**[SmGPCR (AF031196)] complete cDNA sequence (1-2002 nt).**

```
5'-CGGTTTTACT CTTGTGATTT GTTGCATGAT TCTCATATCC ATTTCAACTG GTGTTAACAA
   ACAGTATCAT TATTTGTTAT GCAGTCACCG AATTCTTAGC ACTCGGTTAC ATGTTTGTAT
   TCATTTGTTA ACCAAGTTAC CTTTTAATTC TGAAGTTACA ATGAAACTGA AATATCCATT
   ACAGAATAGC AGAGGAGATT TAACATTCGA AGAAATACAA ACAATTTGAA AGCACATTGT
   CAAAAAACTA ACAATACATT TGACCTAATG CAATATATAA ACAAAACGTC GCTAAATTCC
   AGTGTTATTC CAGATTCCCT AATCAAATCA TGGATCCTCT CGAACCCTAT AATAAAATGG
   ACCATATCAT TATTTTTAAT TATAGCCACT GGGACTACAT TCTTCGGAAA TCTGTTAATA
   ATTCTTGCTT TTATAACAAA TAGTCGCTTA CGGCGTATAA CGGATCAATA CATTGTTTCG
   CTTGCTGTGG CTGATCTACT TGTATCGGTC TTAGTGCTGC CTCTGGCTAT CGTTCGTCAG
   AATTTAGGTT ATTGGCCATT CGAATCTGAT CGATTGTGTC AGTTTTGGTT ATCCGCTAAT
   ATTGTATTAT GTATGGCATC TATACTAAAT TTATGTTGTA TATCTCTGGA TAGATATATT
   GCAATCAGCC GTCCAATGAA ATATTTCACA AAGCGTACTC GTTTTACAGC CTCTACAATG
   ATAGCTGTTG CTTGGATATT ACCACTTATA ACTATGCTTC TCCCCTTTGT TGGTGGGAAT
   CAGCACACTT TGGGACTTGG ATCATGTCAT ATCACATATA ATAAAGCATA TAGGATTTAT
   TCTTCAATTG TTGGATTTTT CGGACCATTT CTGCTAATTG CTTATATATA TTTACGAGTG
   TTTTGGATAA TAAAGCATCG TTTAAAGGTC TTACAAATTA CAAACATTAA ACTTTCCTCA
   CTCAAAAAAC CTAAATCACA TATTAAAGCG ACACGTAAGC CAGCTCCGAT AATAATAAAT
   TTGCAACAAG TTTGGGAAAA TATTAAAGGA AAGATTGGCA AAGTAAATAT TTTACGAAAT
   CAGTCATCTA AATCAAAGAA TACTTGCCCA TATAGTGGCC ATTTTTTCCA TTCTGATGAA
   AATGGATGTA ATCAAATTTA TGCGAGTTGT TTACTCAACC ACAAACATCC CTTTGATTCT
   TTTAGAAGTG ATGATATTCT TAATGAAGAA ATCGATAGGA ATTATAGACA AAAAATATTA
   GATATTAAGA TGAACAAAAC AGATAGAGAA TTTTCTTTGA CATCTCAAGA TCACGTTGAA
   TTAGATTTCC CCGAACCAAC AGATCAAAAA CTAAAAGTAA TACAAATATT TAATATTTCA
   ACGTCAGAGA AAGAATTACA AAAGGCGAAT TGGGAACCAA CACCAAATGT AACATTGAAA
   AGTGAAACAA AACAACATCA ACACGATTAC TTACATAGAA GTAATGAAGA ACAAAAGAAT
   AAACTAGGAG ACAAACAATT AATTAATTCT TCGAAATTAC CATCATTAAC ATCATTTAGT
   ATAGATCATA TAAAGAAGCT TCGACGATTG AAATCTTTAC CAAATTTTAA AACAGAATAC
   AATAAAACAG AACAATCACA AAAGAATCTA ATTTATTATC GACATGATCA TTTCATATTT
   CATAGGGAAC AAAAAACTGC ATTAATTTTA GCTGCAGTTG TTGGATGTTT TACAATATGT
   TGGTTTCCAT TTTTTATATG TATTATTGGT GAAGCTATAT GTGATTGTCA ATATTCAAAT
   ACTATAATTA CATTTGTTTC ATGGTTAGGT TATTTTAATA GTATTTGTAA TCCATTTATA
   TATGCATTTT TTAAAAAAGA ATATGCCAAA TCATTTAAAT ATATTATTCA AGTTAACAAA
   TGGAATATTA AAAGCTATAA CCATAAATAA CTGTTATTTT TAAATAAGAA ATAAATCTTT
   TATTACGAAA AAGAAAAAAA AA-3'
```

GTGTTTTG-------AACAAAAAACTGCA: **this is the <u>full</u> third intra-loop (il3) <u>sequence</u> within SmGPCR. It is located between 898-1701 (i.e. 804 bp). It's sequence as following:**

```
5'-GTGTTTTGGA TAATAAAGCA TCGTTTAAAG GTCTTACAAA TTACAAACAT TAAACTTTCC
   TCACTCAAAA AACCTAAATC ACATATTAAA GCGACACGTA AGCCAGCTCC GATAATAATA
   AATTTGCAAC AAGTTTGGGA AAATATTAAA GGAAAGATTG GCAAAGTAAA TATTTTACGA
   AATCAGTCAT CTAAATCAAA GAATACTTGC CCATATAGTG GCCATTTTTT CCATTCTGAT
   GAAAATGGAT GTAATCAAAT TTATGCGAGT TGTTTACTCA ACCACAAACA TCCCTTTGAT
   TCTTTTAGAA GTGATGATAT TCTTAATGAA GAAATCGATA GGAATTATAG ACAAAAAATA
   TTAGATATTA AGATGAACAA AACAGATAGA GAATTTTCTT TGACATCTCA AGATCACGTT
   GAATTAGATT TCCCCGAACC AACAGATCAA AAACTAAAAG TAATACAAAT ATTTAATATT
   TCAACGTCAG AGAAAGAATT ACAAAAGGCG AATTGGGAAC CAACACCAAA TGTAACATTG
   AAAAGTGAAA CAAAACAACA TCAACACGAT TACTTACATA GAAGTAATGA AGAACAAAAG
   AATAAACTAG GAGACAAACA ATTAATTAAT TCTTCGAAAT TACCATCATT AACATCATTT
   AGTATAGATC ATATAAAGAA GCTTCGACGA TTGAAATCTT TACCAAATTT TAAAACAGAA
   TACAATAAAA CAGAACAATC ACAAAAGAAT CTAATTTATT ATCGACATGA TCATTTCATA
   TTTCATAGGG AACAAAAAAC TGCA -3'
```

CCCGAACC AACAGAT---- CAAAAAAC TGCA this highlighted region (**372bp**) of the third intracellular loop (il3) is selected to be amplified by the **RPB1(NotI)** and **RPF2(EcoRI)** primers (see below their sequences).

The primers that have been designed to amplify a fragment of SmGPCR/il3 are:

**A sense RPF2(EcoRI) sequence: 5'-*CC* "GAATTC" ATG<u>CCCGAACCAACAGA</u>-3'**

**An antisense RPB1(NotI): 5'-*TA* "GCGGCCGC" <u>TGCAGTTTTTTGTTC</u>-3'**

where

*EcoR*I sequence is: **5'-G/AATTC-3'**

and

*Not*I sequence is: **5'-GC/GGCCGC-3'**

These primers will:

(a) Amplify a portion (372bp) of the entire il3 (804bp) of the SmGPCR (2002bp).This segment is 389 bp (after double digestion).

(b) Introduce *EcoR*I and *Not*I restriction sites (for integration with an expression vector pET30a).

A close examination of the primers will tell:
The total length of each primer is 25nucleotides.

**The sense (RPF2) primer consists of:**

- **(a)** 14 nucleotides (underlined) that <u>are **identical** to the 5' end of the 5'-3' (sense) template</u>, which will flank and anneal to the antisense template strand during PCR reaction.
- **(b)** A start codon ATG (highlighted).
- **(c)** The EcoRI sequence (toward the 5' end of this primer) and
- **(d)** CC nucleotides (as GC clamp).

**The antisense (RPB1) primer** is written as 5'-3' mode for convenience. However, if it is written as it should be (i.e. 3'-5'), it appears as:

**3'-<u>CTTGTTTTTTGACGT</u> "CGCCGGCG" AT-5'**

and it consists of:

- **(a)** 15 nucleotides (underlined) that are <u>**complement** to the 3' end of the 5'-3' (sense) template</u>, and will anneal to the sense template during PCR reaction.
- **(b)** The NotI sequence and
- **(c)** Two extra nucleotides.

## A procedure to write the sequence of the primers

(1) Get the 5'-3' sequence of the DNA template strand.

(2) **(A)** For the sense primer, it is so easy. You just determine the selected region of the template you wish to amplify and select 15-18 nucleotide sequence at the beginning of that region (i.e. same orientation 5'-3'), this is your sense primer. **(B)** If you wish to add a restriction sequence, T7 RNA polymerase promoter sequence, ..etc, add that sequence in 5'-3' mode before your primer sequence (i.e. at its 5' end). The start codon (i.e. ATG) should be located between the restriction and your target sequences. You may add C, G, CC, or GG (CG clamp) to stabilize your primer. Never add ≥ 3C or 3G as this will stabilize nonspecific annealing.

(3) For the antisense primer, you should remember that its should be written as 5'-3' (at the end). For now, don't bother yourself, we will modify it later.

Thus, select the last 14-18 nucleotide sequence of the3' end of you selected target (5'-3' sense template DNA strand), write the complement sequence of that sequence. This is 3'-5' orientation of the primer. Add any needed restriction sequence to its right end (i.e. toward 5' end) but in its 3'-5' orientation too. Do the same for the stopping codon (if needed).

Finally, re-orient the sequence of this primer to be written as 5'-3' style.

Look at the above example and try to apply this procedure. Hope you will find it easy.

********** *********** ********* **********