Protein Dynamics

DOI: 10.1002/anie.200803378

Isotope-Coded Affinity Tags with Tunable Reactivities for Protein Footprinting**

*Eric S. Underbakke, Yimin Zhu, and Laura L. Kiessling**

Fundamental to our understanding of biological systems is the ability to define interaction surfaces and conformational changes in dynamic protein complexes, yet this objective remains a challenge. Protein footprinting is a powerful solution to this problem. Methods such as hydrogen–deuterium exchange,[1] limited proteolysis,[2] and radiolytic cleavage[3] assess changes in protein surface exposure by measuring the susceptibility of the polypeptide backbone to modification. A complementary approach is to use side-chain modification. Because of their unique nucleophilicity, cysteine (Cys) residues at select positions can report on solvent accessibility and local chemical environment.[4] The reactivity of a Cys residue can be measured to define interaction surfaces at individual amino acid resolution. Moreover, methods have been developed for rapid generation of a library of Cys variants for a protein of interest.[4a]

In principle, Cys footprinting is amenable to mapping accessibility changes for diverse applications, including monitoring folding, ligand binding, or protein interaction interfaces. For such applications, the accessibility of a residue must be correlated with its rate of reaction. Unfortunately, the intrinsic reaction rate of Cys residues can vary widely.[4a,5] Surface or deeply buried residues can undergo alkylation at rates too high or low, respectively, to yield useful information. As a consequence, a limited number of Cys substitutions can serve as reporters. We reasoned that electrophilic footprinting reagents with diverse reactivities could increase both the utility and information content of Cys footprinting. Here, we report the design and synthesis of isotope-coded affinity tag (ICAT) reagents with a range of alkylation rates. We demonstrate that these reagents can be used to map the accessibility of protein residues in diverse environments. Most notably, they can be used to footprint a protein–protein interaction.

Our approach relies on the use of electrophilic ICAT reagents (Figure 1).[4a] ICAT reagents are trifunctional molecules composed of an electrophilic moiety, an affinity tag, and a stable isotope (e.g., ^{12}C or ^{13}C) mass tag.[6] Pairs of heavy and light ICAT reagents can be used to measure relative label incorporation at a given position using mass spectrometry. The affinity tag allows selective enrichment of ICAT-labeled peptides from a complex mixture. Traditional footprinting approaches are limited by protein size and sample complexity, because they are incompatible with high background signals. Because of its capacity for extensive sample refinement, we envisioned that the ICAT approach would have tremendous potential for investigating protein–protein interactions in complex milieu.

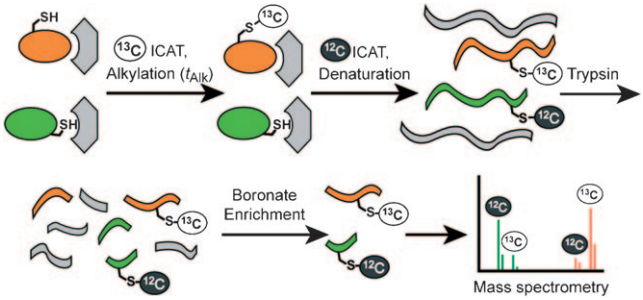


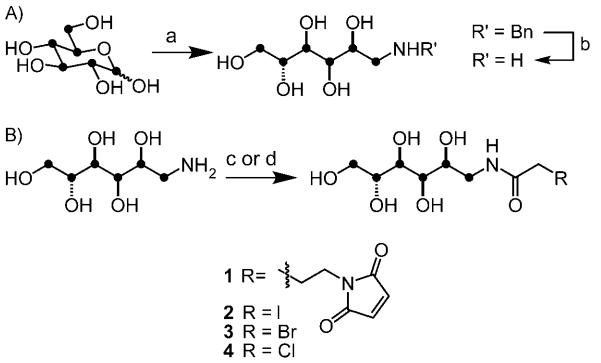
Figure 1. Cys alkylation rates are determined by labeling with a heavy ^{13}C ICAT for time t_{Alk} and then with a light ^{12}C ICAT. After labeling, samples are digested with trypsin, modified peptides are enriched on affinity resin, and the resulting mixtures analyzed by mass spectrometry. Relative abundance of the heavy and light labels at various t_{Alk} define alkylation rate profiles.

Protein footprinting requires an ICAT reagent that is both small and water-soluble. A carbohydrate-based affinity tag that binds tightly to boronate affinity resin, confers excellent water solubility, and serves as the isotope-coding portion of the reagent has been described.[4a] This ICAT reagent appeared to meet our needs because it had been used to detect changes in Cys accessibility in folded proteins. Unfortunately, we found that it had several drawbacks. Specifically, it possesses a chloroaromatic substituent, which complicates sample analysis by mass spectrometry. Additionally, we found its production to be inefficient and low-yielding. To address these issues, we used glucamine as our key building block. From this readily accessible precursor, a series of ICAT reagents (**1**–**4**) was assembled (Scheme 1). Specifically, the ^{12}C compounds were generated in one step from commercially-available glucamine by acylation. The heavy reagents were synthesized from ^{13}C-labeled glucose. Reductive amination with benzylamine afforded *N*-benzylglucamine,[7] which was subjected to palladium-catalyzed hydrogenolysis to yield ^{13}C-labeled glucamine.[8] Acylation of heavy glucamine afforded the ^{13}C-ICAT reagents **1**–**4** in good overall yields.

[*] E. S. Underbakke, Dr. Y. Zhu, Prof. L. L. Kiessling
Departments of Chemistry and Biochemistry
University of Wisconsin, Madison, Madison, WI 53706 (USA)
Fax: (+1) 608-265-0764
E-mail: Kiessling@chem.wisc.edu

[**] This research was supported by the NIH (GM55984). E.S.U. thanks the Molecular Biosciences Training Grant for support (GM07215). We thank R. E. Grant for synthetic assistance and Dr. E. E. Carlson for helpful discussions.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200803378.



Scheme 1. Conditions: a) $BnNH_2$, borane·pyridine, $(CH_3)_3CCO_2H$, $MeOH/H_2O$ (1:1), 40 °C, 81 %. b) $Pd(OH)_2/C$, $(CH_3)_3CCO_2H$, MeOH, 55 °C, 96 %. c) 4-Maleimidobutyric acid, HBTU, HOBt, DIEA, DMF, RT, 4 h, 56 %. d) $(RCH_2CO)_2O$, DMF, 0 °C, 12 h. Yields: **2**, 81 %; **3**, 89 %; **4**, 86 %. Dots denote ^{12}C for light ICAT and ^{13}C for heavy ICAT. HBTU = *O*-(benzotriazol-1-yl)-tetramethyluronium hexafluorophosphate, HOBt = 1-hydroxybenzotriazole, DIEA = diisopropylethylamine.

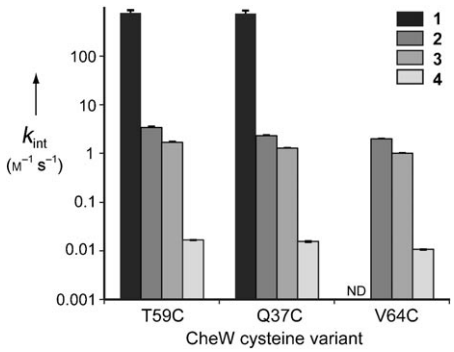


Figure 2. The intrinsic rates of alkylation for ICAT reagents **1**–**4**. ICAT reagent pairs were designed to afford a broad range of alkylation rates. Intrinsic alkylation rates (k_{int}) were determined using representative Cys variants of CheW under denaturing conditions (4 M guanidinium chloride). Data were fit to a first-order exponential decay to derive rate constants. Error bars represent the standard deviation of three independent rate determinations.

ICAT reagents **1**–**4** are based on a common protein modification strategy: the alkylation of Cys residues with maleimide or α-haloacetamide derivatives. We anticipated that a direct comparison of the intrinsic reaction rates of the relevant electrophiles would provide fundamental guidelines for protein footprinting and protein modification, in general. Accordingly, we determined the intrinsic alkylation rates (k_{int}) of compounds **1**–**4** by footprinting representative Cys variants of a denatured protein. These conditions were chosen to provide rates representative of fully solvent-exposed Cys residues.

Our choice of footprinting target, the adaptor protein CheW, was driven by an interest in the role of protein–protein interactions in *Escherichia coli* chemotactic signal transduction.[9] Several Cys variants (Q37C, T59C, and V64C) of CheW were produced, and footprinting reactions were performed under denaturing conditions. Alkylation was initiated with heavy (^{13}C) ICAT and quenched at various timepoints with dithiothreitol; samples were then counter-labeled with light (^{12}C) ICAT (see Supporting Information). The k_{int} values determined for iodoacetamide **2** (ranging from 3.4–2.0 $M^{-1}s^{-1}$) are similar to a value determined independently (Figure 2).[4a] The rates of reaction of bromoacetamide **3** were approximately two-fold lower (1.7–1.0 $M^{-1}s^{-1}$), while the rates of reaction of chloroacetamide **4** were approximately 100-fold lower (0.017–0.011 $M^{-1}s^{-1}$). We also devised a reagent that reacts more rapidly; the maleimide-ICAT **1** provided k_{int} values ranging from 780 to 750 $M^{-1}s^{-1}$. Thus, the k_{int} values of our ICAT kinetic series cover approximately five orders of magnitude, ranging from the extremely rapid **1** to the relatively slow **4**.

To survey the range of reactivity of Cys in a folded protein, we footprinted CheW variants in which Cys residues were engineered into diverse environments (Figure 3a). In principle, such substitutions can perturb native protein folding and function, but Cys replacement is generally conservative.[10] Indeed, Cys substitution was well tolerated at the five residues probed; the phosphotransfer activity of chemotaxis signaling complexes reconstituted with these

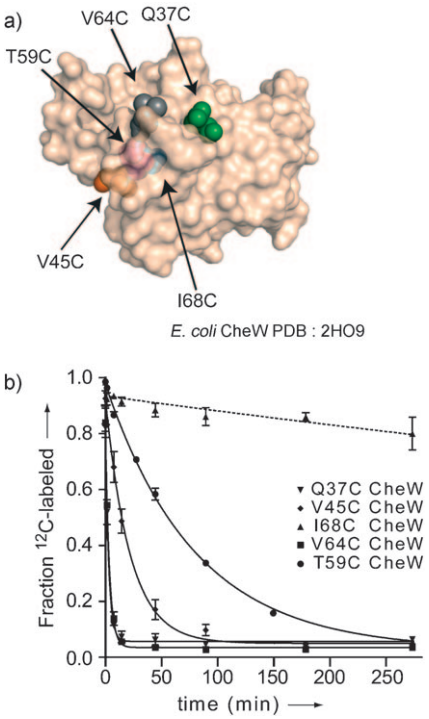


Figure 3. The rates of alkylation of Cys residues in different environments. a) Cys variants of CheW were generated at residues with different degrees of solvent accessibility. b) The five Cys variants of CheW were footprinted in the native, folded state (26 °C, pH 8.0). Footprinting timecourses were fit to a first-order exponential decay. Due to incomplete alkylation, I68C data are connected by a dotted line approximating the fit. Error bars represent the standard deviation of three independent timecourse experiments.

CheW variants was preserved (data not shown). These CheW Cys variants were pooled and footprinted with bromoacetamide reagent **3**, the ICAT reagent with intermediate reactivity. Alkylation rates ranged from 1.0 $M^{-1}s^{-1}$ for the highly solvent-exposed Q37C to less than 0.001 $M^{-1}s^{-1}$ for the core residue I68C (Figure 3b). In the I68C CheW

variant, the rate of alkylation was too low to fit to a kinetic model; therefore, this alkylation rate represents an upper limit. Overall, the selected CheW residues exhibit reactivities ranging over four orders of magnitude. These reaction rate differences are consistent with previous studies in which Cys alkylation rates within native, folded proteins can span three to six orders of magnitude.[4a,5]

Using standard reagents, researchers have noted that buried residues are problematic for footprinting.[12] The extremely long reaction times required are impractical, and they also risk protein instability and reagent cross-reactivity. We envisioned that the rapidly-alkylating reagent **1** could overcome the resistance of buried residues to alkylation. To test this hypothesis, we focused on Ile68, which resides in the core of CheW (Figure 4a).[11a] As expected, attempts to

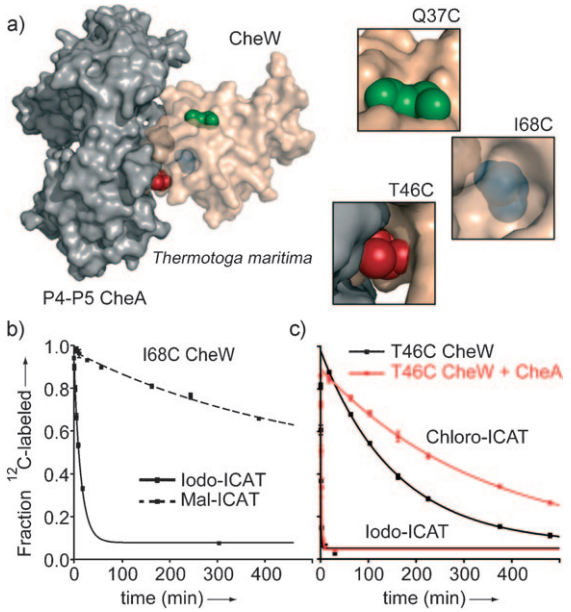


Figure 4. Footprinting CheW Cys residues with ICAT reagents that span a broad range of reactivity. a) The structure of *T. maritima* CheW bound to a truncation of CheA[11b] was used to identify an *E. coli* CheW residue buried by folding (Ile68), engaged in a protein–protein interaction (Thr46), or unaffected (Gln37). b) Timecourse for the alkylation of buried CheW residue I68C (pH 8.0, 26 °C) with iodoacetamide-based ICAT **2** or the more reactive agent **1**. c) CheA*-induced differences in alkylation rates of CheW surface residue T46C probed with **2** or **4** (pH 8.0, 26 °C). Error bars represent the standard deviation of three independent footprinting timecourses. Error bars smaller than the timepoint symbol are not shown.

footprint this residue with either iodoacetamide **2** or bromoacetamide **3** were unsuccessful (Figure 3b, 4b). In contrast, alkylation with reagent **1** was complete within two hours, and a measurable alkylation half-life ($t_{1/2}$) of 11 min was obtained. Thus, reagent **1** provides the means to monitor changes in the accessibility of a deeply buried residue.

Previous attempts to use Cys alkylation to detect protein–protein interactions have highlighted the difficulty of detecting dynamic interactions using the standard alkylating agent iodoacetamide.[13] Surface residues that define interaction regions often alkylate rapidly. Additionally, protein binding typically decreases Cys reactivity less than five-fold,[4c] and these subtle changes are below the detection limits of current methods.[13] Indeed, applications of Cys alkylation to define protein–protein interactions are rare. In one instance, iodoacetamide alkylation was used to map a high affinity antibody–antigen interaction,[4a] but this method failed with a weaker protein–protein interaction ($K_D \approx 1$ μM).[13] In another instance, a weak protein–protein interaction was detected, but only when an extremely high concentration (> 1 mM) of the protein binding partner was employed.[4c] We reasoned that an effective approach would be to modulate the rate of alkylation, so that subtle rate changes in the reaction of surface residues could be amplified and detected. Our fundamental studies of Cys alkylation rates suggested that ICAT reagent, **4**, with a relatively low k_{int} could yield valuable results.

To test this strategy, we examined a protein–protein complex involved in signal transduction: the interaction of CheA with CheW. CheW mediates a functionally important bridging interaction between the transmembrane chemoreceptors and the kinase CheA.[14a] Ternary complex formation couples histidine kinase activity to the ligand occupancy of chemoreceptors.[14] The structure of the *Thermotoga maritima* CheW with a portion of CheA has been determined,[11] thereby providing a context for interpreting footprinting results. The dissociation constant of this complex ($K_D \approx 17$ μM)[14a] is in the range typical for those of dynamic protein–protein interactions involved in signal transduction. Cys residues were engineered at a CheW position predicted to be buried by CheA binding (T46C) and one that should be unaffected (Q37C).[11b] These CheW Cys variants were pooled and footprinted in the presence and absence of a His_6-tagged, Cys-free CheA variant (CheA*). Using the ICAT affinity tag, the alkylated products of protease digestion were enriched prior to analysis by mass spectrometry. Because of the high background from the macromolecular binding partner, this enrichment step was essential. As predicted, the $t_{1/2}$ at residue Q37C showed no significant change in the presence of CheA* when footprinted with either reagent **2** or **4** (Supporting Information). In native CheW, Thr46 is surface-exposed; it therefore undergoes rapid alkylation in the presence of reagent **2** ($t_{1/2} = 40$ s, Figure 2c). When CheA* was present, the alkylation rate appeared to low, but the difference was not obvious. In contrast, when reagent **4** is used for footprinting, the addition of CheA results in an increase in the alkylation $t_{1/2}$ that exceeds two hours. Thus, the ICAT reagent **4** reveals that CheW residue Thr46 is buried upon CheA binding. This finding demonstrates the advantages of our approach for footprinting protein–protein interactions.

The ICAT reagents described herein are valuable new tools for protein footprinting. With this toolkit and information about the relative reactivities of each reagent, highly solvent-exposed residues as well as deeply buried residues can serve as reporters. The utility of these reagents is highlighted by our demonstration that protein footprinting can be applied to complexes of modest affinity. This finding is significant because it is the transient, dynamic protein complexes that are of interest in biological processes. We anticipate that our

results will expand the application of ICAT footprinting to map and monitor diverse protein interaction surfaces.

Received: July 11, 2008
Revised: September 3, 2008
Published online: October 31, 2008

Keywords: alkylation · footprinting · isotope-coded affinity tag · mass spectrometry · protein modification

[1] a) C. A. Hughes, J. G. Mandell, G. S. Anand, A. M. Stock, E. A. Komives, *J. Mol. Biol.* **2001**, *307*, 967 – 976; b) M. M. G. Krishna, L. Hoang, Y. Lin, S. W. Englander, *Methods* **2004**, *34*, 51 – 64; c) D. L. Smith, Y. Z. Deng, Z. Q. Zhang, *J. Mass Spectrom.* **1997**, *32*, 135 – 146.

[2] a) S. Dokudovskaya, R. Williams, D. Devos, A. Sali, B. T. Chait, M. P. Rout, *Structure* **2006**, *14*, 653 – 660; b) M. Zhong, L. Lin, N. R. Kallenbach, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2111 – 2115; c) A. Scaloni, N. Miraglia, S. Orru, P. Amodeo, A. Motta, G. Marino, P. Pucci, *J. Mol. Biol.* **1998**, *277*, 945 – 958.

[3] a) G. Z. Xu, M. R. Chance, *Anal. Chem.* **2005**, *77*, 4549 – 4555; b) R. T. Liu, J. Q. Guan, O. Zak, P. Aisen, M. R. Chance, *Biochemistry* **2003**, *42*, 12447 – 12454.

[4] a) J. A. Silverman, P. B. Harbury, *J. Biol. Chem.* **2002**, *277*, 30968 – 30975; b) J. H. Ha, S. N. Loh, *Nat. Struct. Biol.* **1998**, *5*, 730 – 737; c) O. Abramczyk, M. A. Rainey, R. Barnes, L. Martin, K. N. Dalby, *Biochemistry* **2007**, *46*, 9174 – 9186; d) C. P. Johnson, H. Y. Tang, C. Carag, D. W. Speicher, D. E. Discher, *Science* **2007**, *317*, 663 – 666; e) J. L. Apuy, Z. Y. Park, P. D. Swartz, L. J. Dangott, D. H. Russell, T. O. Baldwin, *Biochemistry* **2001**, *40*, 15153 – 15163; f) B. P. Tu, J. C. Wang, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4862 – 4867.

[5] M. H. Jacob, D. Amir, V. Ratner, E. Gussakowsky, E. Haas, *Biochemistry* **2005**, *44*, 13664 – 13672.

[6] a) R. Aebersold, D. R. Goodlett, *Chem. Rev.* **2001**, *101*, 269 – 296; b) S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nat. Biotechnol.* **1999**, *17*, 994 – 999.

[7] T. Yoshida, Y. C. Lee, *Carbohydr. Res.* **1994**, *251*, 175 – 186.

[8] R. C. Bernotas, R. V. Cube, *Synth. Commun.* **1990**, *20*, 1209 – 1212.

[9] a) J. E. Gestwicki, L. L. Kiessling, *Nature* **2002**, *415*, 81 – 84; b) A. C. Lamanna, G. W. Ordal, L. L. Kiessling, *Mol. Microbiol.* **2005**, *57*, 774 – 785; c) M. J. Borrok, E. M. Kolonko, L. L. Kiessling, *ACS Chem. Biol.* **2008**, *3*, 101 – 109.

[10] Cys footprinting can be applied without determining that Cys substitution preserves function, but there are caveats in interpreting the data that result. Specifically, disruptive Cys substitutions can appear as surface residues, as they will most likely result in an unstructured protein region.

[11] a) Y. Li, Y. F. Hu, W. Y. Fu, B. Xia, C. W. Jin, *Biochem. Biophys. Res. Commun.* **2007**, *360*, 863 – 867; b) S. Y. Park, P. P. Borbat, G. Gonzalez-Bonet, J. Bhatnagar, A. M. Pollard, J. H. Freed, A. M. Bilwes, B. R. Crane, *Nat. Struct. Mol. Biol.* **2006**, *13*, 400 – 407.

[12] J. A. Silverman, P. B. Harbury, *J. Mol. Biol.* **2002**, *324*, 1031 – 1040.

[13] A. S. Burguete, P. B. Harbury, S. R. Pfeffer, *Nat. Methods* **2004**, *1*, 55 – 60.

[14] a) J. A. Gegner, F. W. Dahlquist, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 750 – 754; b) J. A. Gegner, D. R. Graham, A. F. Roth, F. W. Dahlquist, *Cell* **1992**, *70*, 975 – 982.

