International Immunopharmacology 55 (2018) 86–97

Contents lists available at ScienceDirect

International Immunopharmacology

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

Engineering of single chain antibodies for solubility

Yaghoub Safdari

Medical Cellular and Molecular Research Center, Golestan University of Medical Sciences, Gorgan 4934174511, Iran
Golestan Research Center of Gastroenterology & Hepatology (GRCGH), Golestan University of Medical Sciences, Gorgan, Iran

ARTICLE INFO

Keywords:
scFv
VH3
VL3
Hydrophobic cores
Solubility

ABSTRACT

In the current study, we introduce a method to design intrinsically soluble single chain antibodies (scFvs) that can be easily produced in bacterial expression systems in a soluble form. Summarily, CDR loops are grafted on framework regions derived from intrinsically soluble subclass 3 (VH3 and VL3) human germline sequences. Framework-donor sequences should carry CDR loops of interest (in terms of canonical classes) and contain special residues in their hydrophobic cores. Recombinant variable fragments resultant from CDR grafting are subjected to 3D modeling, mutated (if necessary), and superposed to parental variable domains. Recombinant type 3 variable domains with the least RMSD (Root-Mean-Square Deviation) values are chosen to constitute scFv moieties. The scFv designed using this method was shown to be soluble when expressed in bacterial cells.

1. Introduction

The use of scFvs is on the rise for several purposes (e.g. tumor imaging and targeted drug delivery) [1]. Fast growing microorganisms such as bacteria are regarded as a cost-effective host for production of recombinant proteins, including scFvs. However, bacterial expression of foreign proteins - especially those with eukaryotic origin- often leads to formation of insoluble aggregations that are known as inclusion bodies [2]. Working on several scFvs, including anti-EGFRvIII scFv, anti-CD20 scFv, cetuximab scFv, nimotuzumab scFv, and a humanized anti-EGFR antibody [3–6], we noticed that some scFvs are intrinsically soluble and easily produced in bacterial cells, while some others tend to form inclusion bodies and need to undergo refolding process to become active. For example, anti-EGFRvIII, anti-CD20 and nimotuzumab scFvs were highly soluble and active while the anti-EGFR antibodies were always insoluble. We hypothesized that some regions or amino acid residues within the variable domains are responsible for intrinsic solubility/insolubility of the scFv fragments. In the search for similar results, we encountered an article published by Ewert and colleagues in 2003 [7]. They had reported that different antibody subclasses exhibit different degrees of stability and solubility, a property that is determined by type and composition of amino acid residues in upper and lower hydrophobic cores of variable domains. We also hypothesized that if framework regions of an inclusion body- forming scFv are substituted by framework regions derived from intrinsically soluble variable domains, the scFv would be converted to a soluble antibody. We tested the hypothesis and found that framework replacement helps improve the solubility. Most of hydrophobic core residues are in framework regions; therefore, framework regions derived from intrinsically stable variable domains should carry a large fraction of stability improving residues. In the current work, we explain step by step how to screen and engineer human germline sequences to engage in CDR grafting technique in order to obtain an intrinsically soluble scFv that can be easily produced in bacterial cells without inclusion body formation. The method eliminates the need for laborious procedures or inefficient protocols (denaturing and refolding of inclusion bodies, use of chaperones, use of chemical compounds, decrease of aeration rate and growth temperature and so forth) [8–13] to obtain correctly–folded active scFvs. Summarily, CDR loops of an inclusion body-forming scFv are grafted on frameworks of human germline sequences that are intrinsically soluble and carry the same type of CDR loops (in terms spatial structure). Schematic overview of protocol is shown in Fig. 1.

2. Materials and methods

2.1. Protein sequence analysis of soluble and insoluble scFvs

Among the scFvs we produced in our laboratory, the anti-EGFR antibody was found to be completely insoluble. We analyzed the amino acid sequence of both VH and VL domains of this scFv to determine what residues may be responsible for inclusion body formation. The sequence of each variable domain was separately entered in IMGT/DomainGapAlign online tool (http://www.imgt.org/3Dstructure-DB/cgi/DomainGapAlign.cgi) to determine to what antibody subclasses the variable domains are more similar. The tool has an option to adjust the search for human germline genes and arranges the sequences based

E-mail addresses: Safdari_14@yahoo.com, dr.safdari@goums.ac.ir.

https://doi.org/10.1016/j.intimp.2017.11.046
Received 23 October 2017; Received in revised form 16 November 2017; Accepted 30 November 2017

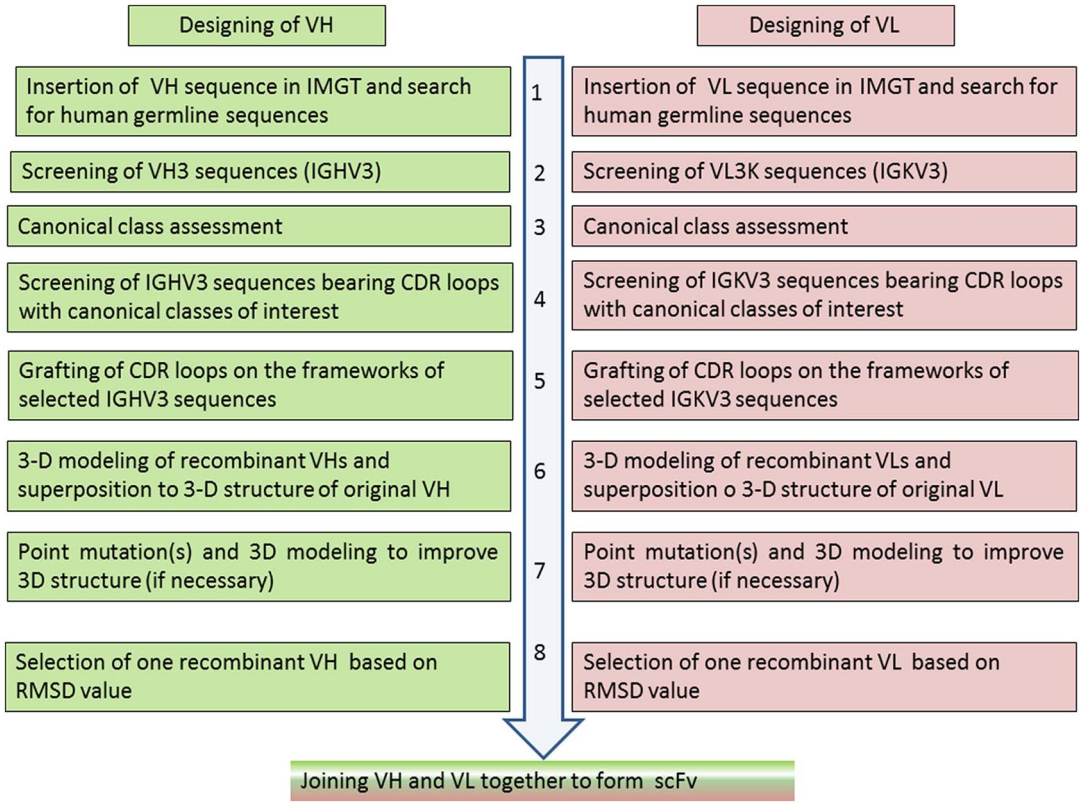

Fig. 1. Schematic overview of the protocol used for screening and engineering variable domains for obtaining a soluble scFv.

on their identities. The types of the arranged sequences provide a rough estimate of the variable domain solubility degree.

2.2. Screening of type 3 human germline sequences

For a given antibody sequence, IMGT/GapDomainAlign tool arranges a number of human germline sequences based on their "Identities" values. Type 3 variable fragments [VH3 and VL kappa3 (Vk3)] have been shown to be more stable and soluble than other types (V_H1a, VH1b, VH2, VH4, VH6, Vk1, Vk2, Vk4, Vλ1, Vλ2, and Vλ3) [7]. Therefore, neglecting the degrees of sequence similarity (Identities), we searched among the IMGT- arranged sequences for type 3 sequences.

2.3. Identification of hydrophobic core residues

Amino acid residues in hydrophobic cores of variable domains have been proven to influence the stability of the molecule by forming intramolecular hydrophobic interactions. To identify hydrophobic core residues, VH and VL sequences were numbered according to AHo numbering scheme (described by Honegger and Pluckthun [7]. AHo numbering scheme is different from Kabat numbering sachem, meaning that amino acid residues numbered by one of these schemes may be given different numbers if they are numbered using the other one (Fig. 2). The main study discussing the role of hydrophobic cores in stability of antibody variable domains is based on AHo numbering scheme [7]. Kabat numbering scheme is the most commonly used antibody numbering schemes. Most of online antibody analysis tools, including canonical assessment tool, use this numbering scheme to analyze antibody sequences. Since the current study includes topics relevant to both hydrophobic cores and CDR canonical classes, we reconciled AHo and Kabat numbering schemes to display what residues in Kabat numbering scheme contribute to hydrophobic core formation (Fig. 2).

2.4. Identification of canonical classes of CDR loops and screening of desired type 3 human germline sequences

The main hypothesis of this study was to use intrinsically soluble frameworks for obtaining an intrinsically soluble scFv from an inclusion body. We assumed that an inclusion body-forming scFv would be converted into a soluble scFv if its frameworks are replaced with intrinsically soluble frameworks. One of the major drawbacks correlated with framework replacement (CDR-grafting technique) is loss of antibody affinity, which happens when the frameworks do not support the grafted CDRs in an efficient manner. To overcome this problem, CDR homology approach has been proposed [14]. The rationale for this approach is based on the hypothesis that if two antibodies carry the same types of CDR loops (in terms of spatial 3D structure), their frameworks would be able to support CDR loops of each other. Therefore, besides the subclass of antibody, we considered the similarity of CDR canonical classes as a criterion to screen framework-donor human germline sequences for CDR grafting. To this aim, canonical classes of all VH3 and VL3 sequences (screened in the first round of selection) were determined using "Chothia Canonical Assignment" online tool (http://www.bioinf.org.uk/abs/chothia.html) and only those carrying the CDR loops of interest (in terms of canonical classes) were allowed to enter the third round of screening.

2.5. 3D-modeling, superposition, and amino acid replacement to obtain humanized variable domains with the minimum RMSD values

CDR loops of cetuximab, which were the same as those of the inclusion body forming scFv (called hscFv), were grafted on the framework regions of type 3 human germline sequences screened in the second round of selection. All the recombinant VH and VL domains were subjected to 3D modeling using PHYRE2 Protein Fold Recognition Server (http://www.sbg.bio.ic.ac.uk/phyre2/). To insure that recombinant scFvs retain the affinity of parental antibodies, CDR loops