

Secreted trimeric viral envelope proteins as a tool for new vaccine design and biochemical assays

Vladimir Pekarik^{1,2}

¹ Department of Physiology, Faculty of Medicine, Masaryk University, Kamenice 5, CZ-625 00 Brno, Czech Republic - European Union

² Central European Institute of Technology (CEITEC), Masaryk University, Kamenice 3, 625 00 Brno, Czech Republic, European Union

* Author to whom correspondence should be addressed; E-Mail: pekarikv@mail.muni.cz; Tel.: +420-54949 7369 / 7432; Fax: +420-5-4521-2044.

Received:3.4.2015 / Accepted:8.12.2015 / Published:1.1.2016

Varies viral envelope proteins are formed by trimeric modules that are highly stable while anchored through the transmembrane domain in the virus membrane. Unfortunately, to raise vaccines against the whole viruses could lead to the production of antibodies against hijacked cellular proteins that can cause adverse effects. Therefore, pure protein preparations are used for vaccines production. It was found that the use of monomeric immunogens is insufficient in induction of neutralising antibodies. Protective antibodies were more often produced when trimeric ENV proteins were used. Here, we describe technical aspects of ENV proteins processing in host cells, the way to construct stable trimeric immunogens, and recent progress with construction of soluble trimers of ENV from Ebola, HIV, rabies, and influenza viruses.

Keywords: lentivirus; influenza; ebola; HIV, foldon; fibritin; GCN4; tetranectin; envelope viruses; GNA lectin; signal sequence;

1. Introduction

Many proteins in cells exist and work in monomeric form however many proteins are known to form dimers to execute their intended role. Transcription factors dimerise as well as growth factor receptors. Another, not as well-known multimeric complexes are trimers. Trimeric proteins are not as common as dimers but they still form substantial number of protein complexes. Among proteins that need to form homo-trimers belongs PCNA (PDB 1AXC), a protein forming part of DNA replication complex. Clathrin forming trimeric units involved in formation of clathrin coated pits during vesicle internalisation. Collagen fibres are another good example of trimeric protein. There are also trimeric signalling molecules such as TNF α [1] and many more.

Most of the viral envelope glycoproteins (ENV GP) are transmembrane trimeric proteins. In order to study biochemical and biophysical

properties of these proteins there is a pressing need to produce recombinant proteins in secreted form. Trimeric soluble ENV GPs are also better immunogens for production of virus neutralising antibodies.

Purpose of this review is to provide an overview of problems and potential solutions to the production of secreted trimeric viral ENV proteins for structural studies, clinical or diagnostic purposes.

2. Virus envelope proteins

Most of the enveloped viruses are characterised by specific envelope proteins that give rise to virus tropism toward specific cell types and also impart specific antigenic properties to a given virus. These proteins are necessary for recognition of specific epitopes on the surface of target cells but also play an important role in the process of virus entry. Being exposed on the surface of the virus they form also an

excellent target for host immune system to produce virus neutralising antibodies. Due to their critical function during virus entry they are a desired target of designing new antiviral drugs or peptides.

Envelope proteins of most enveloped viruses are type I transmembrane proteins forming trimers. Depending on the virus, these trimers are either composed of three single chain monomers (VSV, rabies, Mokola) of the ENV protein or such as in the case of HIV, Ebola or influenza the ENV pre-protein is proteolytically processed into two subunits that often remain attached together through disulphide bridges. Processed dimers then form trimeric ENV protein.

Practically all known viral ENV proteins are glycoproteins. Glycosylation is thought to protect the viruses from antibody recognition and in some cases the polysaccharide chains can be important for recognition of virus receptor proteins. Interestingly, the virus ENV specific glycosylation can be also used for virus concentration for diagnostic purposes or even removal from the blood stream.

Research has shown that a unique lectin protein (*Galanthus nivalis agglutinin*, GNA) from *Galanthus nivalis* (the common snowdrop) has a high affinity for the mannose-rich GPs that are universal constituents on the surface of enveloped viruses leading to the development of ELISA using immobilised snowdrop lectin for detection of envelope proteins of HIV and SIV [2]. Recent case report describes removal of Ebola virus from the blood plasma of a patient by lectin affinity plasmapheresis [3] reducing the amount of circulating viruses over 60 times. Virus carbohydrate specific lectins can provide an invaluable tool in the design of new diagnostic tools for enveloped viruses.

Efficient production of mature ENV protein is rather complex and complicated process. The nascent protein has to be properly targeted to endoplasmic reticulum where signal sequence is removed. Pre-protein is proteolytically cleaved and the glycosylation is initiated. Partially

modified protein is then moved to trans-Golgi network where glycosylation is finished and proteins assembled to proper trimeric forms. Budding vesicles are then transported toward plasma membrane and fusion process initiated. Understanding of each step is crucial to produce recombinant proteins as similar as possible to the native forms for practical purposes.

3. Synthesis of transmembrane proteins

In order to understand how to engineer viral ENV proteins we have to know how transmembrane or secreted proteins are processed and produced in a host cell. In general, transmembrane proteins can be divided into several groups – type I, type II, and multiple span transmembrane proteins (Fig. 1). Type I TM proteins are single span TM proteins with carboxyl terminus (C-) in the cytoplasm of the cells. Type II proteins are proteins with amino (N-) terminus oriented to the cytoplasm and C-terminus exposed to extracellular space. Majority of viral ENV proteins belong to the type I group of TM proteins with N-terminus facing extracellular space or in case of mature virions on the surface of virus particles.

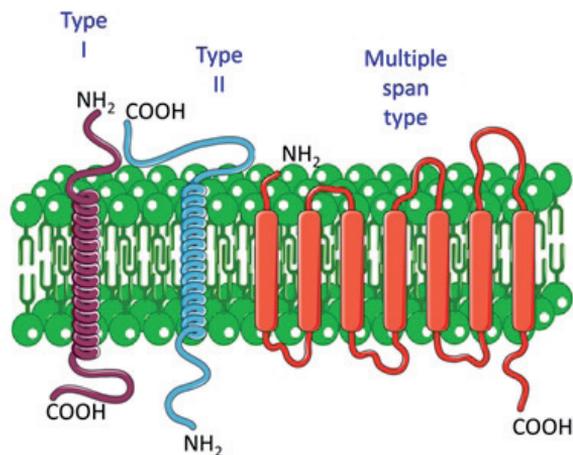


Figure 1: Principal types of transmembrane proteins.

Fig. 1. Structural topology of transmembrane proteins. For single span TM proteins is important orientation of the amino- and carboxy-terminus of the protein.

The proteins are characterised by a signal sequence that is recognised by signal recognition particle that binds hydrophobic sequence at the N-terminus of newly synthesized protein as it emerges from the ribosome and brings the ribosome to the surface of endoplasmic reticulum (ER). Upon docking, the nascent peptide chain is inserted into the translocon channel where it enters into the ER. The synthesis continues up to a moment of reaching transmembrane domain. This domain remains imbedded in the membrane of ER and the remaining part of the protein is synthesised without translocation into lumen of ER forming cytoplasmic tail of the ENV protein.

Proteins in the ER are subject to protein folding, removal of signal sequence and multimerisation. Partially processed proteins are moved through trans-Golgi network where HIV, Ebola and influenza HA proteins are cleaved by furin-like proteases. The proteins are also glycosylated with N-linked oligosaccharides attached in the lumen of ER and further expanded in Golgi. O-linked glycosylation takes place entirely in Golgi. ENV Processing: While ENV proteins of some viruses, namely Rhabdoviridae family (VSV, rabies, Mokola) that are commonly used for pseudotyping lentiviral vectors are not proteolytically cleaved many other viral ENV proteins require further processing.

3.1 Influenza

The major determinant of influenza virus infectivity is hemagglutinin (HA). This protein is produced as a single pre-protein HA0. Despite correct processing (glycosylation, folding, trimerisation) the HA0 is incapable of mediation of membrane fusion. HA0 must be cleaved into HA1 and HA2 by host proteases. Both subunits remain associated as heterodimers of HA1 forming the globular head and the HA2 transmembrane stalk of the protein. HA can be processed either by host protein convertase furin or, as recent progress in research has shown, the activation of HA can occur by

other proteins such as transmembrane serine proteases, secreted serine proteases, plasmin, urokinase, and many other (reviewed in [4]) after the virus release from the host cell.

3.2 HIV

ENV glycoprotein is produced in infected cells in a form of 160 kDa precursor. This pre-protein is subject to signal sequence cleavage, folding and trimerisation. The maturation process is terminated in trans-Golgi network where furin-like protease cleaves 160 kDa precursor into 120 kDa exterior subunit and 41 kDa transmembrane subunit, both remaining associated together. The proteolytically processed mature form is transported to the cell surface membrane and incorporated into budding virions.

The proteolytic cleavage can play an important physiological role. For example uncleaved gp160 membrane bound trimer has nearly indistinguishable structure from GP120/41 but proteolytically processed GPs have greater stability and reduced conformational flexibility [5] important for effective virus recognition and entry.

3.3 Ebola virus (EBOV)

ENV glycoprotein, compared to ENV of influenza or HIV, is interesting due to a fact that the mature pre-protein can be produced only after RNA editing [6]. Translation of unedited RNA leads to the production of smaller 60 kDa precursor glycoprotein that is proteolytically processed into dimeric secreted form of the protein (sGP). The secreted form is thought to play a role in evasion of the host immune system.

Edited RNA (insertion of a single non-template adenosine) is translated into transmembrane glycoprotein that facilitates binding to the host cell receptors and membrane fusion during virus entry. The mature GP1,2 is processed similar way as HIV ENV protein into two associated polypeptides cleaved by intracellular proteinase furin – GP1 (140 kDa) and transmembrane GP2 (25 kDa) that produce membrane bound trimmers [7].

3.3.1 Production of soluble or secreted trimeric GPs.

In order to design a construct for production of recombinant secreted trimeric proteins the individual parts have to be optimised and effectively combined together. Comparison of membrane bound and secreted GPs is in Fig. 2

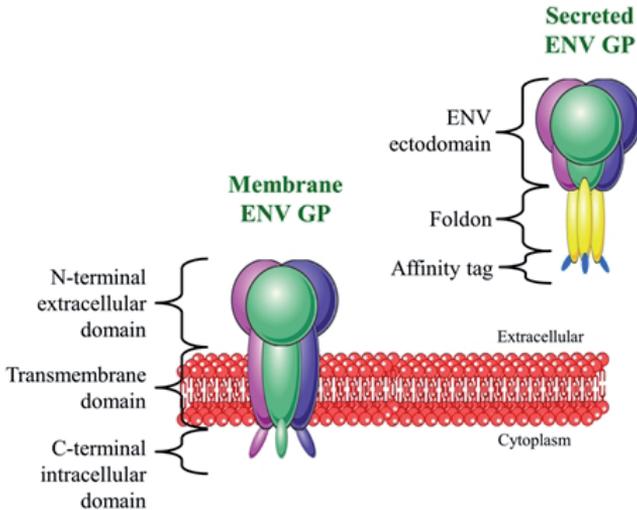


Figure 2: Structure of transmembrane and secreted ENV GP.

Fig. 2. Transmembrane ENV GP is attached in the membrane of the virus. In order to transform ENV proteins into secreted forms, the transmembrane domain has to be removed and replaced with a trimerisation domain. Affinity and recognition tags can be attached to the C-terminus of the protein.

From the figure is evident that from the original GP remains extracellular part (ectodomain). The transmembrane domain has to be removed and replaced with trimerisation motif. In order to facilitate protein purification an affinity tag can be attached to the C-terminus of the protein. An affinity tag can be accompanied also by a recognition tag such as FLAG, HA or any other tag that can be easily distinguished.

In next paragraphs I will describe individual steps that have to be considered during design of any secreted recombinant protein.

4. Signal sequences

Signal sequences are peptides at the N-terminus of nascent protein that direct the growing polypeptide into the lumen of ER. Choice of appropriate signal sequence is very important for efficient production and secretion of the proteins of interest.

Wild-type signal peptides can be used for recombinant protein production however there are many sequences that surpass in efficiency native viral sequences. These sequences are often derived from other transmembrane or secreted proteins. Commonly used signal sequences are derived from human interleukin-2 (IL-2), albumin, trypsinogen-2, and other proteins. Recent progress in cloning of reporter genes has expanded the range of useful signal peptides for sequences derived from secreted luciferases of species *Gaussia* (GLuc), *Metridia* (MLuc) [8]. Surprisingly, it was shown that GLuc signal peptide is probably one of the most effective signal peptides known

so far [9]. It is difficult to explain why a peptide derived from marine organism is far more effective in mammalian expression systems compared to commonly used IL-2 or albumin signal sequences. Useful secreted protein that has found its way to laboratories is placental alkaline phosphatase [10]. This enzyme is normally attached to the outer cell membrane through glycosylphosphatidylinositol (GPI) anchor. Mutation of the C-terminal aminoacids abolishing glycosylphosphatidylinositol site led to production of secreted enzyme [11] SEAP.

Wild-type secretion signals are commonly used but there is still a space for improvement. Zhang et al. [12] has shown that modification of basicity and hydrophobicity of IL-2 secretion signal can dramatically improve protein secretion over 3 times of wild-type (ILco2, ILco3).

Tissue plasminogen activator (tPA) is another secreted protein. Uncharacteristically, the signal sequence cleavage site is formed by proline in position -1. It has been found that small polar

residues with short, neutral side chains, such as alanine (A), glycine (G), serine (S), or threonine (T), are preferred at positions -1. Wang et al. [13] have replaced proline in -1 position of tPA signal sequence with alanine and/or glycine. Replacement of proline with alanine led up to 50 % increase in protein secretion.

the chicken signal is effective in insect cells and its derivative L9 where most of hydrophobic aminoacids in the core region were replaced with leucine is even stronger. Choice of optimal secretion signal can greatly enhance chance of successful secretion of recombinant proteins into cell culture media.

Protein	Signal sequence	ID	Citation:
<i>Gaussia</i> luciferase	MGVKVLFALICIAVAEA//KP	ABW97718	[9]
<i>Metridia</i> luciferase	MDIKVVFTLVFSALVQA//KS	AAR17541.1	[8]
Albumin (human)	MKWVTFISLLFLFSSAYS//RG	NP_000468	[14]
Interleukin-2 (human)	MYRMQLLSCIALSLALVTNS //AP	NP_000577.2	[12]
ILco2	MRRMQLLLIALSLALVTNS //		[12]
ILco3	MRRKMQLLLLIALSLALVTNS //		[12]
Trypsinogen-2 (human)	MNLLLILTFVAAAVA//AP	NP_001290343	
tPA	MDAMKRGLCVLLLCGAVFVSP //SQ	NP_000921.1	[13]
tPA (P – A)	MDAMKRGLCVLLLCGAVFVS A//SQ		[13]
CD5	MPMGSLQPLATLYLLGMLVASCLG //		WT
CD5modified	MPMGSLQPLATLYLLGMLVAS VLA//		[15]
SEAP	MLGPCMLLLLLLLGLRLQLSLG//II	NP_001623.3	
CLSP	MRSLLILVLCFLPLAALG //KV	ACL81619.1	[17]

Table 1: Secretion signal sequences.

A good choice of signal sequence can greatly improve yield of secrete proteins. Kober et al. [14] have screened many various signal sequences for efficient production of recombinant antibodies and among tested sequences found leaders derived from human albumin and azurocidin the most effective.

Another signal of cluster of determinant CD5 protein found on subset of IgM secreting B-cells was used to improve the expression of HIV ENV [15]. In another work, Wen et al. [16] have replaced signal sequences from native hepatitis C virus glycoproteins with computer designed sequence and sequences derived from human tissue plasminogen (tPA) and GLuc. All recombinant proteins were produced more effectively when carrying the new signal sequences.

For many purposes it is more effective to produce recombinant proteins in insect cells (*Spodoptera frugiperda*, Sf9). Tsuchiya et al. [17] have tested signal sequence derived from chicken lysozyme and several synthetic sequences based on the chicken lysozyme signal peptide (CLSP). The authors have found that

Sequences in bold represent wild-type or consensus motives. Shading represents substitutions in consensus sequence, usually improving function of the signal sequence. Different shadings of proteins indicate individual protein groups. The cleavage site of the signal peptide is indicated by // symbol.

5. Codon optimisation:

Aminoacids in proteins are defined by triplets of bases in DNA called codons. Certain aminoacids are coded by a single codon (Met) but some aminoacids can be coded by as many as 6 different codons (Arg, Leu, Ser). Despite the fact that all living organisms utilise 64 available codons (43) each organism utilise codons for a given aminoacid with different frequencies meaning that bacteria will use different codons for coding the same protein than mouse or human cells. In order to improve recombinant protein expression in ectopic expression system the codon composition has to be optimised for cells that produce the protein. Even though, the

viruses are produced from eukaryotic cells the codon usage might still be suboptimal and can pose a limiting factor in recombinant protein production. This was clearly demonstrated for HIV envelope protein subunit gp120 [15] where codon optimisation have greatly enhanced expression of mature protein.

6. Trimerisation motives

Because the viral GPs are firmly anchored in the cell membrane and the sterical confinement stabilises protein trimers the earlier attempts to produce native soluble GPs was met with many problems.

The breakthrough came with the introduction of trimerisation domains derived from other proteins naturally forming trimers in the cells. The most commonly used trimerisation domains are derived from GCN4, which is a transcription factor responsible for the derepression response upon amino acid starvation in yeast [18]. The transcription factor under normal circumstances forms dimers but the leucine zipper structure responsible for oligomerisation was modified to favour trimers formation [19] and the domain was designated GCN4pII.

Another, even more often used, trimerisation domain is derived from bacteriophage T4 protein fibrin. This protein forms whiskers at the base of viral head and its structure is coiled-coil trimer [20]. Fibrin motif was used to produce recombinant trivalent single-chain variable fragment antibody directed against rabies virus glycoprotein with improved neutralizing potency [21] in comparison to bivalent antibodies. The design takes an advantage of trimeric structure of rabies GP. Trimeric antibody can at the same time block all three subunits of the GP leading to enhanced avidity and greater neutralising potential. The trimeric antibody was shown to have apparent affinity constant 75-fold time higher compared to single chain antibody.

Similar mechanism to improve the function of recombinant proteins targeting trimeric receptors was used in trimerization of murine TNF ligand family member LIGHT that has increased the cytotoxic activity against the FM3A mammary carcinoma cell line [22].

The use of trimerization domains for vaccine

development can be hampered by the fact that these domains are inherently strongly immunogenic. The immunogenicity can be reduced by protein engineering of foldon sequence that introduces glycosylation sites [23]. Four sites for N-linked glycosylation were introduced into GCN4-based isoleucine zipper and modified trimerization domain (IZN4) was fused to HIV ENV or influenza HA. IZN4 strongly reduced the antibody responses against the IZ, but did not affect the antibody titres against HIV ENV or HA.

Naturally occurring trimerization domains might be replaced by synthetic constructs based on rational design. Mason and Arndt [24] have defined parameters of coiled coil domains that has been used for rational design of new trimerization domains [13]. The authors have found that coupling of HA or gp120 to MTQ domain leads to 75 % production of these proteins in stable trimers. Interestingly, MTI domain is less efficient forming only up to 56 % of proteins in trimeric form.

Tetranectin, a trimeric plasminogen-binding protein with an alpha-helical coiled coil motif [25] belongs to a class C-type lectins. The trimerisation potential of tetranectin domain was used to target death receptor 4 (DR4) with constructed trivalent antibody [26]. Tetranectin itself can be used as a design scaffold to produce proteins with high affinity toward trimeric targets. Random mutagenesis and phage selection was used to modify structure of loop 1 and loop 4 of C-type lectin-like domain (CTLD) in order to develop selective antagonist of tumor necrosis factor α (TNF α) [27].

It is difficult to determine what trimerisation domain works better because the direct comparisons are scarce. Yang et al. [28] have systematically compared GCN4 and fibrin foldons. They found that both domains can be effectively used to produce stable HIV gp140 trimers. The fibrin construct was more stable to heat and reducing conditions than the GCN4 construct. In general, the fibrin foldon is the most commonly used trimerisation motif but direct comparison of GCN4, fibrin, tetranectin and synthetic foldons and their derivatives shall be conducted.

Foldon	Aminoacid sequence	Derived from	Note	Citation
IZ	RMKQIEDKIEEILSKIYHIENEIARIKLLIGER	GCN4		[23]
IZN4	NGTGRMKQIEDKIENITSKIY^{NIT}NEIARIKLLIGNRT	GCN4	Glycosylation	[23]
GCN4pII	RMKQIEDKIEEILSKIYHIENEIARIKLLVGER	GCN4		[29]
GCNpII	KQIEDKIEEILSKIYHIENEIARIKLLIGEV	GCN4		[30]
Fd	GYPEAPRDGQAYVRKDG EWVLLSTFL	Fibrin		[23]
Fd	GSGYIPEAPRDGQAYVRKDG EWVLLSTFLG	Fibrin		[31]
Fd	GYPEAPRDGQAYVRKDG EWVLLSTFL	Fibrin		[32]
Fd	GSGYIPEAPRDGQAYVRKDG EWVLLSTFLG	Fibrin		[33]
Fd-E	GSGYIPEAPRDGQC^{YVR}CDGEWVLLSTFLG	Fibrin	Disulfide	[33]
Fd	GYPEAPRDGQAYVRKDG EWVLLSTFL	Fibrin		[21]
MTQ	GSGG^IKEE^IAKI^KEE^QAKI^KE^KIAE^IE^KRIA^EI^EK^RIAG^GCC	Synthetic	75 % trimer	[13]
MTI	GSGG^IKEE^IAKI^KEE^QAKI^KE^KIAE^IE^KRIA^EI^EK^RIAG^GCC	Synthetic	56 % trimer	[13]
Tetranectin	LKSRDLTSLQEVALLKEQQALQTVCL	Tetranectin		[25]

Table 2: Trimerisation foldons.

Different colours of shadings correspond to individual protein groups. Bold sequences indicate peptide consensus. Red letters in IZN4 indicate introduced glycosylation sites.

7. Linker sequences:

The assembly of protein domains into functional whole has to respect certain criteria on spatial dimension, topological constraints, protein folding etc. These criteria can be met by the design of peptide linkers that connect individual functional domains. Careful linker design can improve folding, solubility, bioactivity and stability of fusion proteins [34]. Use of flexible linker can be recommended when the joined domains require a certain degree of movement or interaction. The length can be optimized to achieve appropriate separation of the functional domains, or to maintain necessary inter-domain interactions. Rigid linkers can be on the other hand applied when a strict domains separation is required. For the design of secreted trimeric glycoproteins were, with a measure of success, used flexible linkers based predominantly on the (Gly-Gly-Gly-Ser)_n structure. Lu et al. [33] have found that placement of GSGSGS linker between the HA and foldon and placement of GSGSGSGS linker between the foldon and His tag greatly improves protein trimerisation.

8. Soluble trimeric viral envelope proteins:

The usefulness of secreted ENV proteins has been demonstrated in many previously published papers.

There were attempts to produce ENV proteins in ectopic expression systems such as bacteria. The advantages of *E. coli* are well known genetics and easiness to work with. There are many strains that were engineered in order to improve yield of recombinant proteins. However, bacterial production suffers from many drawbacks. In most cases recombinant proteins are sequestered into inclusion bodies that require strong denaturation to release soluble proteins and artificial refolding protocols have to be applied in order to obtain properly folded proteins. Another problem is associated with a lack of protein glycosylation that can severely hamper the use of recombinant proteins. Nonetheless, *E. coli* produced HA stem domain was used to produce broadly protective influenza vaccines [33]. Fibrin foldon was used as trimerisation motif.

Influenza hemagglutinin is a common target of recombinant protein design. HA ectodomain consist of a head domain (HA1) and a stem domain (HA2) and it is the head domain most of antibodies are raised against. Because the HA head domain evolves with a high mutation

rate, the influenza vaccine has to be regularly updated. The stem domain, on the other hand, is much more conserved, therefore it is reasonable to assume that antibodies raised against stem domain can provide broader protection. However, the production of recombinant stem domain faces several challenges. The domain is not evolved to trimerise as an independent unit. Protein folding occurs as co-translational process therefore is associated with intramolecular formation of disulphide bridges. Absence of the head domain exposes aminoacids that are normally buried under the head domain and can lead to undesired protein agglomeration, insolubility or epitope distortion. These problems were addressed by the group of James Swartz [33] when they used bacterial cell-free translation system to redesign the stem domain that can serve as an immunogen. The trimer formation was accomplished by the use of fibrin foldon attached to the C-terminus of HA ectodomain. The trimeric form was further stabilised by introduction of new intermolecular disulphide bonds. Several exposed hydrophobic residues were mutated. The proper formation of stable trimer was confirmed by reaction with antibodies known to recognise the stem of HA. Interestingly, formation of stable trimers was greatly improved by replacing of Ala14 and Lys18 in the foldon with Cys residues allowing formation of intermolecular disulphide bonds. Introduction of flexible linker in between the foldon and His tag and the HA domain facilitated trimerisation even further. HA was effectively trimerised with GCN4pII domain and the authors have found that trimerisation greatly enhances immunogenicity and stabilises the immunogen [29] [30]. Majority of antibodies are produced against the globular head domain of HA. Recently, a new class of broadly neutralizing anti-influenza virus antibodies that target the stalk domain of the viral hemagglutinin was discovered. Findings of Krammer et al. suggest that a carboxy-terminal trimerization domain is a necessary requirement for the structural integrity of stalk epitopes on recombinant soluble influenza virus hemagglutinin, therefore indispensable for production of novel stalk targeting antibodies [35].

Importance of trimeric proteins stabilisation for effective vaccine development was also confirmed in a case of severe acute respiratory syndrome coronavirus (SARS-CoV) spike protein. SARS-foldon induced a significantly higher titer of neutralizing antibodies compared to monomeric protein [36].

Ebola is extremely dangerous virus therefore the research in most laboratories has to be conducted with isolated parts that are not dangerous in itself. In this way many characteristic of the EBOV GP were deduced. With the use of GCN4 foldon was found that Niemann-Pick C1 protein is essential for EBOV infection [37]. After the virus internalisation, the EBOV GP is cleaved with cathepsin L or B (CatL or B) in the endosome. The role of cathepsin L for ENV cleavage and mechanism of cell entry was investigated with EBOV ectodomain fused to fibrin foldon [31]. Published data suggest that CatL cleavage of EBOV GP exposes its receptor-binding domain, thereby facilitating access to a putative cellular receptor in steps that lead to membrane fusion.

As in previous case, HIV ENV was subject to extensive protein engineering. Stabilisation of soluble HIV trimeric ENV with fibrin foldon was shown already in 2002 [28].

GCN4 connected monomers were able to elicit antibodies of greater neutralizing capacity [38].

Interestingly, trimeric proteins or peptides can be used not only to induce immune response but as direct inhibitors of virus entry. During the virus infection GP120 domain binds to the receptor CXCR4 or CCR5. After binding, gp120 dissociates and the N-terminus of GP41 is exposed. The core structure of GP41 ectodomain consist of N-terminal and C-terminal heptads. N-terminal heptads are inserted into the membrane of target cells enabling fusion between viral and cellular membranes. A peptide derived from N-terminal heptad trimerised with T4 fibrin foldon was shown to be a strong inhibitor of virus fusion [32].

Effective production of stable secreted trimeric ENV proteins can play an important role in discoveries of viral cellular receptors. Rabies virus glycoprotein (RVG) is a 65 kDa

single type I transmembrane trimeric protein. In order to facilitate virus entry the virus glycoprotein has to interact with its cognate receptor. So far there were several candidates as to what is the virus receptor, such as the nicotinic acetylcholine receptor (nACh), the neural cell adhesion molecule (NCAM) and the neurotrophin receptor (p75NTR). Sissoeff et al. [39] have demonstrated that recombinant RVG ectodomain forms unstable trimers that dissociate into monomers in a concentration-dependent manner. C-terminal fusion with the foldon induces stable RVG trimerization, which is concentration-independent. Furthermore, the fibrin foldon maintains the native antigenic structure of the carboxy part of RVGect. Cell binding experiments showed that RVG trimerization is required for efficient interaction with p75NTR.

9. Conclusions

Stable trimeric ENV proteins of different viruses are important tools for production of better vaccines against these pathogens. Foldon stabilised ENV ectodomains can be more easily crystallised allowing better understanding of the virus entry mechanisms. There is a potential to develop new antiviral drugs based on trimeric peptides that interfere with fundamental viral life cycle processes such as entry to the host cells.

Acknowledgments

For pictures creation were used pre-designed motifs provided by SERVIER Medical Art.

Conflicts of Interest

The authors declare they have no potential conflicts of interests concerning drugs, products, services or another research outputs in this study. The Editorial Board declares that the manuscript met the ICMJE „uniform requirements“ for biomedical papers.

References

1. Smith, R.A.; Baglioni, C. The active form of tumor necrosis factor is a trimer. *J Biol Chem* 1987, **262**, 6951-6954.
2. Mahmood, N.; Hay, A.J. An elisa utilizing immobilised snowdrop lectin gna for the detection of envelope glycoproteins of hiv and siv. *J Immunol Methods* 1992, **151**, 9-13.

3. Büttner, S.; Koch, B.; Dolnik, O.; Eickmann, M.; Freiwald, T.; Rudolf, S.; Engel, J.; Becker, S.; Ronco, C.; Geiger, H. Extracorporeal virus elimination for the treatment of severe ebola virus disease - first experience with lectin affinity plasmapheresis. *Blood Purif* 2014, **38**, 286-291.
4. Hamilton, B.S.; Whittaker, G.R.; Daniel, S. Influenza virus-mediated membrane fusion: Determinants of hemagglutinin fusogenic activity and experimental approaches for assessing virus fusion. *Viruses* 2012, **4**, 1144-1168.
5. Haim, H.; Salas, I.; Sodroski, J. Proteolytic processing of the human immunodeficiency virus envelope glycoprotein precursor decreases conformational flexibility. *J Virol* 2013, **87**, 1884-1889.
6. Sanchez, A.; Trappier, S.G.; Mahy, B.W.; Peters, C.J.; Nichol, S.T. The virion glycoproteins of ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc Natl Acad Sci U S A* 1996, **93**, 3602-3607.
7. Volchkov, V.E.; Feldmann, H.; Volchkova, V.A.; Klenk, H.D. Processing of the ebola virus glycoprotein by the proprotein convertase furin. *Proc Natl Acad Sci U S A* 1998, **95**, 5762-5767.
8. Markova, S.V.; Golz, S.; Frank, L.A.; Kalthof, B.; Vysotski, E.S. Cloning and expression of cDNA for a luciferase from the marine copepod metridia longa. A novel secreted bioluminescent reporter enzyme. *J Biol Chem* 2004, **279**, 3212-3217.
9. Knappskog, S.; Ravneberg, H.; Gjerdrum, C.; Trösse, C.; Stern, B.; Pryme, I.F. The level of synthesis and secretion of gaussia princeps luciferase in transfected cho cells is heavily dependent on the choice of signal peptide. *J Biotechnol* 2007, **128**, 705-715.
10. Berger, J.; Hauber, J.; Hauber, R.; Geiger, R.; Cullen, B.R. Secreted placental alkaline phosphatase: A powerful new quantitative indicator of gene expression in eukaryotic cells. *Gene* 1988, **66**, 1-10.
11. Lowe, M.E. Site-specific mutations in the coo-terminus of placental alkaline phosphatase: A single amino acid change converts a phosphatidylinositol-glycan-anchored protein to a secreted protein. *J Cell Biol* 1992, **116**, 799-807.
12. Zhang, L.; Leng, Q.; Mixson, A.J. Alteration in the il-2 signal peptide affects secretion of proteins in vitro and in vivo. *J Gene Med* 2005, **7**, 354-365.
13. Wang, J.Y.; Song, W.T.; Li, Y.; Chen, W.J.; Yang, D.; Zhong, G.C.; Zhou, H.Z.; Ren, C.Y.; Yu, H.T.; Ling, H. Improved expression of secretory and trimeric proteins in mammalian cells via the introduction of a new trimer motif and a mutant of the tpa signal sequence. *Appl Microbiol Biotechnol* 2011, **91**, 731-740.
14. Kober, L.; Zehe, C.; Bode, J. Optimized signal peptides for the development of high expressing cho cell lines. *Biotechnol Bioeng* 2013, **110**, 1164-1173.
15. Haas, J.; Park, E.C.; Seed, B. Codon usage limitation in the expression of hiv-1 envelope glycoprotein. *Curr Biol* 1996, **6**, 315-324.
16. Wen, B.; Deng, Y.; Guan, J.; Yan, W.; Wang, Y.; Tan,

- W.; Gao, J. Signal peptide replacements enhance expression and secretion of hepatitis c virus envelope glycoproteins. *Acta Biochim Biophys Sin (Shanghai)* 2011, 43, 96-102.
17. Tsuchiya, Y.; Morioka, K.; Taneda, I.; Shirai, J.; Yoshida, K. Gene design of signal sequence for the effective secretion of recombinant protein using insect cell. *Nucleic Acids Symp Ser (Oxf)* 2005, 305-306.
 18. Arndt, K.; Fink, G.R. Gcn4 protein, a positive transcription factor in yeast, binds general control promoters at all 5' tgactc 3' sequences. *Proc Natl Acad Sci U S A* 1986, 83, 8516-8520.
 19. Harbury, P.B.; Kim, P.S.; Alber, T. Crystal structure of an isoleucine-zipper trimer. *Nature* 1994, 371, 80-83.
 20. Tao, Y.; Strelkov, S.V.; Mesyanzhinov, V.V.; Rossmann, M.G. Structure of bacteriophage t4 fibrin: A segmented coiled coil and the role of the c-terminal domain. *Structure* 1997, 5, 789-798.
 21. Turki, I.; Hammami, A.; Kharmachi, H.; Mousli, M. Engineering of a recombinant trivalent single-chain variable fragment antibody directed against rabies virus glycoprotein g with improved neutralizing potency. *Mol Immunol* 2014, 57, 66-73.
 22. Ito, T.; Iwamoto, K.; Tsuji, I.; Tsubouchi, H.; Omae, H.; Sato, T.; Ohba, H.; Kurokawa, T.; Taniyama, Y.; Shintani, Y. Trimerization of murine tnf ligand family member light increases the cytotoxic activity against the fm3a mammary carcinoma cell line. *Appl Microbiol Biotechnol* 2011, 90, 1691-1699.
 23. Sliepen, K.; van Montfort, T.; Melchers, M.; Isik, G.; Sanders, R.W. Immunosilencing a highly immunogenic protein trimerization domain. *J Biol Chem* 2015.
 24. Mason, J.M.; Arndt, K.M. Coiled coil domains: Stability, specificity, and biological implications. *ChemBiochem* 2004, 5, 170-176.
 25. Nielsen, B.B.; Kastrop, J.S.; Rasmussen, H.; Holtet, T.L.; Graversen, J.H.; Etzerodt, M.; Thøgersen, H.C.; Larsen, I.K. Crystal structure of tetranectin, a trimeric plasminogen-binding protein with an alpha-helical coiled coil. *FEBS Lett* 1997, 412, 388-396.
 26. Allen, J.E.; Ferrini, R.; Dicker, D.T.; Batzer, G.; Chen, E.; Oltean, D.I.; Lin, B.; Renshaw, M.W.; Kretz-Rommel, A.; El-Deiry, W.S. Targeting trail death receptor 4 with trivalent dr4 atrimer complexes. *Mol Cancer Ther* 2012, 11, 2087-2095.
 27. Byla, P.; Andersen, M.H.; Holtet, T.L.; Jacobsen, H.; Munch, M.; Gad, H.H.; Thøgersen, H.C.; Hartmann, R. Selection of a novel and highly specific tumor necrosis factor alpha (tnfalpha) antagonist: Insight from the crystal structure of the antagonist-tnfalpha complex. *J Biol Chem* 2010, 285, 12096-12100.
 28. Yang, X.; Lee, J.; Mahony, E.M.; Kwong, P.D.; Wyatt, R.; Sodroski, J. Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of t4 bacteriophage fibrin. *J Virol* 2002, 76, 4634-4642.
 29. Weldon, W.C.; Wang, B.Z.; Martin, M.P.; Koutsonanos, D.G.; Skountzou, I.; Compans, R.W. Enhanced immunogenicity of stabilized trimeric soluble influenza hemagglutinin. *PLoS One* 2010, 5.
 30. Lin, S.C.; Huang, M.H.; Tsou, P.C.; Huang, L.M.; Chong, P.; Wu, S.C. Recombinant trimeric ha protein immunogenicity of h5n1 avian influenza viruses and their combined use with inactivated or adenovirus vaccines. *PLoS One* 2011, 6, e20052.
 31. Hood, C.L.; Abraham, J.; Boyington, J.C.; Leung, K.; Kwong, P.D.; Nabel, G.J. Biochemical and structural characterization of cathepsin l-processed ebola virus glycoprotein: Implications for viral entry and immunogenicity. *J Virol* 2010, 84, 2972-2982.
 32. Chen, X.; Lu, L.; Qi, Z.; Lu, H.; Wang, J.; Yu, X.; Chen, Y.; Jiang, S. Novel recombinant engineered gp41 n-terminal heptad repeat trimers and their potential as anti-hiv-1 therapeutics or microbicides. *J Biol Chem* 2010, 285, 25506-25515.
 33. Lu, Y.; Welsh, J.P.; Swartz, J.R. Production and stabilization of the trimeric influenza hemagglutinin stem domain for potentially broadly protective influenza vaccines. *Proc Natl Acad Sci U S A* 2014, 111, 125-130.
 34. Chen, X.; Zaro, J.L.; Shen, W.C. Fusion protein linkers: Property, design and functionality. *Adv Drug Deliv Rev* 2013, 65, 1357-1369.
 35. Krammer, F.; Margine, I.; Tan, G.S.; Pica, N.; Krause, J.C.; Palese, P. A carboxy-terminal trimerization domain stabilizes conformational epitopes on the stalk domain of soluble recombinant hemagglutinin substrates. *PLoS One* 2012, 7, e43603.
 36. Li, J.; Ulitzky, L.; Silberstein, E.; Taylor, D.R.; Viscidi, R. Immunogenicity and protection efficacy of monomeric and trimeric recombinant sars coronavirus spike protein subunit vaccine candidates. *Viral Immunol* 2013, 26, 126-132.
 37. Côté, M.; Misasi, J.; Ren, T.; Bruchez, A.; Lee, K.; Filone, C.M.; Hensley, L.; Li, Q.; Ory, D.; Chandran, K., et al. Small molecule inhibitors reveal niemann-pick c1 is essential for ebola virus infection. *Nature* 2011, 477, 344-348.
 38. Feng, Y.; McKee, K.; Tran, K.; O'Dell, S.; Schmidt, S.D.; Phogat, A.; Forsell, M.N.; Karlsson Hedestam, G.B.; Mascola, J.R.; Wyatt, R.T. Biochemically defined hiv-1 envelope glycoprotein variant immunogens display differential binding and neutralizing specificities to the cd4-binding site. *J Biol Chem* 2012, 287, 5673-5686.
 39. Sissoëff, L.; Mousli, M.; England, P.; Tuffereau, C. Stable trimerization of recombinant rabies virus glycoprotein ectodomain is required for interaction with the p75ntr receptor. *J Gen Virol* 2005, 86, 2543-2552.



The article is freely distributed under license Creative Commons (BY-NC-ND). But you must include the author and the document can not be modified and used for commercial purposes.