STRUCTURE OF THE CHROMOPHORE OF *AEQUOREA* GREEN FLUORESCENT PROTEIN

O. SHIMOMURA
Department of Biology, Princeton University, Princeton, NJ 08544, USA

Received 11 June 1979

1. Introduction

In the bioluminescence of the jellyfish *Aequorea*, the green fluorescent protein (GFP) plays the role of the light emitter [1]. The energy needed for the emission of light is produced in the Ca^{2+}-triggered reaction of the photoprotein aequorin [2] that coexists with GFP in the photogenic organ of the jellyfish. The energy is then transferred to GFP molecules by the Föster-type mechanism [1]. The excited state of GFP thus formed ultimately dissipates the energy in the form of green light (λ_{max} 509 nm). In the absence of GFP, aequorin emits blue light (λ_{max} 470 nm) [1].

Bioluminescence of various other coelenterates apparently also involves similar green fluorescent proteins [3]. GFP of the sea pansy *Renilla*, in addition to *Aequorea* GFP, has been extensively studied [4]. It has been thought that the chromophores of all GFPs are the same [5], although until now information has not been reported concerning the chemical nature of the chromophore from any source. The structural knowledge concerning the chromophore of *Aequorea* GFP reported in the present paper may help to identify the chromophore of other GFPs.

2. Materials and methods

Aequorea GFP was extracted and purified as previously described [1] without separating isoprotein components. The molecular weight of GFP monomer determined by the dodecyl sulfate-acrylamide gel electrophoresis method of Weber and Osborn [6] was 27 000 ± 1000 (gel: 7.5%; standards: trypsin inhibitor, chymotrypsinogen, ovalbumin and bovine serum albumin; pretreatment with dodecyl sulfate: 4 min at 90°C). This value was closely similar to that for *Renilla* GFP [4]. Previously reported values for *Aequorea* [1] probably resulted from insufficient pretreatment of GFP samples with dodecyl sulfate.

Purified GFP (100 mg, $A_{400\,nm}/A_{280\,nm}$ = 1.0) was first denatured by heating at 90°C for 1 min, then digested with 14 mg of papain (Sigma, Type III) in 80 ml of 0.75 M NaCl, pH 6.5, containing 1 mM EDTA, 60 mg of cysteine and 1 mg of 2-mercaptoethanol, at 35°C for 3 h. After acidification to pH 1 with HCl, the mixture was extracted with *n*-butanol. A peptide that contained GFP chromophore in the butanol extract was purified by successive thin-layer chromatography (TLC) on silica gel, first with CH_2Cl_2/methanol/14% ammonia (80 : 20 : 4 by volume), then ethyl acetate/ethanol/acetic acid/water (100 : 28 : 4 : 17), finally with ethyl acetate/ethanol/water/28% ammonia (70 : 34 : 14 : 2). The peptide was located by yellow coloration in ammonia vapor.

Hydrolysis was carried out by heating with 6 N HCl at 105°C for 20 h. Amino acids were analyzed on a Beckman amino acid analyzer. N-Terminal residue of peptide was studied by dinitrophenylation according to Sanger and Tuppy [7]. C-Terminal was studied by the method of hydrazinolysis followed by dinitrophenylation [8], however, with final identification of DNP derivatives by TLC on silica gel with ethyl acetate/ethanol/water/28% ammonia (50 : 15 : 5 : 1).

A model compound was prepared as follows. Glycylglycine (0.33 g) was heated with *n*-butyric anhydride (0.8 ml) and *n*-butyric acid (1 ml) at 160°C for 5 min, followed by the addition of *p*-hydroxybenzaldehyde

 Elsevier/North-Holland Biomedical Press

(0.3 g). After additional heating of 5 min, the mixture was cooled and dissolved in 30 ml of ethyl acetate, then extracted with 1 N HCl. Acidity of the HCl extract was adjusted to pH 1 with NaOH, followed by extraction with fresh ethyl acetate. The product in the ethyl acetate extract was evaporated to dryness, then purified by repeated TLC on silica gel, first with ethyl acetate/ethanol/1 N HCl (100 : 28 : 18), then with ethyl acetate/ethanol/water/28% ammonia (70 : 34 : 14 : 2). The purified product was obtained by extracting the major band (yellow in ammonia vapor) with methanol.

To convert the product into an amide, a portion of dried material was warmed with $SOCl_2$, evaporated to dryness under vacuum, then the residue was treated with an amine. The amide was purified by TLC on silica gel with CH_2Cl_2/methanol (10 : 1).

3. Results and discussion

Papain digestion of *Aequorea* GFP yielded a peptide that contained the chromophore of GFP. This peptide exhibited characteristic absorption spectra (fig.1), and was the only product that absorbed above 300 nm. The peptide had at least two pK_a values: 7.9 (by spectrophotometry) and 4.5 (by partition between *n*-butanol and water).

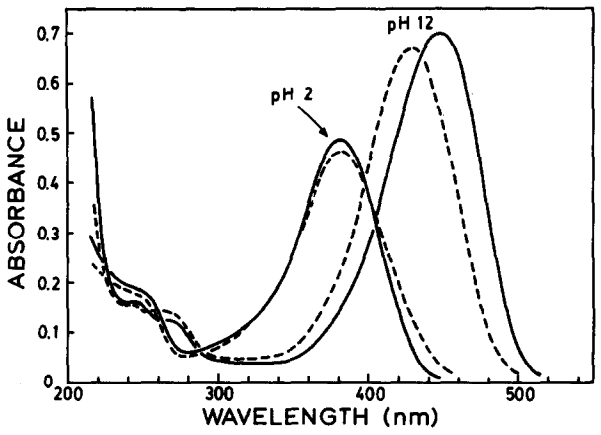

Fig.1. Absorption spectra of the peptide that contained GFP chromophore (solid lines) and of 1-carboxymethyl-2-(*n*-propyl)-4-(*p*-hydroxybenzylidene)-5-imidazolone (structure A, broken lines), both in water. The pH was adjusted with HCl or NaOH.

Acid hydrolysis of this peptide yielded 1 mol each of glycine, glutamic acid, valine, phenylalanine, and *p*-hydroxybenzaldehyde. The last compound was identified by ultraviolet absorption and mass spectrometry.

Catalytic hydrogenation of the peptide in ethanol with Adams catalist caused disappearance of most of the ultraviolet and visible absorption. Acid hydrolysis of the product yielded 1 mol each of glutamic acid and valine, plus approx. 1/2 mol each of glycine, phenylalanine and tyrosine, but no *p*-hydroxybenzaldehyde.

Dinitrophenylation of the peptide indicated that the N-terminal residue is phenylalanine. Hydrazinolysis, in an effort to identify the C-terminal, yielded glutamic acid-α-hydrazide but no free amino acid.

A model compound was synthesized from glycylglycine, butyric anhydride and *p*-hydroxybenzaldehyde. The compound contained a phenolic OH (pK_a 7.9 by spectrophotometry) plus a carboxyl group. The molecular weight was 288 by mass spectrometry. Acid hydrolysis of this compound yielded 2 mol of glycine plus 1 mol of *p*-hydroxybenzaldehyde. Thus, the structure of this compound must be A (fig.2).

The ultraviolet and visible absorption of this model compound closely resembled that of the natural peptide described above (fig.1). The difference of the absorption maxima at pH 12 (natural peptide: 448 nm; A: 430 nm) is probably due to the effect of an ionized carboxyl group. In fact, the amides of A showed closer agreement under the same conditions (morpholine amide: 445 nm; 2,2-diethoxyethylamide: 450 nm).

Based on the evidence described above, the structure of the chromophore of *Aequorea* GFP is proposed as B (fig.2). In structure B, $N-R_1$ represents Glu–Val–NH_2 or Val–Glu–NH_2 with the N atom of

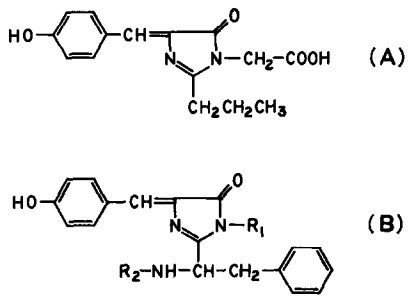

Fig.2. Structures of a model compound (A) and the chromophore of *Aequorea* GFP (B).

the first residue (Glu or Val) incorporated into the imidazolone ring. Although the results of partial hydrolysis favored the sequence Glu–Val–NH_2, rather than Val–Glu–NH_2, a firm conclusion on this point requires further data. The group R_2 represents the rest of the GFP molecule attached by a peptide bond. Partial similarity between structure B and the structure of coelenterazine (coelenterate luciferin) [9,10] may suggest a biogenetic significance.

Acknowledgements

I thank B. Bamman and L. R. Hyde for amino acid analyses, and A. Shimomura for technical assistance. I am indebted to Dr F. H. Johnson for his interest and encouragement. This work was aided by the National Science Foundation Grants PCM76-12301 and PCM78-22959, Public Health Service Grant 1 R01 GM25093-01, and the facilities of the Whitehall and Hartford Foundations at Princeton University and of Friday Harbor Laboratories, University of Washington.

References

[1] Morise, H., Shimomura, O., Johnson, F. H. and Winant, J. (1974) Biochemistry 13, 2656–2662.
[2] Johnson, F. H. and Shimomura, O. (1978) in: Methods in Enzymology vol. 57 (DeLuca, M. A., ed) 271–291, Academic Press, New York.
[3] Morin, J. G. and Hastings, J. W. (1971) J. Cell. Physiol. 77, 313–318.
[4] Ward, W. W. and Cormier, M. J. (1979) J. Biol. Chem. 254, 781–788.
[5] Ward, W. W. and Cormier, M. J. (1978) in: Methods in Enzymology, vol. 57 (DeLuca, M. A., ed), 257–267, Academic Press, New York.
[6] Weber, K. and Osborn, M. (1969) J. Biol. Chem. 244, 4406–4412.
[7] Sanger, F. and Tuppy, H. (1951) Biochem. J. 49, 463–481.
[8] Fraenkel-Conrat, H. and Tsung, C. M. (1967) in: Methods in Enzymology, vol. 11 (Hirs, C. H. W., ed), 151–155, Academic Press, New York.
[9] Shimomura, O. and Johnson, F. H. (1978) Proc. Nat. Acad. Sci. USA 75, 2611–2615.
[10] Cormier, M. J. (1978) in: Bioluminescence in Action (Herring, P. J., ed), 75–108, Academic Press, London.