabeled controls received DMSO in place of dye. The reactions were stopped with a 150- to 300-fold molar excess of 2-ME over dye and acidified with 3 μl of 1 N HCl per 100 μl reaction mix to further prevent undesired alkylation. All incubations were carried out at room temperature in the dark. The typical reaction volume was 100 μl.

PD-10 chromatography

In some cases, the crude labeling mixes were applied to PD-10 columns and eluted with 0.1 M sodium bicarbonate (pH 8.3). Following 2.0 ml of void volume, typically 0.5-ml or smaller fractions were collected. Peaks were located by fluorescence intensity and Coomassie Blue stain intensity on SDS NuPAGE gels, and the peak fractions were pooled.

Electrophoresis

Reduced SDS–polyacrylamide gel electrophoresis (PAGE) in 1.5-mm 14 or 16% Tris–Gly gels, or in Bis-Tris NuPAGE gels with Mes running buffer, was carried out according to the manufacturer’s instructions. Samples were generally diluted in half with sample buffer containing SDS, glycerol, and Tris. Because crude samples contained large amounts of 2-ME, additional 2-ME usually was not added to the SDS–PAGE sample buffer except when running PD-10 eluates. Fronts sometimes were run off the gels to remove unreacted dye and obtain cleaner fluorescence images. Fluorescence images were obtained on an AlphaImager or a ProXpress 2D imager. Gels were then stained with colloidal Coomassie blue, destained, and imaged on the Personal Densitometer.

Two-dimensional electrophoresis was performed employing an IPGphor multiple-sample isoelectric focusing (IEF) device (GE Healthcare, Waukesha, WI, USA) for the first dimension and Bio-Rad’s multiple-gel SDS–PAGE systems (Protean Plus and Criterion Dodeca cells, Hercules, CA, USA) for the second dimension. IEF was performed with 11-cm precast immobilized pH gradient (IPG) strips (Bio–Rad). Sample aliquots were loaded onto dehydrated IPG strips and rehydrated overnight. IEF was performed at 20 °C under the following conditions: 50 V, 11 h; 250 V, 1 h; 500 V, 1 h; 1000 V, 1 h; 8000 V, 2 h; 8000 V, 6 h. The IPG strips were then incubated in 4 ml of equilibration buffer (6 M urea, 2% SDS, 50 mM Tris–HCl [pH 8.8], and 20% glycerol) containing 10 μl/ml tri(2-carboxyethyl)phosphine (Geno Technology, St. Louis, MO, USA) for 15 min at 22 °C with shaking. The unlabeled samples were incubated in another 4 ml of equilibration buffer with 25 mg/ml iodoacetamide for 15 min at 22 °C with shaking to ensure protein S-alkylation. Electrophoresis was performed at 150 V for 2.25 h and 4 °C with precast 8 to 16% polyacrylamide gels in Tris–glycine buffer (25 mM Tris–HCl, 192 mM glycine, and 0.1% SDS [pH 8.3]).

After electrophoresis, the gels were fixed in fix buffer (10% methanol and 7% acetic acid in doubly distilled H2O), stained with SYPRO Ruby (Bio–Rad), and destained in fix buffer. The gels were scanned at a 100-μm resolution with the ProXpress 2D imager using 460-nm excitation and 620-nm emission filters for SYPRO Ruby-stained gels or 480-nm excitation and 530-nm emission filters for BODIPY-labeled protein gels. The exposure was adjusted to achieve a value of approximately 55,000 to
63,000 pixel intensity (16 bits) on the most intense protein spots on the gel.

**Image processing and analysis**

Images acquired from the AlphaImager were stored as TIF images with no compression. The Coomassie stain and corresponding fluorescence images were first cropped to remove extraneous information, and their image sizes (pixel dimensions) were equalized with PhotoShop software (Adobe, San Jose, CA, USA). No other manipulations of the images were performed with PhotoShop. The files were then imported into IPLab (Biovision Technologies, Exton, PA, USA), where a region of interest (ROI) of equal size was defined for each band. Identical ROIs immediately in the vicinity of each band (containing no protein) were also defined and used for subsequent background correction. All ROIs were quantified for intensity, and backgrounds were corrected by subtraction. Specific fluorescence (SF) is defined as the ratio of background-corrected pixel intensity sum across the ROI of the fluorescence band (Pf) and the background-corrected pixel intensity sum across the corresponding ROI of the Coomassie-stained band (Pc) according to the following formula:

\[ SF = \frac{Pf - Bj}{Pc - Bc}. \]

In some cases, as noted in the figure legends, SF of the Cys-containing protein is calculated by subtracting the SF of the non-Cys-containing protein.

**Linearity of labeling**

Equimolar mixtures of enol and myo, ranging from 1 to 10,000 fmol (40 pM–400 nM), were reduced and denatured with 8 M urea, 0.1% Triton X-100, 50 mM Tris (pH 7.5), and 2 mM TCEP for 30 min in 25 μl. The proteins were labeled in duplicate with 750 pmol (30 mM, 75-fold excess over 10 pmol protein thiol) of BODIPY FL-Mal for 2 h, followed by quenching with 5 mM 2-ME for 30 min. All incubations were done at room temperature. Reactions were frozen at −80 °C overnight, and the proteins were run on duplicate 4 to 20% SDS gels. The gels were imaged on a ProXpress 2D imager with 480/30-nm excitation and 535/50-nm emission filters for 60 s. The fluorescent bands were quantified as described above for image processing and analysis, and linear regression was performed on the duplicate intensities using Prism 4.0c (GraphPad, San Diego, CA, USA).

**Results**

**Labeling in thiourea with iodoacetamides versus maleimides**

This study focuses on preseparation labeling with the dye BODIPY FL-Mal. We selected it because we showed previously that the BODIPY family of dyes meets our criteria for good proteomic fluorescence dyes [11]. The structures of BODIPY FL-Iaa and BODIPY FL-Mal are shown in Fig. 1.

We previously showed strong inhibition of labeling with BODIPY TMR cadaverine Iaa in the presence of thiourea [11], corroborating the findings of Galvani and coworkers [14] that thiourea inhibits alkylation of thiol groups with iodoacetamide, and we found that Rhodamine Red C2 maleimide, as well as BODIPY FL-Mal, labeled efficiently in the presence of thiourea. To rule out the possibility that those differences were dye specific, in the current study we compared the maleimide and iodoacetamide forms of a single dye. Labeling was done at two pH levels to address the possible different pH requirements for optimal labeling with iodoacetamides versus maleimides. Throughout this work, we include an internal protein standard containing no Cys to gauge the levels of nonspecific labeling, and we show the Coomassie-stained protein as well as fluorescence images. In some instances where disulfide reducing power is an issue, we use α-lac (containing 8 thiols). The advantage of this protein is that its incompletely labeled forms are observable as bands migrating between unlabeled and fully labeled species, thereby permitting direct visualization of reaction completeness.

We first determined the impact of labeling a mixture of enol (containing 1 thiol) and α-lac with BODIPY FL-Iaa or with BODIPY FL-Mal in 8 M urea or in 7 M urea/2 M thiourea at pH 7.2 or 8.0 (Fig. 2). The dye/thiol ratio in this experiment was 10:1 (2.25 mM dye/225 μM thiol). Reduction was with 2 mM TCEP, giving a 0.88:1 reducer/dye ratio. The maleimide dye labeled efficiently in both chaotrope systems and at both pH levels, as shown by the bright fluorescence of the bands and by the apparent molecular size shift of α-lac. In contrast, the iodoacetamide dye labeled well in urea but was severely inhibited in thiourea, with virtually no labeling at pH 7.2 and very little at pH 8.0. These results confirm that the iodoacetamide, but not the maleimide, sulfhydryl reactivity is negatively

![Fig. 1. Dye structures. Upper: N-(4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-yl)methyl iodoacetamide (BODIPY FL-Iaa). Lower: BODIPY FL-N-(2-aminoethyl)maleimide (BODIPY FL-Mal). (Structures courtesy of Molecular Probes).](image)
affected by the presence of thiourea and that this effect is not due to the structure of the dye molecule.

Effect of TCEP concentration on completeness and specificity of labeling

Clearly, maximum reduction of protein thiols is critical to achieve reproducible labeling. The concentration of reducing agent, therefore, must be optimized without using an excessive amount that might lead to inhibition of labeling. Our previous results suggested interaction of the reducing agent, TCEP, with iodoacetamide and maleimide dyes as well as the importance of the TCEP/dye ratio for complete and specific labeling [11]. To gauge this effect more systematically, we used α-Lac for its eight reducible Cys. Using this protein and bCA (MW 29,000, 0 Cys) as an internal nonspecific labeling control, we examined the effect of TCEP concentration on labeling with a constant dye concentration in the presence of thiourea.

α-Lac was labeled at a concentration of 25 µM in the presence of 12.5 µM bCA in 7 M urea/2 M thiourea at pH 7.2. The difference in protein amounts reflects the difference in the number of potentially reactive amines in each protein that might engage in side reactions with maleimides [15]. Reduction was carried out for 40 min using 0, 0.5, 1, 2, 3, 4, 5, or 10 mM TCEP, followed by labeling with 2 mM BODIPY FL-Mal for 1.5 h (10:1 dye/thiol ratio). The reducing agent/thiol ratio varied from 2.5:1 to 50:1, and the reducing agent/dye ratio varied from 0.25:1 to 5:1. The results are shown in Fig. 3.

At 0.5 and 1 mM TCEP (0.25 and 0.50 TCEP/dye ratios), α-lac displays multiple bands at intermediate positions between the unlabeled and fully labeled bands on the SDS gel, suggesting incomplete modification. Maximum labeling appears at 2 to 3 mM TCEP (Figs. 3 A–C). Very long exposures reveal low labeling (<5%) of bCA regardless of TCEP concentration (quantitative data not shown). Interestingly, labeling of α-lac at 10 mM TCEP (5:1 TCEP/dye ratio) seems to result in electrophoretic behavior similar to that observed at 1 mM TCEP (Fig. 3A), suggesting that interference with labeling occurs at a fivefold or greater excess of TCEP over dye.

Effect of dye concentration on completeness and specificity of labeling

Having found the maleimide-compatible range of TCEP/dye ratios, we investigated the optimal ratio of dye...
to protein thiol. HCA (MW 29,000), containing a single Cys residue, was chosen for this purpose. Equine myo (MW 17,000, no Cys) was included as a negative control to detect nonspecific labeling.

Fig. 4 shows labeling of hCA at 25 μM in the presence of 43 μM myo and 2.5 mM TCEP, with the dye concentration varying from 0.125 to 5 mM (0–200:1 dye/thiol ratio). From the summary graph (Fig. 4C), hCA appears to saturate readily at a dye/thiol ratio of 20:1 (0.5 mM dye) and remain at this level through a dye/thiol ratio of 200:1 (5 mM dye) at 2.5 mM TCEP. Nonspecific labeling appears in significant levels at dye/thiol ratios greater than 100:1 (2.5 mM dye).

In comparing these results with those in Fig. 3, where the dye/thiol ratio was 10:1 and the best results were obtained with 2 mM TCEP concentration (dye concentration was also 2 mM), the results in Fig. 4 indicate saturation at 10-fold or greater excess dye when TCEP concentration was 2.5 mM (Fig. 4C). Higher dye/thiol ratios are well tolerated and may approach 100-fold with little to no nonspecific labeling. Both figures suggest a fine degree of interplay among TCEP, dye, and protein thiols, with a window of optimal concentrations. In light of these data, we routinely use 2 mM TCEP for reduction and 50- to 100-fold dye/thiol ratios for labeling complex mixtures of proteins.

**Effect of pH and reaction time on completeness and specificity of labeling**

To measure completeness of alkylation, α-lac was labeled at 26.8 μM in the presence of 13.4 μM bCA as negative control. Both TCEP and dye were 2 mM, whereas the dye/thiol ratio was 9.3:1. The reactions were buffered at pH 6.5, 7.0, 7.2, 7.5, or 8.0 (Fig. 5). Labeling appeared complete at pH 7.2 or above. There was a trace of fluorescence in the bCA band at pH 8.0 (very long exposure on the AlphaImager showed trace labeling at pH 8.0 and possibly at pH 7.5). Thus, a higher pH seems more efficient but also resulted in some nonspecific labeling. These results show that pH greater than 7.2 but less than 8.0 is best for efficient labeling, with little or no nonspecific labeling.

Because nonspecific reactivity is a concern with all of the maleimide dyes we have examined, we investigated the possibility that a shorter reaction time would minimize side reactions while still allowing complete labeling of thiols. The 1.5-h reaction time used in experiments up to this point was based on our experience with iodoacetamide dyes.
\[ a \]-Lac was labeled (9.5:1 dye/thiol ratio) at 25 \( \mu \)M in the presence of 12 \( \mu \)M BCA as negative control for 30, 60, and 90 min in 2 mM TCEP. We performed the labeling at pH 8.0 because we had established in Fig. 5 that only at this pH was nonspecific labeling detectable.

Fig. 4. BODIPY FL-Mal saturation. Increasing concentrations of dye were incubated with a mixture of hCA (upper band) and myo (lower bands) in panels A and B in 2.5 mM TCEP. (A) Coomassie stained. (B) Fluorescence. (C) Specific fluorescence (SF). SF is expressed as the ratio of the sum of pixel intensities of each hCA fluorescence band and the corresponding hCA Coomassie band minus the SF of myo. Both intensities were background corrected. The abscissa represents the dye/protein molar ratio for myo (0 Cys) and the dye/protein thiol ratio for hCA (1 Cys). Further details are described in Materials and methods.

Fig. 5. pH dependence of BODIPY FL-Mal labeling. BCA (upper bands) and \( a \)-lac (lower bands) were mixed and labeled at the indicated pH values with 2 mM TCEP and 1 mM dye (9:1 dye/thiol ratio). (A) Coomassie stain. (B) Fluorescence image. Details are described in Materials and methods.

\[ a \]-Lac was labeled (9.5:1 dye/thiol ratio) at 25 \( \mu \)M in the presence of 12 \( \mu \)M BCA as negative control for 30, 60, and 90 min in 2 mM TCEP. We performed the labeling at pH 8.0 because we had established in Fig. 5 that only at this pH was nonspecific labeling detectable.
Fig. 6 shows minimal nonspecific labeling at all three time points, even after gross overexposure (data not shown). Moreover, the figure shows that complete labeling under these conditions is achieved within 30 min, with no further increase in specific fluorescence, greatly improving the throughput of the procedure.

**Linearity of labeling**

Finally, using the optimized conditions described above, we labeled increasing amounts of equimolar enol and myo, from 1 fmol to 10 pmol (40 pM–400 nM), with a constant 30 mM BODIPY FL-Mal (75-fold excess over 400 nM enol) in duplicate. Labeled mixtures were run on duplicate SDS gels without further processing, imaged, and quantified as described in Materials and Methods. Fig. 7 demonstrates linear labeling over at least three orders of magnitude of concentration. Linear regression of duplicates yielded a correlation coefficient of 0.9952 with a $P$ value of 1.000 for the residual runs test (maximum significance for linear fit). The coefficient of variation of the replicates ranged from 1.9 to 9.4% ([labeled enol] = 25–10,000 fmol) and 34% for 10 fmol. No significant labeling of myo was observed at any protein level (<1%).

**Impact on 2DGE protein mobility**

To examine the impact of BODIPY FL-Mal modification of a complex protein extract, we performed the 2DGE separation of an *E. coli* extract with and without
labeling (Fig. 8). The gel containing unlabeled proteins was stained postseparation with SYPRO Ruby, and the protein spots were matched against the BODIPY-labeled protein spots using Progenesis software (Nonlinear Dynamics). The vectors defining the matched protein spots were obtained from the spot coordinates. The upper panel of Fig. 8 was generated by superimposition of the vectors on the gel containing the BODIPY-labeled proteins. The length and angle of the vectors summarize the relative difference in mobility between the BODIPY and the unlabeled proteins. Of the 1263 vectors describing proteins matched by the software, the average length was 6.6 ± 4.1 pixels. To put this in perspective, the Fig. 8 insets (A–C) indicate the diameter of a circle containing the average vector length (inset A), the 90% confidence diameter (inset B, 1 SD), and the 95% confidence diameter (inset C, 2 SD).

The regions containing the greatest protein spot densities from both gels were cropped, and only contrast and brightness were adjusted with PhotoShop to match (Fig. 8: lower left, and right panels). These are included to demonstrate the similarities in spot numbers and migration. Protein spot intensities, however, do not necessarily match given that the BODIPY labels reflect the number of Cys residues per protein, whereas the SYPRO Ruby intensities reflect efficiency of dye to protein binding.

Impact on protein solubility during 2DGE

An important question arises regarding the solubility of BODIPY-labeled proteins due to the increase in overall protein hydrophobicity by the modification. This is of particular concern when IEF is complete, when the local
protein concentration is at its maximum, and may result in precipitation and failure to transfer to the second-dimension gel. We exploited the protein fluorescence by performing IEF of a complex protein mixture (extracted from EOL-1 cells) labeled with BODIPY FL-Mal and run in a standard IEF gel (not an IPG strip). The gel was imaged, a lane was cut out, and the proteins were run into a second-dimension SDS gel in typical fashion. After electrophoresis, the IEF region was reimaged to look for retained proteins. No difference was observed between the BODIPY-labeled and unlabeled extracts stained with SYPRO Ruby in either the first-dimension IEF gel or the second-dimension SDS gel (data not shown).

Discussion

Covalent labeling of protein thiols with fluorescent dyes is an attractive strategy for proteomic analyses for a number of reasons. If performed with saturating ratios of dye to thiols (>50:1), it is most likely to result in accurate quantification and reproducible separations. Because most proteomic procedures perform sulfhydryl alkylation to prevent artifactual oxidation, our approach accomplishes this as well but with a fluorescent dye. This leads to a separation pattern that very closely (if not identically) matches the pattern obtained with the conventional sulfhydryl blocking agents provided that the dyes are uncharged. Although saturation labeling with DIGE Cy3 and Cy5 dyes has been shown to be more accurate than postseparation detection with adsorptive dyes for proteomic analyses by 2DGE, in our view its implementation suffers from several difficulties. For example, the labeling protocol by the manufacturer (GE Healthcare) for the maleimide Cy3 and Cy5 saturation dyes recommends a dye concentration that corresponds to a dye/thiol labeling ratio of between 1.7:1 and 15:1, depending on assumptions regarding the average molecular weight of proteins. We have shown that saturation cannot be ensured unless total Cys content of a preparation is known and the dye/thiol ratio exceeds 50:1, at least with the BODIPY and other dyes. In our previous study [11], we investigated the reactivity of various uncharged and zwitterionic (neutral) Cys-reactive dyes. In all cases, including zwitterionic dyes, saturation required at least 40- to 50-fold excess dye over protein thiols. Any differences observed among the dyes were largely in their reactivities with non-Cys-containing protein residues and/or quenching at higher dye excess. Although we did not assay the maleimide Cy3 and Cy5 dyes, it is unlikely that their Cys-specific reactivity will be much different from that of any of these other dyes. This may raise a concern that the DIGE saturation protocol is not truly saturating, thereby affecting its accuracy and reproducibility. A second issue is that Cy3 and Cy5 are charged but neutral by virtue of their quaternary amine ($pK_a > 12$) and sulfonic acid ($pK_a < 1$) moieties [16]. Thus, their impact on the $pI$ of modified proteins will depend on the degree of substitution but will at least compress the IEF separation profiles on a two-dimensional gel, making acidic proteins more basic and making basic proteins more acidic. In general, it is more desirable to increase resolution rather than decrease it and to modify the proteins uniformly and maintain them in a state that most closely resembles their unmodified denatured state. We believe that the best way to accomplish this is to truly saturate proteins with uncharged fluorescent dyes.

To address the suitability for MS identification, the maleimide DIGE dyes have been shown to be compatible with MS analyses—with both modified and unmodified model peptides from a labeling mixture (2:1 dye/thiol) appearing in a matrix-assisted laser desorption/ionization (MALDI) scan [16]—and we also previously demonstrated the suitability of the BODIPY dyes for MALDI MS [11].

Our purpose in pursuing this labeling strategy is to push the limits of conventional 2DGE to the highest level of quantitative accuracy and sensitivity and to do so in the most cost-effective manner. Thus, we have optimized a labeling strategy with off-the-shelf dyes that match our criteria to meet these goals. Previously, we examined the advantages and considerations of preseparation covalent saturation fluorescence labeling of protein thiol groups with an emphasis on iodoacetamide-derivatized dyes [11]. We noted that although labeling with iodoacetamide dyes was severely inhibited in the presence of thiourea, this was not the case with maleimide dyes. However, several maleimide dyes that we surveyed exhibited significant nonspecific labeling in 8 M urea. Saturation was readily achievable with both dye types, and optimization centered on identifying a window of conditions that eliminated nonspecific labeling while still allowing saturation labeling. In the current study, we extended our survey and showed, among other things, that specificity is not significantly affected by choice of chaotrope (8 M urea or 7 M urea/2 M thiourea). Moreover, we showed that with careful choice of dye, dye/thiol ratio, reducer concentrations, pH, and time, conditions can be selected to obtain linear saturation labeling with minimal nonspecific labeling to detect a broad concentration range of proteins.

An important point to emphasize is that the amount of dye necessary to accomplish saturation must be based on the accurate determination of total Cys content for the protein extract to be labeled. We rely on amino acid analysis (AAA) for this estimate, and it remains arguably the most accurate means for quantifying specific amino acid residues and total protein. Labeling with sufficient dye to achieve excess based on AAA estimation of Cys, and performing labeling under conditions that maximally reduce and denature proteins, eliminates the “mass influence” of abundant proteins because in the denatured state (a prerequisite for saturation fluorescence labeling) all Cys residues should be equally reactive regardless of whether they belong to a low-abundance or high-abundance protein; that is, they are likely to be labeled and respond similarly to the excess of dye.
Optimization of labeling with BODIPY FL-Mal in the presence of thiourea

A direct comparison of labeling with the iodoacetamide and maleimide versions of BODIPY FL in urea or urea/thiourea confirmed that the differences we had seen previously between other iodoacetamide and maleimide dyes in the two chaotrope systems were due to the sulfhydryl-selective reactive group. Consistent with our previous observations, labeling by BODIPY FL-Iaa, but not by BODIPY FL-Mal, was strongly inhibited in thiourea. Therefore, we concluded that the maleimide form of BODIPY might allow better labeling efficiency in thiourea with maximum fluorescence yield under optimal reducing and saturating conditions.

As with the iodoacetamid-derivatized dyes that we examined previously, optimization of labeling with maleimide-derivatized dyes involves a matrix of interactive effects and conditions. Using BODIPY FL-Mal, we addressed the effects of TCEP/dye and dye/thiol ratios, of pH, and of reaction time on completeness and specificity of labeling. The critical reactions are the reduction of disulfides by TCEP and the alkylation of free thiols by the maleimide-derivatized dyes involves a matrix of interactive effects and conditions. Using BODIPY FL-Mal, we addressed the effects of TCEP/dye and dye/thiol ratios, of pH, and of reaction time on completeness and specificity of labeling. The critical reactions are the reduction of disulfides by TCEP and the alkylation of free thiols by the maleimide moiety of the dye. Other important reactions that may occur are the binding of dye to TCEP [11,17] and the binding of dye to nonthiol sites on proteins. To ensure quantitative accuracy, the latter reaction is to be avoided; to ensure saturation, the former is to be minimized while still achieving sufficient reducing power for protein disulfide reduction. The apparent interaction of TCEP and dye may be minimized by using 2 mM TCEP to achieve complete protein disulfide reduction. A large excess of TCEP, however, may deplete the dye to levels below the concentration necessary to achieve saturation. In this study, the best results were obtained with a TCEP concentration of 2 mM and a dye/thiol ratio greater than 10:1 for highly purified proteins. For ill-defined complex extracts, we found a more optimal ratio to be approximately 75:1.

The pH and the dye/thiol ratio combine to affect completeness and specificity of labeling. Acylation of a thiol with an iodoacetamide involves nucleophilic displacement of the iodine by the thiolate anion, producing a thioether where the nucleophile (thiolate) reacts with the olefin (maleimide). Maleimides are not known to react with Met or His, but at high pH reactivity with Lys and Tyr are favored over thiols [19,20]. Moreover, at pH above 8.0, maleimides may be hydrolyzed to a mixture of nonreactive maleamic acid isomers. Because alkylation is favored for sulfhydryls and amines (average pKₐ for ε-Lys = 10.8) in the unprotonated state, increasing the pH between 7.0 and 8.0 increased both specific and nonspecific labeling. At pH 7.5 sulfhydrlys will be reactive (unprotonated) approximately 13% of the time and ε-amines will be reactive (unprotonated) approximately 0.05% of the time, whereas at pH 8.0 the reactivity of both groups increases to 68 and 0.22%, respectively. Clearly, maintaining the pH at 7.5 while increasing the dye/thiol ratio and TCEP/dye ratio will yield greater sulfhydryl specificity and drive the reaction to completion. Stated in general terms, the pH range supporting complete and specific labeling for a given dye can be manipulated to some extent by increasing or decreasing the dye/thiol ratio while keeping the TCEP/dye ratio constant. Accordingly, we investigated the optimal pH range that would allow complete labeling with the maleimide dye while minimizing side reactions and found optimal labeling at pH 7.5.

Labeling with BODIPY FL-Mal in thiourea was rapid and relatively insensitive to prolonged reaction times; no differences in completeness or specificity were observed at 30, 60, or 90 min of incubation. This result contrasts with the iodoacetamides in 8 M urea, where we typically labeled for 90 min to ensure saturation.

Linearity of labeling

Our measured limit of detection at a signal/noise ratio of 2:1 in the PerkinElmer imager used in the linearity of labeling experiments was 10 fmol (unpublished observations), which was also the lowest signal measurable from the gels in this experiment (Fig. 7). This amount corresponds to a protein copy number of 3000 molecules from 2 million cells, the number of cells that typically yields 200 µg of protein and the typical two-dimensional gel load (13.3 × 8.7 cm) used in our facility. We have showed that our labeling approach can label this small amount of protein quantitatively and accurately and responds linearly through at least three orders of magnitude of concentration. Moreover, this was accomplished while maintaining specificity for the Cys thiol even in the presence of 75,000-fold excess of dye (10 fmol enol and 750 pmol BODIPY FL-Mal) given that no labeling of the internal myo control was observed at any protein amount (<1%).

Impact of BODIPY modification on protein mobility in 2DGE

We have shown, as expected, that BODIPY-labeled proteins exhibit little (if any) mobility difference with unlabeled, but sulfhydryl-blocked, proteins. This is due to the uncharged nature of the dye that does not alter the protein pI on modification. We might have expected a small impact on size-based mobility (the second dimension); however, because this is dependent on the net binding of SDS to the protein (typically 1.4 g/g protein), unless this ratio is changed by the conjugation of the dye, little (if any) change should be observed, and this is demonstrated in Fig. 8. Moreover, most of the match vectors that define the relative positions of the BODIPY-labeled proteins and their unlabeled...
counterparts fall within the average diameter of the spot itself. Clearly, the dye contributes marginally, if at all, to the two-dimensional mobility of the labeled proteins.

It should be noted that a comparison of the two gels shows unique protein spots in both gels; those that are unique to the SYPRO Ruby-stained gel indicate proteins likely to contain no Cys, whereas the BODIPY proteins that are unique indicate proteins that do not bind the SYPRO Ruby dye. More detailed investigation is required, however, to establish this more definitively.

The coincidence of mobility confers considerable additional advantage when picking the spots for MS identification. Using our approach, one can be assured that the spot designated to be picked will contain the maximum amount of protein, as compared with the “minimal” Cy dyes (DIGE) or even the “saturation” Cy dyes (saturation DIGE), where significant mobility differences occur between labeled and unlabeled protein spots.

**Removal of excess dye**

In many cases, it is desirable to perform the two-dimensional or other analysis on a crude extract to minimize protein loss and to ensure that the relative amounts of all the proteins remain unaltered. Several strategies were found to reduce extraneous fluorescence in crude preparations. After labeling is complete and the reaction has been stopped with 2-ME, any excess dye is present as the 2-ME conjugate and perhaps other components of the reaction mixture. Because the non-protein-associated dye intercalates into the SDS micelle sufficiently so that it migrates to the bottom of the gel, one strategy is to prolong the run time. Although this strategy risks losing low-molecular-weight proteins, the use of high acrylamide concentrations (≥14%) can give a good margin of safety for all but very small proteins. Proteins as small as 10 kDa were still resolved at some distance from the bottom of the gel (note the lowest molecular weight marker in Figs. 5A and 6A). Alternatively, free dye may be readily removed by washing the gel with 10% ethanol prior to imaging. The solution must be changed often or else the free dye may diffuse into the gel and impart a high background to the image. We have found that when this occurs, a useful method for reducing the diffuse fluorescence sometimes evident throughout the gel is acidification of the reaction mix prior to SDS–PAGE. Besides yielding lower background fluorescence in SDS gels, this treatment prevents further N-alkylation and we routinely acidified our samples after stopping with 2-ME.

In some cases, it is desirable to remove the excess dye before further analysis. In general, excess dye could be efficiently removed by gel filtration on PD-10 columns in either sodium bicarbonate at pH 8.3 or citrate at acidic pH (data not shown). Both the concentrated fluorescence at the bottom and the diffuse fluorescence sometimes observed in our gels were effectively removed.

Conventional dialysis of crude-labeled preparations invariably resulted in precipitation of the sample. Dialysis against 8 M urea in Spectrapor Float-A-Lyzer mini-dialysis units avoided precipitation of the sample, but the lengthy exposure to urea at room temperature increases the likelihood of carbamylation with resultant alteration of isoelectric points and horizontal displacement of spots on two-dimensional gels.

This study is an extension of our previous work that surveyed the optimal fluoros and overall conditions for saturation covalent fluorescence labeling in 8 M urea. Because of the importance of thiourea for increasing protein recovery in cell lysates and its widespread use, we undertook this study to counter its confounding influence and optimize the conditions that are compatible with preseparation covalent saturation labeling in the 7 M urea/2 M thiourea chaotrope system. The information presented here provides optimized conditions for saturation labeling with maleimide-derivatized, uncharged fluorescence dyes in the presence of thiourea.

We have extrapolated this performance to a typical proteomic experiment where the Cys content was estimated by AAA. Labeling was then accomplished by the optimized protocol presented here in the presence of a 75-fold excess of BODIPY FL-Mal dye. In our experience, this amount of dye is well below the maximum solubility of the dyes and permits a high degree of quantitative accuracy and dynamic range at a very affordable cost. We are continuing this effort, and our results with complex mixtures of proteins, multiple dyes, and internal protein standards will be published in the future.

**Acknowledgments**

The authors acknowledge the excellent technical assistance in generating the data for Fig. 7 by Susan Stafford of the Biomolecular Resource Facility at the University of Texas Medical Branch. This work was supported in part by NIH/NHLBI contract N01-HV-28184 to the University of Texas Medical Branch (A. Kurosky, principal investigator) and NHLBI SCCOR P50HL083794-01 to the University of Texas Health Science Center at Houston (D. Milewicz, principal investigator). The authors also acknowledge support from Lynx Therapeutics, where the bulk of this work was performed.

**References**


