Experimental Parasitology xxx (2018) 1–7

Contents lists available at ScienceDirect

Experimental Parasitology

journal homepage: www.elsevier.com/locate/yexpr

Red fluorescent protein (DsRFP) optimization for *Entamoeba histolytica* expression

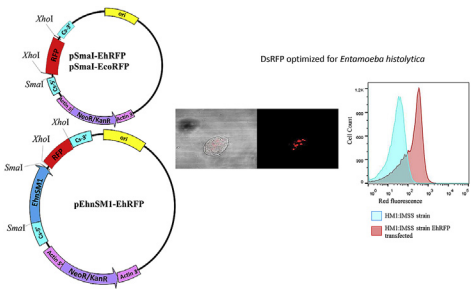
Cindy Jazmín Delgado-Galván [1], Felipe Padilla-Vaca [1], Fátima Berenice Ramírez Montiel [1], Ángeles Rangel-Serrano, Itzel Paramo-Pérez, Fernando Anaya-Velázquez, Bernardo Franco*

Departamento de Biología, División de Ciencias Naturales y Exactas, Universidad de Guanajuato, Noria Alta s/n, Guanajuato, Gto. 36050, Mexico

HIGHLIGHTS

- Molecular tools are available for *Entamoeba histolytica*, except good fluorescent reporter proteins.
- DsRFP codon optimized is a good reporter *in vivo* in this parasite.
- This work will provide efficient protein localization tools for *E. histolytica*.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 13 September 2017
Received in revised form
5 January 2018
Accepted 30 January 2018
Available online xxx

Keywords:
Entamoeba histolytica
Red fluorescent protein
Transient transfection
Reporter protein

ABSTRACT

Entamoeba histolytica genetic organization and genome structure is complex and under intense research. The genome is fully sequenced, and several tools have been developed for the molecular study of this organism. Nevertheless, good protein tracking tags that are easy to measure and image, like the fluorescent proteins are lacking. In this report, we codon-optimized the red fluorescent protein from the coral *Discosoma striata* (DsRFP) for its use in *E. histolytica* and demonstrated functionality *in vivo*. We envision that this protein can be widely used for the development of transcriptional reporter systems and protein-tagging applications.

1. Introduction

Entamoeba histolytica is the causative agent of enteric amoebiasis which can lead to death if not treated properly. This protozoan parasite is one of the leading causes of death among parasitic infections in the world, an estimated 10% of the human population is infected and up to 100,000 patients die each year by this parasite (Ali et al., 2008). While *E. histolytica* can cause invasive disease, the closely related species *Entamoeba dispar* lacks invasive potential *in vivo* despite being able to colonize humans (Ali, 2015).

Pathogenic amoebae have been characterized that they rely on three major virulence factors. (**A**) Cell surface molecules engaged in recognition and adhesion to distinct receptors on host cells: The Gal/GalNAc lectin that recognizes different carbohydrate

* Corresponding author.
E-mail address: bfranco@ugto.mx (B. Franco).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.exppara.2018.01.018
0014-4894/

components on target cells (Singh et al., 2016). (**B**) Several cysteine and acidic proteinases that are released by the amoebae and degrade a variety of host components, such as matrix proteins, mucins, and IgA (Cornick and Chadee, 2017). (**C**) Amoebapores, small protein molecules that form pores in the membranes of target cells and cause depolarization and cell death (Leippe and Herbst, 2004). These molecules are involved in adhesion, immunoevasion, colonization, and invasion, and are associated with responsible for the disease (Padilla-Vaca and Anaya-Velázquez, 2010). These molecules cannot be completely responsible for amoebic virulence because most of them are found in both pathogenic *E. histolytica* strains as well as in the non-pathogenic *E. dispar* (Wilson et al., 2012). Localization, cell trafficking and regulation of these molecules require robust molecular tools to further comprehend their mechanisms and roles in virulence.

To identify genes which expression correlates with the pathogenic or virulent phenotype of *Entamoeba* species/strains or even natural conditions, different approaches that compare the amoebic transcriptomes or proteomes of pathogenic and non-pathogenic species/strains have been performed, but little is known about specific genes that modifies their expression upon exposure to environmental conditions found in the host. One example is the intestinal environment (microbiota mainly) which represents a regulatory element on the virulence of this organism (Padilla-Vaca et al., 1999; Padilla-Vaca and Anaya-Velázquez, 2010).

The intense research on the field of amoebiasis relies on tools that have been developed, such as *in vivo* animal model of infection, cytopathic, cytolytic and cytotoxic assays *in vitro* using different cell lines, trophozoite transfection and gene silencing (Morf and Singh, 2012; Ralston and Petri, 2011).

However, in the literature, there are few references using fluorescent proteins in *E. histolytica*, for two main reasons, one is the high autofluorescence that this organism presents in the green channel, and two, the requirement of oxygen and pH by the green fluorescent protein (GFP) to achieve its maximum fluorescence (Heim et al., 1994). Other attempts have been the use of luciferase, as reported by Gilchrist and colleagues (Gilchrist et al., 1995), but this technique is not useful for protein localization, since the cell must be lysed to quantify the amount of luciferase present in the cell or measured by a PCR-based technique (Gilchrist et al., 1996). In both cases, this approach cannot generate good protein localization reporters. In the close relative *Entamoeba invadens*, luciferase transfection was more successful regarding detection (Singh et al., 2012) since authors found a good correlation between gene expression (northern blot) and luminosity. In a previous report, luciferase expression was also reported for *E. histolytica* and expression relied only on genetic elements from the same organism (SV40 elements were shown to be inactive in this organism), restricting the possibility of adapting current tools for other organisms (Purdy et al., 1994).

There are also some limited examples of the use of the green fluorescent protein (GFP) in this organism. For example, Byekova and collaborators reported the use of lipid-tagged biosensors, which was achieved in a high expression system (Byekova et al., 2010). Also, Ramakrishnan and collaborators reported the implementation of the GFP in a transient transfection assay (Ramakrishnan et al., 2001). However, the fluorescent signal found by Ramakrishnan and collaborators is heterogeneous (determined by flow cytometry) together with a low expression of the fluorescent protein.

Another alternative is the use of antibodies, which are costly or tags that can become fluorescent in the contact with a fluorophore or dye, one such example is the SNAP-tag that has been successfully used in *E. histolytica* (Sateriale et al., 2013) which uses a tag that requires a fluorophore-labeled with O6-benzylguanine derivatives to generate a detectable signal.

All these attempts have highlighted an additional problem: the codon usage of this organism, which may impair the proper expression of fluorescent proteins isolated from different organism's due to the low G + C content in *E. histolytica* genome, challenging the expression of higher G + C containing sequences. Taken together, the need for a good reporter protein, adapted for *E. histolytica* is a necessity to achieve more experimental data on physiological processes in this organism. In the present report, we codon-optimized the DsRFP protein from the coral *Discosoma striata* using previously reported codon usage data of *E. histolytica* (Ghosh et al., 2000; Romero et al., 2000) and successfully demonstrate the correct expression of the protein and its usefulness for protein localization. This monomeric fluorescent protein is ideal for generating small plasmid constructs and the stability is shown here can expand the genetic toolbox for *E. histolytica*.

2. Materials and methods

2.1. Cell culture

We selected a strain that showed high virulence (HM1:IMSS) for the experiments performed in this study (Mendoza-Macías et al., 2009). All strains were grown in TYI-33S media as described previously (Diamond, 1983). *Escherichia coli* cells were grown in LB media (liquid or solid) supplemented with 200 μg/ml ampicillin for the selection of transformants.

2.2. Codon optimization of DsRFP coding sequence

The BBa_E1010 bio part sequence deposited in the iGEM registry (http://parts.igem.org/Protein_coding_sequences/Reporters) was used for *E. coli* or *E. histolytica* codon optimization, which was conducted using the IDT codon optimization tool with a reported previously codon analysis studies (Ghosh et al., 2000; Romero et al., 2000; http://www.idtdna.com/CodonOpt, Fig. S1). In supplementary data (Fig. S1 and supplementary.txt file) we provide the full optimization scheme, sequence alignment of the two sequences and the full sequence of both genes compared between them. The coding sequence was synthesized as gBlocks (Integrated DNA Technologies) for each *E. coli* and *E. histolytica* codon usage sequences.

2.3. Molecular techniques for plasmid construction

Cloning of the DsRFP coding sequence from *E. coli* and the codon-optimized EhRFP was PCR amplified from the gBlocks obtained from IDT using PFX platinum high fidelity polymerase (Invitrogen) with the primers listed in Table 1. Both coding sequence gene fragments were PCR amplification. PCR products were gel purified using PureLink Gel Extraction kit (Invitrogen) for cloning and sequencing. RFP fragments were cloned in the pEhEX plasmid (Nozaki et al., 1998 and Table 2) as follows: Plasmid containing the coding sequence for the sphingomyelinase (SMase 1, EHI_007460) sequence was mutagenized to eliminate the stop codon with a second *Sma*I restriction site (Supplementary information Fig. S1), which was previously cloned as a *Sma*I-*Xho*I PCR fragment (Mendoza-Macías et al., 2010). By doing so, we gained a second *Sma*I restriction site that we used to remove the coding sequence of the SMase 1 gene and generated an expression plasmid for RFP alone or in-frame fusion with the SMase 1 fragment (Supplementary Fig. S2). PCR fragment for the RFP coding sequence was *Xho*I digested and cloned in the 3′ end of the SMase 1 enzyme or in the empty vector (named pSmaI, Table 2 and Supplementary Fig. S2). The coding sequence or DsRFP and the derivatives

Please cite this article in press as: Delgado-Galván, C.J., et al., Red fluorescent protein (DsRFP) optimization for *Entamoeba histolytica* expression, Experimental Parasitology (2018), https://doi.org/10.1016/j.exppara.2018.01.018