Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 372 (2006) 18–25
GENE
SECTION
FUNCTIONAL GENOMICS
www.elsevier.com/locate/gene

A novel yellowish-green fluorescent protein from the marine copepod, *Chiridius poppei*, and its use as a reporter protein in HeLa cells

Hiromi Masuda [a,b,c,*,1], Yasuhiro Takenaka [b,c,1], Atsushi Yamaguchi [d], Satoshi Nishikawa [e], Hiroshi Mizuno [a,b,c]

[a] *VALWAY Technology Center, NEC Soft, Ltd., 1-18-7, Koto-ku, Tokyo 136-8627, Japan*
[b] *Department of Biochemistry, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan*
[c] *Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan*
[d] *Faculty of Fisheries Science, Hokkaido University, 3-1-1 Minato-cho, Hakodate 041-0821, Japan*
[e] *Age Dimension Research Center, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8562, Japan*

Received 5 July 2005; received in revised form 17 November 2005; accepted 24 November 2005
Available online 14 February 2006
Received by T. Sekiya

Abstract

A crustacean gene, encoding for a new class of GFP-like protein, has been isolated from a cDNA library of the deep-sea (benthic) copepod crustacean, *Chiridius poppei*, by expression cloning. The cDNA library was constructed in a pBluescript II vector and screened using a non-UV transilluminator, obtaining a positive clone. The clone consisted of a 781-bp fragment of cDNA with a 660-bp open reading frame, which encoded for a 219-amino acid polypeptide with a calculated molecular mass of 24.7 kDa. The protein was overexpressed in *Escherichia coli*, purified to homogeneity by anion-exchange and size-exclusion chromatographies. The protein, CpYGFP, had excitation and emission maxima at 507 and 517 nm, respectively. CpYGFP existed as a dimer in solution and could be expressed either alone or as a fusion protein in HeLa cells. Dual labeling experiments carried out with CpYGFP-actin and DsRed2-Nuc demonstrated the usefulness of CpYGFP as a reporter in the subcellular localization of actin.

Keywords: Green fluorescent protein; Crustacea; Copepoda; Expression cloning; Actin; DsRed

1. Introduction

The green fluorescent protein from the hydrozoan jellyfish *Aequorea victoria* has been a very versatile and useful marker for studying protein localization, protein trafficking and gene expression in living cells (Ogawa et al., 1995). The special advantage of GFP is that it possesses a chromophore in its primary structure which emits a greenish fluorescence upon irradiation with long UV light. The chromophore is formed post-translationally by an autocatalytic modification of a tripeptide in the primary structure without participation of any external cofactor (Nishiuchi et al., 1998). Besides *Aequorea* GFP, numerous GFP-like proteins have been isolated from hydrozoans and anthozoans of the phylum Cnidaria (Chalfie, 1995; Matz et al., 1999), but not from members of other phyla. Extensive characterization of GFP from *Renilla reniformis*, the bioluminescent sea pansy, has also been carried out (Ward and Cormier, 1979; Ward, 1981). In addition, a colorless, non-fluorescent GFP homologue has also been cloned from *Aequorea coerulescens*, a medusae, of the same genus as *A. victoria* (Gurskaya et al., 2003). Green, yellow, and red fluorescent

Abbreviations: CpYGFP, yellowish-green fluorescent protein of *Chiridius poppei*; GFP, green fluorescent protein; EGFP, enhanced GFP; EYFP, enhanced yellow fluorescent protein; GST-tagged, glutathione *S*-transferase-tagged; DsRed2-Nuc, DiscomaRed2-Nuclear localization signal; CMV, cytomegalovirus; CBB, Coomassie Brilliant Blue; PFA, paraformaldehyde; TCA, trichloroacetic acid; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.

* Corresponding author. Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan. Tel.: +81 29 861 6085; fax: +81 29 861 6159.
E-mail address: hiromi-masuda@aist.go.jp (H. Masuda).
[1] These authors contributed equally.

0378-1119/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2005.11.031

proteins have also been cloned from anthozoan species, notably corals (Verkhusha and Lukyanov, 2004). At the present time, the sequences of more than a hundred GFP-like proteins have been deposited in the GenBank. Phylogenetic analysis of these proteins has provided important insight into the evolutionary history of these proteins.

Recently, the genes of a new class of fluorescent proteins have been cloned from 6 marine copepods of the family Pontellidae (phylum Arthropoda, Class Crustacea) (Shagin et al., 2004). Although some spectrometric characterization of copepod GFP-like proteins was described in the report, protein properties such as stability or pH dependency and applicational probability to monitor the target protein tagged with those GFP-like proteins in living cells were not clearly mentioned. Among other copepods studied, the luciferase genes of *Gaussia princeps* (Verhaegen and Christopoulos, 2002) and *Metridia longa* (Markova et al., 2004) have been cloned, but no fluorescent proteins have been reported. In this paper, we describe the cloning, expression, purification and some characteristic properties of a yellowish-green fluorescent protein, CpYGFP, from the marine copepod, *Chiridius poppei*, family Aetideidae. In addition, data are presented showing the use of CpYGFP as a reporter protein in HeLa cells.

2. Materials and methods

2.1. Collection of copepod, C. poppei

Zooplankton samples were collected from sea water being pumped from a depth of 321 m in Toyama Bay (maximum depth 1 200 m) and distributed to holding tanks in the Hotaruika ("firefly squid") Museum located on shore in Namerikawa, Japan. A fine mesh net placed at the mouth of one of the outflow pipes was used to filter the zooplanktons. Live copepods exhibiting significant fluorescence on exposure to light from Dark Reader (Clare Chemical Research, Inc., Dolores, CO) (http://www.clarechemical.com/), a non-UV transilluminator, were isolated with a transfer pipet, pooled and maintained in test tubes at 4 °C. The copepods were examined under a stereo fluorescence microscope (VB-G25/S20/L11, Keyence, Osaka, Japan), equipped with a "GFP" filter (Excitation BP 470/40, FT 510) and cooled CCD camera (VB-7010, Keyence, Osaka, Japan). After removing the sea water from the test tubes, the copepods were suspended in TRIzol reagent (Invitrogen) and stored at −80 °C.

2.2. Preparation and screening of cDNA expression library

About 300 specimens in 6 ml of TRIzol reagent were used to prepare total RNA. About 7.1 μg of mRNA was purified from 228 μg of total RNA using an Oligotex-dT30 mRNA Purification Kit (Takara Bio, Shiga, Japan). cDNA was synthesized from 5 μg of mRNA using a cDNA Synthesis Kit (Stratagene), digested with *Xho*I, and cloned into *Eco*RV–*Xho*I site of pBluescript II SK(+) vector (Stratagene). The cDNA library was introduced into TOP10 host strain (Invitrogen) by electroporation. About 3×10^5 colonies growing on LB-carbenicillin plates were visually screened using a non-UV transilluminator. A single positive clone was isolated, resuspended in LB medium and subjected to a second screening. The plasmid was purified using QIAprep Spin Miniprep Kit (Qiagen) and sequenced using an ABI PRISM 3100 Gene Analyzer (Applied Biosystems).

2.3. Sequence analysis

Multiple amino acid sequence alignment of CpYGFP with those of other copepod GFP-like proteins was carried out using program CLUSTAL W (ver. 1.60) (Jeanmougin et al., 1998) with default parameters, except for an open gap penalty of 20 and an extended gap penalty of 1.0 (for multiple sequence alignment). For the phylogenetic tree, entire amino acid sequences were aligned by Clustal X (ver. 1.83) (Thompson et al., 1997). The tree was constructed using neighbor-joining method and drawn using program NJPLOT.

2.4. Expression and purification of C. poppei YGFP in Escherichia coli

The coding region of CpYGFP was amplified by *Pyrobest* DNA Polymerase (Takara Bio) from template plasmid isolated by expression screening and cloned into pET101/D-TOPO vector (Invitrogen). CpYGFP protein was expressed in BL21-CodonPlus®(DE3)-RIL *E. coli* host cells (Stratagene) at 37 °C. Cells were harvested at 6 h after 1 mM IPTG induction, resuspended in 20 mM Tris–HCl buffer, pH 8.5, and sonicated at 4 °C. The CpYGFP in the soluble protein fraction was subjected to chromatographic purification using a linear NaCl gradient (0–120 mM) and DEAE sepharose F.F., HiTrap Phenyl HP, and HiLoad 16/60 Superdex 200 pg columns installed on an ÄKTA 10S Explorer System (Amersham Bioscience) (data not shown). Fractions collected were analyzed by SDS-PAGE (12.5%) under reducing conditions.

2.5. Characterization of CpYGFP

To determine thermostability of CpYGFP, 10 μl of purified recombinant CpYGFP (0.58 mg/ml), diluted with 192 μl of 20 mM Tris–HCl, pH 8.5, was incubated in a block-heater over a range of 4–80 °C for 10 min, then cooled on ice for 10 min. The effect of pH on fluorescence intensity was analyzed by mixing 10 μl of the CpYGFP solution with 192 μl of dilution buffer and the mixture was adjusted to give a pH between 3.0 and 13.6 (for details, see legend in Fig. 5). The effect of chemicals on fluorescence intensity was measured by adding various amounts of reagent grade chemicals to 10 μl of CpYGFP solution, mixing and incubating for 10 min at 4 °C. Fluorescence intensity was measured at its peak wavelength of 517 nm using a F-4500 fluorescence spectrophotometer (Hitachi High Technologies, Tokyo). Protein concentration was determined using a Bradford assay kit (Bio-Rad) with bovine serum albumin as a standard. Recombinant *Aequorea* GFP was purchased from Clontech (Cat. No. 632373), and its fluorescent intensity was measured at 507 nm.

2.6. Expression of CpYGFP and co-expression of CpYGFP-tagged actin in HeLa cells

The coding region of CpYGFP was cloned into pcDNA3.2/V5/GW/D-TOPO (Invitrogen), following the manufacturer's instruction manual. The expression vectors pcDNA3.2-CpYGFP, pcDNA3.2-CpYGFP-actin and pDsRed2-Nuc (Clontech) were purified using a QIAfilter Plasmid Maxi Kit (Qiagen). HeLa cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Transfection was performed in a Lab-Tek Chamber Slide (Nalge Nunc) using PolyFect Transfection Reagent (Qiagen). After incubation at 37 °C for 16 h, cells were washed once with PBS and fixed in 4% PFA in PBS for 15 min at 4 °C. Cells were mounted on Gel/Mount (Cosmo Bio, Tokyo) and examined with an Axioskop 2 plus upright microscope (Zeiss), equipped with a Plan-Neofluar 40× objective and a pair of filters, one for observing CpYGFP fluorescence (Zeiss Filter Set No. 17, Excitation BP 485/20, FT 510, Emission BP 515-565) and another for DsRed fluorescence (Zeiss Filter Set No. 15, Excitation BP 546/12, FT 580, Emission LP 590), plus an AxioCam MRc cooled CCD camera (Zeiss). Fluorescent images were generated using AxioVision imaging software (Zeiss).

3. Results and discussion

3.1. Live C. poppei

C. poppei was identified from its morphological characteristics (Fig. 1A–C) and had an average prosome length of 1.70 ± 0.23 mm and the developmental stages ranged from C3 to C6 for males and females. The animal was brightly fluorescent when viewed using a non-UV transilluminator. It did not display any luminescence capability. When suspensions of the copepod were agitated in the dark, no light was observed with the dark-adapted (15 min) eye. Cell-free extracts, prepared by homogenizing the animal in 0.1 M Tris–HCl, pH 8.0, in an all-glass homogenizer at 4 °C and centrifuging at 13,000 ×*g*, did not give light when mixed with greater than 99% pure synthetic *Vargula hilgendorfii* luciferin or synthetic coelenterazine (data not shown).

3.2. Expression and screening of cDNA library, isolation of positive clone and sequence analysis of CpYGFP

Expression of the cDNA library of *C. poppei* yielded 2.9×10^5 independent colonies per ml (1.1×10^7 cfu/μg vector). Screening carried out on approximately 1.5×10^5 colonies using a non-UV transilluminator resulted in the isolation of a single colony with fluorescence activity. The isolated cDNA contained a 781-bp fragment with a 660-bp open reading frame, which encoded for a 219-amino acid polypeptide with a calculated molecular mass of 24.7 kDa. There was a 81-bp untranslated sequence followed by a terminal poly(A) tail after the stop codon. The cDNA sequence for CpYGFP has been deposited in the DDBJ/EMBL/GenBank™ databases with an accession number of AB185173. A homology search of non-redundant protein databases using the BLAST program revealed close amino acid sequence homology of CpYGFP with those of 6 other GFP-like proteins from the copepod species *Pontellina plumata*, *Pontella meadi*, *Labidocera aestiva*, and unclassified *Pontellidae* sp. SL-2003 (Shagin et al., 2004). From multiple amino acid sequence alignment with other GFP-like fluorescent proteins (Fig. 2A), it was presumed that the sequence Gly55-Tyr56-Gly57 formed the chromophore by a post-translational, autocatalytic dehydration–cyclization reaction, followed by an oxidation–dehydrogenation reaction (Cody et al., 1993; Niwa et al., 1996). The amino acid sequence around the putative chromophore region of CpYGFP was highly conserved relative to those of the other copepod proteins. The degree of over all identity between the amino acid sequence of CpYGFP and those of the other copepod proteins ranged from 49% to 54%, being considerably less homologous than among the 6 copepod proteins themselves (63–82%). This agrees with the fact that *C. poppei* belongs to the family Aetideidae, whereas 4 of the other copepods are members of the family Pontellidae. There were no significant similarities (12%) between the amino acid sequences of CpYGFP and *Aequorea* GFP. Based on phylogenetic analysis using the CLUSTAL W program, CpYGFP did not belong to the group of hydrozoan and anthozoan GFPs (data not shown), but was clustered with the copepod proteins. CpYGFP was most closely related to the *Pontellidae* sp. SL-2003 fluorescent protein and did not fall within any branch of the other copepod GFPs (Fig. 2B).

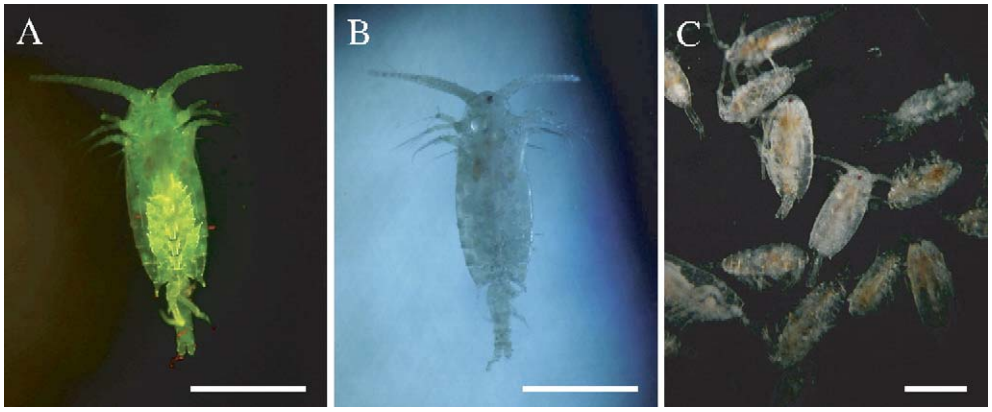

Fig. 1. Stereo microscopic observation of *C. poppei*. A, Image of fluorescing *C. poppei* taken using a "GFP" filter. B, Brightfield image of *C. poppei*. C, Specimens of *C. poppei* immediately before being used for RNA preparation. Magnification: horizontal line in each frame = 1.0 mm.

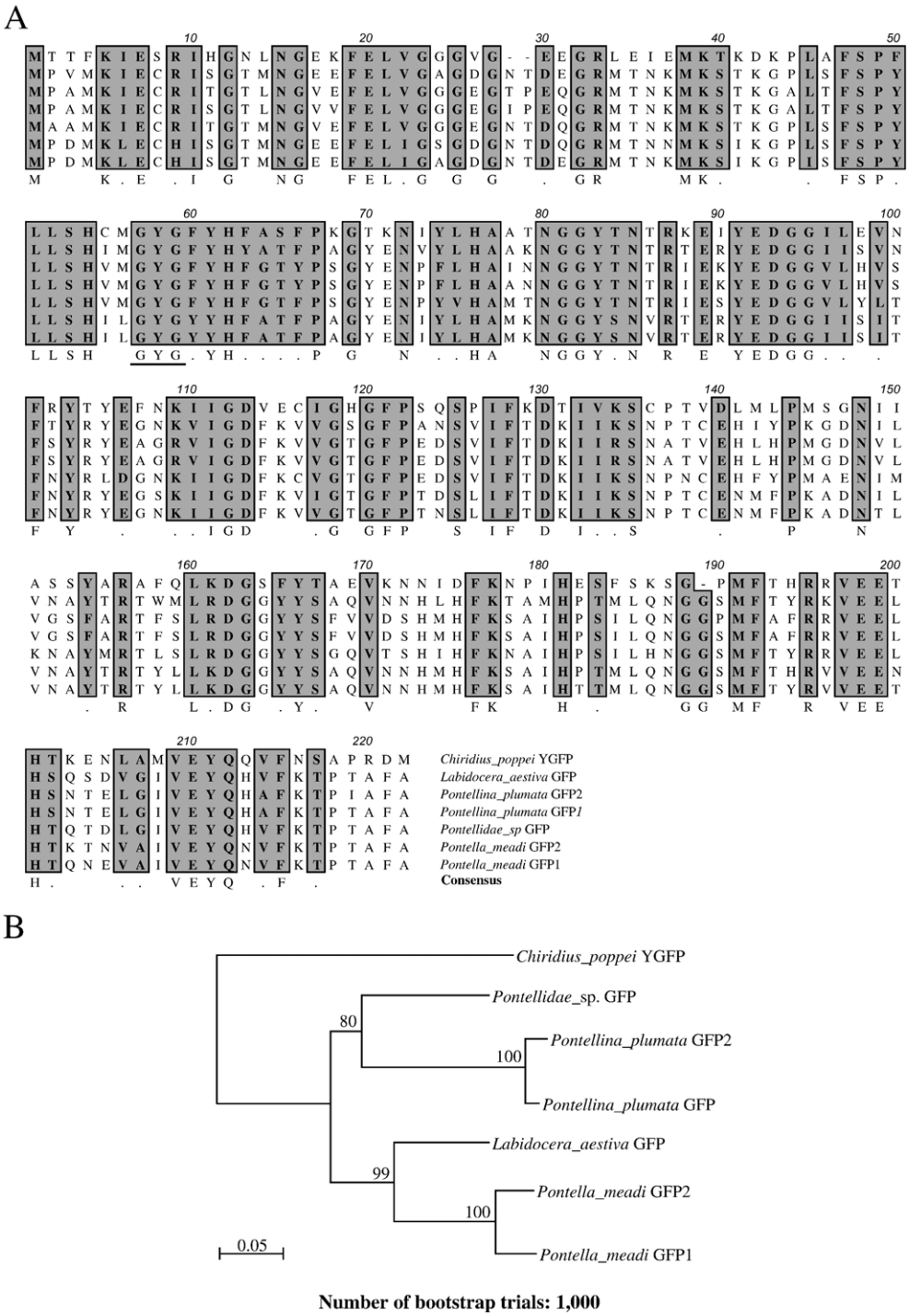

Fig. 2. Amino acid sequence and phylogenetic tree of copepod GFP-like proteins. A, Multiple amino acid sequence alignment of seven GFP-like proteins. Shaded residues and those in the bottom row are identical or similar amino acids. Organisms represented are *C. poppei*, *L. aestiva* (Accession No. AY268073), *P. plumata* (Nos. AY268071 and AY268072), *Pontellidae* sp. SL-2003 (No. AY268076), and *P. meadi* (Nos. AY268074 and AY268075). Underline indicates a putative chromophore region. B, Phylogenetic tree of copepod GFP-like proteins. The neighbor-joining distance tree is depicted by alignment of entire amino acid sequences. The bootstrap support for each tree branch is indicated as the percentage of 1000 bootstrap replicates. Branches were collapsed if bootstrap support was less than 50%.

3.3. Expression and purification of CpYGFP

In preparing enough recombinant CpYGFP for studying its biochemical properties, attempts were made to express CpYGFP both as a GST-tagged and non-tagged protein by employing expression vector pET101-CpYGFP in *E. coli* BL21-CodonPlus®(DE3)-RIL cell. The amount of soluble GST-CpYGFP fusion protein in the lysate of *E. coli* was relatively low (data not shown), probably due to its presence in inclusion bodies. However, non-tagged CpYGFP was found largely in the soluble fraction of the *E. coli* lysate. When the soluble fractions of induced and non-induced crude cell lysates

were analyzed by SDS-PAGE, only the incubation mixture to which IPTG had been added showed a large protein band with a mass of around 25 kDa, which was attributed to newly synthesized CpYGFP (Fig. 3, lane 1). The CpYGFP was subsequently purified to homogeneity by sequential application of anion-exchange chromatography (Fig. 3, lane 2), hydrophobic interaction chromatography (Fig. 3, lane 3), and by gel-filtration (Fig. 3, lane 4) using the ÄKTA Explorer System. The elution of CpYGFP from the column was monitored at 280 and 500 nm. The yield was 2.7 mg of purified CpYGFP from 200 ml of culture medium (13.5 mg/L LB medium).

3.4. Characterization of expressed CpYGFP

Purified recombinant CpYGFP had a weak yellowish-green color in daylight. The fluorescence excitation and emission maxima of 507 and 517 nm, respectively, were unique, when compared to those of 490 and 509 nm for enhanced GFP (GFPmut1) (Cormack et al., 1996), and 512 and 529 nm for enhanced yellow fluorescent protein (GFP-10C) (Ormö et al., 1996) (Fig. 4). When purified recombinant CpYGFP was analyzed by MALDI-TOF MS, the measured molecular mass (desalted, m/z 24,565.0±0.2) was slightly less than the average molecular mass (24,739.6 Da) (K. Suto, personal communication). Although the N-terminal amino acid sequence was not determined, the difference was assumed to be due to the first methionine being removed after translation of CpYGFP by methionyl-aminopeptidase in *E. coli*. Excision of the N-terminal methionine is known to depend largely on the type of amino acid residue in the penultimate position in the polypeptide chain (Hirel et al., 1989). Threonine, the second amino acid at the N-terminus of CpYGFP, is also known to greatly affect methionine processing at a frequency of 89.7±4.5%. The pI was estimated to

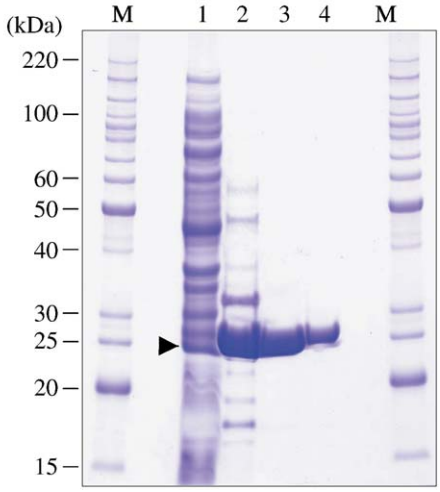

Fig. 3. SDS-PAGE analysis of expressed CpYGFP in *E. coli* BL21-CodonPlus (DE3)-RIL cells. The cells were transformed with pET101-CpYGFP and the fluorescent CpYGFP fractions were pooled and analyzed at each step in the purification by SDS-PAGE (12.5% polyacrylamide gel; reducing conditions). M, BenchMark Protein Ladder (Invitrogen). Lane 1, soluble fraction of *E. coli* cells 6 h after 1 mM IPTG induction. Lane 2, DEAE Sepharose F.F. Lane 3, HiTrap Phenyl HP. Lane 4, HiLoad 16/60 Superdex 200 pg. The gels were stained with CBB. Arrowhead indicates induced CpYGFP protein band.

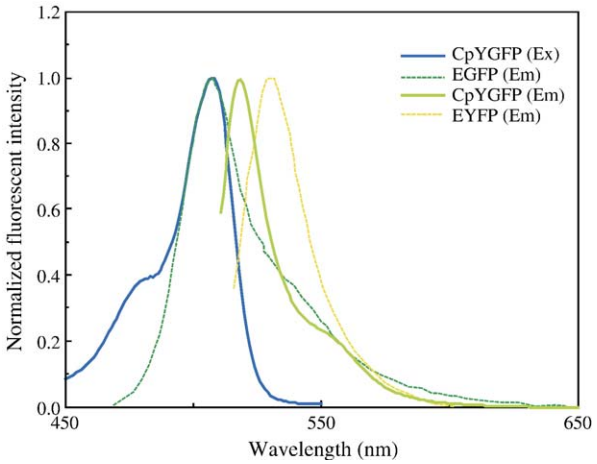

Fig. 4. Fluorescence excitation and emission spectra of recombinant CpYGFP. Normalized excitation (blue line) and emission (yellowish-green line) spectra of CpYGFP in 20 mM Tris–HCl, pH 8.5, are shown with excitation and emission maxima at 507 and 517 nm, respectively. Fluorescence emission spectra of EGFP (green dot line) and EYFP (yellow dot line) are plotted based on a previous report (Cormack et al., 1996; Chalfie and Kain, 1998).

be 6.5 from the pKa's of the amino acid composition of CpYGFP using program GENETYX-MAC (vers. 10.l), and the molecular extinction coefficient was 99,476 M^{-1} cm^{-1} (λ=500 nm).

The oligomeric state of CpYGFP was analyzed by size exclusion chromatography, using a gel-filtration column equilibrated with 20 mM Tris–HCl, pH 8.5. From its retention time, the molecular mass of the protein was estimated to be about 100 kDa, which would make CpYGFP a tetramer under the eluting conditions (data not shown). When the chromatography was repeated using 20 mM Tris–HCl, pH 8.5, containing 200 mM NaCl, CpYGFP eluted as a dimer with a molecular mass of about 50 kDa (data not shown). In dilute solutions native *Renilla* GFP is known to exist as a dimmer of two identical subunits held by non-covalent bonds, whereas *Aequorea* GFP occurs as a monomer (Ward et al., 1982). The above results suggest that CpYGFP can exist as a dimer or tetramer and that the interaction between the dimer partners forming the tetramer complex is largely electrostatic in nature and not very strong.

Because *Aequorea* GFP has been reported to be stable over a wide range of pH's and temperatures, as well as being resistant to denaturing agents (Ward, 1981; Ward and Bokman, 1982), the stability of CpYGFP was examined under similar conditions. Fig. 5A shows the effect of pH on the fluorescence intensity of CpYGFP plotted alongside comparable data (normalized) for *Aequorea* GFP. It is seen that from about pH 5 to 10, the fluorescence intensity curve of CpYGFP is about twice the height of that of *Aequorea* GFP. However, both curves are seen to appear flat and stable between around pHs 6 and 10. The pKa's of the chromophores of *Aequorea* GFP and *Renilla* GFP have been estimated spectrophotometrically to be around 7.9 (Shimomura, 1979) and 8.1 (Ward et al., 1980), respectively. Assuming that the light-emitter of CpYGFP is the ionized phenol group of tyrosine in the chromophore (Niwa et al., 1996) and that the fluorescence intensity curve is a reflection of this ionization, we estimate from the midpoint of the curve of CpYGFP (Fig. 5A) that the pKa of the chromophore is also about 8.0.

Fig. 5B shows the thermostability curve for CpYGFP with fluorescence plotted against increasing temperatures, compared to those (normalized) for *Aequorea* GFP. The fluorescence intensities are seen to be almost the same for both proteins between 25 and 60 °C, declining rapidly to zero around 80 °C. The T_m (temperature at which one-half of the fluorescence intensity is lost) for CpYGFP was around 66 °C, indicating that the protein is relatively stable against heat denaturation.

Fig. 5C shows the effect of various chemical reagents on CpYGFP, which has 3 cysteine residues at positions 55, 116 and 136. To determine whether an inter- or intramolecular disulfide bond is essential for CpYGFP fluorescence, the protein was incubated for 20 min in 20 mM Tris–HCl, pH 8.5, containing either 10 or 100 mM 2-mercaptoethanol or dithiothreitol. As shown in Fig. 5C, neither reagent caused any decrease in fluorescence intensity. The data in Fig. 5C also show CpYGFP to be relatively stable in 8M urea, 50% acetone, 50% ethanol, 50% methanol, 1% SDS, and 1% PFA, but highly or completely unstable in 10% formaldehyde, 10% TCA and 100% solvent.

3.5. Expression of CpYGFP and co-expression of CpYGFP-tagged actin in HeLa cells

To evaluate CpYGFP as a possible reporter protein, the mammalian expression vector, pcDNA3.2-CpYGFP, for expression of CpYGFP alone, was constructed. HeLa cells were transiently transfected and after 16 h of incubation at 37 °C, the transformed cells were analyzed for protein localization using an Axioskop (Zeiss) fluorescence microscope. Aggregated particles were rarely seen. The transfection efficiency was relatively low (10–20%) as judged by the number of fluorescent cells versus total cells. Nonetheless, CpYGFP expressing cells (Fig. 6A) exhibited bright yellowish-green fluorescence which was uniformly distributed throughout the cytoplasm and nucleus. A brightfield image of the same cells is shown in Fig. 6B.

For the CpYGFP-tagged actin experiment, the CMV promoter-driven fusion protein construct pcDNA3.2-CpYGFP-actin was employed. Actin was selected as the fusion partner of CpYGFP because it is capable of undergoing specific intracellular localization and has been extensively studied as a fusion protein with *Aequorea* GFP (Westphal et al., 1997; Robbins et al., 1999). In this respect, Fig. 6C shows HeLa cells expressing *Aequorea* EGFP-actin after transfection with pEGFP-Actin (Clontech). The second marker selected was the anthozoan red fluorescent protein, DsRed, with excitation and emission maxima at 558 and 583 nm, respectively. The plasmid pDsRed2-Nuc, which produces a nuclear-localizing DsRed protein, was used to co-transfect HeLa cells with pcDNA3.2-CpYGFP-actin. Fig. 6D shows an image of co-expressed CpYGFP-actin localized in actin-containing subcellular structures, whereas Fig. 6E shows an image of co-expressed DsRed2-Nuc localized in the nucleus. In the latter case, the image shows little, if any,

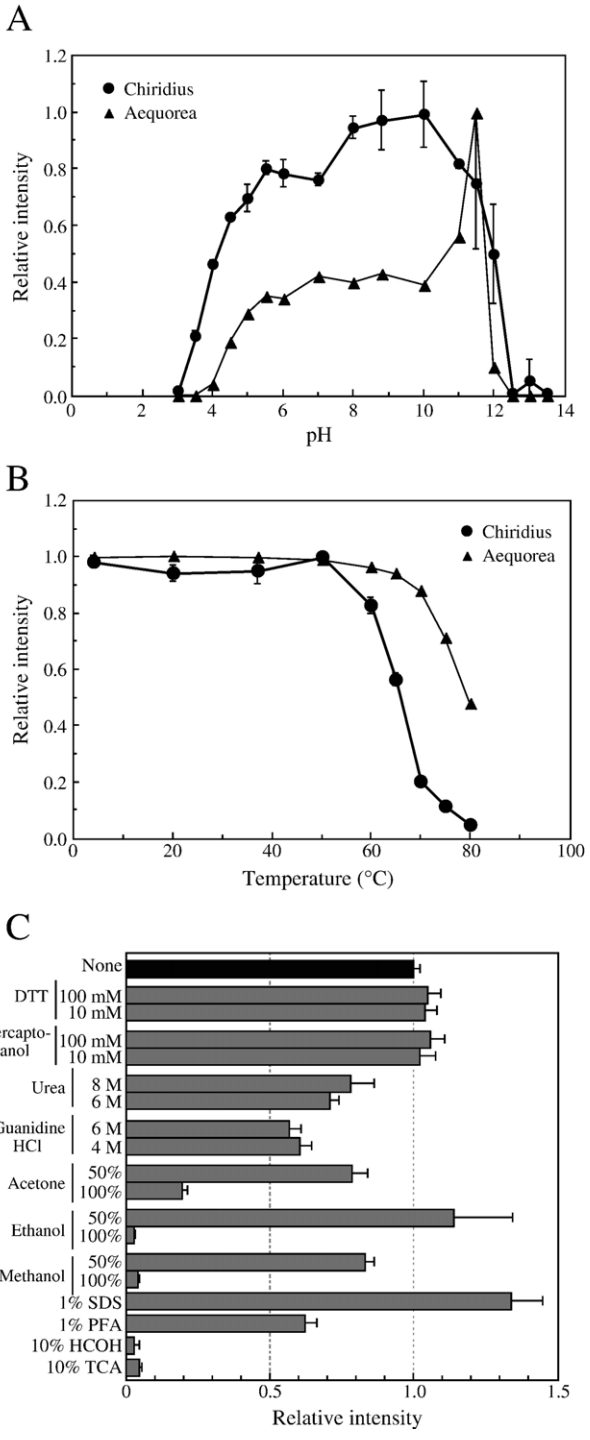

Fig. 5. Characteristic properties of recombinant CpYGFP. A, Plot of CpYGFP fluorescence versus pH, showing CpYGFP in solid circles and recombinant *Aequorea* GFP in solid triangles. The fluorescence intensities of CpYGFP and *Aequorea* GFP were measured at 25 °C after incubation at 4 °C for 10 min in the following buffers: 0.1 M glycine–HCl (pH 3.0–3.5), 0.1 M acetate buffer (pH 4.0–5.5), 0.1 M phosphate buffer (pH 6.0), 0.1 M HEPES buffer (pH 7.0), 0.1 M Tris–HCl buffer (pH 8.0–8.9), 0.1 M carbonate buffer (pH 10.0–11.0), 0.1 M phosphate–NaOH buffer (pH 11.5–13.0) and 0.1 M NaOH (pH 13.6). B, Plot of CpYGFP fluorescence versus temperature, plotted together with data for recombinant *Aequorea* GFP. Both samples were continuously heated at 4–80 °C in 20 mM Tris–HCl, pH 8.5, for 10 min. Fluorescence intensities were measured at 25 °C. C, Chemical stability of CpYGFP. Samples were diluted 20 fold by adding 20 mM Tris–HCl, pH 8.5, containing each of the reagents at the concentrations indicated and incubating at 4 °C for 10 min. Fluorescence intensity was normalized against a control sample containing none of the reagent and diluted 20 fold with 20 mM Tris–HCl buffer, pH 8.5. All values for CpYGFP were determined by three different experiments ($n=3$).

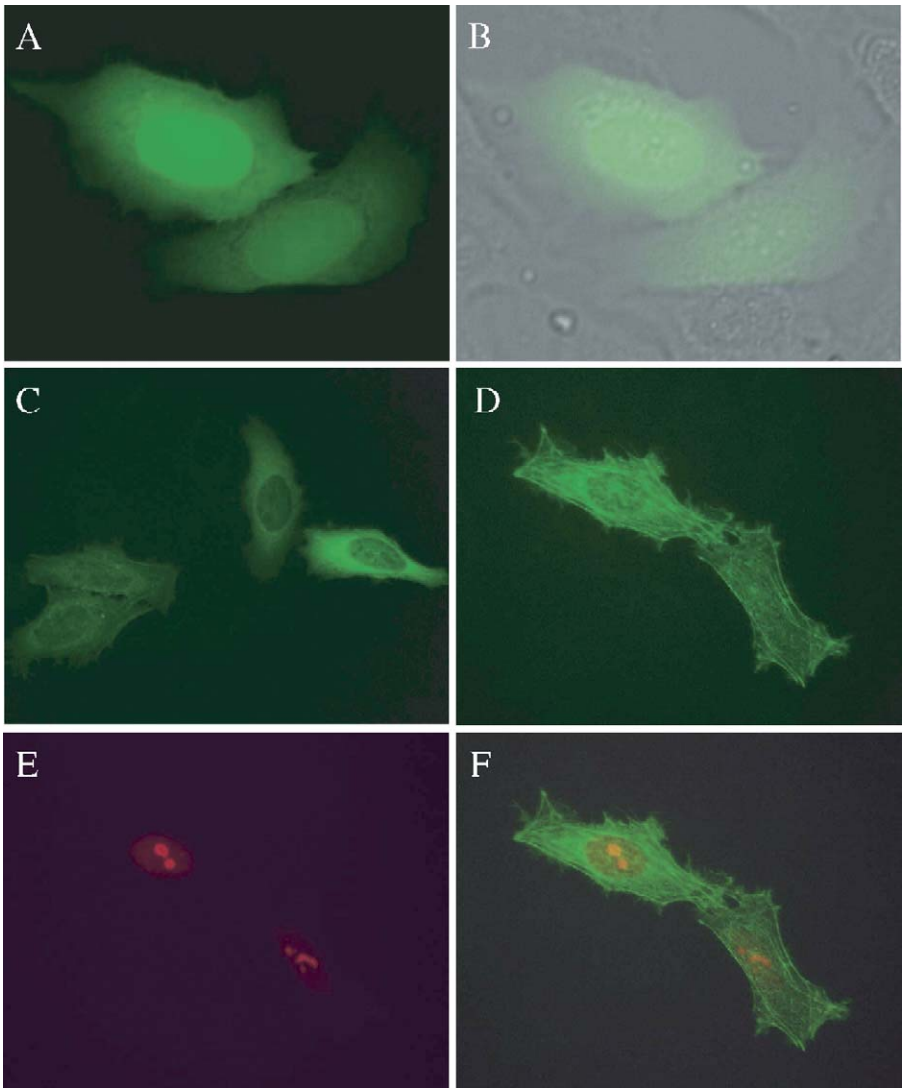

Fig. 6. Expression of CpYGFP, and co-expression of CpYGFP-actin and DsRed2-Nuc, in HeLa cells. A, Fluorescence image of HeLa cells expressing CpYGFP, taken with Zeiss Filter Set No. 17. B, Brightfield image of same HeLa cells in A. C, Fluorescence image of HeLa cells expressing EGFP-actin. D, Subcellular actin localization of CpYGFP-actin in HeLa cells co-expressing DsRed2-Nuc, taken with Zeiss Filter Set No. 17. E, Subcellular nuclear localization of DsRed2-Nuc in HeLa cells co-expressing CpYGFP-actin, taken with Zeiss Filter Set No. 15. F, Merged image of HeLa cells co-expressing CpYGFP-actin and DsRed2-Nuc (from D and E).

detectable fluorescence from CpYGFP-actin. Therefore, employing the two filter sets it was possible to obtain a composite image of actin and nuclear structures in both fixed (Fig. 6F) and living (data not shown) HeLa cells. A further advantage with CpYGFP is that background fluorescence and cell damage are minimized during imaging due to the long wavelengths employed. Preliminary experiments with pcDNA3.2-CpYGFP-tubulin to visualize intracellular tubulin structures also gave similar results (data not shown).

The anthozoan red fluorescent protein, DsRed, is an obligate tetramer with a tendency to form intracellular aggregates when expressed as a DsRed-tagged protein (Lauf et al., 2001). To overcome this problem, some workers have resorted to site-specific mutagenesis to obtain non-aggregating monomeric (Campbell et al., 2002) and pseudo-monomeric (Bulina et al., 2003) forms of the protein. Although CpYGFP did form tetramers and dimers during size-exclusion chromatographies, no apparent aggregation was detected on microscopic examination of the HeLa cells. Instead, CpYGFP-tagged actin appeared to be properly assembled into characteristic actin filaments (Fig. 6D). We assumed that oligomer formation of CpYGFP could be interfered and disrupted by fusing actin protein at C-terminus of CpYGFP or by the formation of actin filaments without loosing its fluorescent activity. Second possible explanation is that CpYGFP-actin oligomer was properly integrated into actin polymer without giving any conformational influence on it. CpYGFP-tagged with tubulin also gave a proper localization in microtubules (data not shown). These results support that fluorescent proteins that have tendency to form oligomer can also be applied to the use in fusion protein studies as seen in the DsRed-tagged proteins. However, for the further application of CpYGFP to other proteins as a reporter, it should be monomerized by site-directed or random mutagenesis without changing its fluorescent properties. The possibility exists that once the crystal structure is known, CpYGFP may be mutated

so that it can exist as a monomer in solution. Several mutant proteins of CpYGFP expressed in *E. coli* are currently under investigation.

It is noteworthy that CpYGFP has an unique fluorescent excitation and emission maximum at 507 and 517, respectively, which are midpoints of those of EGFP and EYFP, and also apparently distinguishable from those of other all copepod fluorescent proteins reported previously (Ex; 480-491, Em; 500-511) (Shagin et al., 2004). Structural basis for CpYGFP will explain how it is different from other fluorescent proteins and how it emits yellowish-green fluorescence. Other advantage of CpYGFP is that it is smaller (219-amino acids) than other fluorescent proteins (222-amino acids for other 6 copepod proteins, 238- and 225-amino acids for *Aequorea* GFP and DsRed, respectively). With these special features, CpYGFP has the potential to be an optional protein for an acceptor or donor in the FRET assay, not only as a reporter protein in living cells as shown in Fig. 6. In conclusion, the present results demonstrate that CpYGFP may find valuable use as a reporter in future cell biology studies.

Acknowledgements

This work was supported in part by Special Coordination Funds for the Promotion Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology of Japan (H. Mizuno) and matching research funds of AIST and NEC Soft, Ltd. (H. Masuda, H. Mizuno and S. Nishikawa). The authors are greatly indebted to the following individuals: Prof. F. I. Tsuji and Mr. N. Nakura, for critical reading and editing of the manuscript; Prof. T. Ikeda and Prof. M. D. Ohman, for aid in identifying *C. Poppei*; Dr. K. Suto, for mass spectral analysis of recombinant CpYGFP; Drs. N. Ishida and K. Miyazaki, for use of the fluorescence microscope; and Ms. A. Hamada and Dr. T. Umehara, for expert technical assistance.

References

Bulina, M.E., Verkhusha, V.V., Staroverov, D.B., Chudakov, D.M., Lukyanov, K.A., 2003. Hetero-oligomeric tagging diminishes non-specific aggregation of target proteins fused with Anthozoa fluorescent proteins. Biochem. J. 371, 109–114.

Campbell, R.E., et al., 2002. A monomeric red fluorescent protein. Proc. Natl. Acad. Sci. U. S. A. 99, 7877–7882.

Chalfie, M., 1995. Green fluorescent protein. Photochem. Photobiol. 62, 651–656.

Chalfie, M., Kain, S., 1998. Molecular Biology and Mutation of Green Fluorescent Protein: Green Fluorescent Protein: Properties, Applications, and Protocols. Wiley-Liss, New York.

Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G., Ward, W.W., 1993. Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. Biochemistry 32, 1212–1218.

Cormack, B.P., Valdivia, R.H., Falkow, S., 1996. FACS-optimized mutants of the green fluorescent protein (GFP). Gene 173, 33–38.

Gurskaya, N.G., et al., 2003. A colourless green fluorescent protein homologue from the non-fluorescent hydromedusa *Aequorea coerulescens* and its fluorescent mutants. Biochem. J. 373, 403–408.

Hirel, P.H., Schmitter, M.J., Dessen, P., Fayat, G., Blanquet, S., 1989. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. Proc. Natl. Acad. Sci. U. S. A. 86, 8247–8251.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., Gibson, T.J., 1998. Multiple sequence alignment with Clustal X. Trends Biochem. Sci. 23, 403–405.

Lauf, U., Lopez, P., Falk, M.M., 2001. Expression of fluorescently tagged connexins: a novel approach to rescue function of oligomeric DsRed-tagged proteins. FEBS Lett. 498, 11–15.

Markova, S.V., Golz, S., Frank, L.A., Kalthof, B., Vysotski, E.S., 2004. Cloning and expression of cDNA for a luciferase from the marine copepod *Metridia longa*. A novel secreted bioluminescent reporter enzyme. J. Biol. Chem. 279, 3212–3217.

Matz, M.V., et al., 1999. Fluorescent proteins from nonbioluminescent Anthozoa species. Nat. Biotechnol. 17, 969–973.

Nishiuchi, Y., et al., 1998. Chemical synthesis of the precursor molecule of the *Aequorea* green fluorescent protein, subsequent folding, and development of fluorescence. Proc. Natl. Acad. Sci. U. S. A. 95, 13549–13554.

Niwa, H., et al., 1996. Chemical nature of the light emitter of the *Aequorea* green fluorescent protein. Proc. Natl. Acad. Sci. U. S. A. 93, 13617–13622.

Ogawa, H., Inouye, S., Tsuji, F.I., Yasuda, K., Umesono, K., 1995. Localization, trafficking, and temperature-dependence of the *Aequorea* green fluorescent protein in cultured vertebrate cells. Proc. Natl. Acad. Sci. U. S. A. 92, 11899–11903.

Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., Remington, S.J., 1996. Crystal structure of the *Aequorea victoria* green fluorescent protein. Science 273, 1392–1395.

Robbins, J.R., Barth, A.I., Marquis, H., de Hostos, E.L., Nelson, W.J., Theriot, J.A., 1999. *Listeria monocytogenes* exploits normal host cell processes to spread from cell to cell. J. Cell Biol. 146, 1333–1350.

Shagin, D.A., et al., 2004. GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. Mol. Biol. Evol. 21, 841–850.

Shimomura, O., 1979. Structure of the chromophore of *Aequorea* green fluorescent protein. FEBS Lett. 104, 220–222.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Verhaegen, M., Christopoulos, T.K., 2002. Recombinant *Gaussia* luciferase. Overexpression, purification, and analytical application of a bioluminescent reporter for DNA hybridization. Anal. Chem. 74, 4378–4385.

Verkhusha, V.V., Lukyanov, K.A., 2004. The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. Nat. Biotechnol. 22, 289–296.

Ward, W.W., 1981. Properties of the Coelenterate Green-Fluorescent Proteins: Bioluminescence and Chemiluminescence: Basic Chemistry and Analytical Applications. Academic Press, New York.

Ward, W.W., Bokman, S.H., 1982. Reversible denaturation of *Aequorea* green-fluorescent protein: physical separation and characterization of the renatured protein. Biochemistry 21, 4535–4540.

Ward, W.W., Cormier, M.J., 1979. An energy transfer protein in coelenterate bioluminescence. Characterization of the *Renilla* green-fluorescent protein. J. Biol. Chem. 254, 781–788.

Ward, W.W., Cody, C.W., Hart, R.C., Cormier, M.J., 1980. Spectrophotometric identity of the energy transfer chromophores in *Renilla* and *Aequorea* green-fluorescent proteins. Photochem. Photobiol. 31, 611–615.

Ward, W.W., Prentice, H.J., Roth, A.F., Cody, C.W., Reeves, S.C., 1982. Spectral perturbations of the *Aequorea* green-fluorescent protein. Photochem. Photobiol. 35, 803–808.

Westphal, M., et al., 1997. Microfilament dynamics during cell movement and chemotaxis monitored using a GFP-actin fusion protein. Curr. Biol. 7, 176–183.