IJPSR (2013), Vol. 4, Issue 4 (Review Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 25 December, 2012; received in revised form, 21 January, 2013; accepted, 13 March, 2013

EXPRESSION OF FULLY FUNCTIONAL GPCR IN *E. COLI*

D. O. Acheampong*, J. Zhang and M. Wang

Department of Molecular Biology, State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009, China

Keywords:

G-protein, G-protein coupled receptors, Membrane proteins, Inclusion bodies, Autoinduction, Gateway vectors

Correspondence to Author:

Desmond Omane Acheampong
Department of Molecular Biology, State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009, China
Email : do.acheampong@yahoo.com

ABSTRACT: One-third of eukaryotic proteins are associated with membranes and these membrane proteins (MP) represent approximately 50% of pharmacological targets. Researchers are therefore nursing the hope of achieving a higher percentage of about 80% in the near future. The only bottleneck to achieving this target is the experimental challengers associated with the production, purification and crystallization at reasonable cost. Expression of protein in *E. coli* is now very popular, because it is easy to use and also comes with low operational cost. But membrane proteins have been difficult to produce in *E. coli*. The membrane proteins (MP), G-protein coupled receptors (GPCRs) represent the major class of potential drug targets but probably the most difficult class of MPs to bring to three-dimensional studies. This review paper therefore sought to bring to the fore some of the techniques employed by researchers to improve the quality and quantity of GPCR expressed in *E. coli*. A literature search on the effective expression of some important members of the GPCR family in *E. coli* was done. The Expression of neurotensin receptor type1 (NTIS1) in BL21 (DE3) and C41 (DE3) strains of *E. coli*; Autoinduction of NTS1 in BL21 (DE3) with the medium Magicmedia™; Expression in *E. coli* by targeting the inner membrane with fusion partners MBP and TRX (MBP-GPCR-TRX); Expression in *E. coli* by targeting the inclusion bodies with fusion partners GST and TRX; and the use of pDEST17 vectors and C43 strain at 37°C have proved experimentally efficient to express highly functional GPCR.

INTRODUCTION: It is estimated that one-third of eukaryotic protein are associated with membranes [1], and these proteins (membrane proteins) have been established to represent approximately 50% of pharmacological targets [2].

QUICK RESPONSE CODE

IJPSR:
ICV (2011)- 5.07

Article can be accessed online on:
www.ijpsr.com

It is the hope of all researchers in the field of drug discovery that, this number will increase up to 80% of all new drug targets for the next decade [2]. The only difficulty to achieving the above target is the fact that, there are experimental challenges associated with the production, purification and crystallization of membrane proteins (MP). Membrane proteins, especially those from eukaryotes, have been difficult to produce in *E. coli* and therefore have limited the ability to use this inexpensive and robust host to produce protein for biochemical and structural studies [3].

More eukaryotic membrane protein have been produced and crystallized by researchers in recent years, but the majority of these have been produced in more expensive hosts such as yeast or insect cells [4]. This adds to the cost of drug discovery. Membrane protein (MP) expression in bacteria has been experimented, but typically results in inclusion body (IB) formation, resulting in protein that is difficult to denature and refold, or in degradation, with low yields of correctly folded, membrane-inserted protein [5]. Among these membrane proteins (MP), G-protein coupled receptors (GPCRs) represent the major class of potential drug targets but probably the most difficult class of MPs to bring to three-dimensional studies.

GPCRs have been successfully produced by cell-free synthesis and heterologous expression in mammalian and insect cells, in photoreceptor cells of Drosophila, Xenopus and mouse [6]. This review paper seeks to bring to the fore some of the techniques employed by researchers to improve the quality and quantity of GPCR expressed in *E. coli*. Successful expression of high quality and appreciable quantity of fully functional GPCR in *E. coli* will go a long way to cut down cost involved in drug discovery.

MATERIALS AND METHODS: This review was done by compiling references from major databases like PubMed, Science Direct, Google scholar, Scopus, Online journals, Open J Gate, etc.

GPCR in perspective: The G-protein coupled receptors (GPCRs) present a general fold of seven transmembrane helices, with the N terminus pointing outside the cell and the C terminus located in the cytoplasm. In contrast to the helical transmembrane backbone, the sizes of the N and C termini and of the loop of the helices are highly variable [7]. It is a common knowledge that GPCRs are involved in mediating a multitude of physiological and pathophysiolgical processes [8] and therefore very important target when considering drug discovery.

GPCRs are the principal target for about 26.8% of prescription drugs and new drug candidates are continually being developed with GPCRs being the target[9]. GPCRs are associated with heterotrimeric G-protein which consists of α, β and γ subunits with the Gα subunit bound to GDP [7, 8]. As shown in (**Fig. 1**), upon occupancy by the agonist, the receptor conformation is shifted to the active state, causing the heterotrimeric G protein to dissociate. This brings about the displacement of the GDP by GTP on Gα subunit while the Gβγ subunits remain bound. Both Gβγ and GTP bound Gα are thus activated and can stimulate or inhibit effector proteins such as adenylyl cyclase, phospholipases and a variety of ion channels [7].

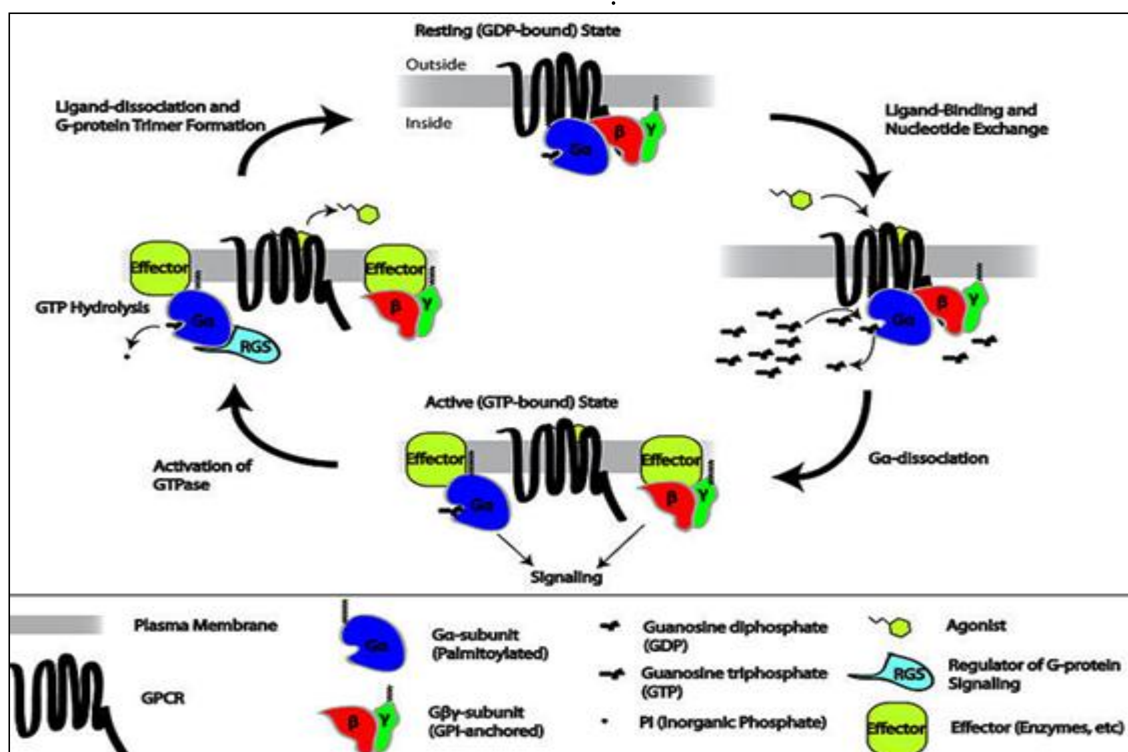


FIG. 1: CARTOON DEPICTING THE HETEROTRIMERIC G-PROTEIN ACTIVATION/DEACTIVATION CYCLE IN THE CONTEXT OF GPCR SIGNALING [10]

Expression of GPCR (NTIS1B) in different strains of *E. coli*: The over-expression of membrane proteins, especially those of eukaryotic source is often toxic to *E. coli*. This usually results in either poor expression levels or may even hamper transformation. The adapted strains C41 (DE3) and C43 (DE3) [11], derived from BL21 (DE3), have allowed the expression of many toxic proteins, particularly membrane Proteins [11]. This may be, at least in part, due to the formation of intracellular membranes during protein induction [12].

In a research carried out by Atrill and his group [13], they examined the expression of NTS1B in four cell lines under three different IPTG concentrations and in the absence of IPTG. The cells were harvested after 30hour of induction. Expression levels were compared by assaying the amount of functional protein as determined using saturation [^{3}H]-NT binding assay as shown in **Table 1**.

All cells lines showed significant leaky expression in the absence of IPTG, despite the presence of the *lac* repressor gene within the expression plasmid. Uninduced expression of NTS1 in DH5α and C43 (DE3) cells approaches that of IPTG-mediated induction. This, in part, may reflect the long period of expression and the growth suppression seen with inducing with IPTG. The highest yield of NTS1 per L of cell culture was seen in BL21 (DE3) cells induced with 0.25mM IPTG (260±32μg/L), and showed relatively high levels of expression across all IPTG concentrations tested.

Expression in C41 (DE3) cells was also relatively high, with the highest yields obtained with 0.25mM IPTG (160±27μg/L). DH5α cell which is traditionally used for NTS1 expression showed the poorest performance, with the highest yield of 75μg/L at 0.1mM IPTG, and a fall of expression beyond this level of IPTG. That is, by changing from DH5α induction at 0.1 mM IPTG to BL21 (DE3), induced with 0.25mM IPTG, a 2.7-fold increase in active protein can be obtained [13].

TABLE 1: EXPRESSION OF FUNCTIONAL ACTIVE RECEPTOR IN DIFFERENT STRAINS

Cell Type	Media	IPTG(mM)	OD600	Yield(μg/L)
DH5α	2xYT	0	2.8 ± 0.2	45 ± 4
DH5α	2xYT	0.1	2 ± 0.3	75 ± 11
DH5α	2xYT	0.25	1.5 ± 0.1	33 ± 6
DH5α	2xYT	0.5	1.3 ± 0.1	25 ± 2
BL21(DE3)	2xYT	0	3.9 ± 0.2	38 ± 9
BL21(DE3)	2xYT	0.1	3.7 ± 0.3	196 ± 19
BL21(DE3)	2xYT	0.25	2.2 ± 0.2	206 ± 32
BL21(DE3)	2xYT	0.5	1.9 ± 0.3	150 ± 29
C41(DE3)	2xYT	0	4.0 ± 0.2	27 ± 4
C41(DE3)	2xYT	0.1	3.9 ± 0.1	101 ± 11
C41(DE3)	2xYT	0.25	2.2 ± 0.3	160 ± 27
C41(DE3)	2xYT	0.5	2.0 ± 0.1	89 ± 6
C43(DE3)	2xYT	0	3.8 ± 0.6	53 ± 6
C43(DE3)	2xYT	0.1	2.5 ± 0.1	69 ± 2
C43(DE3)	2xYT	0.25	1.9 ± 0.3	84 ± 8
C43(DE3)	2xYT	0.5	1.5 ± 0.1	100 ± 13
C43(DE3)	2xYT	1	1.6 ± 0.2	65 ± 6

Autoinduction of GPCR in *E. coli*: By careful formulation of media, Studier [14] developed a reliable protocol for *lac* operon/promoter-dependent autoinduction of genes in *E. coli*. The media contains three carbon sources: glucose, glycerol and lactose. In the initial phase expression is suppressed by glucose and cells grow to a high density.

When glucose is exhausted, lactose-dependent induction begins. Autoinduction has several advantages over IPTG induction: greater biomass, tight expression control, ease of scalability and reduced sample handling. Several commercial formulations based on this method are now available [12].

As shown in **Table 2**, comparison was made between several *E. coli* strains and three autoinduction media formulations: a homemade media based on that defined by Studier[14] and two commercial formulations: Overnight Express™ Autoinduction System 2 from Novagen[12] and Magicmedia ™ from Invitrogen[12]. Autoinduction yielded far more receptor than IPTG-mediated induction, giving a two to fourfold increase over the optimal IPTG concentration for each cell types. Again, the optimal expression strain was BL21 (DE3).

The media formulation that gave the greatest yield of protein was Magicmedia™ with yields in BL21 (DE3) cells reaching 824±22µg/L culture, compared with 650±58µg/L in homemade autoinduction media, 505±18µg/L in Overnight Express Autoinduction System 2 (a maximum of 206±32µg/L was obtained with IPTG-mediated induction. Supply of rare tRNAs (Rosetta2 cells) did not enhance expression over the BL21 (DE3) as shown in Table 2.

TABLE 2: AUTOINDUCTION OF NTS1 EXPRESSION. THREE FORMULATIONS OF AUTOINDUCTION MEDIA COMPARED

Strain Type	Media	OD600	Yield (µg/L)
DH5α	Magicmedia	7.1 ± 1	178 ± 12
BL21	Magicmedia	5.6 ± 0.1	824 ± 22
C41	Magicmedia	6.5 ± 0.2	398 ± 14
C43	Magicmedia	5.2 ± 0.8	369 ± 7
Rosetta 2	Magicmedia	5.9 ± 0.4	704 ± 30
BL21	Studier	5.5 ± 0.3	650 ± 58
C41	Studier	6.2 ± 0.4	430 ± 62
BL21	O/N express	4.6 ± 0.2	505 ± 18
C41	O/N express	4.6 ± 0.4	395 ± 18

GPCR expression in *E. coli*: targeting the inner membrane: In view of the fact that GPCRs are plasma membrane proteins, targeting of recombinant receptors to the inner membrane of the bacterium was initially considered to be the best strategy [Fig 2]. However, in most cases, this leads to severe cell toxicity and low levels of expression. To improve on this, a more efficient insertion into the bacterial inner membrane was done by fusing the GPCR to a protein helper partner. That is, coupling of *E. coli* β-galactosidase (114 kDa) to the N terminus of the human β2-adrenoceptor led to measurable membrane expression [15].

The combination of *E. coli* maltose-binding protein (MBP, 43kDa), used as an N-terminal fusion partner, with *E. coli* thioredoxin A (TRX, 10kDa), added at the C terminus of the GPCR, is particularly well adapted for expression of the NTS1 and cannabinoid CB2 receptor [16].

Membrane expression of the NTS1 receptor has been highly successful and has been applied to automated large-scale purification [13], but the MBP-GPCR-TRX fusion strategy cannot be generally applied without extensive receptor truncations or modification [6].

GPCR expression in *E. coli*: targeting the inclusion bodies: Expression of heterologous proteins in *E. coli* is frequently associated with incorrect folding and accumulation of the recombinant protein in cytoplasmic aggregates named inclusion bodies (IBs) (**Fig. 2**). Targeting of GPCRs to IBs comes with many advantages which include the fact that, IBs are mechanically stable and can be easily isolated from other cell constituents by centrifuging.

Again, they are not toxic to the cell, and they are resistant to proteolytic degradation [6]. This particular strategy was successfully developed first for the rat olfactory OR5 receptor and then several other GPCRs [17] and subsequently improved for human leukotriene BLT1 and the human serotonin 5-HT_{4A} receptors [18, 19].

In most cases efficient production was achieved by employing a fusion partner [18]. GPCRs were mostly attached to a large fusion partner such as the schistosomal glutathione S-transferase (GST, 25 kDa), and had to be truncated at their N terminus [19].

A recent high-throughput effort at large-scale production of more than 100 GPCRs as bacterial IBs revealed that majority of them can be expressed in quantities sufficient for solubilization and purification [20, 21]. A recent study evaluated the efficiency of various fusion partners, namely GST, MBP, TRX and the *E. coli* N-utilization substance A (NusA, 50 kDa), to target GPCRs to IBs. Depending on the culture conditions, GST and TRX were identified as most efficient [6]. The use of β5integrin fragment (β5I, 31 kDa) as a targeting partner has also facilitated the expression of many rhodopsin-like GPCRs at high levels, regardless of their length [337-472 amino acids (aa)], the selectivity of their G protein coupling or the nature of their endogenous ligands [6].

This generic procedure has been successfully applied in the expression of β_3-adrenoreceptor, the vasopressin V2 and V1b and oxytocin OTR receptors, the chemokine CCR5 and CXCR4 and chemokine-like ChemR23 receptors, the ghrelin GHS-1a receptor, canabinoid CB1 receptor and the leukotriene BLT1, BLT2, CysLT1 and CysLT2 receptors, without requiring any optimization of the GPCR coding sequence, the cell culture conditions or the extraction and purification procedures[6].

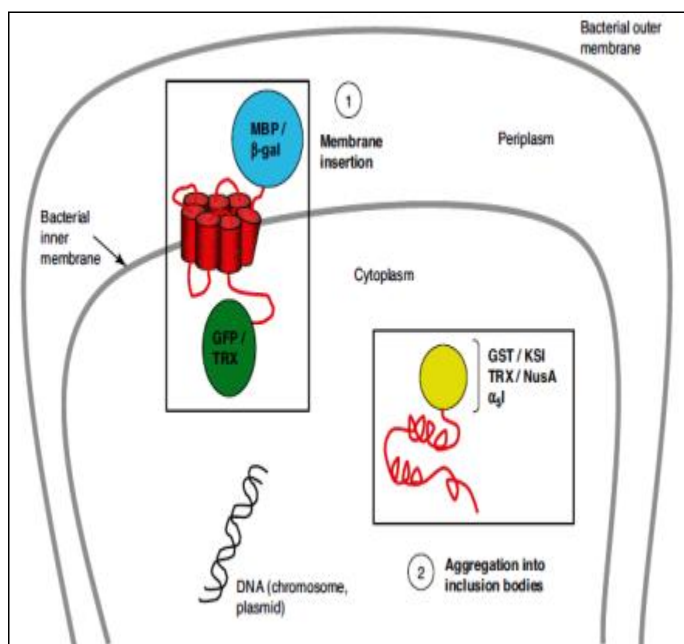
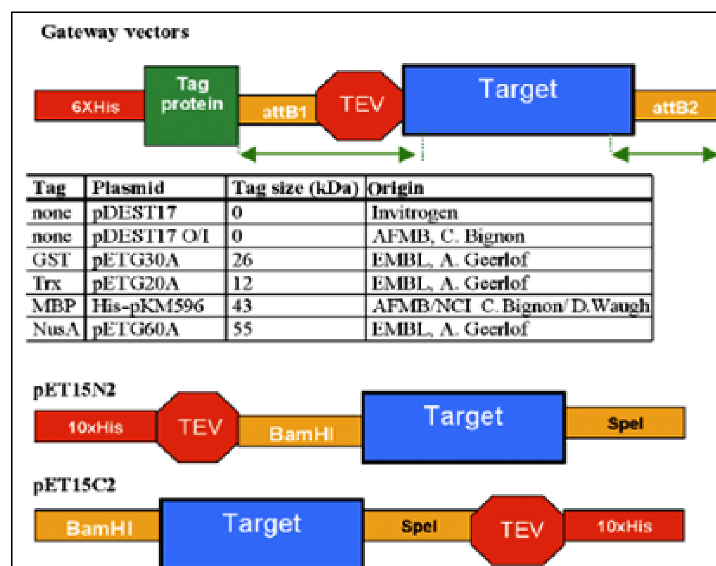


Fig. 2: Strategies for expression of GPCRs in *E. coli* GPCRs can be produced in bacteria either by (1) insertion into the inner membrane or (2) accumulation in inclusion bodies (IBs). In the first case, a protein partner is coupled at the N terminus (MBP, β-gal) or C terminus of the receptor, or at both extremities to target the recombinant protein to the membrane. In the second case, targeting to IBs is favored by coupling the N terminus of another fusion partner [6].

Effective vectors for GPCR Expression in *E. coli*: In a study by Kerstin Michalke and his group [22], two classes of T7-based vectors were used: two pET vectors with His10 tags at either terminus (**Fig. 3**) and a TEV cleavage site and a series of Gateway vectors [23]. The Gateway cloning used the same entry vector (pDONOR 201) further recombined with six different destination vectors, all containing a His6 tag and a TEV cleavage site. As shown in (Fig. 3) two of them did not have insertion between the His6 tag and the TEV site, whereas in the other four a protein (MBP, TRX, GST or NusA) was inserted in between. Out of the 100 receptors selected for the study [22], a subset of 16 targets was subcloned in the eight vectors. Following these preliminary expression results, a second subset of 38 targets was subcloned in the pET15N2 vector and in the four Gateway vectors. The remaining set of 46 targets was subcloned in the pET15N2 vector only. In all, a total of 326 constructs were generated for their expression studies.



Tag	Plasmid	Tag size (kDa)	Origin
none	pDEST17	0	Invitrogen
none	pDEST17 O/I	0	AFMB, C. Bignon
GST	pETG30A	26	EMBL, A. Geerlof
Trx	pETG20A	12	EMBL, A. Geerlof
MBP	His-pKM596	43	AFMB/NCI C.Bignon/D.Waugh
NusA	pETG60A	55	EMBL, A. Geerlof

FIG. 3: SCHEMATIC REPRESENTATION OF THE EXPRESSION VECTOR USED IN THE STUDY [24]

The outcome of their [22] research proved that Gateway vectors, widely used in structural genomics with soluble protein [25], were efficient for GPCR expression as IBs in *E. coli* cytoplasm. They have advantage over pET vectors which are also efficient, to allow high-throughput cloning with about 100% efficiency [22]. Results from the research [22] however showed that to achieve completeness, a large number of vectors and expression conditions are necessary. On the other hand, if cost-effective experiments are the criteria of choice, then pDEST17 vectors, C43 strain and 37°C are the most favorable choice.

CONCLUSION: Generally, structural biology studies require large amounts of the protein on consideration. This can be achieved by employing various methods which have been described by researchers elsewhere[4] but have not been captured in this article. This review article was interested in bringing to light the cost effective means of producing large amounts of membrane protein GPCR, for structural and functional studies, which is an important step to drug discovery. Bacterial expression of heterologous proteins can be used to accomplish this.

Although, this method is cost effective it has its own challenges. Evidence from other Studies has shown that, bacterial expression of GPCR-family has not yielded functional products [26]. Here we have outlined some methods to improve and maximize the yield of *E. coli* expressed GPCR. This is based on the outcome of various studies carried out by researchers in this field of research [6, 13]. The use of alternative strains for expression revealed that BL21 (DE3) cells produced about 2.5-fold increase of the receptor (GPCR) when compared with DH5α, which has been traditionally used for the expression of GPCR [27]. All the cell types used showed leaky expression in the absence of IPTG which could inhibit expression [13].

Autoinduction has been shown to improve the expression of heterologous proteins in *E. coli* [14, 28]. Autoinduction gave largely improved expression than even the highest levels of IPTG-mediated induction [29, 30] and therefore with the right medium in place autinduction can be employed for GPCR expression in *E. coli*. GPCR targeting to IBs expression strategy in *E. coli* has interesting potential in terms of both the amounts produced and general applicability [6, 30].

The fusion partners GST and TRX have been identified as the most efficient for GPCR expression in *E. coli*. The use of β5integrin fragment has also proving to be very efficient. Membrane expression of the NTS1 receptor was highly successful [6] and therefore could be used for effective expression of GPCR. Also according to the research carried out by Kerstin Michalke and his group [22], pDEST17 vectors and C43 strain at 37°C has been proving to be an effective combination for the expression of fully functional GPCR.

REFERENCE:

1. Venter J C, Adams M D, Myers E W, Li P W, Mural R J, Sutton G G, *et al*: The sequence of the human genome. Science 2001; 291: 1304-1351.
2. Terstappen G C and Reggiani A: In silico research in drug discovery. Trends Pharmacol. Sci. 2001; 22: 23-26.
3. White S H: The progress of membrane protein structure determination. Protein Sci. 2004; 13: 1948-1948.
4. Robinson A S and Loll P J: Introduction: current challenges I membrane production. In Production of membrane: Strategies for Expression and Isolation. Wiley-VCH, Weinheim, Germany, (Robinson A. S. ed.), 2011: p424.
5. Link A J and Georgious G: Advances and Challenges in membrane protein expression. AIChE. J. 2007; 53: 752-756.
6. Bane'res J-L, *et al*: New advances in production and functional folding of G-protein-couple receptors. Trends in Biotechnology. 2011; 7: 314-322.
7. Luttrell L M: Review in molecular biology and biotechnology: transmembrane signaling by g protein-coupled receptors. Mol. Biotechnol. 2008; 39: 239-264.
8. Papasaikas P K, Bagos P G, Litou Z I, Promponas V J and Hamodrakas S J: FRED-GPCR: GPCR recognition and family classification server. Nucleic acids Red. 2004; 32: 380-382.
9. Crouch M F and Osmond R I: New strategies in drug discovery for GPCRs: High throughput detection of cellular ERK phosphorylation. Comb. Chem. High Throughput Screen. 2008; 11: 344-356.
10. http://en.wikipedia.org/wiki/File:GPCR_cycle.jpg
11. Dumon-Seignovert L, Cariot G and Vuillard L: The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21 (DE3), C41 (DE3), and C43 (DE3). Protein Expr. Purif. 2004; 37: 203–206.
12. Arechaga I, Miroux B, Karrasch S, Huijbregts R, de Kruijff B, Runswick M J and Walker J E: Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F (1) F (o) ATP synthase. FEBS Lett. 2000; 482: 215–219.
13. Attrill A, *et al*: Improved yield of a ligand-binding GPCR expressed in *E. coli*. for structural studies. Protein Expression and Purification. 2009; 64: 32-38.
14. Studier F W: Protein production by auto-induction in high density shaking Cultures. Protein Expr. Purif. 2005; 41: 207-234.
15. Marullo S, *et al*: Human $_{\beta}$2-adrenergic receptors expressed in *Escherichia coli* membranes retain their pharmacological properties. Proc. Natl. Acad. Sci. U.S.A. 1988; 85: 7551-7555.
16. Tucker J and Grisshammer R: Purification of a rat neurotensin receptor expressed in *Escherichia coli*. Biochem J. 1996; 317: 891-899.
17. Kiefer H, *et al*: Expression of an olfactory receptor in *Escherichia coli*: purification, reconstitution and ligand binding. Biochemistry. 1996; 35: 16077-16084.
18. Bane'res J-L, *et al*. Structure-based analysis of GPCR functions: conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. J. Mol. Biol. 2003; 329: 801-814.
19. Bane'res J-L, *et al*: Molecular characterization of a purified 5-HT_4 receptor. J. Biol. Chem. 2005; 280: 20253-20260.
20. Michalke K, *et al*: Mammalian G protein-couple receptor expression in *Escherichia coli*: I High throughput large-scale production as inclusion bodies. Anal Biochem. 2009; 386: 147-155.
21. Michalke K, *et al*: Mammalian G protein-coupled receptor expression in *Escherichia coli*: II Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human

parathyoroid hormone receptor 1. Anal. Biochem. 2009; 401: 74-80.

22. Michalke K, *et al*: Mammalian G-protein-coupled receptor expression in *Escherichia coli*: I. High-throughput large-scale production as inclusion bodies. Analytical Biochemistry. 2009; 386: 147-155.
23. Walhout A J, *et al*: GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. Methods Enzymol. 2000; 328: 575-592.
24. Thompson M D, *et al*: Pharmaco genomics of G protein-coupled receptor signaling: insights from health and disease. Methods Mol. Biol. 2008; 448: 77-107.
25. Vincentelli R, *et al*: Medium-scale structure genomics: strategies for protein expression and crystallization. Acc. Chem. Res**.** 2003; 36: 165-172.
26. McCusker E C, Bane S E, O' Malley M A and Robinson A S: Heterologous GPCR expression: a bottleneck to obtaining crystal structures. Biotechnol. Prog. 2007; 23: 540-547.
27. Grisshammer R and Henderson R: Expression of a rat neurotensin receptor in *Escherichia coli*. Biochem. J. 1993; 295: 571-576.
28. Furukawa H and Haga T: Expression of functional M2 muscarinic acetylcholine receptor in *Escherichia coli*. Biochem. J. 2000; 127: 151-161.
29. McCuskey E C, *et al*: Heterologous GPCR expression: a bottleneck to obtaining crystal structures. Biotechnol. Prog. 2007; 23: 540-547.
30. Mouillac B and Bane'res J-L: Mammalian membrane receptors expression as inclusion bodies in *Escherichia coli*. Methods Mol. Biol. 2010; 601: 39-48.

How to cite this article:
Acheampong DO, Zhang J and Wang M: Expression of fully functional GPCR in *E. coli*. *Int J Pharm Sci Res* 2013; 4(4); 1356-1362.