Published in final edited form as:

Angew Chem Int Ed Engl. 2017 November 06; 56(45): 14007–14010. doi:10.1002/anie.201706697.

Laser-initiated Radical Trifluoromethylation of Peptides and Proteins and Its Application to Mass Spectrometry-Based Protein Footprinting

Ming Cheng, Bojie Zhang, Weidong Cui, and Michael L. Gross

Department of Chemistry, Washington University in St. Louis, One Brookings Drive, Saint Louis, Missouri 63130

Abstract

We describe a novel, laser-initiated radical trifluoromethylation for protein footprinting and establish its broad residue coverage. $\bullet\text{CF}_3$ reacts with 18 of 20 common amino acids including Gly, Ala, Ser, Thr, Asp, Glu that are relatively “silent” with $\bullet\text{OH}$. This new approach to footprinting is a bridge between trifluoromethylation in materials and medicinal chemistry and structural biology and biotechnology. Its application to a membrane protein and to myoglobin show that the approach is sensitive to protein conformational change and solvent accessibility.

Graphical abstract

We report a novel, laser-initiated trifluoromethylation method for protein footprinting. The trifluoromethyl radical reacts with 18 of 20 common amino acids. This highly reactive radical can probe the conformational changes and the solvent accessibility of proteins, thus serving as a “footprinter” with high hydrophobicity. These results bode well for novel applications of trifluoromethylation as a structural-biology tool.

Keywords

trifluoromethylation; radical reactions; protein footprinting; membrane proteins; mass spectrometry; FPOP

Protein footprinting with reactive species ($\bullet\text{OH}$,^[1] carbenes^[2]) is becoming effective in structural proteomics because it modifies proteins before they undergo conformational change. Although $\bullet\text{OH}$ reacts with all amino acids,^[3] rate constants span nearly three orders of magnitude, indicating that less-reactive residues (e.g., Gly, Ala, Asp, Asn, Ser, Thr) are not measurably footprinted whereas reactive sulfur and aromatic amino acids are preferentially modified. Using epitope mapping as an application, we see that Phe and Cys, which are highly reactive with $\bullet\text{OH}$ have the lowest probability to be in an interface, whereas Gly, Ala, Asp, Asn, Ser, Thr are likely candidates for an epitope.^[4] Thus, no single reactive species will be a panacea for the wide variety of protein footprinting requirements, necessitating development of other reactive species.

Protein footprinting can give coarse-grained protein structure particularly when NMR, X-Ray, and cryoEM fail.^[5] Hydrogen deuterium exchange is a reversible and nearly universal approach to footprinting^[6] but requires fast workup to minimize back exchange. Complementary are chemical reagents that irreversibly react with specific amino-acid side chains,^[7] but reactions are slow and not general. Reactive-species footprinting (prototype is $\bullet\text{OH}$) provides general and fast reactions.^[8] Fast photochemical oxidation of proteins (FPOP) (Fig.S1, SI p 3) affords a ready opportunity for broad-based implementation especially if the number of reactive species can be expanded to give broad coverage.^[9]

Trifluoromethylation (CF_3) may serve this need. Its chemistry in pharmaceuticals, agrochemicals, and specialty materials^[10] show that the addition of CF_3 can affect profoundly physical properties, reactivity, biological activity, solvency, and stability. Although fluorine is nearly absent in the biosphere and CF_3 is difficult to incorporate, recent advances in trifluoromethylation may be transferrable to biomolecules and attain high importance in protein biochemistry.^[11]

Radical trifluoromethylation via photolysis can be traced back to Golitz.^[12] Since then, the groups of Meijere,^[12] Langois,^[13] MacMillan,^[14] Baran^[15] and others,^[16] have increased the scope and generality of trifluoromethylation. Macmillan et al.^[14] showed that radical trifluoromethylation can occur with photoredox-catalysis. These approaches use metals and organic solvents, not the medium of biology. Taking a lead from Baran,^[17] we report a means to generate and utilize $\bullet\text{CF}_3$ in aqueous buffer solutions by using a pulsed laser suitable for fast protein footprinting.

Specifically, we implemented laser-initiated trifluoromethylation, demonstrating footprinting a peptide and three proteins by using the FPOP, a platform that can follow protein conformational changes, protein folding, in-cell labeling, and amyloid formation, identify protein/protein interfaces, and accommodate new footprinting reagents.^[18]

We begin with trifluoromethylation of neuropeptide Y 18-36 (NPY18-36) as a model. We actualized the transformation by exposing NPY18-36 to laser irradiation in the presence of H_2O_2 and the water-soluble salt, NaSO_2CF_3 (**Entries 5, 6, 7, 8**, Table 1, SI page 4), but found no detectable CF_3 -products in the absence of either peroxide or laser irradiation (**Entries 2, 4**). We also removed buffer anions, but to no avail (Compare **Entries 1, 3**). With systematic screening, we ultimately found conditions where peptide modification is extensive.

The key to pulsed formation of $\bullet\text{CF}_3$ is the hydroxyl radical, formed by photolysis of $\text{HO}\cdot\text{OH}$ (pathway **a**). $\bullet\text{OH}$ oxidizes CF_3SO_2^- to $\text{CF}_3\text{SO}_2\bullet$, which rapidly collapses to generate $\text{CF}_3\bullet$ via an entropy-driven loss of SO_2 . Given that the concentration of CF_3SO_2^- is ~ 300 times that of $\bullet\text{OH}$,^[19] the $\bullet\text{OH}$ is readily quenched by the oxidation. The electron-deficient $\bullet\text{CF}_3$ then reacts with the protein (pathway **b**) by addition of CF_3 or by formation of a protein radical that is quenched by another CF_3 . In a minor pathway (**c**), the $\bullet\text{OH}$ abstracts $\text{H}\bullet$ from the protein to produce a protein radical that either couples with $\bullet\text{CF}_3$ to provide a CF_3 -modified product, or reacts with $\text{HO}\bullet$ or dissolved oxygen to generate oxygen-oxidization (see SI Page 6), but CF_3 -substitutions dominate. An inquiry into mechanism is forthcoming but is beyond the scope of this announcement. If these pathways are correct, any method that uses OH radicals (e.g., synchrotron radiation) should be able to utilize this chemistry.

Reaction energetics, available at least for gas-phase processes, offer perspective. $\text{CF}_3\bullet$ reacts exothermically with methane to abstract an $\text{H}\bullet$. Furthermore, reaction of $\bullet\text{CF}_3$ with $\bullet\text{CH}_3$ (mimic for a protein radical) is also highly exothermic (details in SI, page 29).

We investigated three proteins-myoglobin, vitamin K epoxide reductase (VKOR) membrane protein and beta-lactoglobulin. The optimized conditions allow $\bullet\text{CF}_3$ to react with native proteins, even in detergent media. The sites were identified by LC-MS/MS after tryptic digestion by using a canonical proteomics strategy. On the basis of product-ion spectra (MS/MS), $\bullet\text{CF}_3$ reacts with 18 different amino-acid residues (Fig. 2b, SI Page 11–28 for MS/MS data).

Taking advantage of the unique mass shift of +67.9870 detected by MS/MS, we illustrate an unexpected CF_3 substitution on Gly-276 of tryptic peptide 267–283 from the membrane protein VKOR (Fig. 2c). Remarkably, direct Csp3 C-H trifluoromethylation occurs at Gly, Ala, Val, Ile, and Leu, showing a unique trifluoromethylation without Lewis-acid catalysis^[20] or a cross-dehydrogenative coupling (CDC).^[21] Given that the side chains of Gly, Ala, Ser, Thr, Asp, Glu are relatively $\bullet\text{OH}$ silent, $\bullet\text{CF}_3$ fulfills our goal to complement OH footprinting with a broadly reactive reagent with widely different properties.

We next determined whether $\bullet\text{CF}_3$ shows reactivity patterns consistent with protein structure and solvent accessibility (i.e., is a good “footprinter”?). Membrane proteins are challenging for footprinting because they are insoluble in water without detergent, and with detergent, only a fraction of residues are solvent accessible. VKOR in pH 7.4 PBS buffer with 0.03% DDM detergent is soluble. Photolysis in the presence of NaSO_2CF_3 leads to outer membrane footprinting at 10 residues (Y204, Y207, W208, P210, L271, A272, A274, G276, Y277,

P278) in the fused extra-membrane, Trx-like domain (Fig. 2a).^[22] The selectivity for the solvent-exposed region of VKOR is due to the water solubility of the radical-precursor salt, which cannot permeate the transmembrane domains even though $\bullet\text{CF}_3$ is hydrophobic. Membrane proteins are today's challenge because they occur in heterogeneous media and are water insoluble. Incorporation of fluorinated alkyl groups in membrane-soluble reagents should allow inner membrane footprinting.

We then asked whether $\bullet\text{CF}_3$ provides a meaningful footprint of protein conformational change. We choose holo/apo-myoglobins (hMb/aMb) because they are structurally different, as demonstrated many times. Accurate masses of tryptic modified and unmodified peptides in extracted-ion chromatograms (EIC, see SI, p 8–10) are evidence that trifluoromethylation affords reliable outcomes. Apo/hole comparisons reveal similar extents of CF_3 -modification within the A–E, G and H helices, but dramatically different modification extents for regions 80–96 (hMb 6.6% vs aMb 89.0%), which is a helix (F-Helix) only for the holo (Fig. 3a). The results are consistent with top-down HDX^[23] and NMR.^[24] Readily detected modifications in regions 1–16 (hMb 23% vs. aMb 19%), 32–45 (hMb 3.1 % vs. aMb 4.0 %), 119–133 (hMb 2.6% vs. aMb 11%), 134–139 (hMb 3.8% vs. aMb 2.8%) are controls. Product-ion spectra, show a total of 17 modified residues, and those of histidine are consistent with a transition from unstructured to structured for aMb to hMb, a pattern also found in iodination footprinting.^[25] For other aromatic residues, we found relatively extensive labeling at W7, W14 and F43, indicating good $\bullet\text{CF}_3$ radical reactivity toward aromatic rings (Fig. 3b).

With a single CF_3 replacement for F43 or H36, there is a significant shift to longer retention times in reversed-phase HPLC (Fig. 4), owing to the increased hydrophobicity of CF_3 -modified peptides. This sharply contrasts with the hydrophilic reagents of other footprinting. Indeed, the hydrophobic $\bullet\text{CF}_3$ points to footprinting within transmembrane domains by hydrophobic precursors of $\bullet\text{CF}_3$ inter alia equilibrated into a membrane and submitted to FPOP. A challenging example is locating steroid binding sites in receptor proteins, recently done by affinity photolabeling.^[26]

Our structural-proteomics approach introduces for the first time the bioorthogonal CF_3 group onto side chains of proteins, setting the stage for applications in biophysics and biotherapeutics, perhaps even for intransigent human membrane proteins. Further, because $\bullet\text{CF}_3$ can be produced in biological buffers, one may be able to synthesize fluorine-containing proteins for positron emission tomography (PET) tracers.^[27] Second, we can extend the sulfinate chemistry to design isotope-labeling and cross-linking reagents that can be readily detected by mass spectrometry and other approaches (e.g., F-NMR). Currently, we are investigating the mechanism and generality, of protein trifluoromethylation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to PMI for the software for data processing, to Prof. Weikai Li for a sample of VKOR, and to the US National Institutes of Health (P41GM103422) for support of this work.

References

1. Wang L, Chance MR. *Mol Cell Proteomics*. 2017; 16:706. [PubMed: 28275051]
2. a) Jumper CC, Bomgardner R, Rogers J, Etienne C, Schriemer DC. *Anal Chem*. 2012; 84:4411. [PubMed: 22480364] b) Zhang B, Rempel DL, Gross ML. *J Am Soc Mass Spectrom*. 2016; 27:552. [PubMed: 26679355]
3. Xu G, Chance MR. *Chem Rev*. 2007; 107:3514. [PubMed: 17683160]
4. Ramaraj T, Angel T, Dratz EA, Jesaitis AJ, Mumey B. *Biochimica et Biophysica Acta*. 2012; 1824:520. [PubMed: 22246133]
5. a) Zhong M, Lin L, Kallenbach NR. *Proc Natl Acad Sci*. 1995; 92:2111. [PubMed: 7892233] b) Aebersold R, Mann M. *Nature*. 2003; 422:198. [PubMed: 12634793]
6. Pirrone GF, Iacob RE, Engen JR. *Anal Chem*. 2015; 87:99. [PubMed: 25398026]
7. a) Mendoza VL, Vachet RW. *Mass Spectrom Rev*. 2009; 28:785. [PubMed: 19016300] b) Leitner A. *Chem Sci*. 2016; 7:4792.
8. a) Gomez GE, Monti JL, Mundo MR, Delfino JM. *Anal Chem*. 2015; 87:10080. [PubMed: 26348271] b) Manzi L, Barrow AS, Scott D, Layfield R, Wright TG, Moses JE, Oldham NJ. *Nat Commun*. 2016; 7:13288. [PubMed: 27848959]
9. a) Maleknia SD, Brenowitz M, Chance MR. *Anal Chem*. 1999; 71:3965. [PubMed: 10500483] b) Hambly DM, Gross ML. *J Am Soc Mass Spectrom*. 2005; 16:2057. [PubMed: 16263307] c) Watson C, Janik I, Zhuang T, Charvátová O, Woods RJ, Sharp JS. *Anal Chem*. 2009; 81:2496. [PubMed: 19265387]
10. Müller K, Faeh C, Diederich F. *Science*. 2007; 317:1881. [PubMed: 17901324]
11. Furuya T, Kamlet AS, Ritter T. *Nature*. 2011; 473:470. [PubMed: 21614074]
12. Göllitz P, de Meijere A. *Angew Chem*. 1977; 12:892. *Angew Chem Int Ed*. 1977; 16:854.
13. Langlois BR, Laurent E, Roidot N. *Tetrahedron Lett*. 1991; 32:7525.
14. Nagib DA, MacMillan DW. *Nature*. 2011; 480:224. [PubMed: 22158245]
15. Ji Y, Brueckl T, Baxter RD, Fujiwara Y, Seiple IB, Su S, Blackmond DG, Baran PS. *Proc Natl Acad Sci*. 2011; 108:14411. [PubMed: 21844378]
16. a) Natte K, Jagadeesh RV, He L, Rabeah J, Chen J, Taeschler C, Ellinger S, Zaragoza F, Neumann H, Brückner A, Beller M. *Angew Chem*. 2016; 128:2832. *Angew Chem Int ed*. 2016; 55:2782. b) Li L, Mu X, Liu W, Wang Y, Mi Z, Li CJ. *J Am Chem Soc*. 2016; 138:5809. [PubMed: 27137478] c) Lefebvre Q, Hoffmann N, Rueping M. *Chem Commun*. 2016; 52:2493. d) Sladojevich E, McNeill F, Börgel J, Zheng S-L, Ritter T. *Angew Chem*. 2015; 127:3783. *Angew Chem Int Ed*. 2015; 54:3712. e) Ye Y, Küenzi SA, Sanford MS. *Org Lett*. 2012; 14:4979. [PubMed: 22984900]
17. Fujiwara Y, Dixon JA, O'Hara F, Funder ED, Dixon DD, Rodriguez RA, Baxter RD, Herle B, Sach N, Collins MR, Ishihara Y, Baran PS. *Nature*. 2012; 492:95. [PubMed: 23201691]
18. a) Li KS, Chen G, Mo J, Huang RYC, Deyanova EG, Beno BR, O'Neil SR, Tymiak AA, Gross ML. *Anal Chem*. 2017; 89:7742. [PubMed: 28621526] b) Lu Y, Zhang H, Niedzwiedzki DM, Jiang J, Blankenship RE, Gross ML. *Anal Chem*. 2016; 88:8827. [PubMed: 27500903] c) Li KS, Rempel DL, Gross ML. *J Am Chem Soc*. 2016; 138:12090. [PubMed: 27568528] d) Li Z, Moniz H, Wang S, Ramiah A, Zhang F, Moremen KW, Linhardt RJ, Sharp JS. *J Biol Chem*. 2015; 290:10729. [PubMed: 25752613] e) Espino JA, Mali VS, Jones LM. *Anal Chem*. 2015; 87:7971. [PubMed: 26146849] f) Stocks BB, Sarkar A, Wintrobe PL, Konermann L. *J Mol Biol*. 2012; 423:789. [PubMed: 22940366] g) Chen J, Rempel DL, Gau BC, Gross ML. *J Am Chem Soc*. 2012; 134:18724. [PubMed: 23075429]
19. Niu B, Zhang H, Giblin D, Rempel DL, Gross ML. *J Am Soc Mass Spectrom*. 2015; 26:843. [PubMed: 25712620]

20. Kuninobu Y, Nagase M, Kanai M. *Angew Chem*. 2015; 35:10401. *Angew Chem Int Ed*. 2015; 54:10263.
21. Chu L, Qing FL. *Chem Commun*. 2010; 46:6285.
22. Li W, Schulman S, Dutton RJ, Boyd D, Beckwith J, Rapoport TA. *Nature*. 2010; 463:507. [PubMed: 20110994]
23. Pan J, Han J, Borchers CH, Konermann L. *J Am Chem Soc*. 2009; 131:12801. [PubMed: 19670873]
24. Eliezer D, Wright PE. *J Mol Biol*. 1996; 263:531. [PubMed: 8918936]
25. Chen J, Cui W, Giblin D, Gross ML. *J Am Soc Mass Spectrom*. 2012; 23:1306. [PubMed: 22669760]
26. Yip GM, Chen ZW, Edge CJ, Smith EH, Dickinson R, Hohenester E, Townsend RR, Fuchs K, Sieghart W, Evers AS, Franks NP. *Nat Chem Biol*. 2013; 9:715. [PubMed: 24056400]
27. Ametamey SM, Honer M, Schubiger PA. *Chem Rev*. 2008; 108:1501. [PubMed: 18426240]

Figure 1.
Proposed mechanism for trifluoromethylation

Figure 2.

Site reactivity for a native protein higher order structure. **(a)** Scheme for fast trifluoromethylation of VKOR membrane proteins. CF₃-modified sites are in orange. **(b)** Venn diagram comparing •OH-reactive residues (in blue) with •CF₃-reactive residues (in yellow); residues in overlap are reactive to both OH and CF₃ radicals. •OH-silent residues are in yellow crescent region. (2c). A key mass shift of +68 Da is consistent with CF₃ - modification on Gly-276 of peptide 267–83.

Figure 3.

(a) Comparison of modification extents of constituent peptides from hMb and aMb. Significant differences in modification occur in the region represented by peptide 80–96, marked in red. hMb contains a helix region (F helix) in the region of peptide 80–96 whereas aMb doesn't form a F helix (b) Modification ratios at the residue level (His, Phe, and Tyr).

Figure 4.
EIC of wildtype and CF₃-modified peptide 32–45 of aMb