Computers in Biology and Medicine 76 (2016) 24–29

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Effect of linker length and residues on the structure and stability of a fusion protein with malaria vaccine application

Shabnam Shamriz, Hamideh Ofoghi *, Nasrin Moazami

Department of Biotechnology, Iranian Research Organization for Science and Technology, P.O. BOX: 3353-51111, Tehran, Iran

ARTICLE INFO

Article history:
Received 17 February 2016
Received in revised form
16 June 2016
Accepted 16 June 2016

Keywords:
Fusion protein
Linker
Malaria vaccine candidate

ABSTRACT

Background: Recombinant protein technology has revolutionized the world of biology and medicine. Following this progress, fusion protein technology, as a novel innovation, has opened new horizons for the development of proteins that do not naturally exist. Fusion proteins are generated via genetically fusing two or more genes coding for separate proteins, thus the product is a single protein having functional properties of both proteins. As an indispensable element in fusion protein construction, linkers are used to separate the functional domains in order to improve their expression, folding and stability.
Method: We computationally fused an antigen and an adjuvant together using different linkers to obtain a two-domain fusion construct which can potentially act as an oral vaccine candidate against malaria. We then predicted the structures computationally to find out the probable folding of each domain in the designed construct.
Results: One of the fusion constructs was selected based on the highest value for *C*-score. Ramchandran Plot analysis represented that most residues were fallen in favorable regions.
Conclusion: Our in silico analysis showed that $(GGGGS)_3$ linker confers the best structure and stability for our target fusion protein.

1. Introduction

Proteins have found wide applications in pharmacy and industry. In the beginning, protein production was indeed protein purification from natural sources which was difficult and time consuming. In 1982, the first recombinant human insulin called humulin was manufactured by Eli Lilly and colleagues and approved by FDA [1]. Humulin was the first therapeutic recombinant protein for human use. Since then, a significant increase has been seen in the number of recombinant products. After reproducing natural proteins by recombinant DNA technology, efforts have been focused toward developing de novo proteins that do not exist in nature and are called fusion proteins [2]. Fusion proteins are generated by genetically fusing two or more genes coding for separate proteins. The product is a novel single protein that has functional properties of both moieties. In the structure of fusion proteins, especially those acting as drug, one part provide molecular binding while the others have certain functions such as toxicity [3]. Fusion proteins have wide applications in both research and industry areas including protein purification, imaging and drug delivery [4,5]. The combination of functionally different domains in one molecule simplifies their manufacturing and also delivery in drug targeting purposes. Fused domains will have identical distributions and moreover novel biomolecules with new functions, which do not naturally exist, can be created. These advantages make fusion proteins attractive [2]. However in some cases, the fusion of unrelated domains makes their manufacturing difficult since the component domains are not compatible which can lead to misfolding and aggregation. Most importantly, it will be difficult to control the amount or dose of each domain. Despite all these drawbacks, it is still worthy to deal with fusion proteins. In designing fusion proteins a number of parameters have to be considered. Sometimes proteins need a free N- or C-terminus for their activity. For example if a free N-terminus is required, the second domain should be connected at the C-terminus [6]. Orientation has also a high impact on functionality especially when fusion proteins require a specific end [7]. The selection of suitable linkers to fuse different domains is another crucial element in construction of fusion proteins. Direct fusion of distinct domains may result in misfolding and aggregation of domains or impairing their bioactivity [8]. Linkers in fusion proteins can improve expression, folding and stability, bioactivity and target delivery [9]. Linkers have been inspired from natural multi-domain proteins. Many linkers with various length, sequences and structures have been developed based on the knowledge of

* Corresponding author.
E-mail addresses: Shamriz.sh@irost.ir (S. Shamriz), Ofoghi@irost.ir (H. Ofoghi), moazami@irost.ir (N. Moazami).

http://dx.doi.org/10.1016/j.compbiomed.2016.06.015
0010-4825/

natural linkers [10,11]. Linkers are divided into three groups: flexible, rigid and cleavable linkers [9]. Flexible linkers permit certain degrees of movements, thus are applied to join functional domains that require interdomain interactions. To provide flexibility, this kind of linkers should be rich in small and polar amino acids such as serine or non-polar ones like glycine [10,12]. When the spatial separation of domains is critical for their stability or bioactivity, rigid linkers are applied to separate the functional domains and keep a fixed distance between them [13,14]. While flexible and rigid linkers are stable in vivo, cleavable linkers are susceptible to protease cleavage or disulfide-bond reduction which simplify the cleavage of linkers and release of fusion domains [15]. The rational selection of linkers is based on the properties of the linker regarding the target fusion protein. Web servers and software such as “Linker” have been created to generate linkers for protein fusions.

Here, the structure of a fusion protein with vaccine application is reported. To achieve a vaccine structure containing both adjuvant and antigen simultaneously, functional domains were fused together using different flexible linkers. The designed structures were analyzed to select the best linker leading to the most flexible and stable fusion protein structure.

Malaria is one of the most common infectious diseases threatening the lives of millions of people annually. Over one million people die from malaria each year, mostly the children under the age of 5 years and pregnant women. Malaria is commonly associated with poverty and is responsible for an enormous economic burden due to increased healthcare costs, missed education, losing the ability to work and therefore loss of productivity and effects on tourism. Malaria infection is caused by five parasite species belonging to the genus *Plasmodium* including *falciparum*, *berghei*, *vivax*, *malariae* and *ovale*, and transmitted by the bites of infected female *Anopheles* mosquitoes. *Plasmodium falciparum* is the deadliest of malaria species. Malaria infection causes flu-like symptoms such as fever, headache, fatigue and abdominal discomfort; however, if suitable drugs are not administered quickly, the infection can result in neurological deficits such as coma and severe anemia which are the signs of severe disease and can lead to death. Current malaria control strategies are based on early diagnosis and treatment combined with mosquito control through bed nets and insecticides. However, the resistance of *Plasmodium* to many of the currently used antimalarial drugs and also mosquito resistance to insecticides have posed new challenges to current prevention policies. Therefore, further efforts are needed to develop novel intervention strategies in order to reduce the burden of malaria morbidity and mortality. An effective vaccine can be a key tool to support malaria control and eradication. Molecular technique improvements during last decades have led to advancement of malaria vaccine candidates' development. Efforts have been made to develop vaccines against different stages of parasite life cycles including pre-erythrocytic, erythrocytic and transmission stages. Pre-erythrocytic stage vaccines are aimed to block the passage of parasites through the liver and subsequent erythrocytic stage infection, thus totally prevent clinical disease and transmission.

Malaria invasion into both mosquito and vertebrate is mediated by a protein called Cell-traversal protein for ookinetes and sporozoites (CelTOS). Studies have revealed that the presence of functional CelTOS protein is required for malaria infection and target mutation of its gene reduces the infectivity of parasite and sporozoite. Moreover, studies have shown that due to highly conserved sequence of CelTOS among the *Plasmodium* species, immunization with *P. falciparum* CelTOS may protect the mice against *Plasmodium berghei* sporozoites, leading to this conclusion that in contrast to anti-circumsporozoite protein (CSP) immunity response which is species-specific, targeting the immune response to CelTOS protein may result in protection against different species of *Plasmodium*. Therefore, CelTOS antigen can play as an interesting candidate for the development of pre-erythrocytic vaccines [16,17].

In order to increase or modify the vaccine immunogenicity, vaccines are often administered with adjuvants. The selection of adjuvants is based on the desired immune response. Using IL-2 cytokine as adjuvant to mount immune responses against malaria has several advantages. (1) Stimulation of T cells to secrete INF-γ. (2) Activation of T cells to secrete cytokines such as IL-10. IL-10 inhibits the activity of macrophages and also secretion of IL-1, IL-12 and TNF. IL-12 acts as inducer for INF-γ and its inhibition leads to reduction of INF-γ production. Overall, repressive role of IL-10 prevents inflammation which is one of the consequences of immune responses to malaria vaccines and infection [18–20]. In this paper, a fusion protein consisting of *P. falciparum* CelTOS (PfCelTOS) antigen and human interleukin-2 (IL-2) as an adjuvant is designed to develop a fusion protein which could be applied as a vaccine candidate against malaria.

2. Methods

2.1. Fusion-protein design

Amino acid fragments of PfCelTOS (PfCelTOS UniProt id: Q53UB8) and human IL-2 (human IL-2 UniProt id: P60568) were linked together directly and also by using different flexible linkers including $(G)_8$, (GGGGS) and $(GGGGS)_3$. Since the N-terminal of IL-2 and C-terminal of PfCelTOS are critical to preserve their stability

Table 2
Phyre2 prediction and analysis of secondary structure.

Protein	α-helix (%)	β sheets (%)	Coils (%)
IL-2-PfCelTOS	66	5	21
IL-2-(GGGGS)-PfCelTOS	65	5	20
IL-2-$(GGGGS)_3$-PfCelTOS	65	5	24
IL-2-$(G)_8$-PfCelTOS	64	5	21

Table 1
Properties of the designed constructs determined by ProtParam.

Protein	IL-2-PfCelTOS	IL-2-$(GGGGS)_3$-PfCelTOS	IL-2-$(G)_8$-PfCelTOS	IL-2-(GGGGS)-PfCelTOS
Molecular weight	34487.4	35433.2	34943.8	34802.7
Theoretical pI	5.11	5.11	5.11	5.11
Total no. of negatively charged residues (Asp + Glu)	38	38	38	38
Total no. of positively charged residues (Arg + Lys)	28	28	28	28
Formula	$C_{1548}H_{2444}N_{390}O_{476}S_{11}$	$C_{1581}H_{2495}N_{405}O_{494}S_{11}$	$C_{1564}H_{2468}N_{398}O_{484}S_{11}$	$C_{1559}H_{2461}N_{395}O_{482}S_{11}$
Instability index	46.52	48.01	47.80	46.85
Aliphatic index	95.52	91.06	93.09	93.99
GRAVY	−0.135	−0.151	−0.141	−0.140