

เราสามารถเอาหมายเลข M15077.1 หรือ copy ลำดับนิวคลีโอไทด์ (DNA) ในรูป FASTA (เริ่มจากเครื่องหมาย >) ก็ได้ มา paste ลงในช่องของ ORF Finder

FASTA ▾

## P.pyralis (firefly) luciferase gene, complete cds

GenBank: M15077.1

[GenBank](#) [Graphics](#)

```
>M15077.1 P.pyralis (firefly) luciferase gene, complete cds
CTGCAGAAATAACTAGGTACTAAGCCCGTTTGTGAAAAGTGGCCAAACCCATAAATTTGGCAATTACAAT
AAAGAAGCTAAAATTGTGGTCAAACCTACAAACATTTTTATTATATACATTTTAGTAGCTGATGCTTATA
AAAGCAATATTTAAATCGTAAACAACAAATAAAATAAAATTTAAACGATGTGATTAAGAGCCAAAGGTCC
TCTAGAAAAAGGTATTTAAGCAACGGAATTCCTTTGTGTTACATTCTTGAATGTCGCTCGCAGTGACATT
AGCATTCCGGTACTGTTGGTAAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCCATTCTATCC
TCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTTCCTGGAACA
ATTGCTTTTGTGAGTATTTCTGTCTGATTTCTTTTCGAGTTAACGAAATGTTCTTATGTTTCTTTAGACAG
ATGCACATATCGAGGTGAACATCACGTACGCGGAATACTTCGAAATGTCCGTTTCGGTTGGCAGAAGCTAT
GAAACGATATGGGCTGAATACAAATCACAGAATCGTCGTATGCAGTGAAAACCTCTCTTCAATTCTTTATG
CCGGTGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTAAGC
ACCCTCGCCATCAGACCAAAGGGAATGACGTATTTAATTTTAAAGGTGAATTGCTCAACAGTATGAACAT
TTCGCAGCCTACCGTAGTGTTTGTTCCTCAAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAATTA
CCAATAATCCAGAAAATTATTATCATCGATTCTAAAACGGATTACGAGGGATTTCAGTCCGATGTACAGCT
```

ลองใช้วิธีใส่ accession number ดูครับ ก็เอา M15077.1 ใส่ลงไป  
สำหรับ Search Parameters ต่าง ๆ ปล่อยไว้ตามค่า default ที่โปรแกรมกำหนดมาให้  
ไม่ต้องเปลี่ยนอะไรครับ

The screenshot shows a web interface for sequence analysis. It is divided into three main sections: 'Enter Query Sequence', 'Choose Search Parameters', and 'Start Search / Clear'. In the 'Enter Query Sequence' section, the text 'M15077.1' is entered into a text box, with a pink arrow pointing to it. Below this are 'From:' and 'To:' input fields. The 'Choose Search Parameters' section contains several settings: 'Minimal ORF length (nt):' is set to '75' with a pink arrow; 'Genetic code:' is set to '1. Standard' with a pink arrow; 'ORF start codon to use:' has three radio button options, with 'ATG' only selected and a pink arrow pointing to it; and 'Ignore nested ORFs:' is an unchecked checkbox. The 'Start Search / Clear' section at the bottom has a 'Submit' button highlighted with a pink rectangle and a 'Clear' button next to it.

**Enter Query Sequence**

Enter accession number, gi, or nucleotide sequence in FASTA format:

M15077.1

From: To:

**Choose Search Parameters**

Minimal ORF length (nt): 75

Genetic code: 1. Standard

ORF start codon to use:

- ☒ "ATG" only
- ☐ "ATG" and alternative initiation codons
- ☐ Any sense codon

Ignore nested ORFs: ☐

**Start Search / Clear**

Submit Clear

From To ใช้กรณีเราต้องการให้ค้นหา ORF เฉพาะช่วงนิวคลีโอไทด์ที่ต้องการเพียงบางส่วน เช่นจากนิวคลีโอไทด์ที่ 150 ถึง 1120 เป็นต้น

Minimal ORF length ตั้งไว้ 75 หมายความว่า ถ้าเจอ ORF แต่สั้นนิดเดียวไม่ถึง 75 นิวคลีโอไทด์ ก็ไม่น่าใช้โปรตีน เพราะสั้นเกินไป

Genetic code เป็นรหัสพันธุกรรมจาก DNA RNA เป็นโปรตีน ในสิ่งมีชีวิตบางชนิดอาจใช้รหัสที่แตกต่างจากมาตรฐาน ก็เลือกได้ แต่ปกติจะใช้ Standard

ORF start codon ใช้ ATG only  
ซึ่งเป็นแบบที่พบในสิ่งมีชีวิตทั่ว ๆ ไป

กด Submit ได้เลยครับ

The screenshot shows the 'Enter Query Sequence' and 'Choose Search Parameters' sections of a web application. A pink arrow points to the text input field containing 'M15077.1'. Another pink arrow points to the 'Minimal ORF length (nt):' dropdown menu, which is set to '75'. A third pink arrow points to the 'ORF start codon to use:' section, where the 'ATG only' radio button is selected. A fourth pink arrow points to the 'Genetic code:' dropdown menu, which is set to '1. Standard'. At the bottom, the 'Submit' button is highlighted with a pink rectangle. The 'Start Search / Clear' section contains 'Submit' and 'Clear' buttons.

พบ ORF ทั้งหมด 19 ORF โดย ORF ที่ยาวที่สุดคือ **ORF 3** เริ่มต้นจากนิวคลีโอไทด์ที่ 763 จบที่ 1092 รวม 330 นิวคลีโอไทด์ คิดเป็นกรดอะมิโน (amino acid = aa) 109 อะมิโน ( $=330/3 = 110$  แต่ codon สุดท้ายเป็น stop codon ไม่นับเป็นกรดอะมิโน)

**Open Reading Frame Viewer** Help

**P.pyralis (firefly) luciferase gene, complete cds**

ORFs found: 19 Genetic code: 1 Start codon: 'ATG' only

M15077 Find:  Tools Tracks ?

(U) ORFfinder\_2.10.75743218

M15077.1: 1..2.4K (2,387 nt) Tracks shown: 2/5

**ORF3 (109 aa)** [Display ORF as...](#) [Mark](#)

```
>orf3
MNISQPTVVFVSKKGLQKILNVQKKLPPIQKIIMDSKTDYQGFSMYTF
VTSHLPFGFNEYDFVPSFDRDKTIALIMNSSGSTGLPKGVLPHTACV
RFSHARYVY
```

**ORF3**   Marked set (0)

BLAST Database:

[Go back to the submitting page...](#)

Mark subset... Marked: 0 Download marked set as

| Label       | Strand | Frame | Start      | Stop        | Length (nt   aa) |
|-------------|--------|-------|------------|-------------|------------------|
| <b>ORF3</b> | +      | 1     | <b>763</b> | <b>1092</b> | <b>330   109</b> |
| ORF4        | +      | 1     | 1204       | 1452        | 249   82         |
| ORF8        | +      | 2     | 1694       | 1909        | 216   71         |
| ORF2        | +      | 1     | 535        | 744         | 210   69         |
| ORF9        | +      | 2     | 2066       | 2251        | 186   61         |
| ORF5        | +      | 1     | 1870       | 2031        | 162   53         |
| ORF11       | -      | 1     | 1463       | 1305        | 159   52         |
| ORF19       | -      | 3     | 498        | 352         | 147   48         |
| ORF1        | +      | 1     | 304        | 447         | 144   47         |
| ORF15       | -      | 2     | 343        | 212         | 132   43         |

โปรแกรม ORF Finder นอกจากจะค้นหา ORF ให้เราแล้ว ยังแปลรหัสจาก DNA เป็น โปรตีนให้เราด้วย ตามที่เห็นในกรอบ (MNISQPTVV...RYVV) จะเห็นกรดอะมิโนตัวแรก คือ M หรือ methionine ซึ่งเป็น start codon ในตัวด้วย (ATG หรือ AUG ใน RNA)

**ORF3 (109 aa)**    [Display ORF as...](#)    [Mark](#)

```
>|c||ORF3
MNISQPTVVVFSKKGLQKILNVQKKLPPIQKIIIMDSKTDYQGFQSMYTF
VTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACV
RFSHARYVV
```

| Label       | Strand |
|-------------|--------|
| <b>ORF3</b> | +      |
| ORF4        | +      |
| ORF8        | +      |
| ORF2        | +      |
| ORF9        | +      |
| ORF5        | +      |
| ORF11       | -      |
| ORF19       | -      |
| ORF1        | +      |
| ORF15       | -      |

**ORF3**    Marked set ( 0 )

[SmartBLAST](#)    [SmartBLAST best hit titles...](#)

**BLAST**    [BLAST](#)

BLAST Database:  
UniProtKB/Swiss-Prot (swissprot) ▼

เพื่อพิสูจน์ว่า ลำดับกรดอะมิโน ที่ได้จาก ORF3 นี้ เป็นโปรตีน Luciferase ของหิ่งห้อยจริง ๆ หรือไม่ ก็ทำได้โดยการนำลำดับกรดอะมิโนนี้ ไปเปรียบเทียบกับฐานข้อมูลโปรตีน โดยใช้โปรแกรม BLAST ที่เราเรียนมาแล้ว โดยกดปุ่ม BLAST ซ้ายล่างได้เลย

ORF3 (109 aa)

Display ORF as...

Mark

>|c|ORF3  
MNISQPTVVVFSKKGLQKILNVQKKLPPIQKIIIMDSKTDYQGFQSMYTF  
VTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACV  
RFSHARYVV

Mark subset...

M

| Label | Strand |
|-------|--------|
| ORF3  | +      |
| ORF4  | +      |
| ORF8  | +      |
| ORF2  | +      |
| ORF9  | +      |
| ORF5  | +      |
| ORF11 | -      |
| ORF19 | -      |
| ORF1  | +      |
| ORF15 |        |

ORF3

Marked set ( 0 )

SmartBLAST

BLAST

SmartBLAST best hit titles...

BLAST

BLAST Database:

UniProtKB/Swiss-Prot (swissprot) ▼



กดปุ่ม BLAST แล้ว โปรแกรมก็จะดึงลำดับกรดอะมิโนมาลงในโปรแกรม BLAST  
ในรูป FASTA ให้เลย (สังเกตจากมีเครื่องหมาย > )  
เราก็กดปุ่ม BLAST สีนํ้าเงินล่างซ้ายอีกที

BLAST® » blastp suite

Standard Protein BLAST

blastn **blastp** blastx tblastn tblastx

BLASTP programs search protein databases using a protein query. [more...](#)

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#)

>lc|ORF3\_M15077.1:763:1092 unnamed protein product  
MNISQPTVVFYSKKGLQKILNVQKLPPIQKIIMDSKTDYQGFQSMYFVTSH  
LPPGFNEYDFVPESFD  
RDKTIALIMNSSGSTGLPKGVALPHRTACVRFSHARYVV

Query subrange [?](#)

From

To

Or, upload file   [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

**Choose Search Set**

Database  [?](#)

Organism [Optional](#)  ☐ exclude [Add organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown. [?](#)

Exclude [Optional](#) ☐ Models (XM/XP) ☐ Non-redundant RefSeq proteins (WP) ☐ Uncultured/environmental sample sequences

**Program Selection**

Algorithm

☒ blastp (protein-protein BLAST)

☐ PSI-BLAST (Position-Specific Iterated BLAST)

☐ PHI-BLAST (Pattern Hit Initiated BLAST)

☐ DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)

Choose a BLAST algorithm [?](#)

**BLAST** Search database swissprot using Blastp (protein-protein BLAST)

☐ Show results in a new window

ผลการใช้ BLAST เปรียบกรดอะมิโนที่ได้จากการแปลรหัส DNA จากยีน กับกรดอะมิโน  
 ในฐานข้อมูลโปรตีน ได้เป็นแบบนี้ เรียงลำดับตามความเหมือนกันจากมากไปหาน้อย  
 ลองคลิกดูข้อมูลบนสุดครับ

BLAST® » blastp suite » results for RID-08U9YHKT01R Home Recent Results Saved Strategies Help

[Edit Search](#) [Save Search](#) [Search Summary](#) [How to read this report?](#) [BLAST Help Videos](#) [Back to Traditional Results Page](#)

**Job Title** Protein Sequence

**RID** [08U9YHKT01R](#) Search expires on 02-11 08:24 am [Download All](#)

**Program** BLASTP [Citation](#)

**Database** swissprot [See details](#)

**Query ID** lc|Query\_53159

**Description** lc|ORF3\_M15077.1:763:1092 unnamed protein product

**Molecule type** amino acid

**Query Length** 109

**Other reports** [Distance tree of results](#) [Multiple alignment](#) [MSA viewer](#)

**Filter Results**

**Organism** only top 20 will appear ☐ exclude

Type common name, binomial, taxid or group name

[Add organism](#)

**Percent Identity**  to  **E value**  to  **Query Coverage**  to

[Filter](#) [Reset](#)

**Descriptions** [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

**Sequences producing significant alignments** [Download](#) [New](#) [Select columns](#) [Show](#) 100 [?](#)

☒ select all 39 sequences selected [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [New](#) [MSA Viewer](#)

|                                     | Description                                                                                                   | Scientific Name                      | Max Score | Total Score | Query Cover | E value | Per. Ident | Acc. Len | Accession                |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------|--------------------------------------|-----------|-------------|-------------|---------|------------|----------|--------------------------|
| <input checked="" type="checkbox"/> | RecName: Full=Luciferin 4-monooxygenase; Short=Luciferase [Photinus pyralis]                                  | <a href="#">Photinus pyralis</a>     | 222       | 222         | 97%         | 2e-70   | 100.00%    | 550      | <a href="#">P08659.1</a> |
| <input checked="" type="checkbox"/> | RecName: Full=Luciferin 4-monooxygenase; Short=Luciferase [Luciola cruciata]                                  | <a href="#">Luciola cruciata</a>     | 124       | 124         | 97%         | 1e-33   | 54.72%     | 548      | <a href="#">P13129.1</a> |
| <input checked="" type="checkbox"/> | RecName: Full=Luciferin 4-monooxygenase; Short=Luciferase [Aequatic lateralis]                                | <a href="#">Aequatic lateralis</a>   | 122       | 122         | 97%         | 1e-32   | 54.72%     | 548      | <a href="#">Q01158.1</a> |
| <input checked="" type="checkbox"/> | RecName: Full=Luciferin 4-monooxygenase; Short=Luciferase [Luciola mingrelica]                                | <a href="#">Luciola mingrelica</a>   | 113       | 113         | 97%         | 2e-29   | 51.40%     | 548      | <a href="#">Q26304.1</a> |
| <input checked="" type="checkbox"/> | RecName: Full=Luciferin 4-monooxygenase; Short=Luciferase [Photuris pensylvanica]                             | <a href="#">Photuris pensylv...</a>  | 93.6      | 93.6        | 97%         | 2e-22   | 40.57%     | 545      | <a href="#">Q27757.2</a> |
| <input checked="" type="checkbox"/> | RecName: Full=Long chain acyl-CoA synthetase 9, chloroplastic [Arabidopsis thaliana]                          | <a href="#">Arabidopsis thali...</a> | 48.0      | 48.0        | 87%         | 8e-07   | 29.20%     | 601      | <a href="#">G0CAP9.1</a> |
| <input checked="" type="checkbox"/> | RecName: Full=Nonribosomal peptide synthetase 2; Short=NPRS 2; AltName: Full=Ferrirociclin synthetase; Alt... | <a href="#">Fusarium gramin...</a>   | 42.0      | 42.0        | 93%         | 3e-04   | 34.62%     | 4841     | <a href="#">I1RN14.1</a> |
| <input checked="" type="checkbox"/> | RecName: Full=Fatty acyl-CoA synthetase B; AltName: Full=Long chain fatty acyl coenzyme A-synthetase 2; S...  | <a href="#">Dictyostelium dis...</a> | 41.2      | 41.2        | 81%         | 4e-04   | 29.90%     | 673      | <a href="#">Q1ZXQ4.1</a> |



ข้อมูลบนสุด บอกว่า ตรงกันกับโปรตีนที่ชื่อเต็ม Full ว่า

**Luciferin 4-monooxygenase**

ชื่อสั้น ๆ Short ว่า **Luciferase**

โดยมาจากสิ่งมีชีวิต ชื่อวิทยาศาสตร์ว่า *Photinus pyralis* (หิ่งห้อยชนิดหนึ่ง)

Descriptions Graphic Summary **Alignments** Taxonomy

Alignment view Pairwise [Restore defaults](#) Download

39 sequences selected

[Download](#) [GenPept](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

**RecName:** Full=Luciferin 4-monooxygenase; Short=Luciferase [Photinus pyralis]

Sequence ID: [P08659.1](#) Length: 550 Number of Matches: 1

Range 1: 118 to 223 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

| Score         | Expect | Method                       | Identities    | Positives     | Gaps      |
|---------------|--------|------------------------------|---------------|---------------|-----------|
| 222 bits(566) | 2e-70  | Compositional matrix adjust. | 106/106(100%) | 106/106(100%) | 0/106(0%) |

Query 1 MNISQPTVVFVSKKGLQKILNVQKKLPPIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFN 60  
MNISQPTVVFVSKKGLQKILNVQKKLPPIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFN  
Sbjct 118 MNISQPTVVFVSKKGLQKILNVQKKLPPIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFN 177

Query 61 EYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVRFSHAR 106  
EYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVRFSHAR  
Sbjct 178 EYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVRFSHAR 223

[Download](#) [GenPept](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

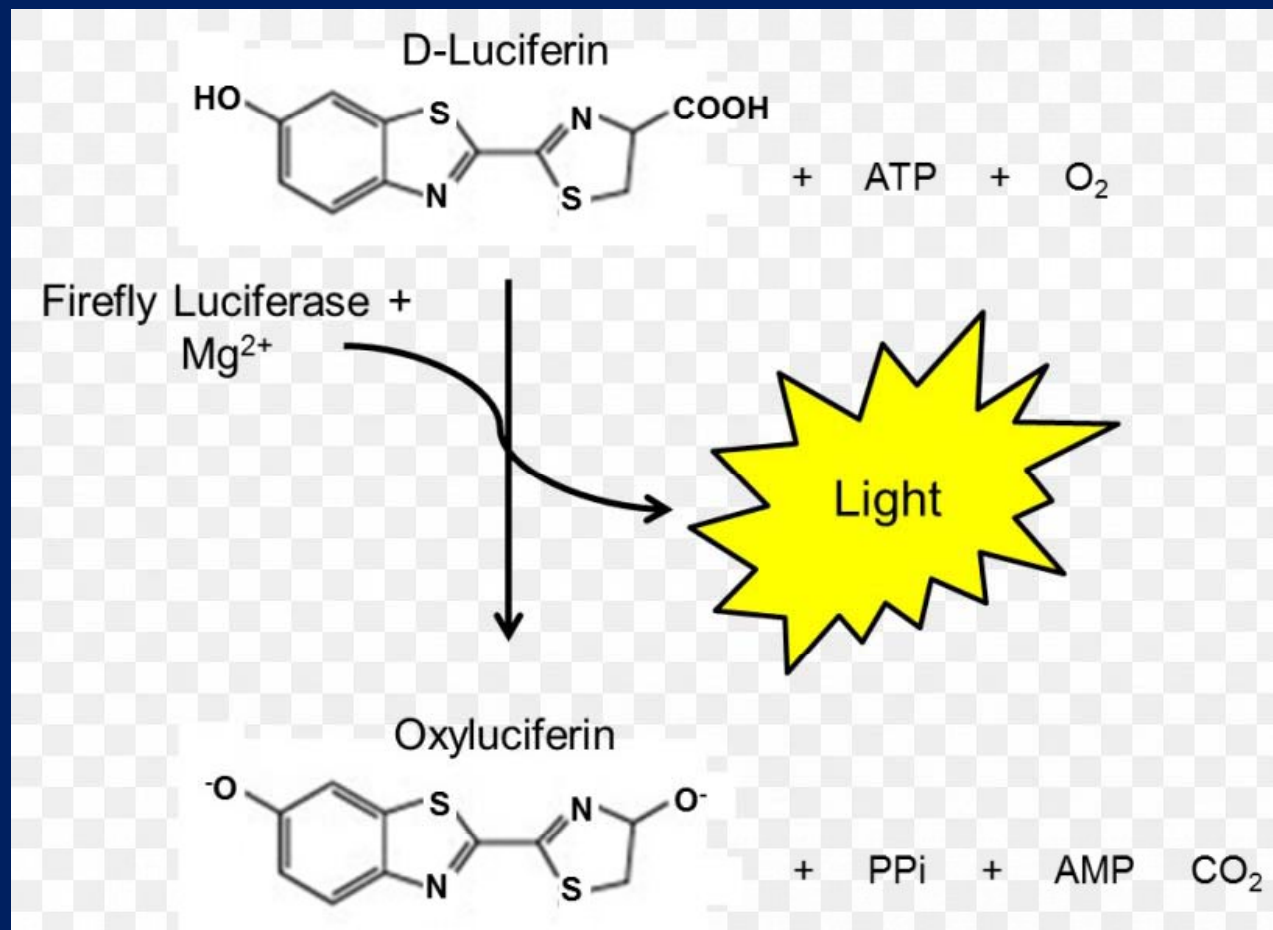
**RecName:** Full=Luciferin 4-monooxygenase; Short=Luciferase [Luciola cruciata]

Sequence ID: [P13129.1](#) Length: 548 Number of Matches: 1

Range 1: 120 to 225 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

ทบทวนความรู้กันอีกที หิ่งห้อยเปล่งแสงได้ เพราะผลิตสาร Luciferin แล้วเติมออกซิเจนเข้าไป โดยใช้พลังงานจาก ATP และเร่งปฏิกิริยาโดยเอนไซม์ Luciferase และปล่อยพลังงานออกมาในรูปแสงสว่าง

Luciferin เป็นสารอินทรีย์มีโครงสร้างตามในภาพ ส่วน Luciferase เป็นเอนไซม์ซึ่งเป็นโปรตีนนั่นเอง



หากนิสิตสนใจว่านักวิทยาศาสตร์ที่ค้นพบยีน Luciferase ครั้งแรกในปี 1987 นั้นเขาทำ  
อย่างไร ก็อ่านได้ในวารสาร Molecular and Cellular Biology ครับ  
จะเอาลิงก์ PDF ใส่ไว้ให้ในหน้าเว็บรายวิชาครับ

MOLECULAR AND CELLULAR BIOLOGY, Feb. 1987, p. 725-737  
0270-7306/87/020725-13\$02.00/0  
Copyright © 1987, American Society for Microbiology

Vol. 7, No. 2

## Firefly Luciferase Gene: Structure and Expression in Mammalian Cells

JEFFREY R. DE WET,<sup>1+</sup> KEITH V. WOOD,<sup>2</sup> MARLENE DeLUCA,<sup>2</sup> DONALD R. HELINSKI,<sup>1</sup>  
AND SURESH SUBRAMANI<sup>1\*</sup>

*Departments of Biology<sup>1</sup> and Chemistry,<sup>2</sup> University of California, San Diego, La Jolla, California 92093*

Received 25 August 1986/Accepted 7 November 1986

The nucleotide sequence of the luciferase gene from the firefly *Photinus pyralis* was determined from the analysis of cDNA and genomic clones. The gene contains six introns, all less than 60 bases in length. The 5' end of the luciferase mRNA was determined by both S1 nuclease analysis and primer extension. Although the luciferase cDNA clone lacked the six N-terminal codons of the open reading frame, we were able to reconstruct the equivalent of a full-length cDNA using the genomic clone as a source of the missing 5' sequence. The full-length, intronless luciferase gene was inserted into mammalian expression vectors and introduced into monkey (CV-1) cells in which enzymatically active firefly luciferase was transiently expressed. In addition, cell lines stably expressing firefly luciferase were isolated. Deleting a portion of the 5'-untranslated region of the luciferase gene removed an upstream initiation (AUG) codon and resulted in a twofold increase in the level of luciferase expression. The ability of the full-length luciferase gene to activate cryptic or enhancerless promoters was also greatly reduced or eliminated by this 5' deletion. Assaying the expression of luciferase provides a rapid and inexpensive method for monitoring promoter activity. Depending on the instrumentation employed to detect luciferase activity, we estimate this assay to be from 30- to 1,000-fold more sensitive than assaying chloramphenicol acetyltransferase expression.