Complementary DNA Coding Click Beetle Luciferases Can Elicit Bioluminescence of Different Colors

KEITH V. WOOD,* Y. AMY LAM, HOWARD H. SELIGER, WILLIAM D. MCELROY

Eleven complementary DNA (cDNA) clones were generated from messenger RNA isolated from abdominal light organs of the bioluminescent click beetle, *Pyrophorus plagiophthalamus*. When expressed in *Escherichia coli*, these clones can elicit bioluminescence that is readily visible. The clones code for luciferases of four types, distinguished by the colors of bioluminescence they catalyze: green (546 nanometers), yellow-green (560 nanometers), yellow (578 nanometers), and orange (593 nanometers). The amino acid sequences of the different luciferases are 95 to 99 percent identical with each other, but are only 48 percent identical with the sequence of firefly luciferase (*Photinus pyralis*). Because of the different colors, these clones may be useful in experiments in which multiple reporter genes are needed.

NEARLY ALL OUR KNOWLEDGE OF beetle luciferases is derived from studies of a single species, the North American firefly *Photinus pyralis*. Comparative studies with other beetle luciferases have been hampered because of limited availability of the other species. Evolutionarily, beetle luciferases are unrelated to any of the other groups of luciferases that have been studied biochemically (*1*). Little is known about the luciferases from other beetles except that they all catalyze the production of various colors of light through the oxidative decarboxylation of beetle luciferin (*2*). Since the substrates of the luminescent reaction are the same in all these beetles, the different colors must be due to differences in the structure of the enzymes (*3*).

Recently we cloned a cDNA that codes for the luciferase of *P. pyralis*, and have shown that it can be used to express bioluminescence in *Escherichia coli*. We report here the cloning of cDNAs that code for several new luciferases from a bioluminescent click beetle, *Pyrophorus plagiophthalamus*. This beetle is unusual because it can emit bioluminescence of a wide range of colors from a single species. The expression products in *E. coli* of the cDNAs derived from this beetle are able to produce green, yellow-green, yellow, and orange light. As determined from the nucleotide sequences of the clones, the amino acid sequences of these click beetle luciferases are highly conserved among one another, but diverge from the sequence of the firefly luciferase. Taxonomy indicates that the click beetle luciferases probably are the most evolutionarily distant of the beetle luciferases from the firefly luciferase (*4*). This distance is reflected by differences in their chemical properties.

Pyrophorus plagiophthalamus is a large beetle with two sets of light organs. One set, on the dorsal surface of the head, emits light that is greenish but the exact color varies between individual beetles of the species, ranging from green (548 nm) to yellow-green (565 nm). The other set, at the anterior of the abdomen, generally emits light of a longer wavelength than the head organs but also varies between individuals ranging from green (547 nm) to orange (594 nm) (*5*). We converted mRNA isolated from the abdominal light organ of 60 beetles to cDNA and inserted this into a specialized lambda cloning vector, Lambda ZAP (*6*). The ability to convert this modified lambda vector into a bacterial expression plasmid (Bluescript) through an in vivo process allowed us to screen the cDNA library by two methods (*7*). In the phage form of the library, we screened with antibody to firefly luciferase that cross-reacts with the click beetle luciferases (*8*) and isolated four full-length clones that expressed bioluminescence in *E. coli*. A portion of the cDNA library was converted into the plasmid form, and we screened this for bioluminescence in the bacterial colonies. Bioluminescence can be initiated in colonies of *E. coli* expressing luciferase by adding luciferin to the media (*9*). Seven more cDNA clones were isolated by this method. It was determined visually that of the eleven clones, one produced green light, one produced yellow-green light, six produced yellow light, and three produced orange light.

Immunoblot analysis confirmed the production of full-length click beetle luciferase in *E. coli*. Despite some of these clones being detected with antibody to firefly luciferase during the library screening of plaques, we could not detect the gene products in blots made directly with *E. coli* lysates. The expression of bioluminescence was improved by transferring the cDNA clones into a plasmid vector incorporating the *tac* promoter (*10*). A lysate from *E. coli* expressing the green-emitting luciferase from this vector was partially purified. After gel electrophoresis and blotting, a single antigenic band was revealed that comigrated with the native click beetle luciferase. Subsequently one cDNA clone from each of the four color-emitting groups was sequenced. An open reading frame was revealed in each that could potentially code a protein, the sequence of which correlated with the entire length of the sequence for firefly luciferase. Thus the complete protein coding regions of the click beetle luciferases were apparently contained within their cDNA clones.

Expression of bioluminescence from the *tac* vector yielded sufficient intensity, upon addition of luciferin to the media, to allow measurement of the spectral distribution from intact cells (Fig. 1). This confirmed the visual assignment of the 11 cDNA clones into four color groups. For each of the four colors, the bioluminescence spectrum is a single peak qualitatively similar to the spectra of the native click beetle luciferases (*3*). The range of colors from the clones is representative of the full range measured from the abdominal light organs of living beetles. However, there are colors emitted by the beetles, within the extremes of this range, that do not correspond to any of the clones (*5*). Thus other luciferase genes may not have been isolated. Spectra of the luciferases were also measured from partially purified preparations obtained from lysates of the *E. coli* expressing the cDNA clones.

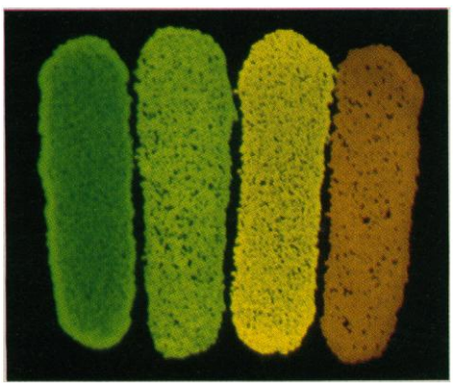

Fig. 1. Bioluminescence from colonies of *E. coli* expressing the click beetle luciferases. Four streaks of *E. coli*, each consisting of hundreds of colonies, show the four colors of bioluminescence emitted by the different luciferases. The colonies were grown on nitrocellulose filters layered on top of nutrient agar. To initiate the bioluminescent reaction, the filters were removed from the agar and soaked with 1 m*M* luciferin in 100 m*M* sodium citrate, *p*H 5.0. The photograph was produced from a 2-s contact exposure of the colonies onto Ektachrome 64.

K. V. Wood, Y. A. Lam, W. D. McElroy, Department of Chemistry, M-001, University of California, San Diego, La Jolla, CA 92093.
H. H. Seliger, Department of Biology, Johns Hopkins University, Baltimore, MD 21218.

*To whom correspondence should be addressed.

Between *p*H 6 to 7, the spectra of these preparations were indistinguishable from those of intact cells. At *p*H 8 there was a slight broadening of the spectra for the green- and yellow-emitting luciferases. The firefly luciferase shows a large spectral shift between *p*H 6 to 8. At *p*H 8 its spectral maximum is at 560 nm, which shifts to 615 nm (red) at *p*H 6 with a decrease in the quantum yield (*11*).

The sequences of the different click beetle luciferases are highly similar (Fig. 2). The open reading frame of each of the sequenced cDNA clones potentially codes a 543-residue polypeptide. Comparisons of the derived amino acid sequences show a 95 to 99% identity between the different color-emitting luciferases. Thus the number of amino acids that are responsible for the differences in the color is small. Because variation in color results directly from differences in the primary structures of the luciferases, specialized posttranslational modifications or unusual microenvironmental effects are not necessary to account for the color variation in the living beetles.

Comparison of the sequences of click beetle luciferases with that of firefly luciferase shows a low similarity. Alignment of their deduced amino acid sequences reveals that the various click beetle and the firefly luciferases are 48% identical (Fig. 3). Six gaps in the alignment of one to two amino acids in length account for most of a seven–amino acid difference in the lengths of the open reading frames between the firefly and click beetle luciferases. No regions in the alignment show especially high sequence similarity, thus giving little indication that particular regions have been conserved because of catalytic or structural constraints on the enzymes. An exception to this is in the last three amino acids which, for the firefly luciferase, have been shown to be necessary for translocation into peroxisomes (*12*). Given the close functional similarity of these enzymes, it is almost certain that the click beetle luciferases are also located in peroxisomes.

Firefly luciferase has historically been used as a bioluminescent reporter of chemical events associated with adenosine triphosphate (ATP) metabolism (*13*). With the cloning of its cDNA, this luciferase has also recently found application as an effective reporter of genetic events (*14, 15*). Its principal advantages are that (i) the initial polypeptide derived from the mRNA requires no posttranslational modifications for enzymatic activity; (ii) the luminescent reaction can be measured with high sensitivity; (iii) the assay of the gene product is rapid and does not use substrates requiring special precautions (such as radioactive isotopes or chemically unstable compounds); and (iv) gene expression may be detected without disruption of living tissue. Compared with the conventionally used assay of chloramphenicol acetyltransferase (CAT) for gene activity, firefly luciferase is assayed in minutes as opposed to hours, and is 100 to 1000 times more sensitive (*15*).

The cDNAs coding for the click beetle luciferases also have these features, and, as they can be distinguished by color, may be useful in situations where multiple reporters are desirable. Expression in exogenous hosts should differ little between these luciferases because of their sequence similarity. Also, since the colors do not shift near physiological *p*H, the different luciferases can be distinguished in vivo as well as in vitro. Thus the click beetle luciferases may provide a dual reporter system that can allow two different promoters to be monitored within a single host, or for different populations of

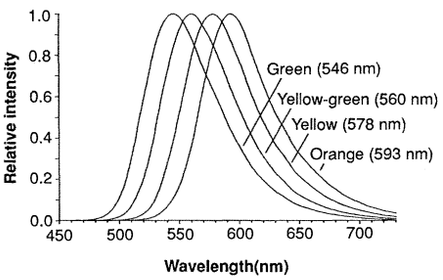

Fig. 2. Spectra of the click beetle luciferases, from intact *E. coli* immersed in luciferin, show four overlapping peaks of nearly even spacing. The maximum intensity of each spectrum has been normalized. The spectra were measured on a Fastie-Ebert–type grating spectrometer (*2*) and were corrected for the spectral sensitivity of the photomultiplier tube and for time-dependent variation in the intensity of the luminescent reactions. Wavelength positions were calibrated against the spectral lines of a mercury vapor lamp. Virtually identical spectra were obtained from lysates of these *E. coli* between *p*H 6 to 8. The lysates were prepared from cells grown to middle or late log-phase growth at 37°C, and then for 2 hours at 30°C with isopropylthiogalactoside (IPTG) added to 1 m*M*. The cells were washed and resuspended in approximately 1/150 volume of the culture with 100 m*M* potassium phosphate, *p*H 7.0, 2 m*M* EDTA, and 35% of saturation $(NH_4)_2SO_4$. After lysis by sonication and removal of the debris by centrifugation, $(NH_4)_2SO_4$ was added to 53% of saturation and the precipitate was dissolved in 1/15 volume of the lysate with 100 m*M* potassium phosphate, *p*H 7.0, 2 m*M* EDTA, and 50% glycerol. The spectra were measured from 10 μl of this solution diluted 100-fold with 50 m*M* 2-(*N*-morpholino)ethanesulfonic acid, 50 m*M* 2-(*N*-morpholino)propanesulfonic acid (MOPS), 50 m*M* tricine, 5 m*M* $MgSO_4$, 1 m*M* EDTA, 0.1 m*M* luciferin, 1.5 m*M* ATP, 1 m*M* NaF, 0.2 mg of bovine serum albumin per milliliter, and 10% glycerol. The spectra were measured at *p*H 6.0, 7.0, and 8.0.

```
FF   .EDA..IKK..A.FY....G....Q.HK.MKRYALV.GTI.FT.AHI.VN.T.A.Y..MSVR..EAMKRY.LNTNHRIVV.S..SLQ..M.   91
GR  MMKREKNVVYGPEPLHPLEDLTAGEMLFRALRKHSHLPQ--ALVDVYGEEWISYKEFFETTCLLAQSLHNCGYKMSDVVSICAENNKRFFVP  90
YG  ........I.............................  .....F.D.SL.......A.............N..............I.  90
YE  ........I...........K.................  .I...F.D.SL.......A.............N..............I.  90
OR  ........I...........K.................  .I...F.D.SL.......A.............N..............I.  90

FF  VLG.LF..VA...A.DI.NER..LNS.N..Q.TV..VS.KG.Q.I.N..KKLPI.QK...M.SKTDYQ.FQ.MYT.VT.HLPP.F.EYD.V.  183
GR  IIAAWYIGMIVAPVNEGYIPDELCKVMGISRPQLVFCTKNILNKVLEVQSRTDFIKRIIILDAVENIHGCESLPNFI-SRYSDG-NIANFKP  180
YG  ..............S.............K..I.................N.........T.............. ..... .......  180
YE  ..............S.............K..I.................N.........T.............. ..... .......  180
OR  ..............S.............K..I.................N.........T.............. ..... .......  180

FF  ESF.RDKTI.L.MN...S.......ALP..TA...FS..R..IF.N.I..DTAI.SVV..H.G..MFTT...LIC.F..VLMY..EE.L..R  275
GR  LHYDPVEQVAAILCSSGTTGLPKGVMQTHRNVCVRLIHALDPRVGTQLIPGVTVLVYLPFFHAFGFSINLGYFMVGLRVIMLRRFDQEAFLK  272
YG  ............................Q.I...........A.....................................................  272
YE  ............................Q.I...........EA..........V.........................E.....  272
OR  ............................Q.I...........EA..........V........G.................E.....  272

FF  SL...KIQ.ALL..TLFS.FA..T.I......N.H.IAS.G...S...G.AVA..FH.....Q.Y....T...ILITPEGDD.P.AV.K.V.  367
GR  AIQDYEVRSVINVPAIILFLSKSPLVDKYDLSSLRELCCGAAPLAKEVAEIAVKRLNLPGIRCGFGLTESTSANIHSLRDEFKSGSLGRVTP  364
YG  ..........................................V.....................G.............  364
YE  ........IV..................................................................G.............  364
OR  ........IV..................................................................G.............  364

FF  FFE..VV.LD...T..V..R....VR...IMS.....P...NAL..K........IA.W......FI...L.S......Y........S...  459
GR  LMAAKIADRETGKALGPNQVGELCIKGPMVSKGYVNNVEATKEAIDDDGWLHSGDFGYYDEDEHFYVVDRYKELIKYKGSQVAPAELEEILL  456
YG  ........................V.....................................................................  456
YE  ........................V.....................................................................  456
OR  ....................................K.....................................................................  456

FF  QH.N.F.AG.A.L..DD.....A.V..LEH..TM.E..IV..V.SQ.TTA.K.....V...EV.KGL...LDA..IR.I.I.AK.GGK...  550
GR  KNPCIRDVAVVGIPDLEAGELPSAFVVIQPGKEITAKEVYDYLAERVSHTKYLRGGVRFVDSIPRNVTGKI-TRK--E-LLKQLLEKSSKL  543
YG  ..........................K..... ... . ..... ... . .....  543
YE  ..... ... . ..... ... . .....  543
OR  ..... ... . ..... ... . .....  543
```

Fig. 3. Alignment of the amino acid sequences of the click beetle and the firefly luciferases is shown to emphasize sequence differences. The sequence information is derived from the open reading frames of the corresponding cDNA clones. The identity of each luciferase sequence is indicated at the right of each line by a two letter code: FF, firefly; GR, green-emitting click beetle; YG, yellow-green–emitting click beetle; YE, yellow-emitting click beetle; and OR, orange-emitting click beetle. Only the sequence for the green-emitting click beetle luciferase is shown in entirety. Gaps in the alignment of this sequence are indicated by hyphens. Other luciferase sequences have letter designations only at sites where they differ from the green-emitting luciferase; where the sequences are the same there is a period. Numbers on the right indicate the position of the amino acid at the end of each line. Abbreviations for the amino acid residues are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

cells to be observed simultaneously. The ability to distinguish each of the luciferases in a mixture, however, is limited by the width of their emissions spectra. From calculations based solely on the overlap of the spectra of the green- and orange-emitting luciferases, one luciferase in a mixture should be detectable in the presence of a 25-fold excess of the other.

REFERENCES AND NOTES

1. J. W. Hastings, *J. Mol. Evol.* **19**, 309 (1983).
2. H. H. Seliger, J. B. Buck, W. G. Fastie, W. D. McElroy, *J. Gen Physiol.* **48**, 95 (1964).
3. H. H. Seliger and W. D. McElroy, *Proc. Natl. Acad. Sci. U.S.A.* **52**, 75 (1964).
4. R. A. Crowson, *The Biology of the Coleoptera* (Academic Press, London, 1981).
5. W. H. Biggley, J. E. Lloyd, H. H. Seliger, *J. Gen. Physiol.* **50**, 1681 (1967).
6. *Pyrophorus plagiophthalamus* were collected in Jamaica and frozen in liquid N_2. Polyadenylated RNA was isolated from the abdominal organ; 1 μg was converted to cDNA and inserted into Lambda ZAP to yield 700,000 recombinants, 5.5% of which could express luciferase antigens in *E. coli*. Of 18 phages chosen from the unamplified library by their reactivity with antibody to firefly luciferase, 4 could express bioluminescence in *E. coli*. The library was screened directly for bioluminescent activity from five portions that were amplified and converted into the Bluescript plasmid. Several bioluminescent colonies from each portion were identified by their ability to darken x-ray film; seven were determined as arising from independent cDNA clones. From two of the portions, two sets of colonies were judged as arising from independent clones based on widely different luminescent intensities, which were subsequently confirmed by restriction mapping.
7. J. M. Short, J. M. Fernandez, J. A. Sorge, W. D. Huse, *Nucleic Acids Res.* **16**, 7583 (1988).
8. G. Wienhausen and M. DeLuca, *Photochem. Photobiol.* **42**, 609 (1985).
9. K. V. Wood and M. DeLuca, *Anal. Biochem.* **161**, 51 (1987).
10. The plasmid pKW9 contains the *trp-lac* hybrid promoter (*tac*) and ribosome binding site (RBS) derived from pDR540 (Pharmacia LBK Biotechnology), and a transcriptional terminator derived from pAD9. The replicon portion of the vector is derived from pUC19. The cDNA clones were excised from Bluescript with the Bam HI and Xho I sites of the vector, and inserted into these sites in pKW9 between the *tac* promoter/RBS and the transcriptional terminator. Although the Bam HI site in pKW9 is immediately adjacent to the RBS, the 5′ untranslated regions of the cDNAs and a portion of the polylinker region of Bluescript removed the translational initiation codon from the RBS by at least 50 bp. The first ATG encountered downstream of the *tac* promoter is the proper initiation codon for each of the cDNAs, thus reducing spurious initiations and the production of fusion proteins.
11. W. D. McElroy and H. H. Seliger, *Arch. Biochem. Biophys.* **88**, 136 (1960).
12. S. J. Gould, G. A. Keller, S. Subramani, *J. Cell Biol.* **105**, 2923 (1987); *ibid.* **107**, 897 (1988); S. J. Gould *et al.*, *J. Cell Biol.*, in press.
13. M. DeLuca and W. D. McElroy, Eds., *Methods in Enzymology* (Academic Press, Orlando, FL, 1978), vol. 57.
14. J. R. de Wet, K. V. Wood, D. R. Helinski, M. DeLuca, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7870 (1985); D. W. Ow *et al.*, *Science* **234**, 856 (1986); H. T. Tatsumi, T. Masuda, E. Nakano, *Agric. Biol. Chem.* **52**, 1123 (1988); P. K. Howard, K. G. Ahern, R. A. Firtel, *Nucleic Acids Res.* **16**, 2613 (1988); A. G. DiLella *et al.*, *ibid.*, p. 4159.
15. J. R. de Wet, K. V. Wood, M. DeLuca, D. R. Helinski, S. Subramani, *Mol. Cell. Biol.* **7**, 725 (1987); S. J. Gould and S. Subramani, *Anal. Biochem.* **175**, 5 (1988).

16. We thank S. T. Swanson for assistance with the library screening and D. R. Helinski for supplying pAD9 and his comments on the manuscript. Supported by NSF grant DMB 8603776 and Office of Naval Research grant N00014-86-K-0267. This work is dedicated to the memory of Marlene DeLuca, a leader in the field of bioluminescence.

6 December 1988; accepted 7 March 1989

Reexamination of the Three-Dimensional Structure of the Small Subunit of RuBisCo from Higher Plants

STEFAN KNIGHT, INGER ANDERSSON, CARL-IVAR BRÄNDÉN

The structure of L_8S_8 RuBisCo (where L is the large subunit and S is the small subunit) from spinach has been determined to a resolution of 2.8 ångstrom by using fourfold averaging of an isomorphous electron density map based on three heavy-atom derivatives. The structure of the S subunit is different from that previously reported for the tobacco S subunit in spite of 75 percent sequence identity. The elements of secondary structure, four antiparallel β strands and two α helices, are the same, but the topology and direction of the polypeptide chain through these elements differ completely. One of these models is clearly wrong. The spinach model has hydrophobic residues in the core between the α helices and β sheet as well as conserved residues in the subunit interactions. The deletion of residues 49 to 62 that is present in the *Anabaena* sequence removes a loop region in the spinach model. The positions of three mercury atoms in the heavy-atom derivatives agree with the assignment of side chains in the spinach structure.

CHAPMAN *et al.* (*1*) HAVE RECENTLY described the tertiary structure of plant RuBisCo, the key enzyme (*2*) in the Calvin cycle of carbon dioxide fixation in photosynthesis. Their model is based on an electron density map to 2.6 Å of the L_8S_8 molecule from tobacco. We have determined the structure of L_8S_8 RuBisCo from spinach to 2.8 Å resolution and find very significant differences in the structure of the S subunit compared with the reported tobacco structure. Since there is 75% identity between the amino acid sequences of these two polypeptide chains, they are expected to have similar tertiary structures.

Crystals of spinach RuBisCo that diffract to 1.7 Å resolution were grown from solutions of the activated form of the enzyme with a bound transition-state analogue (*3*). These crystals contain one-half the L_8S_8 molecule in the asymmetric unit. There is a local noncrystallographic fourfold axis through the molecule, which has approximate 422 symmetry. X-ray data were collected on the synchrotron radiation source in Daresbury, United Kingdom, for the native enzyme and three heavy-atom derivatives. An initial electron density map was calculated with the use of isomorphous phase angles. These were refined by real-space averaging (*4*) around the local fourfold axis. Data collection procedures and phasing statistics have been briefly described (*5*).

The final electron density map was of very good quality, as would be expected by fourfold averaging of an electron density map based on three heavy-atom derivatives. Almost all of the side chains could easily be identified from the known sequences of the spinach S and L chains (*6, 7*), which comprise 123 and 475 residues, respectively. The sequence of the S subunit, which was determined by amino acid analysis (*6*), contains only one Cys residue, Cys 112. However, two independent determinations of the amino acid content of the spinach small subunit (*8*) made in different laboratories have shown that there are three Cys residues per subunit. Furthermore, almost all of the small subunits from higher plant RuBisCo for which the sequences are known contain three Cys residues at positions 41, 77, and 112. We therefore conclude that in all probability the spinach small subunit also contains Cys residues at these three positions. Our electron density map also strongly supports Cys side chains at these positions; the side-chain electron densities are appropriate for Cys (Fig. 1b).

We first built the L chain (*5*) using the known structure of L_2 RuBisCo from *Rhodospirillum rubrum* (*9*). We found, in agreement with the work on the tobacco enzyme (*1, 10*), that higher plant L chains have a structure that is quite similar to that of the bacterial enzyme (*9*) except at the carboxyl terminal. The arrangement of the L subunits in the spinach enzyme into four L_2 dimers

Swedish University of Agricultural Sciences, Uppsala Biomedical Center, Department of Molecular Biology, S-751 24 Uppsala, Sweden.