3810 *Biochemistry* **2009,** *48,* 3810–3812
DOI:10.1021/bi900327f

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Rapid Report

Mass Spectrometry-Based Structural Dissection of Fluorescent Proteins†

Luis A. Alvarez,‡ Fabienne Merola,‡ Marie Erard,‡ and Filippo Rusconi*,§

‡*Laboratoire de Chimie Physique, Université Paris-Sud 11, UMR 8000 CNRS, F-91405 Orsay Cedex, France, and* §*Laboratoire de biophysique, Muséum national d'Histoire naturelle, CNRS UMR5153-INSERM U565-MNHN USM0503, 57 rue Cuvier, Case postale 26, F-75231 Paris Cedex 05, France*

Received February 25, 2009. Revised Manuscript Received March 12, 2009

ABSTRACT: Fluorescent proteins (FPs) are essential for live cell studies using fluorescence microscopy. To date, the molecular basis for FPs' irreversible photobleaching and the nature of the associated photoproducts are a matter of debate. Mass spectrometry, which should be an ideal technique for the structural dissection of FPs, cannot be harnessed efficiently due to their extreme resistance to trypsinolysis, due to the compactness of the barrel structure containing the chromopeptide. We devised a mild endoproteolysis procedure that affords a peptide mass fingerprint almost totally covering the sequence, thus allowing high-resolution mass spectrometric investigations of the protein structure.

Since their first demonstration as universal and specific fluorescent probes for live cell imaging (*1*), fluorescent proteins (FPs)[1] have stimulated a constantly renewed interest. This is due to their evolutive nature, as genetically encoded markers, and also to the progressive elucidation of their remarkable properties, as complex, optically active biopolymers (*2*). Among the latter, it has become clear that all FPs are able to convert reversibly or irreversibly between different chemical forms, possessing different photophysical properties, as a function of time, either because of posttranslational processing, as an effect of light exposure, or due to changes in the chemical environment (*3*). On one hand, this has led to innovative cell imaging techniques, such as pulse−chase studies of protein expression, trafficking, and interactions (*4*−*6*) or super-resolution optical microscopy (*7*). On the other hand, light-induced FP conversions pave the way to photobleaching, which remains the main drawback of FP usage as compared to quantum dots or the best organic dyes. Deciphering the relationships between the molecular chemistry of FPs and their changing photophysical properties is therefore a major issue in the ongoing development of FP-based techniques.

Detailed atomic-resolution structures, as obtained from X-ray crystallography, have revealed an extreme diversity in the posttranslational chemistry of FPs and its strong dependence on peptide sequence (*10*−*14*).

The nature of FP photoproducts is currently a matter of debate and appears to depend strongly on the photodynamic conditions (*8* and references cited therein, *9*). The covalent changes responsible for photophysical perturbations are not necessarily restricted to the chromopeptide itself, as exemplified by the photoinduced decarboxylation of Glu 222 in *Aequora victoria* green fluorescent protein (*av*GFP)(*10*). Considering the wealth of new FP variants produced each year by genetic engineering and combinatorial methods, a full structural characterization of their end products clearly requires large-scale, if not high-throughput, analytical techniques.

The modern tools of mass spectrometry and proteomics are ideally suited to this aim. However, their use in the case of FPs has until now remained limited, due to the strong resistance of FPs to proteolytic attack, an obligatory step in the mass spectrometry dissection of protein structure. Indeed, because of the compactness of their β-can folding pattern, FPs display an extremely high resistance to tryptic cleavage (*15*), thus yielding unsatisfactory sequence coverage. The few reports successfully harnessing mass spectrometry used extremely harsh physicochemical conditions and dealt with only the chromopeptide (*16*−*19*). We report here on a new, efficient, and general method for achieving the proteolytic digest of FPs under mild conditions, without resorting to in-gel protocols. The method was initially developed for the analysis of the γ-ray-induced oxidation products of enhanced cyan fluorescent protein (CFP) but may likely be applied to all FPs derived from currently known natural sources.

In our studies of CFP's susceptibility to γ-irradiation, we observed a fluorescence loss that we suspected was due in part to alterations of residues located outside of the chromopeptide. A number of previous reports have established that methionyl, tyrosinyl, and tryptophanyl residues are particularly susceptible to radical oxygen species (*20*−*22*). Such residues occur in strands forming the barrel of the fluorescent protein. These strands are not easily

†L.A.A. is supported by a fellowship from the French Ministry of Research (MESR).

*To whom correspondence should be addressed. E-mail: rusconi@mnhn.fr.

[1]Abbreviations: *av*GFP, *Aequorea victoria* green fluorescent protein; FP, avGFP-like fluorescent protein; CFP, cyan fluorescent protein; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; ESI-MS, electrospray ionization mass spectrometry.

Published on Web 3/13/2009

isolated using conventional protein chemistry protocols, even when harsh sample preparation methods are used (*16*−*19*). Besides, harsh sample preparation protocols used throughout the literature could induce side alterations of the protein structure and might have resulted, in the case presented here, in artifactual data (*23*). Therefore, we set out to devise a mild proteolysis procedure that would lead to a sufficiently wide mass spectrometry-based sequence coverage of the protein. The results reported here deal with the unmodified CFP (that is, unirradiated CFP).

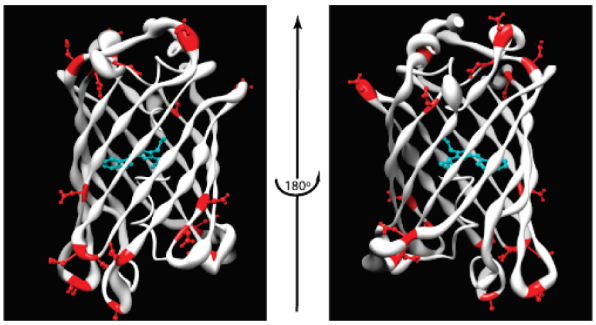

FIGURE 1: EndoAspN cleavage sites allow efficient CFP proteolysis and proper subsequent ESI-MS sequence coverage. CFP three-dimensional worm structure representation with the width being proportional to the solvent-accessible surface area of residues (*25*) (red residues, Asp).

One first attempt was to submit CFP to a conventional in-gel tryptic cleavage, which afforded a sequence coverage of at most 50% with peptides located between positions [4−45] and [80−140], which correspond to the N-terminal half of the protein primary structure (Figure 1 of the Supporting Information). This result left us dissatisfied because the residues that might have undergone chemical modification by reactive oxygen species are located for the most part in the C-terminal half of the protein.

As a second attempt, we studied the three-dimensional structure of the CFP [Protein Data Bank entry 1OXD (*24*)] and related proteins and noticed the systematic presence of accessible aspartyl residues in the flexible loops connecting the β-sheet strands forming the β-barrel (Figures 1 and 3). The EndoAspN proteolytic enzyme cleaves the peptide bond at the N-terminal side of aspartyl residues, making it a good candidate for cleavage of these loops. When using the standard protocol supplied by the manufacturer, no cleavage could be detected by SDS−PAGE analysis. To

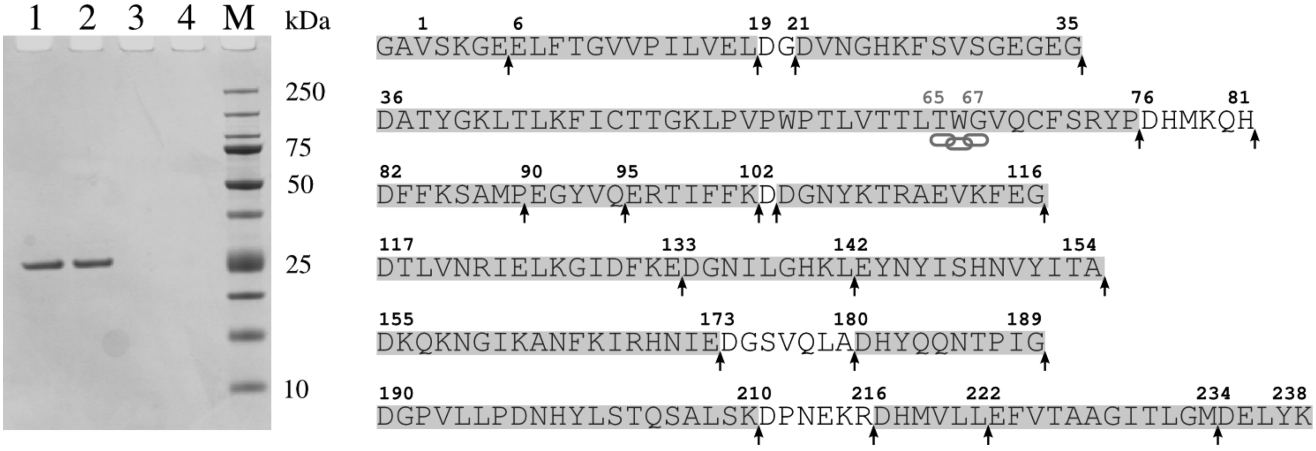

FIGURE 2: SDS−PAGE of the CFP sample before and after endoproteolysis: lane 1, unprocessed CFP; lanes 2−4, CFP solutions processed as described in the text with E/S ratios of 0, 10, and 20, respectively; lane M, molecular mass markers (Bio-Rad). The right panel is a schematic representation of the protein coverage obtained by mass spectrometry: gray highlighting, observed peptides; arrows, cleaved positions. The chromophore-engaged residues are numbered and indicated with links (see Table 1 of the Supporting Information).

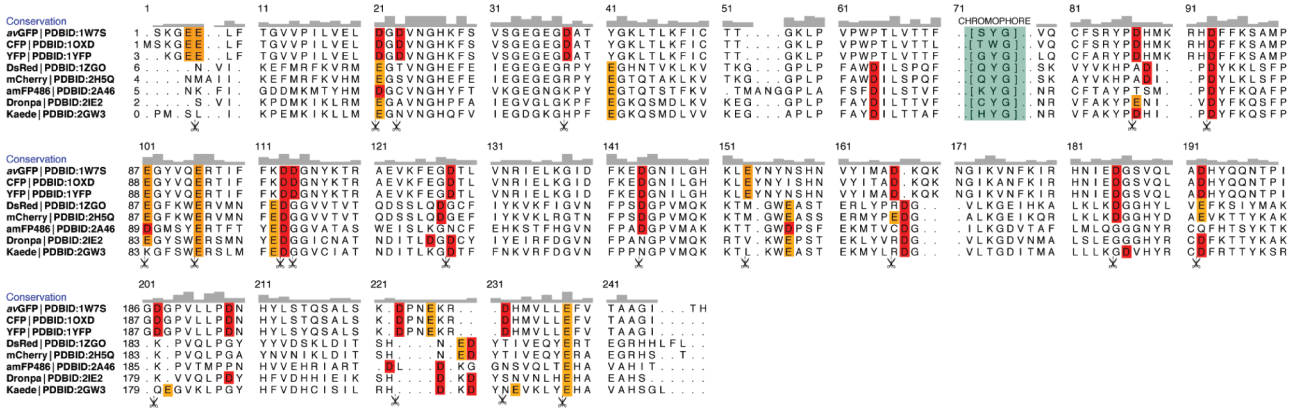

FIGURE 3: Sequence alignment derived from the three-dimensional structural alignment of eight fluorescent proteins with Asp residues located in the flexible loops connecting the β-sheets forming the barrel highlighted in red. When an Asp residue at a given position was not found in all the sequences, an equivalent Glu residue was found in the neighboring sequence and is highlighted in orange. The location of that Glu residue in a flexible loop was ascertained by looking at the corresponding protein's crystallographic data. Scissors indicate the positions at which cleavage occurred using EndoAspN as the cleaving enzyme of CFP.

destabilize the CFP structure to make it more susceptible to EndoAspN cleavage, the protein was first brought to 10% acetonitrile (v/v) and heated for 5 min at 90 °C and promptly dropped in an ice−water bath before being supplemented with an additional 10% acetonitrile. The enzyme was then added (7 ng/μL final EndoAspN concentration; 1/20 E/S ratio). Cleavage occurred at 37 °C overnight. Both the heat shock and the high E/S ratio were necessary for an efficient disruption of the β-barrel. Figure 2 shows the results obtained. The left panel shows that the CFP is effectively hydrolyzed under our conditions (lanes 3 and 4), and the right panel demonstrates that the obtained sequence coverage (91%) is suitable for high-resolution mass spectrometric dissection of the FPs' structure (see also Table 1 of the Supporting Information showing, for each peptide, the corresponding calculated and observed masses along with the relative intensities of the peaks).

In our experiments, chromopeptide [36−75] is observed with a disulfide bond linking cysteinyl residues at positions 48 and 70 (ion at m/z 1120.80, $z = 4$). Although not producing a peak as intense as those of other species in the peptide mass fingerprint spectrum, this ion could easily be sequenced by gas phase fragmentation. We found an almost complete b-series sequence (b2, b3, b5, and b7−b12) from the left-end residue (Asp 36) up to Ile 47 (hence stopping right at the disulfide bond-engaged Cys 47). Likewise, the y-series contained ions y3−y5, partly confirming the C-terminal sequence of the chromopeptide, up to Phe 71 (hence stopping right before encountering the other disulfide bond-engaged Cys 70). Finally, as a confirmation of the chromopeptide nature, the immonium ions generated for this peptide contained rather intense signals for leucine/isoleucine, tryptophan, and proline, the latter being represented above usual frequencies in the peptide (Figure 2 of the Supporting Information).

Interestingly, when the E/S ratio was increased to 1/15, unspecific cleavage occurred at glutamyl residues (positions 6, 90, 95, 142, and 222). This observation is of great interest, because the sequence alignment of fluorescent proteins from a number of organisms (Figure 3) shows that when an aspartyl residue at a given position is not found in all the sequences, an equivalent glutamyl residue is present in the neighboring sequence in its stead. Therefore, an increased E/S ratio should prove useful when performing molecular dissections of fluorescent proteins from a variety of organisms.

In conclusion, this report describes a mild method for liquid-based endoproteolysis of fluorescent proteins that is suitable for CFP and which should also be applicable to the great many fluorescent proteins from a number of organisms.

ACKNOWLEDGMENT

The mass spectrometric work was performed at the Muséum national d'Histoire naturelle.

SUPPORTING INFORMATION AVAILABLE

Supplementary methods, supplementary Table 1, and supplementary Figures 1 and 2. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) *Science 263*, 802–805.
2. Shaner, N. C., Patterson, G. H., and Davidson, M. W. (2007) *J. Cell Sci. 120*, 4247–4260.
3. Remington, S. J. (2006) *Curr. Opin. Struct. Biol. 16*, 714–721.
4. Terskikh, A., Fradkov, A., Ermakova, G., Zaraisky, A., Tan, P., Kajava, A. V., Zhao, X. N., Lukyanov, S., Matz, M., Kim, S., Weissman, I., and Siebert, P. (2000) *Science 290*, 1585–1588.
5. Patterson, G. H., and Lippincott-Schwartz, J. (2002) *Science 297*, 1873–1877.
6. Demarco, I. A., Periasamy, A., Booker, C. F., and Day, R. N. (2006) *Nat. Methods 3*, 519–524.
7. Fernandez-Suarez, M., and Ting, A. Y. (2008) *Nat. Rev. Mol. Cell Biol. 9*, 929–943.
8. Kirber, M. T., Chen, K., and Keaney, J. F. (2007) *Nat. Methods 4*, 767–768.
9. Sinnecker, D., Voigt, P., Hellwig, N., and Schaefer, M. (2005) *Biochemistry 44*, 7085–7094.
10. van Thor, J. J., Gensch, T., Hellingwerf, K. J., and Johnson, L. N. (2002) *Nat. Struct. Biol. 9*, 37–41.
11. Shu, X. K., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., and Remington, S. J. (2006) *Biochemistry 45*, 9639–9647.
12. Barondeau, D. P., Kassmann, C. J., Tainer, J. A., and Getzoff, E. D. (2007) *J. Am. Chem. Soc. 129*, 3118–3126.
13. Barondeau, D. P., Kassmann, C. J., Tainer, J. A., and Getzoff, E. D. (2006) *J. Am. Chem. Soc. 128*, 4685–4693.
14. Adam, V., Lelimousin, M., Boehme, S., Desfonds, G., Nienhaus, K., Field, M. J., Wiedenmann, J., McSweeney, S., Nienhaus, G. U., and Bourgeois, D. (2008) *Proc. Natl. Acad. Sci. U.S.A. 105* 18343–18348.
15. Chiang, C. F., Okou, D. T., Griffin, T. B., Verret, C. R., and Williams, M. N. V. (2001) *Arch. Biochem. Biophys. 394*, 229–235.
16. Pouwels, L. J., Zhang, L. P., Chan, N. H., Dorrestein, P. C., and Wachter, R. M. (2008) *Biochemistry 47*, 10111–10122.
17. Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K., and Tsien, R. Y. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 11990–11995.
18. Pakhomov, A. A., Pletneva, N. V., Balashova, T. A., and Martynov, V. I. (2006) *Biochemistry 45*, 7256–7264.
19. Tretyakova, Y. A., Pakhomov, A. A., and Martynov, V. I. (2007) *J. Am. Chem. Soc. 129*, 7748–7749.
20. Houée Levin, C., and Sicard Roselli, C. (2001) in *Radiation chemistry of proteins*, pp 553−584, Elsevier, Amsterdam..
21. Blouquit, Y., Duchambon, P., Brun, E., Marco, S., Rusconi, F., and Sicard-Roselli, C. (2007) *Free Radical Biol. Med. 43*, 216–228.
22. Mozziconacci, O., Mirkowski, J., Rusconi, F., Pernot, P., Bobrowski, K., and Houee-Levin, C. (2007) *Free Radical Biol. Med. 43*, 229–240.
23. Lippincott, J., and Apostol, I. (1999) *Anal. Biochem. 267*, 57–64.
24. Bae, J. H., Rubini, M., Jung, G., Wiegand, G., Seifert, M. H. J., Azim, M. K., Kim, J. S., Zumbusch, A., Holak, T. A., Moroder, L., Huber, R., and Budisa, N. (2003) *J. Mol. Biol. 328*, 1071–1081.
25. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) *J. Comput. Chem. 25*, 1605–1612.