How additives influence the refolding of immunoglobulin-folded proteins in a stepwise dialysis system

Spectroscopic evidence for highly efficient refolding of a single-chain Fv fragment

Mitsuo Umetsu,1,2 Kouhei Tsumoto,1,* Masaki Hara,1 Kumar Ashish,1 Shuichiro Goda,1 Tadafumi Adschiri,2 and Izumi Kumagai1,*

1Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, Aobayama 07, Aoba-ku, Sendai 980-8579, Japan

2Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, 2-1-1, Katahira, Aoba-ku, Sendai 980-8577, Japan

*Author to whom correspondence should be addressed.

IK, Tel: +81-22-217-7274; fax: +81-22-217-6164; e-mail: kmiz@mail.cc.tohoku.ac.jp;

KT, Tel and fax: +81-22-217-7276; e-mail: tsumoto@mail.cc.tohoku.ac.jp

Running Title: How additives influence the refolding of immunoglobulin-folded proteins
Abbreviations used: bis-ANS, 4,4’-dianilino-1,1’-binaphthyl-5,5’-disulfonic acid dipotassium salt; β-ME, 2-mercaptoethanol; CD, circular dichroism; C\textsubscript{L}, light chain of immunoglobulin constant regions; DTNB, 5,5’-dithiobis(2-nitrobenzoic acid); FT-IR, Fourier transform infrared; F\textsubscript{v}, fragment of immunoglobulin variable regions; GSH, reduced form of glutathione; GSSG, oxidized form of glutathione; GuHCl, guanidine-hydrochloride; HyHEL-10, mouse anti-lysozyme monoclonal antibody; IPTG, isopropyl β-d-thiogalactopyranoside; MALDI-TOF, matrix-assisted laser desorption/ionization–time; PAGE, polyacrylamide gel electrophoresis; SA, Sinapic acid; scF\textsubscript{v}, single-chain F\textsubscript{v} fragment; SDS, sodium lauryl sulfate; TFA, trifluoroacetic acid; V\textsubscript{H}, heavy chain domain in F\textsubscript{v}; V\textsubscript{L}, light chain in F\textsubscript{v}
Abstract

The gradual removal of the denaturing reagent guanidine-HCl (GuHCl) using stepwise dialysis with the introduction of an oxidizing reagent and L-arginine resulted in the highly efficient refolding of various denatured single-chain Fv fragments (scFvs) from inclusion bodies expressed in *Escherichia coli*. In this study, the influence of the additives on the intermediates in scFv refolding was carefully analyzed on the basis of the stepwise dialysis, and it was revealed that the additive effect critically changes the pathway of scFv refolding. Circular dichroism and tryptophan fluorescence emission spectroscopies demonstrated that distinct secondary and tertiary structures were formed upon dialysis from 2 M GuHCl to 1 M GuHCl; and 4,4’-dianilino-1,1’-binaphthyl-5,5’-disulfonic acid dipotassium salt (bis-ANS) binding analysis indicated that the addition of L-arginine to the stepwise dialysis system effectively stabilized the exposed hydrophobic area on the scFv. Quantification of the free thiol groups in the scFv by means of Ellman’s assay revealed that there was a particular stage in which most of the free thiol groups were oxidized and that adding an oxidizing reagent (the oxidized form of glutathione, GSSG) at that stage was important for complete refolding of the scFv. The particular stage depended on the nature of the refolding solution, especially on whether L-arginine was present. Spontaneous folding at the 1 M GuHCl stage resulted in a structure in which a free thiol group accessed to the proper one for correct disulfide linkage; however, the addition of L-arginine resulted in the formation of a partially folded intermediate without disulfide linkages. Mass spectrometry experiments on alkylated scFv were carried out at each stage to determine the effects of L-arginine. The spectroscopic studies revealed two different pathways for scFv refolding in the stepwise dialysis system, pathways that depended on whether L-arginine was present. Controlled coupling of the effects of GSSG and L-arginine led to the complete refolding of scFv in the stepwise dialysis.
Introduction

The internal disulfide linkage in the immunoglobulin fold is of particular importance for most proteins in the immunoglobulin superfamily because this linkage has a critical influence on the stability of these proteins (1,2). The linkage, which is highly conserved, connects two β-sheets in a sandwich structure (3–6). In the case of antibody molecules, Goto and Hamaguchi were the first to analyze intrachain disulfide-bond formation in the constant fragments of the immunoglobulin light chain (C_L), demonstrating the contribution of the disulfide linkage to the stabilization and folding of the C_L domain (1,7). For the variable fragment in antibody molecules (Fv), which is composed of heavy and light chain domains (V_H and V_L, respectively), the immunoglobulin fold is known to be partially and irreversibly denatured to an aggregate under reducing conditions (2); and, furthermore, the instability of the Fv that is missing its disulfide linkages has been proved by natural antibodies missing a disulfide linkage (8) and replacement of the essential cysteinyl residue by serine residue (9). Recently, the folding kinetics and aggregation behavior of the single-chain Fv fragment (scFv) after removal of the disulfide linkages have been analyzed by fluorescence spectroscopy and a combination of mass spectrometry and an H/D exchange technique (10,11). In the folding of the disulfide-containing scFv in vitro, the presence of the disulfide linkage accelerates the independent folding of the V_H and V_L domains, and then the domains are associated with each other to form the correct quaternary structure. For the scFv with a domain containing no cysteinyl residues, the disulfide-free domain refolds more slowly than the disulfide-containing domain in the same scFv, and refolding requires an interface on the domain with a native disulfide linkage to form a functional structure (10). It has also been noted that when V_H folds prior to interaction with the V_L domain, the native structure results; whereas when the stable V_L domain is formed before V_H, aggregate formation is likely (10).

Antibody molecules are one of the key proteins for diagnostics in medicine and various
researches in the field of proteomics (12,13). Bacterial expression systems for recombinant proteins, especially systems using Escherichia coli, have been utilized to produce tailor-made antibody molecules such as scFv and Fab fragments (14,15). However, the expression of antibodies in E. coli is usually limited to the periplasm because the reducing condition in the cytoplasm rules out the formation of disulfide linkages which is crucial for the immunoglobulin fold. Furthermore, the secretory expression of most scFv and Fab fragments involves insoluble aggregated forms called inclusion bodies in the periplasm (14,16), and some antibody fragments cannot be expressed even as inclusion bodies (17). Consequently, a few other systems have been developed for the production of functional antibodies from E. coli: for example, direct production of functional scFv in the cytoplasm by the coexpression of disulfide-bond chaperones (18), and refolding of recombinant antibody fragments from inclusion bodies in vitro (19).

The coexpression system is a recent development, but the refolding approach has been utilized for a decade to renature scFv and Fab fragments from their aggregated forms (2,19–22). The general refolding system is based on dilution of denatured protein in a refolding buffer, and thiol–disulfide interchange reaction is applied to the system in order to increase the yield of refolding. However, the dilution method does not always result in a high refolding yield, and a more efficient method is needed.

Recently, we proposed an improved refolding system for completely unfolded antibody molecules with reduced cysteiny1 residues, a system based on the gradual removal of the denaturing reagent by means of stepwise dialysis (23). The refolding yield and the recovery of biological activity are greater than 95% in the case of various scFvs. Disulfide exchange and suppression of protein aggregation are controlled by the addition of the oxidized form of glutathione (GSSG) and L-arginine at appropriate stages (23). The high refolding yield, however, has not yet been evaluated in the view of physical and chemical studies, which are critically important for elucidating the stepwise folding process that leads to the completely functional scFv and for understanding how
GSSG and L-arginine influence the folding pathway.

In this paper, we present spectroscopic studies on the refolding process of the scFv from a mouse anti-lysozyme monoclonal antibody, HyHEL-10, in our stepwise dialysis system. We measured circular dichroism (CD) and tryptophan fluorescence emission spectra for the scFv solution at each stage to observe the conformational changes in the secondary and tertiary structures. We also used 4,4’-dianilino-1,1’-binaphthyl-5,5’-disulfonic acid dipotassium salt (bis-ANS) analysis to probe hydrophobic-area formation on scFv during the dialysis; and the analysis indicated a correlation between aggregation and the formation of an exposed hydrophobic area on scFv. Quantification of the free thiol groups in scFv, by means of Ellman’s assay and mass spectrometry experiments, indicated that the pathway of disulfide-bond formation in the presence of L-arginine was essentially different from the pathway in the solution containing only GuHCl. The secondary structure and the number of free thiol groups for an aggregated scFv in the stepwise dialysis system were estimated by Fourier transform infrared (FT-IR) spectroscopy and Ellman’s assay, and we discuss the factors leading to the formation of insoluble aggregates at each stage. The folding pathway for scFv depended strongly on the additives (GSSG and L-arginine), and the addition of these reagents at a particular stage enabled scFv to be refolded in high yield with recovery of biological activity.
Experimental Procedures

Materials. The plasmid containing the HyHEL-10 scFv (VL–linker–VH–(His)₆) cDNA was constructed by insertion of the NcoI–XbaI fragment of pUTN3 (24) into the NcoI–XbaI-digested plasmid pUT7 (25). Transformed E. coli BL21 (DE3) cells were incubated in 2 × YT medium at 28 °C, and expression of the HyHEL-10 scFv under the control of the T7 promoter was induced by adding 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). The harvested cells were centrifuged and suspended in 50 mM Tris-HCl (pH 8.0) buffer with 200 mM NaCl. After sonication, the suspension was centrifuged at 5800 g for 30 min at 4 °C. The pellet was suspended several times in the 50 mM Tris-HCl (pH 8.0) buffer with 4% Triton X-100 and 200 mM NaCl to remove nonspecifically adsorbed proteins, and the solution was centrifuged again at 5800 g for 30 min at 4 °C (23). After the removal step was repeated, the pellet was washed in water to remove Triton X-100, and then it was solubilized in the 50 mM Tris-HCl (pH 8.0) buffer with 6 M GuHCl and 200 mM NaCl. The solubilized scFv was refined by means of (1) a metal-chelate chromatography column that interacted with the histidine tag in scFv and (2) a gel filtration chromatography column (Sephacryl S-200).

Refolding of HyHEL-10 scFv by stepwise dialysis (23). The concentration of the unfolded HyHEL-10 scFv was adjusted to 7.5 μM in the 50 mM Tris-HCl (pH 8.0) buffer with 6 M GuHCl, 200 mM NaCl, and 1 mM EDTA, and the fragment was reduced by the addition of 2-mercaptoethanol (β-ME) at a 50-fold molar excess relative to the protein. After β-ME was removed by dialysis against the same Tris-HCl buffer without β-ME, the unfolded scFv was refolded by gradual removal of GuHCl by means of stepwise dialysis from 6 M to 0 M through 3 M, 2 M, 1 M, and 0.5 M by one of the following four methods (Figure 1): (a) the concentration of GuHCl was decreased by stepwise dialysis without any additives; (b) GSSG (50-fold molar excess relative to protein) was added at the 1 M and 0.5 M GuHCl stages; (c) 400 mM L-arginine was
added at the 1 M and 0.5 M GuHCl stages; and (d) GSSG (50-fold molar excess relative to protein) and L-arginine (400 mM) were added at the 1 M and 0.5 M GuHCl stages.

**Absorption, CD, fluorescence, and FT-IR spectroscopy experiments.** Absorption spectra were measured on a U-3000 spectrophotometer (Hitachi Ltd. Japan) in a 1-cm cuvette. CD spectra were measured on a J-720w spectropolarimeter (Jasco Inc. Japan): path length, 1.0 mm; bandwidth, 1.0 nm; resolution, 0.1 nm; response time, 8 s; scan speed, 5 nm/min; number of scans, 4. The spectra of the scFv solutions at each stage of the dialysis were directly measured, but the CD spectra of the solution containing aggregated particles were measured after the removal of the particles by centrifugation.

Fluorescence spectra were recorded for each stage with an RF-5300PC spectrofluorophotometer (Shimadzu Co. Japan) in a 1-cm quartz cuvette. Tryptophan fluorescence emission spectra were measured for the scFv solution at a concentration of 0.5 μM with an excitation wavelength of 280 nm, and bis-ANS fluorescence spectra were measured at an excitation wavelength of 395 nm 10 min after bis-ANS was mixed within scFv in a 1:10 ratio (bis-ANS, 0.38 μM; scFv, 3.8 μM). The scFv solutions containing aggregated particles were utilized after the removal of the particles by centrifugation.

Diffuse reflectance FT-IR spectra were obtained on a FTS-165 (Bio-Rad Lab. Inc. USA); the aggregated form of scFv was dried under vacuum and then mixed with KBr at about 1% (w/w).

**Ellman’s assay (26).** Ellman’s assay was applied to the scFv solutions at each stage and the scFv denatured by the dialysis from the scFv solutions at each stage to 6 M GuHCl solution (100 mM potassium phosphate–HCl, pH 5, 1 mM EDTA). Each scFv solution was diluted to 1.5 μM with the Tris-HCl buffer keeping each GuHCl concentration after aggregated particles had been removed by centrifugation, and then 50 μL of 10 mM 5,5’-dithiobis(2-nitrobenzoic acid) (DTNB) solution (50 mM potassium phosphate buffer, pH 7.0, 1 mM EDTA) was added to 1.25 mL of the diluted scFv solution. The absorbance at 412 nm was monitored after the addition of DTNB, and
the number of the free thiol groups in a scFv was calculated from the extinction coefficient of nitrothiophenol, which is generated by reaction of DTNB with a thiol group. The exact extinction coefficient of nitrothiophenol was estimated in each solution by reaction with the reduced form of glutathione (GSH) because the coefficient depended on the concentrations of GuHCl and L-arginine.

For quantification of the free thiol groups in the aggregated scFv in the stepwise dialysis, the aggregated particles were collected by centrifugation at each stage so that the insoluble particles aggregated at the different stage did not become mixed; e.g., the dialysis from 1 M GuHCl to 0.5 M was carried out after the removal of the aggregated particles formed at the 1 M stage. The aggregated scFv at each stage was dissolved in 6 M GuHCl at pH 5 before the assay was carried out.

**Alkylation of the scFv at each stage in the stepwise dialysis.** An iodoacetic acid solution (50 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA) was added to the scFv solution obtained from each stage (final concentrations: iodoacetic acid, 1 mM; scFv, 5 μM), and the pH was adjusted to 8 with a NaOH solution. After 30 min, the alkylated-scFv solution was dialyzed to 2 M GuHCl to remove the residual iodoacetic acid and uniform the concentration of GuHCl. For alkylation of the scFv in the denatured state, the scFv solutions from each stage were dialyzed to 6 M GuHCl (100 mM potassium phosphate–HCl, pH 5, 1 mM EDTA), and then the iodoacetic acid solution was added to the denatured scFv solutions.

**Mass spectrometry.** Mass spectra were measured on a REFLEX III matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometer (Bruker Analytische, GmbH, Germany) equipped with a nitrogen laser (337 nm). Sinapic acid (SA) was applied as a matrix and was dissolved to saturation in water/acetonitrile (2:1, v/v) containing 0.067% trifluoroacetic acid (TFA). Sample solutions from each stage were mixed with the SA-saturated solution in a 1:1 (v/v) ratio, and then 1 μL of the mixed solution was loaded onto the sample target. After
co-crystallization on the target, the crystals were washed 2 times with 2 μL of water containing 0.1% TFA to remove GuHCl and other salts. Analysis was performed in positive and linear modes with an accelerating voltage of 27 kV, and 200 scans were averaged. The spectra obtained were calibrated externally using the [M + H+] ions from two protein standards: cytochrome c from horse heart (m/z 12360.08) and bovine trypsin (m/z 23311.53).
Results

Aggregate growth in the stepwise dialysis system. We used the absorbance at 320 nm as a probe for suspension of scFv in the stepwise dialysis system (Figure 2). No suspension was observed in solutions at GuHCl concentrations higher than 2 M. A rapid increase in the 320-nm absorbance was observed at the 1 M GuHCl stage only for the simple GuHCl solution, that is, the solution without the GSSG and L-arginine additives (Figure 2); this increase indicated aggregation of scFv in the solution. In contrast, no scFv aggregates formed at the 1 M GuHCl stage in the solutions with GSSG or L-arginine (Figure 2). These additives suppressed aggregation in the dialysis from 2 M to 1 M. At the 0.5 M GuHCl stage, aggregation was observed in all the solutions except for the optimal GuHCl solution, which contained both GSSG and L-arginine (Figure 2), though the additives promoted the folding of scFv to some extent. The aggregation of scFv at the 0.5 M GuHCl stage of stepwise dialysis could be suppressed only by the addition of both GSSG and L-arginine.

Formation of secondary and tertiary structures. Figure 3 shows the CD spectra of scFv in the simple GuHCl solution at each stage of the stepwise dialysis. In the 6 M GuHCl solution, the CD spectrum of scFv showed no characteristic CD signals derived from secondary structure in the region of 215–225 nm and instead showed a rather large trough with a tail extending beyond 220 nm, which implies random structure (Figure 3a). Spectra with extended tails were also observed for the 3 M and 2 M GuHCl solutions, and no secondary structure was observed (Figure 3b,c). For the solutions of scFv in 1 M, 0.5 M, and 0 M GuHCl, the CD spectra were measured after insoluble scFv aggregates were removed by centrifugation. Although the CD spectra still indicated random structure in the 1 M and 0.5 M GuHCl solutions, the random component got small around 210–220 nm, and a new trough appeared at 218 nm (Figure 3d,e).
The 218-nm trough could be derived from a polypeptide with β-strand structure. In the 0 M GuHCl solution, in which scFv was folded, the 218-nm trough was also observed, and there was no CD signal for random structure (Figure 3f). Therefore, the secondary structure of scFv was mainly formed upon dialysis from the 2 M GuHCl solution, and the refolded scFv has a structure rich in β-strands. The CD spectrum of scFv at the 1 M GuHCl stage was also measured in the GuHCl solution with GSSG (Figure 4). Comparison of the spectra in the GuHCl solution with those in GuHCl–GSSG solution clearly showed that more β-strand structure was formed in the presence of GSSG. Because L-arginine has strong CD intensity in the measured region, we could not measure the CD spectra of scFv in the GuHCl solutions with L-arginine.

The folding of scFv in the dialysis system was also explored by means of the change in the tryptophan fluorescence emission maximum (Figure 5). The scFv has six tryptophan residues (four in V_H, and two in V_L), and all the residues except for Trp95 in V_H are buried inside the native structure (27). The tryptophan molecules in scFv showed an intense emission with a maximum at 353 nm in the 6 M GuHCl solution, and the emission was slightly blue-shifted in the 3 M and 2 M GuHCl solutions. These results indicate that the tryptophans were exposed to the hydrophilic solvent owing to the fact that scFv was unfolded at GuHCl concentrations higher than 2 M, even though the environment around tryptophans was a little changed by the decreasing of GuHCl concentration. In all the refolding solutions, a large blue-shift to 346 nm was observed as the GuHCl concentration changed from 2 M to 1 M, and the fluorescence intensity also decreased at this same stage (data not shown). The blue-shift and accompanying decrease of intensity were also observed in the dialyses from 1 M to 0.5 M GuHCl, and from 0.5 M to 0 M. Decreasing the GuHCl concentration to less than 2 M drastically changed the environment around the scFv tryptophan residues. These results, along with the CD spectra (Figure 3), indicate that the secondary and tertiary structures in scFv began forming at the same stage (1 M GuHCl stage). In dialysis systems in which GSSG or L-arginine was applied, the tryptophan molecules showed
emission changes similar to those observed in the simple GuHCl solutions, except for the fact that the tryptophan emissions were more blue-shifted at the 1 M and 0.5 M GuHCl stages in the solutions with GSSG or L-arginine (Figure 5). This result suggests that scFv folded more promptly in the GuHCl solution with GSSG or L-arginine than it did in the simple GuHCl solution.

In order to probe the hydrophobicity of scFv, we measured the fluorescence intensity of bis-ANS in the presence of scFv by adding bis-ANS to the scFv–GuHCl solution at each stage of the dialysis (Figure 6). The weak bis-ANS fluorescence in the 6 M GuHCl solution was equal in intensity to the fluorescence in the scFv-free GuHCl solution, which indicates that no hydrophobic area formed on scFv, owing to its completely unfolded stage (28,29).

In contrast, bis-ANS became highly fluorescent as the concentration of GuHCl was decreased to 3 M and 2 M (Figure 6). This result, which differed from the CD and tryptophan emission results (Figures 3 and 5), indicates that a hydrophobic area on scFv was formed without discernible secondary and tertiary structures even at the high GuHCl concentrations of 3 M and 2 M. The bis-ANS fluorescence was further enhanced at the 1 M GuHCl stage, during which the secondary and tertiary structures were forming. The hydrophobic area on scFv was most exposed to the solvent at the 0.5 M GuHCl stage in all the refolding solutions, except for the simple GuHCl solution, in which the bis-ANS had similar fluorescence intensity at GuHCl concentrations of 1 M and 0.5 M (Figure 6).

Comparison of the bis-ANS result with the observation of 320-nm absorbance indicates a connection between the extent of the hydrophobic area on scFv and aggregation of scFv at the 0.5 M GuHCl stage (Figures 2 and 6). The hydrophobic area on scFv did not expand from the 1 M GuHCl stage in the simple GuHCl solution, despite the fact that the bis-ANS fluorescence intensity at the 1 M GuHCl stage was comparable to the intensities in the other three solutions (GuHCl–GSSG, GuHCl–L-arginine, and optimal GuHCl solutions). The conformational change at the 0.5 M GuHCl stage may be insufficient in the simple GuHCl solution because there is nothing
to stabilize the hydrophobic area. At the 0 M GuHCl stage, at which GSSG and L-arginine were removed as well as GuHCl, the size of the hydrophobic area on scFv rapidly decreased in all cases (Figure 6). This result suggests that the \( V_L \) and \( V_H \) domains become stacked at the final stage.

**Quantification of the free thiol groups in scFv at each stage in the stepwise dialysis.**

The free thiol groups in untreated and denatured (6 M GuHCl) scFv were quantified at each stage by means of Ellman’s assay. The pH of the denaturing solution was adjusted to 5.0 in order to prevent formation of disulfide linkage in the denaturing process. The average number of free thiol groups in scFv at the 6 M GuHCl stage was 3.0 in spite of scFv having four thiol groups, indicating that a few thiol groups were naturally oxidized even in the 6 M GuHCl solution. When the GuHCl concentration was decreased from 6 M to 3 M, the number of free thiol groups in scFv decreased from 3 to 1.5 for the untreated scFv (Figure 7a), whereas the same number of free thiol groups were detected in the scFv denatured from the 6 M and 3 M GuHCl stages (Figure 7b). A similar discrepancy in the number of free thiol groups at the 3 M GuHCl stage was also observed at the 2 M GuHCl stage, although the difference was smaller than that at the 3 M GuHCl stage. This result reveals that the scFv structure in the 3 M and 2 M GuHCl solutions hinders DTNB access to the thiol groups, which supports the bis-ANS results suggesting that some structure was formed at the 3 M and 2 M GuHCl stages.

In the dialysis from 2 M GuHCl, oxidation of free thiol groups was significantly dependent upon the composition of the refolding solutions. In the GuHCl solution without additives (Figure 7a), the free thiol groups in the 2 M GuHCl solution were spontaneously oxidized by disulfide-bond formation only in the dialysis from 2 M to 1 M; and consequently, some free thiol groups remained even at the 0 M GuHCl stage. These results indicate that natural oxidization cannot make a sufficient number of disulfide linkages in scFv. For the GuHCl solutions with GSSG, Ellman’s assay was used after complete exclusion of the oxidizing reagent by dialysis. The assay showed that the addition of GSSG left few free thiol groups at the 1 M and 0.5 M GuHCl
stages (Figure 7b). We attempted to detect the binding of GSSG to scFv by mass spectrometry, but the scFv without GSSG was enhanced in the GuHCl–GSSG solution at the 1 M GuHCl stage (data not shown). Therefore, the addition of GSSG enabled sufficient disulfide-bond formation in scFv at the 1 M GuHCl stage.

Ellman’s assay also distinguished an essentially distinct disulfide-bond formation pathway in the presence of L-arginine (Figure 7c). In the presence of L-arginine, the results differed from those observed for the simple GuHCl and GuHCl–GSSG solutions (Figure 7a,b): the thiol groups showed no disulfide-bond formation behavior at the 1 M GuHCl stage, and instead the number of thiol groups increased and reached a maximum at the 6 M GuHCl stage (Figure 7c). Considering the tryptophan emission results (Figure 5), this result indicates that the folding of scFv occurred without disulfide linkage in the presence of L-arginine at the 1 M GuHCl stage. The increase in the amount of free thiol groups might be caused by a drastic conformation change from random structure to native-like secondary and tertiary structures. Instead, the free thiol groups were naturally oxidized at the 0.5 M GuHCl stage, but at the 0 M GuHCl stage as many free thiol groups remained as in the simple GuHCl solution (Figure 7a,c).

Disulfide-bond formation in the optimal GuHCl solution occurred by the same pathway as in the GuHCl–GSSG solution (Figure 7b,d), and mass spectrometry showed enhancement of scFv without GSSG at the 1 M GuHCl stage (data not shown). Therefore, the addition of GSSG can lead to sufficient disulfide-bond formation between free thiol groups in the stepwise dialysis even in the presence of L-arginine. Although we carried out SDS–polyacrylamide gel electrophoresis (PAGE) for all the refolding solutions in order to detect interchain disulfide-bond formation, we observed little scFv dimer or trimer in the gel (data not shown). Furthermore, attempt of digestion with endoproteinase Lys-C detects no incorrect disulfide linkages for the scFv at each stage (data not shown). Hence, the thiol groups in scFv formed a correct intrachain linkage in all cases in this study.
Reactivity of iodoacetic acid with scFv at each stage in the stepwise dialysis. In order to verify the results of the bis-ANS and Ellman’s assays, we measured the MALDI-TOF mass spectra of scFv alkylated by iodoacetic acid at each stage of the dialysis. The structure of scFv prevented access of DTNB to thiol groups in the 3 M and 2 M GuHCl solutions, and few of the thiol groups formed disulfide linkages at the 1 M GuHCl stage in the presence of L-arginine. Figure 8a shows the MALDI-TOF mass spectra for the scFv fragment that was alkylated by the addition of iodoacetic acid directly to the scFv solution at the 6 M, 3 M, and 2 M GuHCl stages. To crystallize each alkylated scFv sample under the same conditions, we dried the alkylated scFv on the target after GuHCl concentration was decreased to 2 M by dialysis. For the scFv alkylated in the 6 M GuHCl solution, we observed unmethylated, dimethylated and tetramethylated scFv fragments (I, II, and III in Figure 8a-1), and the band of the dimethylated scFv was enhanced. However, the band of the unmethylated scFv was more intense at the 3 M and 2 M GuHCl stages than at the 6 M stage, and the band intensity was inversely proportional to the GuHCl concentration at which scFv was alkylated (Figure 8a-2,a-3).

In contrast, the spectra of the scFv alkylated after denaturation from each GuHCl stage by the dialysis to the denaturing solution (6 M GuHCl, pH 5.0) were the same (Figure 8b). In light of the results of Ellman’s assay, this result can be explained by the existence of a structure around the thiol groups that hinders the access of the alkylating agent in the 3 M and 2 M GuHCl solutions.

MALDI-TOF mass spectra were also measured for scFv that had been alkylated after denaturation from the 1 M and 0.5 M GuHCl stages in the presence of L-arginine (Figure 9). Remarkably, the spectrum for the scFv alkylated after denaturation from the 1 M GuHCl stage containing L-arginine was identical to that from the 2 M GuHCl stage (Figure 9a,b), whereas the intensity of the unmethylated scFv fragment increased at the 0.5 M GuHCl stage (Figure 9b,c). This result, which is consistent with the results of Ellman’s assay, indicated that most of the disulfide linkages in scFv were formed not at the 1 M GuHCl stage but at the 0.5 M stage (Figure
Structural analysis of the aggregated form of scFv at the 1 M and 0.5 M GuHCl stages. FT-IR spectra were measured on the insoluble aggregated form of scFv at each stage (Figure 10). The aggregated scFv in the simple GuHCl solution at the 1 M GuHCl stage had a broad band at around 1685 nm that is derived from the amide groups in the β-turn structure; but a weak signal at around 1620 nm, which is assigned to the β-strand structure, was also observed. The aggregated scFv at the 1 M GuHCl stage had no β-strand structure. In contrast, the aggregated scFv in the GuHCl–GSSG solution at the 0.5 M GuHCl stage showed relatively intense bands at around 1620 nm (Figure 10), indicating the existence of β-strand structure in the aggregated scFv at the 0.5 M GuHCl stage.

We also measured the spectrum of the aggregated scFv in the GuHCl–L-arginine solution at the 0.5 M GuHCl stage and recognized β-strand formation as well (data not shown). Therefore, scFv formed an aggregate with native-like secondary structure when the GuHCl concentration was reduced from 1 M to 0.5 M in the GuHCl–GSSG and GuHCl–L-arginine solutions.

Table 1 shows the number of free thiol groups in the scFv aggregates at the 2 M, 1 M, 0.5 M, and 0 M stages. The aggregated scFv at the 1 M GuHCl stage in the simple GuHCl solution showed a value between those of the soluble fractions at the 2 M and 1 M stages in the simple GuHCl solution. This result indicates that the number of disulfide linkages in the aggregated scFv at the 1 M GuHCl stage did not provide sufficient stabilization at the GuHCl concentration of 1 M.

At the 0.5 M GuHCl stage, scFv formed an aggregate even in the presence of GSSG or L-arginine; however, there was a significant difference in the number of the free thiol groups in the aggregate. The aggregated scFv in the GuHCl–GSSG solution showed the same number of thiol groups as the soluble fraction at the 1 M and 0.5 M GuHCl stages, but the number of thiol groups in the aggregate in the GuHCl–L-arginine solution was midway between the values for the soluble fractions at the 1 M and 0.5 M GuHCl stages. These results indicate that the scFv in the
GuHCl–GSSG solution formed an aggregate independently of thiol group behavior, whereas, the aggregation of scFv in the GuHCl–L-arginine solution strongly depended on the structure of scFv without disulfide linkages. Therefore, the factors leading to the formation of aggregates at the 0.5 M GuHCl stage in the GuHCl–GSSG solution were essentially different from the factors in the GuHCl–L-arginine solutions.

We also measured the SDS–PAGE without a reducing reagent for all the scFv aggregates and detected no interchain disulfide linkages (data not shown). Hence, interchain disulfide linkages, which can lead to the aggregation, were not formed even in the aggregated forms in this study.
Discussion

The formation of disulfide linkages in the stepwise dialysis system. For the folding of scFv in the stepwise dialysis system, we have provided abundant spectroscopic data on the formation of secondary and tertiary structures, hydrophobic areas, and disulfide linkages; the folding processes are illustrated in Figure 11. When the GuHCl concentration was reduced from 6 M to 2 M, we observed a structure in which the thiol groups were inaccessible to DTNB and iodoacetic acid, even though the structure had no distinct secondary or tertiary structure (Figure 11a).

In the GuHCl solution without additives (i.e., the simple GuHCl solution), the CD spectrum showed β-strand formation at a GuHCl concentration of 1 M (Figure 3), and the formation of tertiary structure was also observed at the same stage in tryptophan fluorescence emission experiments (Figure 5). Interestingly, most of the thiol groups in scFv were naturally oxidized at the 1 M GuHCl stage, and no further oxidation of the residual thiol groups to disulfide linkages occurred at the stages after 1 M (Figure 7a). Goto and Hamaguchi’s regeneration study of the reduced constant fragment has demonstrated that the proximity of a pair of cysteinyl residues to each other is essential for the formation of a disulfide linkage and that the thiol groups are not easily oxidized in the interior of protein molecules (30).

In our study, disulfide-bond formation was strongly coupled to the formation of secondary and tertiary structures and resulted from the proximity of a native pair of cysteinyl residues in a native-like intermediate that formed at the 1 M GuHCl stage (Figure 11b). At the stages after 1 M, steric constrains prevented oxidation of the residual thiol groups to disulfide linkages. Therefore, the introduction of thiol-disulfide interchange reaction by GSSG at the 1 M GuHCl stage was best suited for making the correct disulfide linkages since each native pair of cysteinyl residues in the
VH and VL domains were in close proximity to each other at this stage.

The importance of the disulfide linkage in the immunoglobulin fold has been demonstrated in a large number of studies (2,10,19), and recent studies on disulfide-free antibody suggest that the stabilization of the disulfide-free antibodies requires amino acid exchange to compensate for the loss of folding stability upon cleavage of the disulfide linkage, although a functional disulfide-free scFv can be obtained (31–33). A mutation study on the scFv of the antibody hu4D5-8 against the extracellular domain of human epidermal growth factor receptor-2 indicates that the wild-type 4D5 scFv with a disulfide linkage in each of VH and VL domains has a higher midpoint of denaturation in the equilibrium transition than the variants lacking one or two disulfide linkages (10), which means that the 4D5 scFv with the native disulfide linkages is folded at a higher GuHCl concentration than the variants.

In our study, the trough derived from the β-strand structure that was observed in the CD spectrum in the presence of GSSG at the 1 M GuHCl stage was more distinct than that observed in the CD spectrum of the simple GuHCl solution (Figure 4). The change in the tryptophan fluorescence emission maximum also showed more scFv folding in the GuHCl–GSSG solution than in the simple GuHCl solution (Figure 5). The sufficient disulfide-bond formation in scFv caused by the addition of GSSG accelerated the formation of the immunoglobulin fold (Figure 11b). In the simple GuHCl solution at the 1 M GuHCl stage, scFv formed an aggregate with more free thiol groups than in the soluble scFv, whereas no aggregated scFv was formed in the GuHCl–GSSG solution at the same stage (Table1).

Ramm et al. have reported that the 4D5 scFv variant with no disulfide linkage in either domain (VL−VH−) is prone to aggregation and, furthermore, that the 4D5 scFv variant with a disulfide linkage in VL but not in VH (VL+VH−) is less stable than the VL−VH− variant (10). Therefore, in our study, the lack of aggregation of the HyHEL-10 scFv at the 1 M stage upon the addition of GSSG was due to the immediate formation of native disulfide linkages; the
simultaneous formation of the disulfide linkages in both domains might be also a factor in suppressing aggregation at the 1 M GuHCl stage. Possible factors leading to aggregation in the folding of scFv could be inappropriate disulfide-linkage formation and non-native association of the constituent polypeptide chains. Our results suggest that the native disulfide linkages inhibit non-native association of the constituent polypeptide chains at the 1 M GuHCl stage. In the simple GuHCl solution, the slow formation of the disulfide linkages by natural oxidation allowed non-native association, which led to the formation of aggregated particles (Figure 11b). The difference between the rate of disulfide-bond formation in the V<sub>L</sub> domain and that in the V<sub>H</sub> domain might also lead to aggregation at the 1 M GuHCl stage.

**Effect of L-arginine on the folding of scFv in the stepwise dialysis system.** In the GuHCl–GSSG solution, the essential factor that suppressed aggregation of scFv at the 1 M GuHCl stage was the immediate formation of native disulfide linkages by GSSG, linkages that stabilized a native-like intermediate state (Figure 11b). In contrast, few of the disulfide linkages in the GuHCl–L-arginine solution were formed at the 1 M GuHCl stage, despite the fact that there was no aggregation of scFv (Figures 2 and 7c). Although the CD spectra of scFv could not be measured in the presence of L-arginine, owing to the strong CD activity of L-arginine, the changes in the tryptophan fluorescence emission spectra in the dialysis system showed that the scFv with a partially folded structure but no disulfide linkages was stabilized at the 1 M GuHCl stage by L-arginine (Figure 11b).

L-Arginine has been used in many refolding systems to suppress protein aggregation (19,34–37), and the reagent is considered to be a labilizing agent that facilitates correct folding by destabilizing incorrectly folded structures (19,33–35). However, the theoretical basis for the mechanism by which L-arginine suppresses aggregation remains undetermined.

In the simple GuHCl solution without GSSG or L-arginine, the factor that led to scFv aggregation at the 1 M GuHCl stage was non-native association of the polypeptide chains, which
would be inhibited by immediate disulfide-bond formation in the $V_L$ and $V_H$ domains. Considering that scFv was partially folded but had no disulfide linkages and that the hydrophobic area on scFv increased in size at the 1 M GuHCl stage in the GuHCl–L-arginine solution (Figures 5, 6, and 7c), the addition of L-arginine may have inhibited association between the partially folded elements so that the thiol groups were unable to approach each other (Figure 11b).

At the 0.5 M GuHCl stage, in which exposure of the hydrophobic area on scFv was at a maximum (Figure 5), the aggregation of scFv in the GuHCl–GSSG solution could be suppressed by the addition of L-arginine (Figure 2). This result supports the proposal that the association between the partially folded elements was inhibited in the presence of L-arginine, since the correct disulfide linkages in scFv had already been formed by GSSG at the 1 M GuHCl stage (Figure 7c,d). In contrast, scFv was aggregated at the 0.5 M GuHCl stage in the GuHCl–L-arginine solution without GSSG. At this stage, most of the thiol groups in the soluble scFv were naturally but not completely oxidized to disulfide, and, furthermore, some free thiol groups remained in the aggregated scFv, as in the case of the folding in the simple GuHCl solution (Table 1). This thiol group behavior implies that the native pair of cysteinyl residues approached each other at the 0.5 M GuHCl stage in the GuHCl–L-arginine solution but that slow disulfide-bond formation by natural oxidation led scFv to form aggregates (Figure 11b). Ahn et al. have suggested that L-arginine supplements the role of GuHCl because they found that the optimal GuHCl concentration for refolding decreased when L-arginine was added (36). Therefore, the combination of GuHCl and L-arginine may be necessary for inhibition of the association between the partially folded elements. Without the formation of native disulfide linkages, the association in the scFv could not be inhibited even by the combination of 0.5 M GuHCl and 0.4 M L-arginine. Attempts to elucidate the mechanism by which the addition of L-arginine inhibits the association between the partially folded elements are in progress.

Highly efficient refolding of scFv in the optimal stepwise dialysis system.

We
achieved a high refolding yield with our stepwise dialysis system by adding GSSG and L-arginine at the appropriate stages (the 1 M and 0.5 M GuHCl stages, respectively). Our spectroscopic studies emphasize that the addition of an oxidizing reagent is necessary at the stage in which the native pairs of cysteinyl residues approach each other, owing to the formation of secondary and tertiary structures (at the 1 M GuHCl stage). In our previous attempts (23), the addition of GSSG at the 1 M GuHCl stage resulted in complete recovery of the biological activity of the soluble scFv at the 0 M GuHCl stage, whereas adding GSSG at other times resulted in recoveries of less than 60%.

From our present study, we can interpret the reduced recovery of activity as being the result of steric constraints that prevented GSSG from accessing the thiol groups. In the optimal GuHCl solution, GSSG and L-arginine were added simultaneously at the 1 M and 0.5 M GuHCl stages. Considering that the disulfide-bond formation pathway in the optimal GuHCl solution was similar to that in the GuHCl–GSSG solution (rather than to that in the GuHCl–L-arginine solution), the effect of GSSG was more immediate than that of L-arginine. Therefore, the formation of native disulfide linkages at the 1 M GuHCl stage by the addition of GSSG and inhibition of the association between the partially folded elements by the presence of L-arginine at the 0.5 M stage enabled complete recovery of biological activity and no aggregation in the optimal stepwise dialysis system. In addition, the current study shows that the presence of L-arginine delayed disulfide-bond formation (Figure 7c), which suggests that we should introduce GSSG at the 0.5 M GuHCl stage if L-arginine is already contained in the refolding buffer. Hence, the appropriate stage for adding GSSG is not necessarily the same as that for adding L-arginine; however, the timings of the introduction of GSSG and L-arginine are strongly coupled each other.

Recently, we succeeded in approving the refolding efficiency for interleukin-21 (IL-21) by the introduction of reducing reagent (GSH) at the 1 M and 0.5 M GuHCl stages which induces thiol–disulfide interchange reaction (38), which is different from the result of scFv that the introduction of reducing reagent decreases the refolding efficiency (23). Considering that IL-21
has an α-helix as a secondary structure (38), the possibility that the addition of GSH in the stepwise dialysis system improves the yield of correct folding is strongly dependent upon the native structure. The refolding process with thiol–disulfide interchange reaction in the stepwise dialysis system will be reported in the future.

**Conclusion**

Various spectroscopic data were used to elucidate the folding process of scFv from the denatured and reduced state in the stepwise dialysis system. Disulfide-bond formation occurred at the particular stage when the secondary and tertiary structures were formed, and the formation of correct disulfide linkages stabilized the native-like intermediate at the same stage. The addition of GSSG at that stage was most effective for making the native disulfide linkage because a native pair of cysteiny1 residues were in proximity to each other in the native-like intermediate. L-Arginine may have inhibited association between the constituent polypeptide chains or the partially folded elements in scFv, but that inhibition depended on the combination of GuHCl and L-arginine. The combination of 1 M GuHCl and 0.4 M L-arginine was especially good for stabilizing the partially folded intermediate with reduced thiol groups, and as a result, the free thiol groups formed disulfide linkages through a different pathway in the GuHCl–L-arginine solution than in the GuHCl–GSSG solution. The highly efficient refolding of scFv in the presence of GSSG and L-arginine (the optimal GuHCl solution system) was achieved by the formation of native disulfide linkages at the 1 M GuHCl stage and inhibition of association between the partially folded elements at the 0.5 M stage. In general, the refolding of scFv is carried out by dilution of the denatured protein in a refolding buffer with a redox reagent (GSH/GSSG) and L-arginine, and the timing of the addition of these additives has not been controlled. In this study, however, we demonstrated the importance of controlling the additive effects and correctly timing their introductions. We have found that
applying the optimal stepwise dialysis system is also effective for other immunoglobulin-folded proteins such as Fab and various receptors. This work will be reported in the future.

Acknowledgements

This work was supported by a Grant-in-Aid for the Original Industrial and Technology R&D Promotion Program from the New Energy and Industrial Technology Development Organization (NEDO) of Japan and by a Scientific Research Grant (no. 14780502) from the Ministry of Education, Science, Sports, and Culture of Japan.
References


<table>
<thead>
<tr>
<th></th>
<th>2 M</th>
<th>1 M</th>
<th>0.5 M</th>
<th>0 M</th>
</tr>
</thead>
<tbody>
<tr>
<td>simple GuHCl</td>
<td>no agg.</td>
<td>1.4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(2.3)</td>
<td>(0.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GuHCl–GSSG</td>
<td>no agg.</td>
<td>no agg.</td>
<td>0.3</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(2.3)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td></td>
</tr>
<tr>
<td>GuHCl–L-arginine</td>
<td>no agg.</td>
<td>no agg.</td>
<td>2.1</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(2.3)</td>
<td>(3.0)</td>
<td>(1.5)</td>
<td></td>
</tr>
</tbody>
</table>

The values in parentheses correspond to the number of free thiol groups in the soluble fraction; no agg. = no aggregation; — = could not be measured.
**Figure captions**

Figure 1. The stepwise dialysis system.

Figure 2. Change in the absorbance at 320 nm for the scFv solution at each stage in the stepwise dialysis: simple GuHCl solution (solid squares), GuHCl–GSSG solution (open triangles), GuHCl–L-arginine solution (solid triangles), and optimal GuHCl solution (open squares).

Figure 3. CD spectra of scFv in the simple GuHCl solution: (a) 6 M, (b) 3 M, (c) 2 M, (d) 1 M, (e) 0.5 M, and (f) 0 M. At the 0 M GuHCl stage, the concentration of scFv was low, owing to aggregation, and therefore the CD spectrum was measured after the solution was concentrated by ultrafiltration.

Figure 4. CD spectra of scFv at the 1 M GuHCl stage in the simple GuHCl solution (solid line) and the GuHCl–GSSG solution (dashed line).

Figure 5. Changes in the tryptophan fluorescence emission maximum for scFv at each stage in the stepwise dialysis: simple GuHCl solution (solid squares), GuHCl–GSSG solution (open triangles), GuHCl–L-arginine solution (solid triangles), and optimal GuHCl solution (open squares). The concentrations of scFv were adjusted to 0.5 μM.

Figure 6. Change of bis-ANS fluorescence intensity (fluorescence maximum, ~500 nm) in the scFv solutions at each stage in the stepwise dialysis: simple GuHCl solution (solid squares), GuHCl–GSSG solution (open triangles), GuHCl–L-arginine solution (solid triangles), and optimal GuHCl solution (open squares).
Figure 7. Change in the amount of free thiol groups in scFv at each stage in the stepwise dialysis: (a) simple GuHCl solution, (b) GuHCl–GSSG solution, (c) GuHCl–L-arginine solution, and (d) optimal GuHCl solution; untreated scFv (open circles) and scFv denatured in 6 M GuHCl solution (solid circles).

Figure 8. MALDI-TOF mass spectra of alkylated scFv in the simple GuHCl solution: (a-1) 6 M, (a-2) 3 M, and (a-3) 2 M; after denaturation from (b-1) 6 M GuHCl, (b-2) 3 M GuHCl, and (b-3) 2 M GuHCl; unmethylated fragment (I), dimethylated fragment (II), and tetramethylated fragment (III).

Figure 9. MALDI-TOF mass spectra of alkylated scFv after denaturation from GuHCl: (a) 2 M, (b) 1 M, and (c) 0.5 M in the GuHCl–L-arginine solution; unmethylated fragment (I), dimethylated fragment (II), and tetramethylated fragment (III).

Figure 10. FT-IR spectra of the aggregated particles of scFv in the simple GuHCl solution at the 1 M stage (solid line) and in the GuHCl–GSSG solution at the 0.5 M stage (dotted line).

Figure 11. The scFv folding pathway in the stepwise dialysis (a) from 6 M through 3 M to 2 M GuHCl and (b) from 2 M to 0 M GuHCl.
7.5 μM scFv
50 mM Tris-HCl (pH 8.0)
10 mM β-ME
1 mM EDTA
200 mM NaCl
6 M GuHCl

Dialysis against

6 M GuHCl → 3 M GuHCl → 2 M GuHCl → 1 M GuHCl → 0.5 M GuHCl → 0 M GuHCl

Addition
(a) simple GuHCl
(b) GuHCl–GSSG
(c) GuHCl–L-arginine
(d) optimal GuHCl

375 μM GSSG
400 mM L-arginine
375 μM GSSG
400 mM L-arginine

Fig. 1
Fig. 2

Absorbance at 320 nm vs. GuHCl concentration ([M]).
Fig. 3
Fig. 4
Fig. 5
Fig. 8
Fig. 10
Fig. 11

(a)

6 M GuHCl  →  3 M GuHCl  →  2 M GuHCl

(b)

2 M GuHCl  →  1 M GuHCl  →  0.5 M GuHCl  →  0 M GuHCl

oxidation

In the presence of L-arginine

inappropriate disulfide linkage

Aggregated form

non-native association of the partially folded elements

Aggregated form

stacking of domains

Aggregated form

Aggregated form

Aggregated form

Fig. 11
How additives influence the refolding of immunoglobulin-folded proteins in a stepwise dialysis system: Spectroscopic evidence for highly efficient refolding of a single-chain Fv fragment
Mitsuo Umetsu, Kouhei Tsumoto, Masaki Hara, Kumar Ashish, Shuichiro Goda, Adschiri Tadafumi and Izumi Kumagai

*J. Biol. Chem.* published online January 7, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M212247200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/early/2003/01/07/jbc.M212247200.citation.full.html#ref-list-1