Protein Expression and Purification 75 (2011) 192–203

Contents lists available at ScienceDirect

Protein Expression and Purification

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

A high-throughput protein refolding screen in 96-well format combined with design of experiments to optimize the refolding conditions

Vincent Dechavanne [a,*], Nicolas Barrillat [a], Frederic Borlat [a], Aurélie Hermant [b], Laurent Magnenat [c], Mikael Paquet [d], Bruno Antonsson [a,**], Laurent Chevalet [a]

[a] Geneva Research Center, Merck Serono S.A., Department of Protein and Cell Sciences, 9 chemin des mines, 1202 Geneva, Switzerland
[b] Geneva Research Center, Merck Serono S.A., Department of Auto Immune and Inflammatory Disease, 9 chemin des mines, 1202 Geneva, Switzerland
[c] Geneva Research Center, Merck Serono S.A., Department of Protein Engineering and Antibody Technologies, 9 chemin des mines, 1202 Geneva, Switzerland
[d] Geneva Research Center, Merck Serono S.A., Department of Lead Discovery Technologies, 9 chemin des mines, 1202 Geneva, Switzerland

ARTICLE INFO

Article history:
Received 21 July 2010
and in revised form 7 September 2010
Available online 17 September 2010

Keywords:
Design of experiments (DOE)
Inclusion bodies
Refolding screen
IL-17F
SDF-1α
BCA-1
GM-CSF
C5a

ABSTRACT

Production of correctly folded and biologically active proteins in *Escherichia coli* can be a challenging process. Frequently, proteins are recovered as insoluble inclusion bodies and need to be denatured and refolded into the correct structure. To address this, a refolding screening process based on a 96-well assay format supported by design of experiments (DOE) was developed for identification of optimal refolding conditions. After a first generic screen of 96 different refolding conditions the parameters that produced the best yield were further explored in a focused DOE-based screen. The refolding efficiency and the quality of the refolded protein were analyzed by RP-HPLC and SDS–PAGE. The results were analyzed by the DOE software to identify the optimal concentrations of the critical additives. The optimal refolding conditions suggested by DOE were verified in medium-scale refolding tests, which confirmed the reliability of the predictions. Finally, the refolded protein was purified and its biological activity was tested *in vitro*. The screen was applied for the refolding of Interleukin 17F (IL-17F), stromal-cell-derived factor-1 (SDF-1α/CXCL12), B cell-attracting chemokine 1 (BCA-1/CXCL13), granulocyte macrophage colony stimulating factor (GM-CSF) and the complement factor C5a. This procedure identified refolding conditions for all the tested proteins. For the proteins where refolding conditions were already available, the optimized conditions identified in the screening process increased the yields between 50% and 100%. Thus, the method described herein is a useful tool to determine the feasibility of refolding and to identify high-yield scalable refolding conditions optimized for each individual protein.

Introduction

Functional and structural studies of proteins often require large amounts of correctly folded, pure, and active protein. Bacterial expression systems, mostly in *Escherichia coli*, are cheap, simple and frequently used to produce large quantities of recombinant proteins. However, production of foreign proteins in *E. coli* frequently leads to the formation of inclusion bodies (IB).[1] Inclusion bodies are insoluble protein aggregates composed of precipitated unfolded or incorrectly folded protein molecules devoid of biological activity. The major challenge is to efficiently convert this inactive and insoluble protein into soluble, correctly folded and active product [1,2]. To achieve this, the protein needs to be submitted to an elaborated solubilization, refolding and purification procedure [3–5]. After a complete reduction and linearization of the protein from the inclusion bodies, the *in vitro* refolding involves several crucial steps. The covalent disulfide bonds must be generated correctly and the molecule must adopt the native secondary, tertiary, and quaternary structure. Defining conditions that promote refolding of a chemically denatured protein into its native conformation is often empirical, time consuming and frequently yields low amounts of correctly refolded protein.

In recent years, several high-throughput protein refolding methods have been developed [6,7]. These methods are based on dilution, dialysis or solid phase immobilization for the renaturation process [8]. Dialysis and dilution based methods along with the use of various additives have been reported to improve recovery of

* Corresponding author.
** Corresponding author.
E-mail addresses: vincent.dechavanne@merckserono.net (V. Dechavanne), bruno.antonsson@merckserono.net (B. Antonsson).

[1] *Abbreviations used:* DOE, design of experiments; IL-17F, interleukin 17F; SDF-1α/CXCL12, stromal-cell-derived factor-1; BCA-1/CXCL13, B cell-attracting chemokine 1; GM-CSF, granulocyte macrophage colony stimulating factor; SPA, scintillation proximity assay; SEC, size exclusion chromatography; IB, inclusion bodies; PEG, polyethylene glycol; DTT, dithiothreitol; IPTG, isopropyl-β-D-thiogalactopyranoside; TFA, trifluoroacetic acid; FCS, fetal calf serum; PI3K gamma, phosphatidylinositol-3 kinase gamma.

1046-5928/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2010.09.008

refolded proteins [6]. Other methods described to increase refolding yields include pulse renaturation processes [9], size exclusion chromatography and adsorption chromatography [10]. However, identifying the optimal refolding conditions can be difficult and time consuming.

Several fractional factorial refolding screens (Novagen iFOLD™, Hampton Research Foldit™, and AthenaES QuickFold™) are available commercially [11]. Despite this, several difficulties still remain. The screens only identify additives affecting the refolding, but optimal concentrations and cooperative interactions between additives are not determined. Furthermore, the method for determination of refolding yields is frequently not well defined and/or unreliable.

Different publications have shown that physical parameters such as pH, ionic strength, and temperature may have a tremendous effect on the efficacy of the *in vitro* folding process [12]. Aggregation is presumably the major process competing with correct refolding. Therefore a logical strategy to improve refolding yield is to prevent aggregation by adding low molecular weight compounds that interfere with unwanted protein–protein interactions.

A wide range of chemical additives have been described to prevent or reduce misfolding of proteins during the refolding process. L-Arginine is one of the most widely used additives for refolding of proteins. Its effects as enhancer of *in vitro* protein refolding were discovered a decade ago [1]. L-Arginine, which contains a guanido group, does not destabilize the native folded structure to the same extent as guanidine chloride, and at a concentration between 0.2 and 1 M, frequently increases the renaturation yield significantly [13]. An additional mechanism of L-arginine as suppressor of aggregation was recently proposed [14].

Polyethylene glycol (PEG) is another additive frequently used to improve correct structure formation by inhibiting aggregation [15]. Several studies have suggested that the success of PEG as a refolding additive depends on maintaining a specific molar ratio of PEG to protein [16]. Salt concentration plays an important role for the solubility of the protein during refolding. Mostly, NaCl or KCl is used at concentrations between 0 and 500 mM. For several proteins, nonionic as well as ionic and zwitterionic detergents have been found to have a favorable effect on the renaturation process [17]. Very efficient protein folding has been achieved by using a mixture of detergents (e.g. Triton X-100 or lauryl-maltoside) and phospholipids [18]. Non-detergent sulfobetaines (NDSBs) have shown to have a positive synergistic interaction with reductants and to decrease protein aggregation [19].

Proteins that possess disulfide bonds present additional challenges, since the folding of the protein is dependent upon correct disulfide bond formation. Disulfides formed between the wrong residues lead to aggregated or misfolded proteins. A low concentration of EDTA is frequently recommended to prevent metal-catalyzed air oxidation of cysteines, which could result in wrong disulfide bridge formation [5]. Addition of a mixture of reduced (RS^-) and oxidized (RSSR) forms of low molecular weight thiol reagents such as glutathione (GSH/GSSG), cysteine/cystine, cystamine/cysteamine and dithiothreitol (DTT/ODTT) usually provides the appropriate redox potential to allow formation and reshuffling of correct disulfide bridges [20]. The mechanism of disulfide bond formation has been treated by Gilbert [21]. The optimal ratio of the redox mixture in the refolding buffer is protein-specific. The ratio of GSH to GSSG should correspond to a redox potential (E'_0) that is compatible with protein disulfide formation [22]. For refolding experiments, various ratios of reduced to oxidized reagents at 0.2–5 mM are frequently used.

Protein concentration has a major impact on the refolding yield and is one of the key factors for successful renaturation. Because of the increased risk of aggregation at high protein concentrations, renaturation often has to be performed at a high dilution [23]. However, refolding at low concentrations is generally not economical for large scale protein production. Thus, many proteins can be folded at concentration near 1 mg/ml under optimized conditions [24,25].

Recently, the design of experiments (DOE) approach was used to optimize the refolding of a human cytokine fusion protein from *E. coli* inclusions bodies [26]. The DOE approach allows for the evaluation of the effect of additives over a wide concentration range, as well as their interactions, in a very efficient manner. This tool was successfully used to improve the refolding yield by 57% for the protein progenipoietin-1 [26]. By using this approach, interactions between the two additives cysteine and urea were identified, which could not be detected with an empirical refolding process.

There is no rational method for predicting conditions that will promote the refolding of any given protein. In this study, we have developed a generic 96-well refolding screen combined with a protein specific DOE-based focused secondary screen. The refolding yield was determined by a generic RP-HPLC analytical method for identification of the optimal refolding conditions. To set-up and validate the refolding screen, five different proteins were used, interleukin 17-F (IL-17F) which is a 30 kDa disulfide-linked homodimer, stromal-cell-derived factor-1 (SDF-1α/CXCL12), an 8-kDa CXC chemokine, B cell attractant chemokine 1 (BCA-1/CXCL13), anaphylatoxin C5a, a small protein of 74 residues [27–30] and granulocyte macrophage colony stimulating factor (GM-CSF), a 15 kDa cytokine [31,32].

Materials and methods

Materials and reagents

Tris (hydroxymethyl) aminomethane, D-sucrose, L-arginine, tri-sodium citrate, citric-acid, sodium chloride, polyethylene glycol 3300, and dithiothreitol (DTT) were obtained from Fluka. Reduced glutathione (GSH) and oxidized glutathione (GSSG) were purchased from Acros Organics. Zwittergent 3-14, ethanolamine and NDSB-256 were from Calbiochem. Guanidium hydrochloride was purchased from ABCR GmbH & Co. Source 30 RPC chromatography media, columns and the AKTA Explorer 100 equipped with an air detector were from GE-Healthcare Lifesciences. SE-HiCap Fractogel (M) was supplied by Merck KGaA. Multi Screen Deep Well Solvinert low binding Hydrophilic PTFE plates were from Millipore. The Cornerstone Design of experiment software (version 4.2.2) was purchased from P&P Informations technologie GmbH.

Protein expression and preparation of inclusion body extracts

All proteins described in this paper were expressed in *E. coli* and extracts prepared following the protocol below. The expression vectors were derived from the pET30 backbone vector (Novagen) and carried the gene of interest downstream of the T7/lac promoter. The plasmids were transformed into competent BL21 [DE3] *E. coli* cells. 5 L of culture medium containing 0.1 mM kanamycin were inoculated and, after induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at OD_{600} = 10, grown for 3.5 h at 37 °C. The cells were harvested and centrifuged at 10,000×*g* for 60 min at 4 °C (RC5C centrifuge, Sorvall instruments). Wet cell paste (100 g) was suspended in 500 ml 100 mM Tris–HCl buffer pH 8.0 containing 20 mg/L of DNase. Cells were broken by three passages through a French Press at 1600 bars. The suspension was centrifuged at 10,000×*g* for 60 min at 4 °C (RC5C centrifuge). The inclusion body pellet containing the recombinant protein was washed with 500 ml water two times, solubilized in 500 ml 0.1 M Tris/HCl, pH 8.0, containing 6 M guanidine/HCl and 10 mM