

Photoinduced, Family-Specific, Site-Selective Cleavage of TIM-Barrel Proteins

Nicola Floyd,[†] Neil J. Oldham,[‡] Christopher J. Eyles,[§] Stephen Taylor,^{||} Dmitry A. Filatov,[⊥] Mark Brouard,[§] and Benjamin G. Davis^{*†}

Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford, OX1 3TA, U.K., Department of Chemistry, University of Oxford, Physical and Theoretical Chemistry Laboratory, South Parks Road, OX1 3QZ, U.K., The School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, U.K., Computational Biology Research Group, University of Oxford, Oxford, OX1 3RE, U.K., and Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, U.K.

Received April 1, 2009; E-mail: ben.davis@chem.ox.ac.uk

Proteins have long been known to be affected by UV irradiation in processes ranging from amino acid side chain degradation, the reduction of disulfide bridges, aggregation, and even to consequent loss of biological activity.^{1–3} However, natural examples of clean, site-selective photocleavage are rare. To the best of our knowledge, the only reported examples are fluorescent proteins Kaede and EosFP that experience cleavage (between Phe61, His62 at 400 nm) and require prior formation of a specific chromophore.^{4,5}

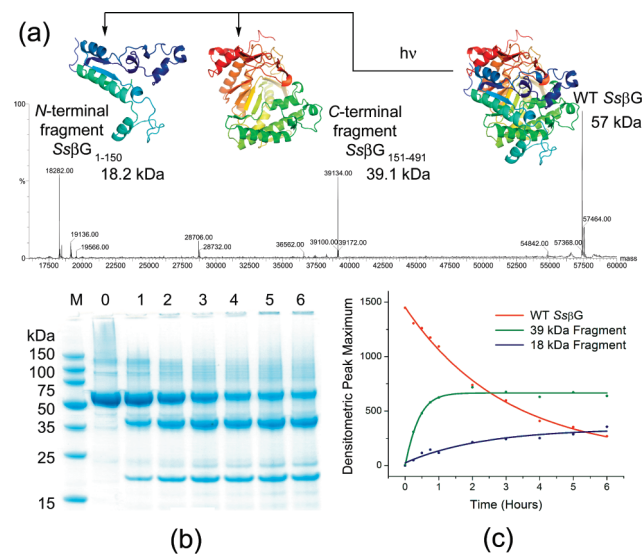
A more general, clean, photoinduced cleavage mechanism, with potential biotechnological utility in protein preparation⁶ or proteomic analyses,^{7,8} is therefore also of fundamental interest. Indeed, while a number of site-selective photolytic methods requiring the addition of extraneous cleavage agents have been investigated,^{9–14} these are to date poorly selective and low yielding. Use of unnatural amino acids, such as *o*-nitro-Phe,^{15,16} only allows photocleavage in up to ~30%.¹⁷ Unexpectedly, while investigating the photochemical modification of proteins, we have discovered a clean, photocleavage reaction localized to TIM-barrel proteins from family 1 of glycosylhydrolases (GH-1).

UV irradiation of archetypal GH-1 protein from the hyperthermophilic archaeon *Sulfolobus solfataricus* (*SsβG*)¹⁸ gave fragment product profiles that were sharp and clean (Scheme 1), suggesting protein cleavage at a single site.¹⁹ To probe reactive transitions precise frequency light (193–355 nm) was generated using (Nd:YAG)-pumped tunable dye and ArF-excimer (193 nm) lasers. No cleavage was observed in the absence of UV light or at wavelengths outside the range ~240–310 nm.

Cleanly formed fragment masses (18.2, 39.1 kDa) were determined by ESI-MS (Scheme 1). *N*-Terminal sequencing revealed the latter to be “blocked” by an unnatural, unreactive residue and the former to share the sequence found at the *N*-terminus of *SsβG*, indicating a cleavage site between residues His150 and Trp151. *SsβG* mutants containing alterations at this junction (H150A, W151A, H150F) remained intact and did not fragment, giving additional evidence for this specific cleavage site and highlighting the need for specific residues and not simply for potential functional (e.g., aromatic His150→Phe) group equivalents.

Bioinformatic analysis reveals that the His-Trp (HW) diad is widespread and, indeed, is repeated at positions His424Trp425 in *SsβG*. Since this site did not dissociate on exposure to UV light, we considered that an extended series of residues bordering His150 and Trp151 must contribute to the remarkable regioselectivity of this cleavage reaction. To probe the underlying molecular contributions, short peptide fragments of *SsβG*, expected to have little inherent conformation but correct primary sequence, were synthesized: H-HWP-NH₂ and

Scheme 1. Photoinduced Site Selective Cleavage of TIM-Barrel GH-1 Proteins^a



^a (a) ESI-MS after 1 h irradiation; (b) SDS-PAGE time course analysis 1–6 h; (c) densitometry determined reaction course.

H-LNMYHWPLPL-NH₂, representing residues 150–152 and 146–155 of *SsβG*, respectively. Neither fragmented under UV light.

These results implicated conformation and secondary/tertiary structure as key determinants in fragmentation. Ala-scanning was used to probe a 3.5 Å sphere of residues around the cleavage site: *SsβG* mutations Y149A, P152A, F222A, and W433A prevented fragmentation, while alteration of more remote residues (Q18A, R79A, N81A, L153A, N205A, E206A, V210A, and E387A) had no effect. Modeling of the cleavage site (Figure 1) reveals that essential residue groups Y149, F222, W433, and P152 cradle residue W151 in an hydrophobic, π -rich environment.

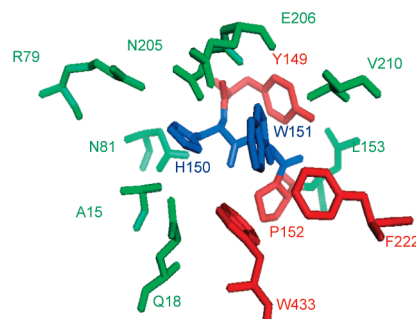


Figure 1. *SsβG*'s photocleavage site (blue) and associated essential (red) and nonessential (green) residues. Model based on PDB 1gow.

[†] CRL, University of Oxford.

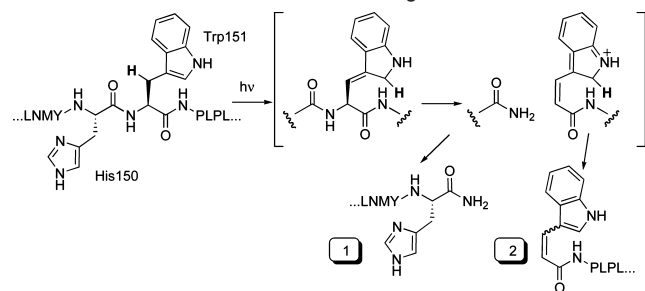
[‡] University of Nottingham.

[§] PTCL, University of Oxford.

^{||} Computational Biology, University of Oxford.

[⊥] Plant Sciences, University of Oxford.

Scheme 2. Putative Photoinduced Cleavage Mechanism



Digestion-MS/MS analysis allowed accurate mass and formulae determination of the C-terminal amino acid of the 18.2 kDa fragment and N-terminal amino acid of the 39.1 kDa fragment giving results consistent with His-amide **1** and enamide **2** (Scheme 2). Lack of an N-terminal amine in **2** is also consistent with its lack of reactivity under Edman degradation. MS analyses of the b1 ions of **2** and 3-indoleacrylic acid were identical; both showed characteristic loss of CO followed by loss of HCN from the indole ring. The presence of indolylenamide in **2** was also confirmed by UV spectroscopy (λ_{max} 340 nm).^{20–22}

Evidence to suggest the involvement of singlet oxygen in the cleavage reaction was not found. Exclusion of oxygen did not alter reaction rate. After bubbling with argon or addition of scavenging reagent trolox,²³ protein cleavage occurred at the same rate had oxygen not been removed. Reaction rate was also unaffected by D₂O.²⁴ Furthermore, one-photon photosensitized production of singlet oxygen using TMPyP²⁵ at 390 nm, a wavelength known not to cause protein fragmentation, did not result in the cleavage of WT *SsβG*. Together these suggested a mechanism other than one involving singlet oxygen. A plausible sequence consistent with 240–308 nm light-induced cleavage is shown in Scheme 2: suprafacial 1,3-β-H sigmatropic shift,²⁶ α-elimination of fragment **1**, and rearomatization. Nearby critical residues might serve to lower the transition state of energy of the shift and/or provide general acid/base assistance in the elimination. Other mechanisms, e.g., photoactivated electron transfer–elimination,²⁷ cannot be excluded.²⁸

The generality of the photocleaved His-Trp motif was investigated through various methods. Using the motif's atomic coordinates, all proteins with known 3D structure containing similarly spatially arranged HWP-triads were identified (using SPASM²⁹). The most similar, the arylsulfatase from *Pseudomonas aeruginosa*,³⁰ the chalcone reductase from *Medicago sativa*,³¹ and the β-glycosidase from *Thermosphaera aggregans* (*TaβG*),³² were expressed and purified; only the closest *TaβG* (<0.5 Å rmsd, 54% sequence similarity) was cleaved by UV, whereas other proteins were degraded by UV radiation.³³ Daughter fragments (17.2 and 38.1 kDa) revealed a homologous scission between His and Trp. Next, motif-pattern analysis (using MEME with HMM) and sequence alignments (using MUSCLE³⁴ and CAZY³⁵) revealed ~1800 proteins with the highest similarity across a variety of species. Phylogenetic analysis of this entire grouping (using Quick Neighbor Joining algorithm) based on clustering of the full protein sequence revealed branches focused according to the tetrapeptide region that corresponds to Tyr149–Pro152 in WT-*SsβG*. Representative proteins (10) from these branches were evaluated for photocleavage activity. Together these analyses correctly predicted other cleaving proteins (Y/FHWP, e.g., *PfβG*, Y149F-*SsβG*) and revealed branches and motifs with weak (YHWD, e.g., *ReβG*, *CtβG*) and no cleavage activity (HHFD/FHWD, e.g., *SpβG*, *SaM*). Finally, to explore transplantation of the photocleavable motif, we created a GFP-*SsβG*-(His)₆tag fusion protein; Ni-Chromatography allowed affinity purification and then “photorelease” of GFP (see Supporting Information (SI)).

In conclusion, we have discovered an efficient photocleavage reaction present in a natural motif of GH1 proteins that can be used to create a photocleavable tag. The biological function, if any, of this photochemical post-translational processing is not clear. However, early phylogenetic analyses revealed a cluster of photocleavable proteins that are products of isolated genes or genes in operons associated with peptide transport.³⁶

Acknowledgment. We thank Drs. J. McCullagh, R. Procter for technical assistance; Prof. P. Ogilby for helpful discussions; Prof. J. Noel, Drs. M. Kertesz, M. Moracci for plasmids encoding chalcone reductase, arylsulfatase, and *TaβG*, respectively; Dr. S. Kengen for *PfβG*; Prozomix for *ReβG*, *CtβG*, *SpβG*; and the International AIDS Vaccine Initiative (IAVI) for funding.

Supporting Information Available: Experimental procedures, bioinformatics and mass spectrometry data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Permyakov, E. A.; Permyakov, S. E.; Deikus, G. Y.; Morozova-Roche, L. A.; Grishchenko, V. M.; Kalinichenko, L. P.; Uversky, V. N. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 498–503.
- (2) Clymer, D. J.; Geren, C. R.; Ebner, K. E. *Biochemistry* **1976**, *15*, 1093–7.
- (3) Artyukhov, V. G.; Gusinskaya, V. V.; Dvurekova, E. A.; Rubtsov, M. P. *Biofizika* **2007**, *52*, 1001–1007.
- (4) Mizuno, H.; Mal, T. K.; Tong, K. I.; Ando, R.; Furuta, T.; Ikura, M.; Miyawaki, A. *Mol. Cell* **2003**, *12*, 1051–1058.
- (5) Wiedenmann, J.; Ivanchenko, S.; Oswald, F.; Schmitt, F.; Roecker, C.; Salih, A.; Spindler, K.-D.; Nienhaus, G. U. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15905–15910.
- (6) Arnau, J.; Lauritzen, C.; Petersen, G. E.; Pedersen, J. *Protein Expression Purif.* **2006**, *48*, 1–13.
- (7) Gevaert, K.; Vandekerckhove, J. *Electrophoresis* **2000**, *21*, 1145–1154.
- (8) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. *Nucleic Acids Res.* **2003**, *31*, 3784–3788.
- (9) Jones, B. J.; Vergne, M. J.; Bunk, D. M.; Locascio, L. E.; Hayes, M. A. *Anal. Chem.* **2007**, *79*, 1327–1332.
- (10) Kumar, C. V.; Buranaprapuk, A. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2085–7.
- (11) Kumar, C. V.; Thota, J. *Inorg. Chem.* **2005**, *44*, 825–827.
- (12) Buranaprapuk, A.; Kumar, C. V.; Jockusch, S.; Turro, N. J. *Tetrahedron* **2000**, *56*, 7019–7025.
- (13) Hoyer, D.; Cho, H.; Schultz, P. G. *J. Am. Chem. Soc.* **1990**, *112*, 3249–50.
- (14) Suzuki, A.; Hasegawa, M.; Ishii, M.; Matsumura, S.; Tushima, K. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4624–4627.
- (15) England, P. M.; Lester, H. A.; Davidson, N.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11025–11030.
- (16) Endo, M.; Nakayama, K.; Kaida, Y.; Majima, T. *Angew. Chem., Int. Ed.* **2004**, *43*, 5643–5645.
- (17) Peters, F. B. B.; Wang, J.; Schultz, P. G. *Chem. Biol.* **2009**, *16*, 148–52.
- (18) Aguilar, C. F.; Sanderson, I.; Moracci, M.; Ciaramella, M.; Nucci, R.; Rossi, M.; Pearl, L. H. *J. Mol. Biol.* **1997**, *271*, 789–802.
- (19) Indiscriminate cleavage results in diffuse PAGE smears.
- (20) Takai, K.; Sasai, Y.; Morimoto, H.; Yamazaki, H.; Yoshii, H.; Inoue, S. *J. Biol. Chem.* **1984**, *259*, 4452–7.
- (21) Such electron-rich indolylenamides are poor conjugate electrophiles (see ref 22); the 39 kDa fragment of C344SM21IM73IM148IM206IM236IM275IM-280IM383IM439I mutant did not react with ethanethiol, thiophenol, 1-pyrenebutanoic hydrazide, or fluorescein hydrazide.
- (22) Soloshonoka, V. A.; Uekia, H.; Ellisa, T. K.; Yamadaa, T.; Ohfune, Y. *Tetrahedron Lett.* **2005**, *46*, 1107–1110.
- (23) Rasnik, I.; McKinney, S. A.; Ha, T. *Nat. Methods* **2006**, *3*, 891–893.
- (24) Merkel, P. B.; Nilsson, R.; Kearns, D. R. *J. Am. Chem. Soc.* **1972**, *94*, 1030–1.
- (25) Snyder, J. W.; Lambert, J. D. C.; Ogilby, P. R. *Photochem. Photobiol.* **2006**, *82*, 177–184.
- (26) Kropp, P. J.; Fravel, H. G.; Fields, T. R. *J. Am. Chem. Soc.* **1976**, *98*, 840–1.
- (27) Bent, D. V.; Hayon, E. *J. Am. Chem. Soc.* **1975**, *97*, 2612–2619.
- (28) The reaction was unaffected by metals, EDTA, or substrate.
- (29) Kleywegt, G. J. *J. Mol. Biol.* **1999**, *285*, 1887–1897.
- (30) Beil, S.; Kehrl, H.; James, P.; Staudenmann, W.; Cook, A. M.; Leisinger, T.; Kertesz, M. A. *Eur. J. Biochem.* **1995**, *229*, 38–94.
- (31) Bomati, E. K.; Austin, M. B.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *J. Biol. Chem.* **2005**, *280*, 30496–30503.
- (32) Chi, Y. I.; Martinez-Cruz, L. A.; Jancarik, J.; Swanson, R. V.; Robertson, D. E.; Kim, S. H. *FEBS Lett.* **1999**, *445*, 375–83.
- (33) Degradation of noncleavable over 10 h. Optimal cleavable reaction times 2–4 h give background degradation < ~20% (see SI); LASER light allows cleaner 15 min reactions.
- (34) Edgar, R. C. *Nucleic Acids Res.* **2004**, *32*, 1792–1797.
- (35) Cantarel, B.; Coutinho, P.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. *Nucleic Acids Res.* **2009**, *37*, D233–238.
- (36) Alm, E. J.; Huang, K. H.; Price, M. N.; Koche, R. P.; Keller, K.; Dubchak, I. L.; Arkin, A. P. *Genome Res.* **2005**, *15*, 1015–1022.

JA9026105