Chemically Mediated Site-Specific Cleavage of Proteins

Bixun Wang,[†] Michiel Lodder,[†] Jia Zhou,[†]
Teaster T. Baird, Jr.,[‡] Kathlynn C. Brown,[‡]
Charles S. Craik,[‡] and Sidney M. Hecht*,[†]

Departments of Chemistry and Biology
University of Virginia, Charlottesville, Virginia 22901
Department of Pharmaceutical Chemistry
University of California
San Francisco, California 94143

Received January 21, 2000

In contrast to native chemical ligation,[1] which has defined an elegant new strategy for protein synthesis via peptide bond formation at specific sites, there is a paucity of methods for site-specific cleavage of proteins. The cleavage of (semi)synthetic proteins at predetermined sites[2] would facilitate protein engineering, and studies of protein structure and function.[3] Reported herein is a novel strategy for cleaving the protein amide backbone at a single, predetermined site with a simple chemical reagent.

The strategy relies on our recent finding that the 4-pentenoyl group and certain derivatives can be used as protecting groups for N^{α} of the aminoacyl moiety in misacylated tRNAs.[4] Deprotection occurs readily by treatment with aqueous iodine at 25 °C via a presumed iodolactone intermediate. We reasoned that analogous cleavage should take place in proteins containing allylglycine (Figure 1), thus permitting site-specific protein cleavage. While not normally a protein constituent, allylglycine is not dramatically different in physicochemical properties than leucine, isoleucine, and valine; its incorporation into specific sites in proteins may be envisioned by readthrough of a nonsense codon[5] with a misacylated suppressor tRNA[6] activated with allylglycine.

Presently we demonstrate (i) site-specific cleavage of *E. coli* dihydrofolate reductase (DHFR)[7] analogues containing allylglycine at each of three predetermined sites and (ii) iodine-induced conversion of a zymogen (rat trypsinogen[8,9]) to its mature form (trypsin).

DHFR analogues having allylglycine (^{A}Gly) at predetermined sites were prepared by in vitro protein synthesis in the presence of DHFR mRNAs[5e] having a UAG codon in lieu of the codons for Asp27, Val10, or Glu-1. Inclusion of allylglycyl-$tRNA_{CUA}$[10] was essential for the synthesis of full length protein (Figure 2), indicating that allylglycine must be incorporated specifically into DHFR at the site of each UAG codon.[5e]

The ligand binding and catalytic properties of the DHFR analogues were determined.[11] Analogues Val10 ^{A}Gly and Glu-1 ^{A}Gly had the same specific activity and chromatographic properties as wild type, while Asp27 ^{A}Gly was dysfunctional as anticipated.[5e,7] Thus allylglycine could be incorporated into DHFR at positions −1 and 10 without any obvious effect on ligand binding capability or catalytic competence (Supporting Information).

Treatment of the three DHFR analogues containing allylglycine with I_2 afforded cleavage products having the expected sizes (illustrated in Figure 3 for Asp27 ^{A}Gly).[12] The cleavage site was established both by gel and capillary electrophoresis.[13] The cleavage products could be separated readily from unreacted protein by Ni-NTA chromatography (Figure 3B). While I_2-induced cleavage efficiency was expected to vary from site to site based on solvent accessibility of the allyl group, protein secondary structure at the allylglycine site seems to be a more important determinant of cleavage efficiency (Table 1). Presumably, the ease of formation of the cyclic intermediate (Figure 1) is critical to the success of the overall cleavage reaction and certain peptide conformations (e.g. α-helix) are more conducive to the requisite cyclization.[14]

Allylglycine was also introduced into rat trypsinogen at the site at which it is normally cleaved to afford mature trypsin.[8] I_2

* Address correspondence to this author.
[†] University of Virginia.
[‡] University of California.

(1) (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776. (b) Liu, C.-F.; Tam, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 4149.

(2) Lactic acid (e.g., Fahnestock, S.; Rich, A. *Science* **1971**, *173*, 340) and glycolic acid (Chung, H.-H.; Benson, D. R.; Schultz, P. G. *Science* **1993**, *259*, 806) can be incorporated into proteins at discrete sites, affording base-labile ester bonds, but the derived protein analogues contain at least one linkage that can alter the stability and enzymatic activity of the protein (Chapman, E.; Thorson, J. S.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 7151).

(3) A limited number of cleavage reagents have been reported; cleavage typically occurs at defined sequence but multiple sites (see, e.g.: (a) Lawson, W. B.; Gross, E.; Foltz, C. M.; Witkop, B. *J. Am. Chem. Soc.* **1961**, *83*, 1509. (b) Hotez, P. J.; Trang, N. L.; McKerrow, J. H.; Cerami, A. *J. Biol. Chem.* **1985**, *260*, 7343. (c) Marcello, A.; Loregian, A.; De Filippis, V.; Fontana, A.; Hirst, T. R.; Palù, G. *FEMS Microbiol. Lett.* **1996**, *136*, 39).

(4) (a) Lodder, M.; Golovine, S.; Hecht, S. M. *J. Org. Chem.* **1997**, *62*, 778. (b) Lodder, M.; Golovine, S.; Laikhter, A. L.; Karginov, V. A.; Hecht, S. M. *J. Org. Chem.* **1998**, *63*, 794.

(5) (a) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182. (b) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Wacker, D. A.; Lyttle, M. H.; Dix, T. A.; Chamberlin, A. R. *Biochemistry* **1991**, *30*, 5411. (c) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537. (d) Cornish, V. W.; Mendel, D.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 621. (e) Karginov, V. A.; Mamaev, S. V.; An, H.; Van Cleve, M. D.; Hecht, S. M.; Komatsoulis, G. A.; Abelson, J. N. *J. Am. Chem. Soc.* **1997**, *119*, 8166. (f) Short, G. F., III; Golovine, S. Y.; Hecht, S. M. *Biochemistry* **1999**, *38*, 8808.

(6) (a) Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. *J. Biol. Chem.* **1978**, *253*, 4517. (b) Heckler, T. G.; Zama, Y.; Naka, T.; Hecht, S. M. *J. Biol. Chem.* **1983**, *258*, 4492. (c) Heckler, T. G.; Chang, L.-H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M. *Biochemistry* **1984**, *23*, 1468. (d) Baldini, G.; Martoglio, B.; Schachenmann, A.; Zugliani, C.; Brunner, J. *Biochemistry* **1988**, *27*, 7951. (e) Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R.; Diala, E. S. *J. Am. Chem. Soc.* **1989**, *111*, 8013. (f) Robertson, S. A.; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Res.* **1989**, *17*, 9649. (g) Noren, C. J.; Anthony-Cahill, S. J.; Suich, D. J.; Noren, K. A.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Res.* **1990**, *18*, 83. (h) Hecht, S. M. *Acc. Chem. Res.* **1992**, *25*, 545.

(7) (a) Baccanari, D.; Phillips, A.; Smith, S.; Sinski, D.; Burchall, J. *Biochemistry* **1975**, *14*, 5267. (b) Johnson, L. F. In *Folates and Pterins*; Blakley, R. L., Benkovic, S. J., Eds.; Wiley: New York, 1984; Vol. 1, p 581. (c) David, C. L.; Howell, E. E.; Farnum, M. F.; Villafranca, J. E.; Oatley, S. J.; Kraut, J. *Biochemistry* **1992**, *31*, 9813. (d) Lee, H.; Reyes, V. M.; Kraut, J. *Biochemistry* **1996,** *35*, 7012.

(8) The rat trypsinogen used in this study is a variant of native rat trypsinogen[9] having the peptide sequence MGHHHHHHGGG^{A}G in place of the wild-type activation peptide. A UAG codon was included at position −1, immediately prior to the authentic trypsin coding region. The hexahistidine moiety facilitated purification of the derived trypsinogen analogue on Ni-NTA agarose. (Janknecht, R.; de Martynoff, G.; Lou, J.; Hipskind, R. A.; Nordheim, A.; Stunnenberg, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 8972.)

(9) (a) Craik, C. S.; Largman, C.; Fletcher, T.; Roczniak, S.; Barr, P. J.; Fletterick, R.; Rutter, W. J. *Science* **1985**, *228*, 291. (b) Graf, L.; Craik, C. S.; Patthy, A.; Roczniak, S.; Fletterick, R. J.; Rutter, W. J. *Biochemistry* **1987**, *26*, 2616. (c) Perona, J. J.; Hedstrom, L.; Wagner, R. L.; Rutter, W. J.; Craik, C. S.; Fletterick, R. J. *Biochemistry* **1994**, *33*, 3252.

(10) Allylglycyl-$tRNA_{CUA}$ was prepared and characterized in analogy with other misacylated tRNAs (Supporting Information).[6] To demonstrate the versatility of this misacylated tRNA, substitution at positions 10 and 27 in DHFR was carried out using rabbit reticulocyte lysate,[5e] while Glu-1-^{A}Gly was elaborated in an *E. coli* S30 system.[5f]

(11) Catalytic competence was determined by oxidation of NADPH (monitored by the decrease in absorption at 339 nm).[7a] The ability of DHFR analogues to bind to Ni-NTA agarose and methotrexate-agarose was used to determine whether each analogue (i) contained a hexahistidine moiety at the N terminus and (ii) folded in a fashion similar to wild type, respectively.

(12) The optimal iodine concentration for protein cleavage differed slightly among DHFR analogues containing ^{A}Gly at different positions (Table 1). Wild-type DHFR, and analogues Asp27Val and Glu-1Val gave no cleavage products under the same conditions (Table 1).

(13) For DHFR analogue Glu-1^{A}Gly, I_2 treatment afforded a protein that comigrated with an authentic standard by high-resolution PAGE and native capillary electrophoresis.

(14) The actual distances between the relevant carboxamide backbone residues and side chain carbon atoms determined crystallographically for DHFR support this interpretation (data not shown).

10.1021/ja0002262 CCC: $19.00
Published on Web 07/12/2000

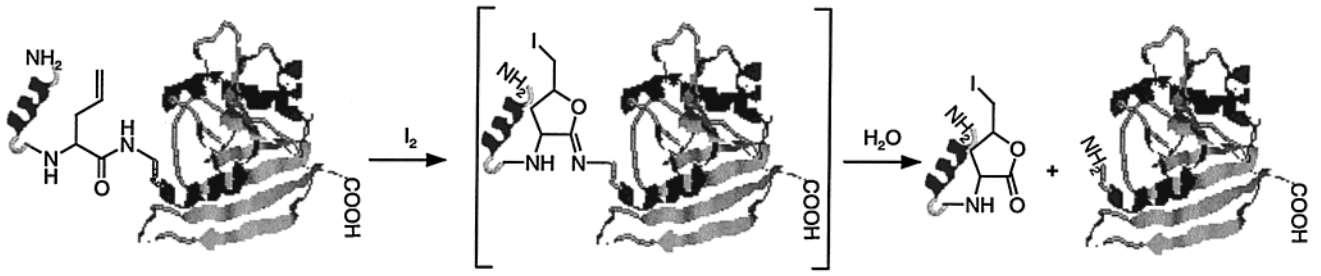

Figure 1. Scheme illustrating how treatment of an allylglycine-containing protein with I_2 can lead to site-specific cleavage of the protein backbone.

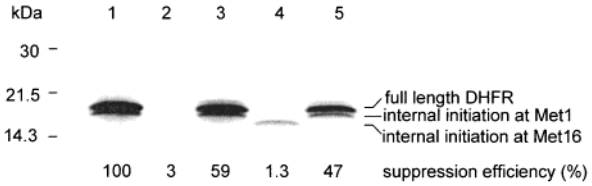

Figure 2. Autoradiogram of a 15% SDS-polyacrylamide gel illustrating the in vitro synthesis of DHFR analogues containing allylglycine at predetermined sites. Protein synthesis[5] was carried out using [^{35}S]-methionine with wild-type mRNA (lane 1), or mRNA containing UAG at position 27 (lanes 2 and 3) or position 10 (lanes 4 and 5) and suppressor tRNAs as noted. Lane 1, no suppressor tRNA; lane 2, unacylated $tRNA_{CUA}$; lane 3, allylglycyl-$tRNA_{CUA}$; lane 4, unacylated $tRNA_{CUA}$; lane 5, allylglycyl-$tRNA_{CUA}$. Suppression efficiencies are shown below each lane; the yield of wild type was approximately 20 μg/mL of incubation mixture.

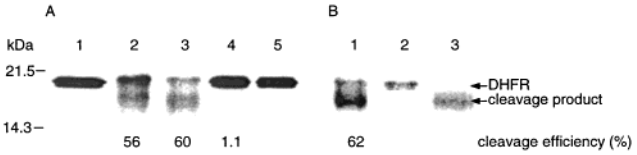

Figure 3. Iodine-induced cleavage of DHFR analogue Asp27^{A}Gly. The purified protein was treated with I_2 in 100:15 water−THF at 25 °C for 30 min. Panel A: lane 1, Asp27 ^{A}Gly without I_2 treatment; lane 2, Asp27 ^{A}Gly + 1.0 mM I_2; lane 3, Asp27 ^{A}Gly + 1.5 mM I_2; lane 4, wild-type DHFR + 1.5 mM I_2; lane 5, wild-type DHFR without I_2 treatment. Panel B: Asp27 ^{A}Gly was treated with 1.2 mM I_2 (lane 1), then applied to a Ni-NTA column. The (uncleaved) protein that was retained by the column was analyzed in lane 2; the nonretained cleavage product appears in lane 3. Cleavage efficiencies are shown below each lane. The N-terminal peptide cleavage product could not be observed, plausibly due to peptidases present in the protein-synthesizing system.

Table 1. Iodine-Induced Backbone Cleavage Efficiency of DHFR Analogues with Allylglycine (^{A}Gly) Located in Different Secondary Structure Contexts

DHFR analogue	optimal [I_2] for cleavage (mM)	optimal cleavage eff (%)	secondary structure in which ^{A}Gly was incorporated	solvent accessibility[b]
wild type		nd[a]		
Asp27^{A}Gly	1.2	62	α-helix	+
Val10^{A}Gly	2.0	11	hydrogen bonded turn	++
Glu-1^{A}Gly	1.5	41	extended strand[b]	+++
Glu-1Val		nd[a]		+++

[a] Not detectable. [b] Inferred from the crystal structure of DHFR,[7c,d] assuming that the modified protein structures are fundamentally analogous to those of wild type. Increased solvent accessibility is denoted by increasing numbers of pluses.

treatment should thus result in the chemical activation of trypsin from its inactive precursor. In common with wild-type rat trypsinogen, the in vitro synthesized trypsinogen containing ^{A}Gly (^{A}Gly-Tg) had similar minimal tryptic activity. However, the trypsinogen containing ^{A}Gly could not be converted to active trypsin through proteolysis mediated either by enteropeptidase or trypsin since the zymogen activation sequence was not present.[8]

Table 2. Comparison of Kinetic Parameters Obtained from Trypsin Generated by Iodine Treatment of Trypsinogen Containing Allylglycine (^{A}Gly-Tg) and Wild-type Trypsin (Tn[a])

sample	k_{cat} (S^{-1})	K_M (μM)	k_{cat}/K_M ($\mu M^{-1} s^{-1}$)
Tn[a]	44.1 ± 1.6	12 ± 2	3.7 ± 0.8
Tn + L[b]	42.3 ± 1.3	10.5 ± 1.4	4.0 ± 0.4
Tn + L + THF	42.5 ± 2.4	14.1 ± 2.8	3.0 ± 0.5
Tn + L + THF + I_2	41.2 ± 2.7	11.6 ± 2.9	3.5 ± 0.7
^{A}Gly-Tg + THF + I_2	42.1 ± 1.7	11.4 ± 1.6	3.7 ± 0.5

[a] Synthesized in vivo as rat trypsinogen and subsequently activated with enterokinase (1:20, w/w). Active Tn was purified by sequential hydrophobic and affinity column chromatography methods. [b] Wild-type Tn was combined with rabbit reticulocyte lysate (L) and subjected to the same treatment used to prepare ^{A}Gly-Tg.

Wild-type trypsinogen and trypsinogen analogues containing either allylglycine or valine at the normal cleavage site were assayed for trypsin activity,[15] before and after I_2 treatment. As shown (Table 2), only the allylglycine-containing trypsinogen could be activated with I_2. The kinetic parameters of this trypsin were measured and were comparable to that of wild type. The presence of trypsin activity following iodine treatment supports the chemical mechanism proposed in Figure 1, since a free N-terminal amino group in trypsin at the authentic cleavage site[16] is a prerequisite for enzymatic activity.[17] Additionally, a synthetic tripeptide having the same sequence as trypsinogen analogue ^{A}Gly-Tg at the putative cleavage site (Gly^{A}GlyIle) was shown explicitly to undergo the anticipated cleavage reaction in good yield (Supporting Information).

These experiments establish the feasibility of a novel strategy for site-specific cleavage of protein backbones by elaboration of proteins containing allylglycine at predetermined sites. The activation of trypsin from trypsinogen by I_2 treatment is the first report of activation of a zymogen by a chemical treatment; this approach has potential utility for the study of time-dependent processes mediated by trypsin or other proteins having inactive precursors. Other possible applications include the systematic removal of individual protein domains of putative higher order structures.

Acknowledgment. We thank Dr. James Landers and Braden Giordano for assistance with capillary electrophoresis. This work was supported at the University of Virginia by Research Grant CA77359 from the National Cancer Institute and at UC San Francisco by Research Grant MCB9604379 (C.S.C.) and by a fellowship (T.T.B.) from the National Science Foundation.

Supporting Information Available: Experimental details of protein preparation, cleavage, and assay (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0002262

(15) Zimmerman, M.; Ashe, B.; Yurewicz, E. C.; Patel, G. *Anal. Biochem.* **1977**, *78*, 47.
(16) Pasternak, A.; Liu, X.; Lin, T.-Y.; Hedstrom, L. *Biochemistry* **1998**, *37*, 16201.
(17) (a) Scrimger, S. T.; Hofmann, T. *J. Biol. Chem.* **1967**, *242*, 2528. (b) Huber, R.; Bode, W. *Acc. Chem. Res.* **1978**, *11*, 114.
(18) Varshney, U.; Lee, C.-P.; RajBhandary, U. L. *J. Biol. Chem.* **1991**, *266*, 24712.