BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

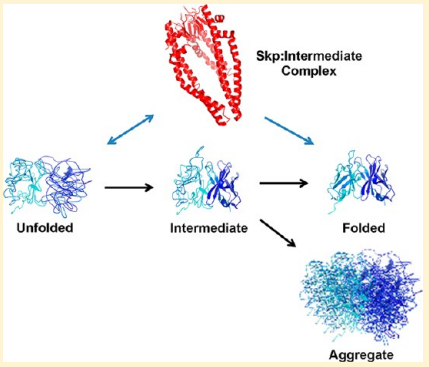
The Skp Chaperone Helps Fold Soluble Proteins *in Vitro* by Inhibiting Aggregation

Kevin C. Entzminger,[†] Christine Chang,[§] Ryan O. Myhre,[§] Katie C. McCallum,[‡] and Jennifer A. Maynard*[,§]

[†]Departments of Chemistry, [‡]Microbiology, and [§]Chemical Engineering, University of Texas at Austin, Austin, Texas 78712, United States

Supporting Information

ABSTRACT: The periplasmic seventeen kilodalton protein (Skp) chaperone has been characterized primarily for its role in outer membrane protein (OMP) biogenesis, during which the jellyfish-like trimeric protein encapsulates partially folded OMPs, protecting them from the aqueous environment until delivery to the BAM outer membrane protein insertion complex. However, Skp is increasingly recognized as a chaperone that also assists in folding soluble proteins in the bacterial periplasm. In this capacity, Skp coexpression increases the active yields of many recombinant proteins and bacterial virulence factors. Using a panel of single-chain antibodies and a single-chain T-cell receptor (collectively termed scFvs) possessing varying stabilities and biophysical characteristics, we performed *in vivo* expression and *in vitro* folding and aggregation assays in the presence or absence of Skp. For Skp-sensitive scFvs, the presence of Skp during *in vitro* refolding assays reduced aggregation but did not alter the observed folding rates, resulting in a higher overall yield of active protein. Of the proteins analyzed, Skp sensitivity in all assays correlated with the presence of folding intermediates, as observed with urea denaturation studies. These results are consistent with Skp acting as a holdase, sequestering partially folded intermediates and thereby preventing aggregation. Because not all soluble proteins are sensitive to Skp coexpression, we hypothesize that the presence of a long-lived protein folding intermediate renders a protein sensitive to Skp. Improved understanding of the bacterial periplasmic protein folding machinery may assist in high-level recombinant protein expression and may help identify novel approaches to block bacterial virulence.



Molecular chaperones are of fundamental interest to biotechnology, microbiology, and protein folding due to their innate ability to guide nascent proteins down productive folding pathways. In the complex, crowded environment of a cell, many proteins require chaperones to prevent off-pathway folding and aggregation events in order to achieve their final active form.[1] These chaperones fall into two general categories: foldases, which catalyze chemical reactions such as disulfide or peptidyl−prolyl bond isomerization, and holdases, which act to isolate aggregation-prone polypeptides. Enhanced understanding of protein folding allows high-level commercial expression of recombinant proteins,[2] contributes to development of therapeutics to treat protein misfolding diseases such as Alzheimer's and Parkinson's,[3] facilitates drug development,[4] and can even form the basis of anticancer vaccines.[5]

The seventeen kilodalton protein (Skp) is a molecular chaperone resident in the *Escherichia coli* periplasm, an aqueous compartment between the inner and outer membranes. Most proteins destined for secretion or transport to the outer membrane pass through this compartment, rendering it a major site for maturation of recombinant therapeutic proteins such as antibodies[6] and virulence factors such as such as the *Shigella* IcsA protein.[7] The periplasm is an oxidizing environment, and the only location within *E. coli* where disulfide bonds can spontaneously form. Here, Skp, in conjunction with the protease DegP, is genetically redundant with the well-known outer membrane protein (OMP) chaperone SurA. Skp is implicated in OMP folding by capturing nascent OMPs as they emerge from the Sec translocon and protecting them from the aqueous periplasm until transfer to the outer membrane BAM insertion complex.[8]

The functional form of Skp is a trimer composed of a central β-barrel-like structure from which three flexible α-helical "tentacles", which adopt a coiled coil-like topology, extend 60 Å and serve to give the protein a jellyfish-like appearance.[9,10] Skp interacts stably with a number of OMPs,[11,12] apparently based on a conserved pattern of exposed hydrophobic residues found in partially folded β-sheet-containing proteins,[12] driven by electrostatic and hydrophobic interactions,[13] as opposed to a specific amino acid sequence such as the Aro-X-Aro motif recognized by SurA.[14] NMR studies have directly observed Skp encapsulating a partially unfolded OmpA β-barrel domain, while the folded OmpA periplasmic domain protruded beyond Skp's tentacles in a folded form.[15] Structurally, Skp is

Received: March 31, 2012
Revised: May 30, 2012
Published: May 31, 2012

   dx.doi.org/10.1021/bi300412y | *Biochemistry* 2012, 51, 4822−4834

homologous to the eukaryotic and archaeal cytoplasmic chaperone prefoldin, which is comprised of six nonidentical subunits arranged in a jellyfish orientation and exhibits specificity for actin and tubulin monomers.[16]

Skp was initially identified as a chaperone with broader client protein specificity using a phage display approach in which a poorly expressed scFv-gpIII fusion was coexpressed with an *E. coli* chromosomal library.[17] Subsequently, Skp has been shown to enhance expression of numerous scFvs,[18−20] in addition to larger Fab fragments[21] and intact immunoglobulins.[6] More recently, Skp has been shown to aid in the transport and folding of bacterial virulence factors, such as the IscA protein of *Shigella*, for which a Skp deletion strain is unable to form plaques on a cell monolayer, and uropathogenic *E. coli*, which are unable to form the adhesive pili used in establishing infection without Skp.[7,22] Skp immunoprecipitation experiments with *E. coli* lysate revealed a broad spectrum of potential Skp OMP and soluble client proteins, biased for acidic proteins to complement Skp's basic nature (10.3 isoelectric point).[23]

However, little is known about the mechanism by which Skp enhances folding of soluble proteins or the basis for its client protein selectivity. Enhanced mechanistic knowledge would provide insight into Skp's native role in *E. coli* homeostasis and pathogenesis as well as aiding the engineering of client protein-specialized Skp variants for recombinant protein expression, analogous to similar efforts with GroEL.[24] Using a panel of scFvs, we performed *in vivo* expression experiments to classify these as Skp-sensitive and -insensitive proteins. Next, using purified scFv and Skp proteins, we performed *in vitro* aggregation and folding assays in the presence and absence of stoichiometric quantities of Skp. Skp inhibited aggregation but did not affect the folding rates of Skp-sensitive scFvs, while the rates of Skp-insensitive scFvs were unaffected. We then characterized the biophysical and biochemical stabilities of the Skp-sensitive and -insensitive proteins to identify biophysical indicators of client protein selectivity. Here, client proteins with folding intermediates, as determined using equilibrium urea denaturation and *a priori* calculated aggregation propensity, were more sensitive to Skp during refolding. By analogy to its role with OMPs, we propose Skp also acts as a holdase for soluble proteins by sequestering aggregation-prone folding intermediates.

EXPERIMENTAL PROCEDURES

Phage Display and Phage ELISA. To assess Skp's effects on scFv phage display, two antibody and one T cell receptor scFv, consisting of the two immunoglobulin variable regions linked by a flexible $(Gly_4Ser)_4$ peptide, were cloned into the pMoPac24 phagemid via directional *Sfi*I−*Sfi*I cloning.[19] This phagemid expresses the scFv fused to a *c*-terminal gpIII protein to facilitate M13 phage display and coexpresses the Skp chaperone in the same operon. Expression constructs are depicted in Figure 1. The scFvs used include the ovalbumin peptide-MHC specific DO11.10 single-chain T cell receptor [Roy, B., and Maynard, J. A., unpublished data], the protective-antigen-specific 14B7 scFv,[20] and the fluorescein-specific 4-4-20 scFv.[17] To provide control phagemids without Skp coexpression, mutagenic oligonucleotides were used to introduce three tandem stop codons in the Skp leader sequence, with the resulting phagemid termed pMoPac24SkpStop. The three scFvs were then cloned into pMoPac24SkpStop via directional *Sfi*I−*Sfi*I cloning as above. The fidelity of all constructs was confirmed by DNA sequencing at the University of Texas Core

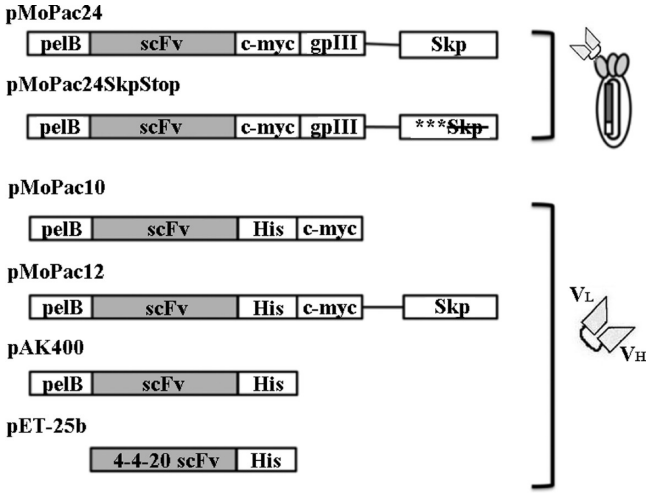


Figure 1. Plasmids used for scFv and Skp coexpression. To produce M13 phage displaying scFv as gpIII fusions, pMoPac24[19] and a variant, pMoPac24SkpStop, containing triplicate tandem stop codons to block Skp expression, were used. To express scFv protein for solubility analysis, pMoPac10, pMoPac12[19] (14B7 and 4-4-20), and pAK400[28] (DO11.10) were used. For large-scale expression and purification, pAK400 (14B7 and DO11.10) and pET-25b (4-4-20, Stratagene) were used. An scFv consists of two immunoglobulin domains (typically variable light, V_L, and heavy, V_H, chains) joined by a flexible 20 amino acid $(Gly_4Ser)_4$ linker. Notations: *pelB*, periplasmic leader sequence; *c-myc*, c-myc peptide tag for antibody detection; *gpIII*, minor M13 phage coat protein III; *Skp*, Skp gene with native periplasmic leader sequence coexpressed with a second ribosome binding site; ***, tandem stop codons to prevent Skp coexpression; *His*, histidine tag to facilitate purification.

Facility. The six phagemids were transformed into *E. coli* XL1-Blue, and scFv-bearing phage produced by M13 superinfection, harvested, and titered as described previously.[25]

To measure the activity of phage-displayed scFvs, phage ELISAs were performed in high-binding 96-well plates (Costar) coated with 1 μg/mL KJ1-26 monoclonal antibody (BD Pharmingen; to assess DO11.10 scFv activity), 16 μg/mL anthrax protective antigen (List Laboratories; 14B7), or 80 μg/mL BSA conjugated to FITC (Thermo Scientific, prepared according to manufacturer's protocols; 4-4-20) overnight at 4 °C in PBS. After a 2 h incubation with blocking buffer (5% nonfat milk in PBS), phage were serially diluted 2-fold in blocking buffer, allowed to equilibrate for 1 h, and washed in triplicate with PBS-0.05% Tween-20. Bound phage was detected with anti-M13 HRP (1:5000 dilution in blocking buffer; Sigma) for 1 h, plates washed as above, and signal developed with tetramethylbenzidine substrate (TMB; Thermo Scientific). The reaction was quenched with 1 N HCl and absorbance read at 450 nm on a SpectraMax M5 microplate reader (Molecular Devices). Data plots were prepared using the average of at least two experiments, including error bars equal to one standard deviation. GraphPad Prism 5 was used to fit data to a three-parameter logistic model and used for representation of all graphic data.

Analytical-Scale scFv Expression and Western Blotting. To compare yields of soluble periplasmic scFv expression, the scFv genes were *Sfi*I−*Sfi*I subcloned into the pMoPac10 and pMoPac12 plasmids, which do not and do coexpress Skp, respectively.[19] After transformation into *E. coli* BL21, small-scale expression experiments were conducted in 2 mL culture volumes. When the cultures reached an OD_{600} of 0.5, 1 mM

dx.doi.org/10.1021/bi300412y | *Biochemistry* 2012, 51, 4822−4834

isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce expression for 3.5 h at room temperature, as previously described.[18] Under these conditions, DO11.10 was entirely insoluble (data not shown). Since this protein has previously been expressed at high soluble levels in our hands with the related pAK400 plasmid [Roy, B., and Maynard, J. A., unpublished data], this construct was used instead of the pMoPac series. In this case, the *Skp* gene was expressed in *trans* from the compatible pBAD30 plasmid. To generate this construct, gene-specific oligonucleotides were used to append Skp with *Nhe*I and *Hind*III restriction sites. The resulting PCR product was digested with *Nhe*I and *Hind*III and ligated into similarly digested pBAD30 plasmid.[26] The pAK400/DO11.10 and pBAD30/Skp (or pBAD30 lacking the *Skp* insert as a negative control) were cotransformed into *E. coli* DH5α, with the scFv expressed in a similar manner as above, except that Skp expression was induced with 0.2% arabinose 30 min prior to DO11.10 induction.

An equal number of cells, based on an OD_{600} measurement, was pelleted, and the soluble periplasmic fraction was harvested by osmotic shock/lysozyme permeabilization of the outer membrane. The sphereoplasts were solubilized with urea to analyze the insoluble periplasmic/total cytoplasmic fraction, which contains both aggregated periplasmic material and improperly processed cytoplasmic material, as described previously.[27] To analyze scFv yields and solubility, 12 μL volume of each OD-normalized fraction was separated by 12% SDS-PAGE and transferred to a PVDF membrane. After serial incubation steps with blocking buffer and antipolyhistidine−HRP conjugate (Sigma) at 1:5000 dilution in blocking buffer, the membrane was washed with PBS-0.5% Tween. Signal was developed with SuperSignal West Dura Extended Duration Substrate (Thermo Scientific), with the resulting image captured on X-ray film. Though differing *E. coli* strains were used for DO11.10 vs 14B7 and 4-4-20, comparisons of fractions within each scFv tested remain valid.

Protein Expression and Purification. To have sufficient protein for *in vitro* studies, maltose binding protein (MBP) and the scFv genes were expressed and purified from the pAK400 plasmid[28] in *E. coli* BL21 in 500 mL culture volumes, as previously described.[18] Following Ni^{2+}-charged immobilized metal affinity chromatography (IMAC) purification of polyhistidine-tagged scFvs, proteins were further purified by analytical gel chromatography using an ÄTKA FPLC system (GE Healthcare) and a Superdex S75 (GE Healthcare) column with a PBS (pH 7.4) running buffer. This isolated the monomeric scFvs, which were used for all subsequent assays. Since pAK400 expression of 4-4-20 resulted in very poor yields (data not shown), the gene was transferred to pET25b(+) (Novagen) to produce cytoplasmic inclusion bodies. The gene was amplified with gene-specific primers, an *Nde*I and *Nhe*I gene fragment containing 4-4-20 was isolated and ligated into similarly digested pET-25b(+). After sequence verification and transformation into *E. coli* BL21(DE3), 4-4-20 was expressed and purified from inclusion bodies as described previously[29] with the following exceptions. Cells were lysed via three passes at 1500 psi through a French press (Thermo FA-078A) in a standard 35 mL cell (Thermo FA-032), and inclusion bodies were collected by centrifugation at 15 000 rpm for 40 min and 4 °C. The inclusion body pellet was washed three times with 20 mL of wash buffer (100 mM NaCl, 10 mM EDTA, 50 mM Tris, pH 8.0) with 0.5% Triton X-100 followed by three times with 20 mL of wash buffer to remove lipids and cellular debris. 4-4-20 was further purified from inclusion bodies using IMAC under denaturing conditions.

Skp was amplified from *E. coli* chromosomal DNA without its native leader sequence and TA-cloned into plasmid pET100D-TopoTA (Invitrogen) to facilitate soluble expression of Skp with an *n*-terminal hexahistidine tag in the cytoplasm. After sequencing and transformation into BL21(DE3), cells were grown at 37 °C in Terrific Broth (TB, Fisher) supplemented with 200 μg/mL ampicillin to an $OD_{600} \sim 0.3-0.6$. At this time, cells were induced with 1 mM IPTG and grown ∼18 h at 25 °C. Cells were harvested and resuspended in 50 mL of cold wash buffer (25 mM HEPES pH 7.4, 500 mM NaCl, and 40 mM imidazole), lysed via French press as above, and purified with IMAC. Size exclusion chromatography performed as above using a Superdex S200 (GE Healthcare) column confirmed purification of the trimeric species. Concentrations of all proteins were determined using the Micro-BCA protein assay kit (Pierce), and purity was confirmed by 12% SDS-PAGE prior to use in all other assays.

***In Vitro* Aggregation Assay.** To perform refolding studies under aggregation-prone conditions, proteins (including hen egg white lysozyme (Amresco) and scFvs) were individually adjusted to 100 μM in denaturation buffer (HBS with 8 M urea and 50 mM DTT at pH 7.4) and allowed to equilibrate for 6 h. To monitor aggregation, unfolded and reduced protein was rapidly diluted 100-fold into HEPES-buffered saline (HBS, 10 mM HEPES, 150 mM NaCl, pH 7.4) alone or HBS containing 1 μM maltose binding protein (MBP) or Skp at concentrations of 1, 3, and 6 μM in a black, clear-bottomed 96-well plate (Costar). Protein aggregation was monitored by absorbance at 360 nm for up to 30 min at room temperature with a SpectraMax M5. Experiments were repeated with at least three separate protein preps for each scFv, with the standard deviations from at least two replicates shown as error bars. Control experiments with varying amounts of protein (0.25−4 or 1−4 μM final concentration) were performed without Skp to demonstrate the concentration dependence of aggregation. To report the concentration dependence, the relative aggregation at each concentration was measured by dividing the absorbance at 360 nm after 15 min by the absorbance at 360 nm after 30 min.

***In Vitro* Folding Assay and ELISA.** To assess folding rates in the absence of aggregation, a similar experiment was performed, with lower protein concentrations. Here, urea-denatured protein (10 μM protein in HBS with 8 M urea and 0.1 mM DTT) was rapidly diluted 100-fold to a final concentration of 100 nM in HBS−20% glycerol no additional proteins, 100 nM BSA, or 100, 300, or 600 nM Skp in a black special optics 96-well plate (Costar) at room temperature. Under these low protein concentrations, scFv aggregation was minimal. Folding was monitored by changes in tryptophan fluorescence emission at 351 nm after excitation at 295 nm for 1.5 h in a SpectraMax M2. As Skp does not contain any tryptophan residues, and a control Skp-only solution did not record a signal, the presence of Skp is not expected to affect the data. Experiments were repeated with at least three separate protein preps for each recombinant protein studied. Folding rates were obtained by fitting the slope to a single exponential using GraphPad Prism 5.

To assess the specific activity of refolded protein, a sample from the 1.5 h time point was added to an ELISA plate coated with 1 μg/mL of the appropriate antigen. ELISAs were performed as above, with anti-polyhistidine HRP (Sigma) used

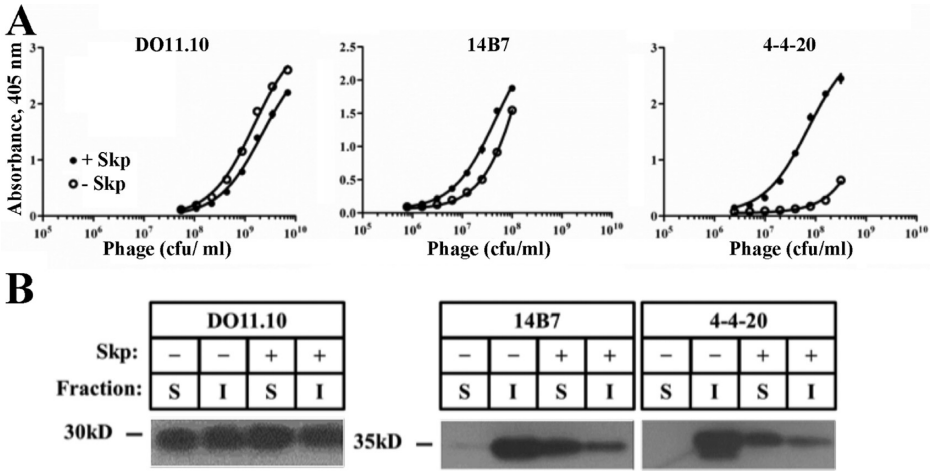


Figure 2. Skp coexpression differentially affects scFv activity and solubility *in vivo*. (A) ELISA results demonstrate the activity of scFvs expressed in a phage display format in the presence (●) or absence (○) of Skp co-overexpression. Phage expressing DO11.10 were produced at much higher levels, enabling titrations to start at a higher concentration. Error bars showing one standard deviation are included. (B) Western blot showing yield and solubility of scFvs in the presence or absence of Skp co-overexpression (S, soluble periplasmic fraction; I, insoluble periplasmic/total cytoplasmic fraction).

for detection. The specific activities were determined by fitting a standard containing untreated protein to a saturation binding curve with GraphPad Prism 5 and reporting signal from folding assay samples as a percentage compared to untreated protein at the same total protein concentration as follows: % activity = (absorbance at 450 nm from 100 nM refolded sample/ absorbance at 450 nm from untreated sample at 100 nM) × 100%. The quality of folded protein was also measured by circular dichroism, 1-anilinonaphthalene-8-sulfonic acid (ANS) binding, and analytical gel filtration to directly compare aggregate and monomer levels, but in general the protein concentration was too low for accurate measurements.

Protein Stability Analyses. To assess the stability of each protein to denaturant and probe the presence of a folding intermediate, urea denaturation curves were measured. Here, DO11.10 or 14B7 at a final concentration of 1 mg/mL was incubated with a series of urea concentrations at room temperature in a black 96-well plate (Costar). Tryptophan fluorescence was monitored by excitation at 280 nm while measuring emission from 310 to 380 nm with a 325 nm cutoff filter. The results are presented at wavelengths experimentally determined to maximize the difference between unfolded and fully folded protein (323 nm for DO11.10; 360 nm for 14B7), as described previously,[30] and are representative of at least two independent trials. The midpoints of unfolding were determined by fitting to a two-state unfolding mechanism (DO11.10) or a three-state unfolding mechanism (14B7) as described previously[31] using GraphPad Prism 5.

To assess the thermal stability of each protein, a small-scale fluorescent analytical technique was used.[32] Purified DO11.10, 14B7 (20 μL each at 10 μM), or an HBS buffer blank was mixed with Sypro Orange (1 μL of a 1:1000 dilution; Molecular Probes) and analyzed in triplicate in PCR tubes. Samples were heated in a 7900HT Fast Real-Time PCR System (Applied Biosystems) at increments of 0.5 °C from 20 to 90 °C. Results were analyzed using SDS.2 (Applied Biosystems) and standard deviations are reported from at least three replicates.

Modeling the Skp:scFv Interaction. To understand the effects of Skp on scFv aggregation and folding rates from a more mechanistic perspective, the experimental data were combined with possible reaction schemes and global fitting using Global Kinetic Explorer version 3.0.[33] In this kinetic simulation program, the user defines a reaction scheme (or model), each step of which is governed by forward and reverse rate constants. Next, the observables for each experiment are defined and adjusted with fitting parameters if necessary. Global Kinetic Explorer then compares time-dependent experimental data with simulated reaction traces, and dynamic simulation is used to vary each of the rate constants and to adjust parameters to globally fit the model to experimental data. Thirty-four models were evaluated by varying the number of folding pathway intermediates from 0 to 3, the number of intermediate aggregation steps from 0 to 3, and the Skp interaction from no interaction to interaction with unfolded or intermediate species. Control and Skp 1:1 kinetic traces from the folding and aggregation assays described above were fitted globally, with the most minimal model able to accurately simulate experimental data selected empirically for each protein tested. To prevent overfitting the data, the final folding and both aggregation steps were assumed to be irreversible, and other steps within the reaction pathways were linked or set at a constant value during fitting. This reduced the final number of fitted rate constants from ten to three: k_{fold} (rate of folding), k_{agg} (rate of aggregation), and k_{bind} (rate of Skp:scFv association). After fitting to the best model, all rate constants were well-defined except for DO11.10 and 4-4-20 k_{agg}.

RESULTS

scFvs Exhibit a Range of Skp Sensitivity during *in Vivo* Expression. The Skp chaperone was initially identified from an *E. coli* genomic library as an open reading frame whose overexpression increased binding activity of the poorly folding 4-4-20 scFv in a phage display format.[17] Since this initial report, Skp co-overexpression has enhanced production of a number of poorly folding or toxic scFvs, Fabs, and even whole IgGs in *E. coli*.[6,18,21] However, in our hands, we noticed that Skp did not increase expression levels of all scFvs, suggesting that Skp exhibits client protein selectivity and that this selectivity may provide insight into the mechanism by which Skp chaperones scFv folding. In order to explore Skp's role in scFv folding, we

first sought to identify a set of scFvs with varying Skp sensitivity. Of the eight scFvs tested, three that exhibited a range of Skp sensitivity *in vivo* were selected for further characterization: DO11.10, 14B7, and 4-4-20.

First, scFvs were expressed as phage-displayed molecules in the presence or absence of plasmid-borne Skp, with phage titers and phage ELISA used to assess the relative activity of scFv displayed on the phage particle. Here, the overall phage titers track roughly with the ELISA sensitivity, in that DO11.10 phage were produced at 2 orders of magnitude higher level than either 14B7 or 4-4-20. In ELISA data (Figure 2A), Skp appears to slightly decrease DO11.10 display (EC_{50} values of 2.2×10^9 with, versus 1.5×10^9 cfu/mL without Skp), which could reflect Skp interactions with the highly expressed DO11.10 inhibiting total DO11.10 yields. Skp co-overexpression modestly enhanced the sensitivity of the 14B7 ELISA (EC_{50} values of 4.8×10^7 with, versus 2.7×10^8 cfu/mL without Skp), while it dramatically enhanced sensitivity in the 4-4-20 ELISA, increasing the signal by about 2 orders of magnitude (EC_{50} values of 7.2×10^7 versus $>10^9$ cfu/mL without Skp), similar to previous observations.[17]

While useful for directed evolution experiments and readily assaying protein activity, phage display requires a fusion between the scFv and a phage coat protein, in this case, gpIII. This fusion alters folding pathways versus a soluble scFv, as scFv-gpIII folding occurs while anchored to the periplasmic face of the bacterial inner membrane. In order to more directly examine the effects of Skp on soluble scFv folding, we compared periplasmic expression of the three scFvs in the presence or absence of Skp co-overexpression to estimate the fraction of scFv present as soluble versus insoluble (and thus inactive) protein.

After small-scale expression, an equal number of cells were harvested and the soluble periplasmic and insoluble periplasmic/total cytoplasmic fractions isolated for analysis by Western blot, with scFvs detected via the c-terminal hexahistidine tag (Figure 2B). Similar to the DO11.10 phage ELISA results, Skp did not appear to affect the overall yields or soluble fraction of DO11.10. In contrast, while Skp had only a modest effect on 14B7 phage display, during periplasmic expression without Skp, the protein appears almost entirely in the insoluble fraction, but with Skp co-overexpression, over half of total protein appeared in the soluble fraction. Phage display of the 4-4-20 scFv was strongly dependent on Skp, accurately predicting Skp's effect on soluble expression, which also converts from entirely insoluble to over half soluble protein. Taken together, these data indicate that DO11.10 is relatively Skp-insensitive, 14B7 is moderately Skp-sensitive, and 4-4-20 is strongly Skp-sensitive, and these three proteins provide a set of scFvs to assess Skp's chaperone activities *in vitro*.

Skp Inhibits Aggregation of Sensitive scFvs *in Vitro*. Because Skp can suppress scFv aggregation during expression *in vivo*, we first assessed Skp's effect on aggregation during refolding of purified soluble protein *in vitro*. Skp, DO11.10, and 14B7 were expressed in *E. coli* and purified via sequential IMAC and size exclusion chromatographic steps to yield >90% pure protein (Figure 3). SDS-PAGE confirmed the expected size of all proteins, and size exclusion chromatography allowed isolation of monomeric scFvs. Periplasmic expression of 4-4-20 produced almost no active protein (data not shown). Instead, this scFv was expressed in the cytoplasm as inclusion bodies and purified by IMAC under denaturing conditions. Circular dichroism analysis was performed on the purified Skp

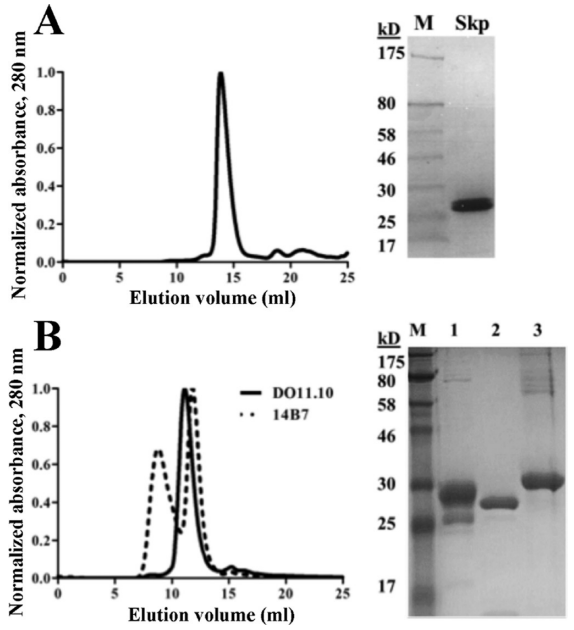


Figure 3. Purification of Skp and scFv proteins. (A) Analytical gel purification of Skp on a Superdex200 column; the elution volume corresponds to the size of a recombinant Skp trimer. The SDS-PAGE analysis shows pure monomer running between 25 and 30 kDa, which is slightly larger than the expected size for this construct (23 kDa including plasmid encoded tags), likely due to skp's very basic isoelectric point. *M*, molecular weight marker, with sizes indicated. (B) Analytical gel purification of scFvs on a Superdex75 column. The DO11.10 scFv expresses exclusively as a monomer; 14B7 purifies as both monomeric and dimeric species. The elution volume of 12 mL corresponds to a molecular weight of 26.9 kDa. The SDS-PAGE gel shows the purity of (1) DO11.10 purified by IMAC followed by FPLC, (2) 14B7 purified as DO11.10, and (3) 4-4-20 solubilized from inclusion bodies with 8 M urea and purified using IMAC.

trimer at multiple concentrations and agreed with previously published spectra,[12] indicating the protein is properly folded (data not shown). Client scFv proteins were first denatured with 8 M urea in the presence of DTT as a reducing agent and then rapidly diluted into an HBS folding buffer with varying molar ratios of Skp or a control (MBP), with the resulting aggregation monitored by light scattering at 360 nm. Lysozyme was used as a reference protein with known Skp sensitivity, exhibiting a dose−response reduction in aggregation as the Skp ratio was increased from 0.3:1 to 2:1 (Figure 4A), similar to previous reports.[10] Conversely, the robustly folding and GroEL-insensitive MBP^{34} was used as a Skp-insensitive negative control and did not exhibit any change in aggregation (Figure 4B). The relative amount of aggregation was reported as a percentage by dividing the aggregation signal (absorbance at 360 nm at the 30 min end point) in the presence of 1:1 ratio Skp:protein by the signal in absence of Skp at the same time point (Table 1).

Consistent with the *in vivo* results, the presence of Skp or MBP control during *in vitro* DO11.10 refolding did not affect aggregation of this protein (Figure 4C). The 14B7 scFv exhibited modest Skp sensitivity, with reduced aggregation in the presence of equimolar ratios of MBP or Skp, with further reductions in the presence of a 2:1 Skp:14B7 ratio (Figure 4D). The 4-4-20 scFv exhibited aggregation in buffer alone, which was dramatically reduced by the presence of Skp (Figure 4E). An additional set of experiments conducted without Skp or

dx.doi.org/10.1021/bi300412y | *Biochemistry* 2012, 51, 4822−4834

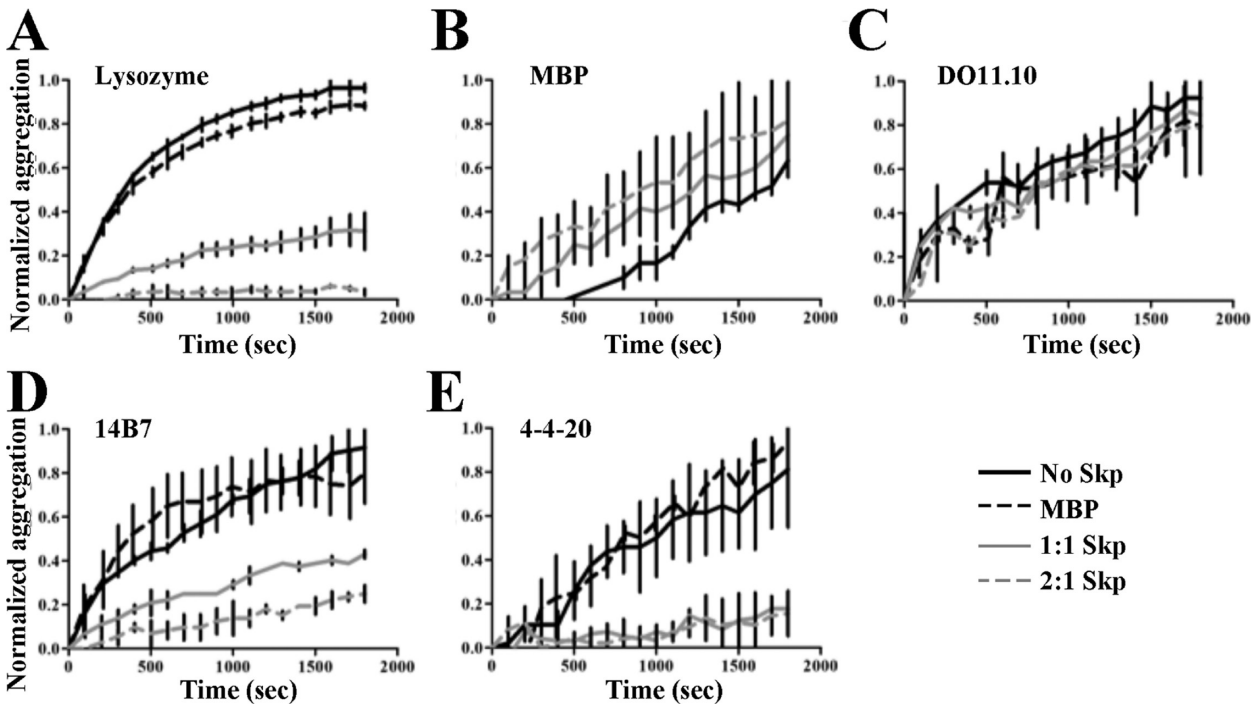


Figure 4. Skp differentially affects the aggregation of scFvs *in vitro*. Denatured protein was diluted into HBS, pH 7.4, and aggregation monitored by light scattering (360 nm) in the presence or absence of Skp or MBP to serve as a nonspecific control. (A) Positive control with lysozyme, which is known to be Skp-sensitive during *in vitro* refolding.[10] (B) Negative control with MBP, which has been previously shown to robustly fold without the aid of other chaperones.[34] (C) DO11.10 shows no change in aggregation in the presence of Skp, while (D) 14B7 and (E) 4-4-20 both show reduced aggregation in the presence of Skp. The relative amounts of aggregation at the final time point are in Table 1. Molar ratios (Skp:protein) are reported in terms of functional Skp trimer.

Table 1. Summary of Skp Effects on Client Protein Aggregation and Folding Rates

protein	% reduced aggregation with Skp[b]	obsd folding rate (10^{-2} min^{-1})[a]	specific activity, %[c] (refolding, %[a])
lysozyme	68 ± 9	0.259 ± 0.026	ND[d]
		(0.290 ± 0.067)	
DO11.10	8.3 ± 12	0.176 ± 0.012	18.7 ± 4.1
		(0.181 ± 0.017)	(16.3 ± 4.0)
14B7	53 ± 6	0.165 ± 0.026	102 ± 2.9
		(0.207 ± 0.020)	(102 ± 10)
4-4-20	78 ± 4	0.153 ± 0.039	ND
		(0.120 ± 0.046)	

[a]In the absence or (presence, one molar ratio) of Skp. [b]The percent reduction in aggregation by one molar ratio Skp as compared to no Skp sample at 30 min, calculated as $(A_{360,\text{no Skp}} - A_{360,\text{Skp}})/A_{360,\text{no Skp}}$. [c]Specific activity of samples was determined by ELISA at the end point of the refolding assay as a percentage compared to untreated protein. [d]ND = not determined.

MBP additives confirmed that aggregation of Skp-sensitive, but not Skp-insensitive, proteins is a function of the client protein concentration (Figure S1). Aggregation of lysozyme, 14B7, and 4-4-20 increases sharply as the concentration of protein after dilution is increased from 1 to 4 μM. Conversely, aggregation of MBP and DO11.10 exhibited no significant change in observed aggregation at the concentrations, as reported previously for MBP.[34] Thus, if Skp can act to reduce the concentration of free scFv during refolding, lower levels of aggregation would be expected. For the Skp-sensitive scFvs, these aggregation profiles demonstrate that Skp assists in protein folding by inhibiting aggregation but does not exclude additional roles, such as accelerated folding. We next sought to examine Skp's effects on folding under nonaggregating conditions.

Skp Does Not Accelerate scFv Folding *in Vitro*. Skp is naturally present in the bacterial periplasm, a compartment devoid of ATP and thus lacking the ability to refold non-native proteins over a thermodynamic barrier.[10] Here, folding and aggregation pathways compete for protein folding intermediates. By analogy to the GroEL chaperonin,[34] Skp could employ two distinct mechanisms to inhibit aggregation of soluble proteins. First, Skp could actively accelerate folding through electrostatic interactions or physical confinement. Alternatively, Skp could act as a holdase, passively preventing aggregation by sequestering exposed, aggregation-prone hydrophobic regions. The first mechanism inhibits aggregation by accelerating the folding rate of bound client proteins, while the latter only slows the aggregation rate. If aggregation is artificially inhibited *in vitro* by low protein concentration, we can distinguish between these mechanisms by observing whether Skp accelerates folding rates. Urea-denatured and reduced protein was rapidly diluted into buffer with or without Skp, with protein folding rates monitored by quenching of tryptophan fluorescence as these residues become buried within the protein core. Since Skp contains no tryptophan residues, its presence does not contribute to the observed signal. The final protein concentration used, 100 nM, was selected to minimize intermolecular aggregation so that the observed folding rate is not affected by changes in aggregation rates. Under these circumstances, if Skp acts primarily as a foldase, the folding rate would be expected to vary in the presence of Skp, while if it acts as a holdase, no change in the folding rate would be expected. Control proteins used in the aggregation and folding assays

dx.doi.org/10.1021/bi300412y | *Biochemistry* 2012, 51, 4822−4834

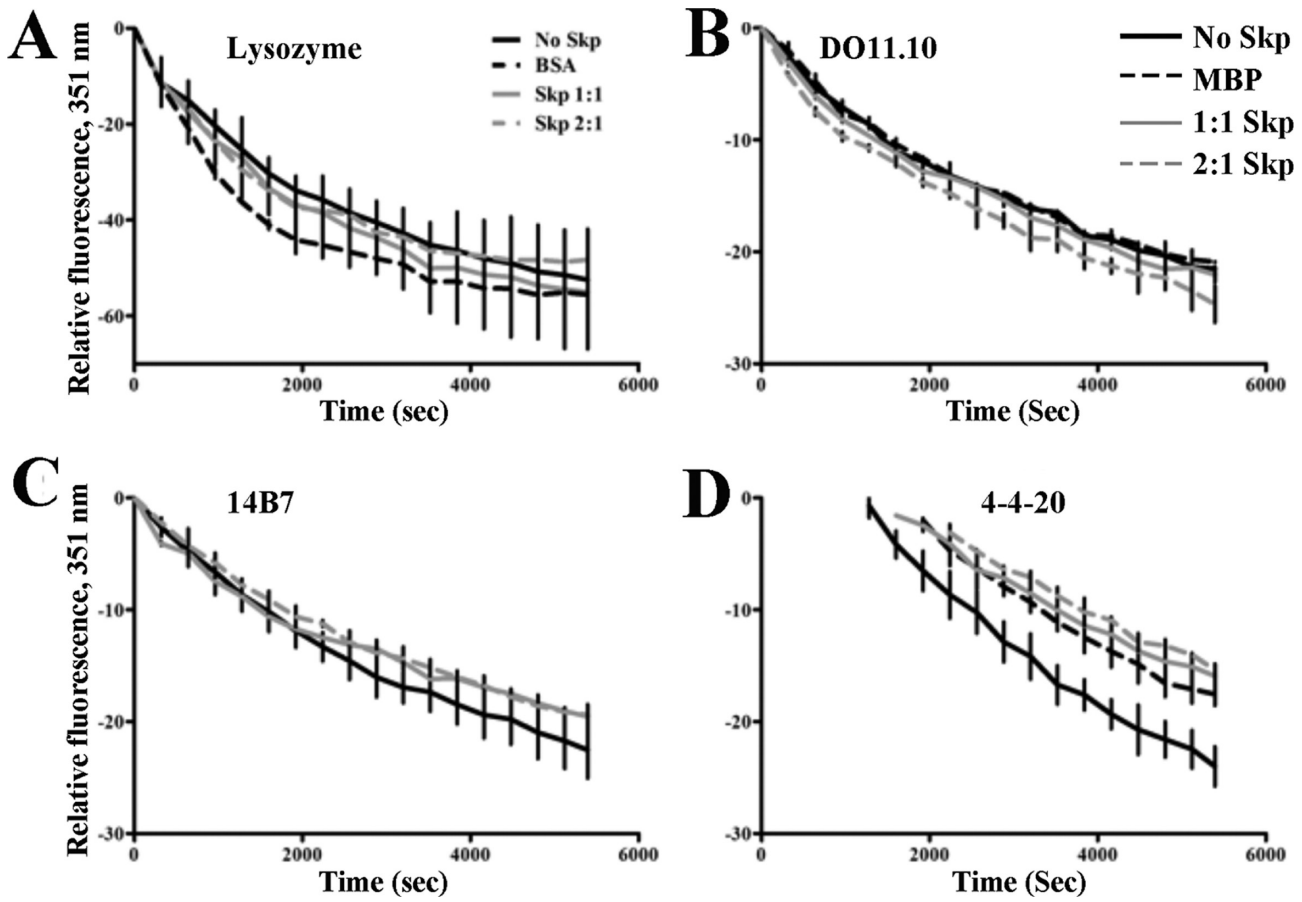


Figure 5. Skp does not affect the folding rates of scFvs *in vitro* under conditions that limit aggregation. Denatured protein was diluted into refolding buffer at a low concentration to minimize intermolecular aggregation, with folding monitored by tryptophan fluorescence in the absence or presence of Skp or BSA as a nonspecific control. (A) Control experiment using lysozyme as a client protein. (B) DO11.10, (C) 14B7, and (D) 4-4-20 folding rates are not significantly altered by the presence of Skp, although signal intensity varies somewhat. Folding rates were obtained by fitting the slope to a single exponential and are provided in Table 1. Molar ratios (Skp:protein) are reported in terms of functional Skp trimer.

(MBP or BSA, respectively) were chosen based upon which gave the most consistent results.

Folding experiments were conducted with protein alone or in the presence of a 1:1 or 2:1 molar ratio of Skp trimer:protein. This data were used to calculate folding rates by fitting fluorescence data to a single-exponential decay rate constant (Table 1). For the control protein lysozyme, all calculated folding rates were within error, indicating that Skp does not increase lysozyme's folding rate in the absence of aggregation (Figure 5A). Similarly, no differences in folding rates were observed for the Skp-insensitive DO11.10 and moderately sensitive 14B7, and all fluorescence traces are very similar (Figure 5B,C). Signal intensities vary among individual 4-4-20 samples, but all calculated folding rates are within error (Figure 5D). Absolute signals vary slightly between runs, but this does not affect the key data which is a function of the slope, or rate of fluorescence decay. The rates of folding in the presence of a BSA control or submolar (0.3:1) to supramolar (2:1) ratios of Skp did not significantly differ from the rate in the absence of Skp for all proteins (data not shown).

To provide a functional measure of protein folding, ELISAs were performed on refolded samples after 1.5 h. Here, the specific activities of samples refolded in the presence of 0:1 or 1:1 molar Skp ratios are presented as a percentage of activity measured from untreated protein at 100 nM. Both refolded DO11.10 and 14B7 samples were within error in all conditions (Table 1), consistent with the kinetic data. Notably, the 14B7 1:1 sample showed no increase in activity over the no Skp sample. The activity of 4-4-20 could not be determined because the signal was below the limit of detection. Since Skp does not affect scFv folding rates or total yields of refolded scFvs under nonaggregating conditions, we conclude its major role in soluble protein folding is to prevent aggregation via holdase-like activities.

Skp-Sensitive scFvs Have Folding Intermediates. To explain the variable sensitivity of different scFv proteins to Skp during *in vitro* refolding, we hypothesized that this sensitivity may correlate with biophysical or folding characteristics of the three scFv proteins. To explore this, we first sought to further characterize the folding pathways of these scFvs, in particular to determine whether the pathway includes a stable folding intermediate, which could appear similar to a partially folded OMP and interact with Skp in a similar manner. Using equilibrium urea denaturation, the DO11.10 denaturation curve fit a two-state unfolding model (Figure 6A). No intermediate was observed for this scFv, suggesting that any potential folding intermediates may be too transient or poorly populated for Skp recognition.

In contrast, 14B7 denaturation showed a clear plateau, and the data fit a three-state model containing a folding intermediate (Figure 6B). The intermediate is well-defined by multiple data points, suggesting the potential for Skp

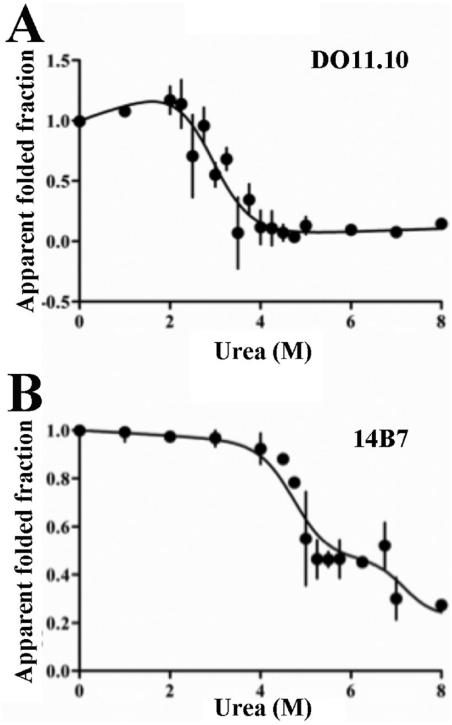


Figure 6. Urea equilibrium denaturation of scFvs were allowed to equilibrate in buffer containing 8 M urea, with the fraction of folded protein detected by tryptophan fluorescence. The fraction of folded scFv is normalized between one (native) and zero (unfolded) as reported previously.[31] (A) DO11.10 displays two-step unfolding with no detectible intermediate. Data were fit to a two-state model (black line), yielding $R^2 > 0.86$. (B) 14B7 unfolding reveals an intermediate. Data were fit to a three-state model (black line), yielding $R^2 > 0.90$. The midpoints of denaturation are provided in Table 2.

recognition of this highly populated intermediate. Because of the poor solubility of 4-4-20, denaturation experiments yielded poor results (data not shown). However, 4-4-20 has an exposed hydrophobic patch thought to be responsible for aggregation by inhibiting the solubility of folding intermediates. When the hydrophilic valine to aspartic acid substitution (V84D) is introduced into the 4-4-20 heavy chain or when its antigen binding complementarity determining loops are transferred from the native murine β sheet framework onto a more soluble human framework, aggregation is greatly reduced.[35,36] Lysozyme aggregation proceeds via amyloid fibril formation, with nucleation dependent on the β-domain,[37] and a folding intermediate has also been observed for hen egg white lysozyme.[38]

Skp-Sensitive scFvs Have Higher Predicted Aggregation Propensities. The ability to predict a protein's aggregation propensity *a priori* is a useful tool that can address aggregation in human disease and industrial protein production.[39] Of a number of available algorithms, Aggrescan was chosen for this study, as it is based on a large set of experimentally derived *E. coli* data in which the relative intracellular solubility of each amino acid is used as a measure of aggregation propensity.[40] Using primary sequence data, Aggrescan assigns these derived values to each individual amino acid within an input sequence, averages a score over a "sliding window" covering the sequence, and normalizes this value to the total number of amino acids, with a more positive value indicating a greater likelihood of aggregation. Aggrescan correctly ranked the relative extent of aggregation observed for all scFvs, and these predictions correlated with relative Skp sensitivities. In particular, it identified 4-4-20 as the most aggregation-prone protein in this study, while DO11.10 and lysozyme were the least (Table 2). Because of the predominantly β-sheet topology of scFvs, aggregation prediction algorithms may be more sensitive to this class of proteins. While lysozyme possesses a similar intrinsic aggregation propensity value as DO11.10, the aggregation profile generated, which is a plot of the aggregation propensity of each amino acid within a protein's primary sequence, was similar to that reported previously.[41] Analysis with the Betascan[42] and FoldAmyloid[43] algorithms, which use slightly different methods of calculating aggregation propensity, resulted in similar predictions (data not shown).

Skp Sensitivity Does Not Correlate with Other Biophysical Characteristics. We further attempted to explore the basis of Skp sensitivity by comparing additional biophysical characteristics of the proteins used in this study. Skp is a highly basic protein with an isoelectric point (pI) of 10.3, with previous reports indicating that the charge on a client protein influences its interactions with Skp.[23] Here, *in vitro* aggregation and refolding assays were performed at near-neutral pH, a condition in which both Skp and most client proteins will be positively charged based on their pI values (Table 2). Under these conditions, Skp has a strong effect on refolding of 14B7 and 4-4-20 scFvs. Moreover, increasing the NaCl concentration to shield potential electrostatic interactions between Skp and its client proteins did not reduce Skp's effect (data not shown). In this work, pI did not correlate with Skp sensitivity, indicating that charge−charge interactions may play a minor role relative to hydrophobic interactions.

Because Skp prevents aggregation, Skp may primarily interact with poorly stable proteins, as these may be subject to increased aggregation levels in the cell and *in vitro* assays. Accordingly,

Table 2. Biophysical Characteristics of Skp Client Proteins Used in This Study[a]

protein	M_w[b] (kDa)	pI[b]	T_m (°C)	midpoint of unfolding (M urea)	folding intermediate	intrinsic aggregation propensity
lysozyme	14.3	9.32	68.5 ± 0.5	ND[c]	yes[38]	−13.7
DO11.10	26.1	7.07	51.3 ± 0.5	3.05 ± 0.12	no	−13.4
14B7	26.1	8.81	68.3 ± 0.6	4.83 ± 0.11 6.74 ± 0.30	yes	−8.6
4-4-20	27.0	8.28	53.0 ± 0.1[61]	4.1[35]	yes[35]	−5.1

[a]Melting points (T_m) were determined using thermal denaturation, while the presence of folding intermediates was determined using urea denaturation. The aggregation propensity was calculated from primary amino acid sequence information using the Aggrescan algorithm; a higher number represents a higher propensity to form amyloid-like aggregates.[40] [b]Theoretical isoelectric point (pI) and molecular weight (M_w) computed using the ExPASy server (http://web.expasy.org). [c]ND = not determined.

dx.doi.org/10.1021/bi300412y | *Biochemistry* 2012, 51, 4822−4834

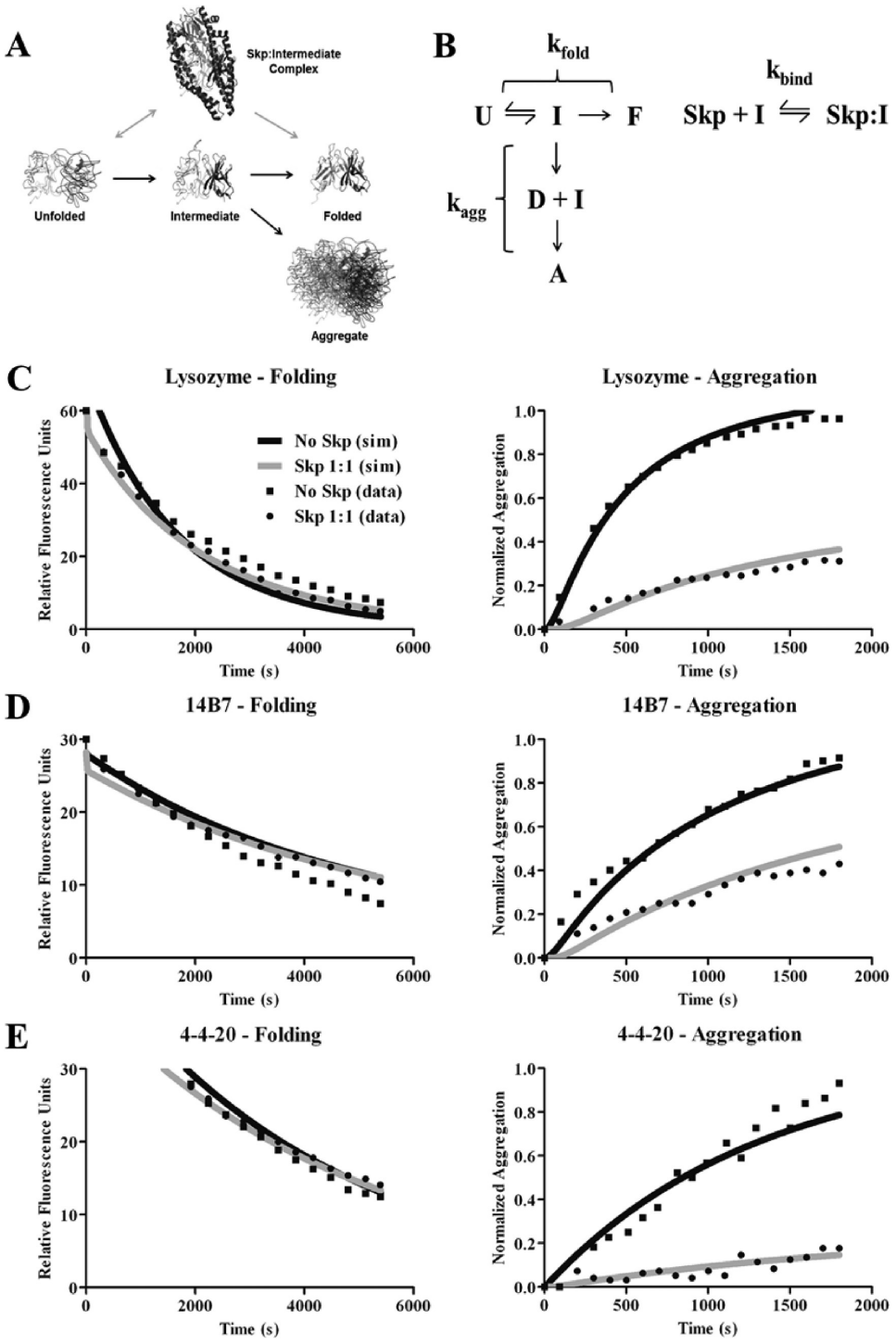


Figure 7. Proposed model depicting Skp's role in soluble protein folding. (A) Robust proteins proceed directly from an unfolded topology (U) to a short-lived, low-populated folding intermediate (I) before proceeding to the native state (F). For proteins with kinetically trapped or highly populated folding intermediates, Skp inhibits aggregate (A) formation through a Skp:intermediate complex (Skp:I), increasing the total amount of folded protein through kinetic partitioning. Image created using Pymol and PDB codes 3NN8 (scFv, blue) and 1U2M (Skp, red). (B) Model of Skp interaction with client proteins used to fit data. Species are labeled as in (A), in addition to a dimeric intermediate species "D." The number of parameters used to fit data was reduced by grouping rates (k) as shown to a total number of three rate constants. (C) Lysozyme, (D) 14B7, and (E) 4-4-20 folding and aggregation assays fit globally to the scheme in (B).

protein stabilities were measured by thermal and urea denaturation (Table 2). Although DO11.10 has the lowest T_m and exhibits poor stability in the presence of urea denaturant, with an unfolding midpoint of 3.0 ± 0.1 M urea, it was insensitive to the presence of Skp both *in vivo* and *in vitro*. In contrast, 14B7 is moderately Skp-sensitive but exhibits much higher stability, with a T_m of 68.3 ± 0.6 °C and two unfolding midpoints at 4.8 ± 0.1 and 6.7 ± 0.3 M urea, which may reflect sequential unfolding of the heavy and light chain domains. No clear correlation was observed between protein

dx.doi.org/10.1021/bi300412y | *Biochemistry* 2012, 51, 4822−4834

stability, as measured by thermal denaturation and Skp effect on aggregation.

A Proposed Mechanism for Skp Activity and scFv Specificity. Building upon the data presented here, we propose a model whereby Skp chaperones soluble protein folding by inhibiting the aggregation of folding intermediates (Figure 7A). In this model, stable proteins with poorly populated or short-lived intermediates such as DO11.10 will proceed directly to the native state, bypassing any interaction with Skp and exhibiting low levels of aggregation. However, proteins with long-lived, kinetically trapped or aggregation-prone intermediates, such as lysozyme, 14B7, and 4-4-20, are expected to interact with Skp via hydrophobic surfaces exposed in the non-native state. Skp may allow the intermediate to fold while isolated within the Skp cavity or Skp may act to reduce the concentration of aggregation-prone intermediates in solution, thereby reducing the rates of this higher-order aggregation process. For the Skp-sensitive proteins studied here, a lower concentration resulted in reduced aggregation in the absence of Skp (Figure S1). If Skp physically isolates partially folded intermediates from each other, collision and aggregation between their respective hydrophobic surfaces are prevented. This mechanism would increase the total amount of folded protein through kinetic partitioning, in an analogous manner to Skp's role during OMP folding.[15]

Modeling of Kinetic Traces Supports Skp-Folding Intermediate Interaction. To quantitatively compare our experimental data with our proposed mechanistic model, the folding and aggregation kinetic traces obtained with and without a 1:1 molar ratio of Skp:protein were fit globally by dynamic simulation by Global Kintek Explorer. By globally fitting all available data directly to a model through dynamic simulation, no assumptions are made and multiple possible models can be rapidly and robustly evaluated.[31] This is more robust than classical model fitting methods, which often require assumptions in order to derive the equations required for direct data fitting. To avoid overfitting our limited data set, the most minimal model that could still accurately describe the data was empirically chosen from 34 models tested. All possible models involving a 0 to 3-step folding pathway, 0 to 3-step aggregation pathway, and Skp interaction varying from none to interaction with folding intermediates were included. Additionally, rates were grouped into three final rate constants (Figure 7B) to reduce the total number of fitted parameters. Remarkably, all Skp-sensitive proteins (lysozyme, 14B7, and 4-4-20) fit the same model, in which folding proceeds via an scFv intermediate that interacts with Skp, with aggregation occurring as a third-order process (Figure 7C–E). Lysozyme aggregation has been previously reported as occurring via third-order kinetics,[44] supporting this model. In contrast, the Skp-insensitive DO11.10 best fit model lacks a Skp–scFv interactions, and aggregation occurs via first-order kinetics (data not shown).

The consistency of models derived from global fitting can be assessed by comparing the calculated rate constants to rates obtained by other fitting methods. The simulated folding rate constant (k_{fold}) values were used to calculate an observed rate of folding (k_{obs}) and compared to the observed rates obtained from direct fitting of the folding assay to single exponentials (as reported in Table 1). Global fitting of lysozyme data gave a k_{obs} of $(0.322 \pm 0.199) \times 10^{-2}$ min^{-1} compared to $(0.259 \pm 0.025) \times 10^{-2}$ min^{-1} by direct fitting using single exponentials, values that are comparable within error. Global fitting of 14B7 data resulted in a k_{obs} of $(0.076 \pm 0.030) \times 10^{-2}$ min^{-1} compared to $(0.165 \pm 0.026) \times 10^{-2}$ min^{-1} by direct fitting. The 14B7 rates differ likely due to the limited data set available to define the model during global fitting. The 4-4-20 observed folding rate was very similar for the two methods: global fitting gave a k_{obs} of $(0.144 \pm 0.028) \times 10^{-2}$ min^{-1} compared to $(0.153 \pm 0.039) \times 10^{-2}$ min^{-1} by direct fitting. Thus, not only are almost all rate constants obtained from these global fits well-defined, but the results are quantitatively comparable to other fitting methods. While these global fitting results are not direct evidence supporting interactions between Skp and scFv folding intermediates, they provide a model that is consistent with all our kinetic data obtained for Skp-sensitive scFvs.

DISCUSSION

Here, we have analyzed the effect of the Skp chaperone on *in vitro* refolding of three scFv proteins and observed a reduced aggregation rate but no change in the folding rate in the presence of stoichiometric amounts of Skp. Examination of the biophysical characteristics of these proteins identified a correlation between concentration-dependent aggregation, a long-lived folding intermediate, and Skp sensitivity, but not other characteristics such as scFv thermal or denaturant stability. These data suggest that Skp affects the folding of soluble proteins by acting as a holdase to reduce the intermediate concentration and thus the overall aggregation rate.

OMP–Skp versus skp–scFv Recognition. Skp appears to interact with client proteins primarily through exposed hydrophobic residues and to a lesser extent, electrostatics, guided by Skp's very basic pI.[9,23] Sequence comparison of Skp and six homologues from Gram-negative Enterobacteriaceae revealed a distinct pattern of conservation at the interior surface of the basket, dominated by residues Phe 25, 42, 86, and 93,[9] which is postulated to facilitate recognition of the hydrophobic β-barrel domain of OmpA in solution.[12] Similarly, the OMP PhoE forms a sufficiently stable complex with Skp to allow copurification by size exclusion chromatography.[45] In contrast, Skp's interactions with scFvs are presumably very transient, precluding our isolation of the complex by gel filtration or chemical cross-linking (data not shown). The β-sheet-rich structure of partially unfolded scFvs may mimic endogenous OMP substrates, allowing for serendipitous recognition by Skp.

Aggregation during *in vivo* and *in vitro* folding is a specific off-pathway process involving folding intermediates.[46,47] The β-barrel domain of OmpA and *n*-terminal domain of PhoE can both be considered trapped folding intermediates, since they are unable to assume their active form in the aqueous periplasm.[45] Similar off-pathway folding intermediates have been identified for scFvs,[48] resulting from premature heavy and light domain association in the artificial single-chain format,[49] non-native disulfide bond formation, or slow cis/trans proline isomerization. Here, we observed a correlation between the presence of a folding intermediate, concentration-dependent aggregation, and Skp sensitivity for all three scFvs tested. For the 4-4-20 scFv, aggregation results from the off-pathway in the folding reaction due to exposed hydrophobic patches[35] which is ameliorated by a V84D amino acid substitution in the variable heavy chain or transfer of complementarity determining residue loops onto the robust human scFv framework. Folding intermediates detectable by urea denaturation are likely to aggregate *in vivo* as well.[50]

Furthermore, aggregation propensities were calculated *a priori* using multiple algorithms that predict the relative

propensity of proteins to form β-sheet-rich amyloid-like aggregates from primary sequence information. Most were able to correctly predict the relative aggregation of scFvs *in vitro*, which correlated well with Skp sensitivity. Amyloid aggregates are rich in amphipathic β-sheets, and protein aggregation in the *E. coli* cell can proceed via an amyloid-like form.[51] Because of the β-sheet-rich nature of scFvs, these algorithms may be useful to not only predict relative aggregation propensities but also to identify Skp-sensitive scFvs for use in laboratory or industrial scFv expression.

The cavity depicted by the Skp crystal structure can accommodate a ∼25 kDa protein, but flexing of the α-helical tentacles may allow larger client proteins.[10] The scFvs used in this study were ∼26 kDa when folded, with a larger hydrodynamic radius for unfolded topologies. If an entire, partially folded scFv is encapsulated by Skp, the tentacles would need to open to enlarge this cavity. Alternatively, scFv folding intermediates may be composed of one correctly folded variable domain and one unfolded variable domain, which could be directly accommodated by the Skp cavity.[52] This situation is suggested by the 14B7 urea unfolding curve, in which the two denaturant midpoints may reflect sequential unfolding of one domain followed by unfolding of the second, more stable, domain (Figure 6).[53] In this case, one domain would reside within the Skp cavity while the other extends beyond the tentacles. This is a direct analogy to the observed Skp−OmpA interaction, in which the partially folded β-barrel domain interacted with Skp, while the folded periplasmic domain resided in the surrounding buffer.[15]

Skp−scFv Interaction in the Context of the Crowded Cell. In these studies, we used *in vitro* refolding as an experimentally tractable approach to begin dissecting the transient interactions between Skp and its soluble scFv client proteins. However, in the context of a cell, the folding environment is quite different. First, the periplasm is highly crowded which can promote aggregation during protein folding[54] but stabilizes proteins once folded through attraction-depletion effects.[55] To help nascent polypeptides fold in this crowded environment, a number of chaperones are present to guide molecules down productive pathways, including disulfide bond isomerases. Properly folded scFvs contain two disulfide bonds, which form during periplasmic folding. Depending on the degree of native conformation in the intermediate recognized by Skp, disulfide bond formation or isomerization, catalyzed by the DsbA and DsbC chaperones respectively, may occur before or after Skp interactions.[56] The experiments presented here included DTT in the initial denaturation buffer to reduce the disulfide bonds, as this resulted in a more pronounced aggregation response.

Second, *in vivo*, scFv and OMP biogenesis occur in the context of multiple related homeostatic mechanisms. Inactivation of Skp by Tn10 transposon insertion induces the heat shock factor σ^E, signaling envelope stress.[57] Extracytoplasmic stress, including accumulation of misfolded OMPs, in turn regulates Skp via the σ^E and Cpx factors.[58] Overexpression of poorly folding scFvs, such as 4-4-20, may exhaust the pool of available Skp, resulting in an effective "Skp deletion" phenotype, OMP misfolding, envelope stress, and induction of σ^E.[17] In effect, the *E. coli* cell may "mistake" scFv folding intermediates for OMP intermediates, serendipitously accommodating these recombinant proteins into established OMP folding pathways. While activation of these regulatory elements will induce expression of chaperones to prevent protein aggregation, the system is not equipped to deal with the magnitude of misfolded proteins encountered during high-level recombinant expression. Coordinated scFv and Skp expression from a single open reading frame ensures stoichiometric levels of Skp are available to chaperone the scFv.

Third, during *in vitro* refolding, all molecules are simultaneously exposed to refolding conditions at supraphysiologic concentrations of the client protein. *In vivo*, scFvs and OMPs will be individually extruded through the inner membrane into the periplasm by the SecYEG machinery in a temporally and spatially isolated manner, anchored to the inner membrane via a leader sequence until protease activity releases the scFv into the periplasm. Skp may interact with client proteins at the inner membrane, as it copurifies with this cellular fraction, or after release into periplasm.[12] What is unclear is how client proteins are released from the Skp cavity in the absence of an ATP-driven conformational change such as that used by GroEL. Since OMPs are delivered to the BAM outer membrane protein insertion complex, OMP release may be driven by high affinity between the OMP−BAM complex or a conformational change that occurs upon Skp−BAM interaction. However, soluble proteins, which associate very transiently with Skp, will require a different release mechanism. If the client protein achieves its final form while complexed with Skp, the burial of exposed hydrophobic residues could reduce the affinity between the two molecules, allowing the client protein to diffuse away. If Skp merely acts to reduce the concentration of free folding intermediates, scFv capture and release may be governed simply by rapid association and dissociation rates characteristic of weak binding affinity, resulting in a transient scFv−Skp interaction. Our future studies will directly evaluate these possibilities.

Skp and Other *E. coli* Chaperones. We have proposed a mechanism whereby Skp sequesters aggregation-prone scFv folding intermediates, passively preventing aggregation and thereby increasing active yields in solution via kinetic partitioning. *In vivo*, the system is more complex, owing to redundancies among periplasmic protein quality control machinery. Skp, in conjunction with the DegP protease, is genetically redundant with the primary OMP chaperone, SurA.[8] However, the mechanism of biochemical redundancy is less clear, as Skp's jellyfish shape and lack of a specific binding motif are very different from SurA, which appears to bind a linear aromatic sequence common within OMPs. Skp instead appears to bind a conformational rather than linear sequence,[45] able to protect unfolded OMPs from degradation by DegP and proteinase K *in vitro*.[59]

While Skp has no eukaryotic homologue, it does share structural homology with the ATP-independent archaeal and eukaryotic chaperone prefoldin, which adopts a similar jellyfish-like shape comprised of six subunits. Archaeal prefoldin binds a wide range of substrates using the tips of its tentacles, while the eukaryotic homologue fully encapsulates actin or tubulin substrates within its cavity.[16] This is very similar to the interaction between Skp and OmpA, as observed with NMR.[15] However, a Skp−scFv interaction has yet to be directly observed and may be challenging, as evidenced by the inability to isolate stable complexes using analytical gel chromatography.

■ CONCLUSIONS

Skp is an important chaperone involved in OMP biogenesis that is being increasingly recognized for its role in assisting soluble protein folding in the bacterial periplasm. Using three

distinct scFvs and *in vitro* refolding assays, our data are consistent with Skp acting as a holdase to prevent aggregation of scFv folding intermediates. Notably, the Skp sensitivity of each scFv correlated with the presence of a folding intermediate as well as *a priori* calculated aggregation propensity. These results provide insight into Skp's endogenous role chaperoning soluble proteins and its growing role in recombinant production of biologics. They also suggest the possibility of engineered Skp variants, able to specifically and efficiently chaperone immunoglobulin client proteins, analogous to efforts to engineer GFP-specific GroEL variants.[24] Finally, since Skp is an important factor in *Salmonella* virulence,[60] among other pathogens,[22] better understanding of its ability to chaperone soluble proteins may suggest novel antibiotic targets.

■ ASSOCIATED CONTENT

Supporting Information

Figure S1 depicting concentration dependence of protein aggregation. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel (512) 471-9188; Fax (512) 471-7060; e-mail Maynard@che.utexas.edu.

Funding

This work was supported by grants from the NIH (AI066239 and GM095638), the Packard Foundation (2005-098), and the Welch foundation (F-1767) to J.A.M.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Eric Boder (University of Tennessee Knoxville) for the 4-4-20 gene and Dr. Benjamin Roy (University of Texas at Austin) for the DO11.10 gene.

■ ABBREVIATIONS

BSA, bovine serum albumin; DTT, dithiothreitol; IPTG, isopropyl β-D-1-thiogalactopyranoside; IMAC, immobilized metal affinity chromatography; MBP, maltose binding protein; OMP, outer membrane protein; scFv, single-chain variable or T cell receptor fragment.

■ REFERENCES

(1) Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis. *Nature 475*, 324−332.

(2) Baneyx, F., and Mujacic, M. (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol. 22*, 1399−1408.

(3) Cummings, C. J., Sun, Y., Opal, P., Antalffy, B., Mestril, R., Orr, H. T., Dillmann, W. H., and Zoghbi, H. Y. (2001) Over-expression of inducible HSP70 chaperone suppresses neuropathology and improves motor function in *SCA1* mice. *Hum. Mol. Genet. 10*, 1511−1518.

(4) Banerji, U. (2009) Heat shock protein 90 as a drug target: some like it hot. *Clin. Cancer Res. 15*, 9−14.

(5) Tan, H., Deng, L., and Yang, Z. (2009) Development of the hsp110-heparanase vaccine to enhance antitumor immunity using the chaperoning properties of hsp110. *Mol. Immunol. 47*, 298−301.

(6) Mazor, Y., Van Blarcom, T., Mabry, R., Iverson, B. L., and Georgiou, G. (2007) Isolation of engineered, full-length antibodies from libraries expressed in *Escherichia coli*. *Nat. Biotechnol. 25*, 563−565.

(7) Purdy, G. E., Fisher, C. R., and Payne, S. M. (2007) IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp and SurA. *J. Bacteriol. 189*, 5566−5573.

(8) Sklar, J. G., Wu, T., Kahne, D., and Silhavy, T. J. (2007) Defining the roles of the periplasmic chaperones SurA, Skp and DegP in *Escherichia coli*. *Genes Dev. 21*, 2473−2484.

(9) Korndörfer, I. P., Dommel, M. K., and Skerra, A. (2004) Structure of the periplasmic chaperone skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nat. Struct. Mol. Biol. 11*, 1015−1020.

(10) Walton, T. A., and Sousa, M. C. (2004) Crystal structure of skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol. Cell 15*, 367−374.

(11) Chen, R., and Henning, U. (1996) A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol. Microbiol. 19*, 1287−1294.

(12) Bulieris, P. V., Behrens, S., Holst, O., and Kleinschmidt, J. H. (2003) Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone skp and by lipopolysaccharide. *J. Biol. Chem. 278*, 9092−9099.

(13) Qu, J., Mayer, C., Behrens, S., Holst, O., and Kleinschmidt, J. H. (2007) The trimeric periplasmic chaperone Skp of *Escherichia coli* forms 1:1 complexes with outer membrane proteins *via* hydrophobic and electrostatic interactions. *J. Mol. Biol. 374*, 91−105.

(14) Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J., and Behrens, S. (2005) The periplasmic chaperone surA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J. Biol. Chem. 280*, 23540−23548.

(15) Walton, T. A., Sandoval, C. M., Fowler, C. A., Pardi, A., and Sousa, M. C. (2009) The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proc. Natl. Acad. Sci. U. S. A. 106*, 1772−1777.

(16) Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U., and Moarefi, I. (2000) Structure of the molecular chaperone prefoldin: unique interaction of the multiple coiled coil tentacles with unfolded proteins. *Cell 103*, 621−632.

(17) Bothmann, H., and Plückthun, A. (1998) Selection for a periplasmic factor improving phage display and functional periplasmic expression. *Nat. Biotechnol. 16*, 376−380.

(18) Maynard, J., Adams, E. J., Krogsgaard, M., Petersson, K., Liu, C. W., and Garcia, K. C. (2005) High-level bacterial secretion of single-chain $\alpha\beta$ T-cell receptors. *J. Immunol. Methods 306*, 51−67.

(19) Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B. L., and Georgiou, G. (2003) Isolation and expression of recombinant antibody fragments to the biological warfare pathogen *Brucella melitensis*. *J. Immunol. Methods 276*, 185−196.

(20) Maynard, J. A., Maassen, C. B. M., Leppla, S. H., Brasky, K., Patterson, J. L., Iverson, B. L., and Georgiou, G. (2002) Protection against anthrax toxin by recombinant antibody fragments correlates with antigen affinity. *Nat. Biotechnol. 20*, 597−601.

(21) Levy, R., Weiss, R., Chen, G., Iverson, B. L., and Georgiou, G. (2001) Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli trxB gor* mutants via the coexpression of molecular chaperones. *Protein Expression Purif. 23*, 338−347.

(22) Wagner, J. K., Heindl, J. E., Gray, A. N., Jain, S., and Goldberg, M. B. (2009) Contribution of the periplasmic chaperone Skp to efficient presentation of the autotransporter IcsA on the surface of *Shigella flexneri*. *J. Bacteriol. 191*, 815−821.

(23) Jarchow, S., Lück, C., Görg, A., and Skerra, A. (2008) Identification of potential substrate proteins for the periplasmic *Escherichia coli* chaperone Skp. *Proteomics 8*, 4987−4994.

(24) Kawe, M., and Plückthun, A. (2006) GroEL walks the fine line: the subtle balance of substrate and co-chaperonin binding by GroEL. A combinatorial investigation by design, selection and screening. *J. Mol. Biol. 357*, 411−426.

(25) Sidhu, S. S., and Weiss, G. A. (2004) Constructing phage display libraries by oligonucleotide-directed mutagenesis, in *Phage Display: A*

Practical Approach (Clackson, T., and Lowman, H. B., Eds.) pp 27−42, Oxford University Press, New York.
(26) Guzman, L., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol. 177*, 4121−4130.
(27) Schaefer, J. V., and Plückthun, A. (2010) *Antibody Engineering*, Vol. 2, 2nd ed., Springer, New York.
(28) Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R., and Plückthun, A. (1997) Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods 201*, 35−55.
(29) Sambrook, J., and Russell, D. W. (2001) Purification of expressed proteins from inclusion bodies, in *Molecular Cloning: A Laboratory Manual* (Argentine, J., Ed.) 3rd ed., p 15.49, Cold Spring Harbor Laboratory Press, New York.
(30) Pace, C. N., Shirley, B. A., and Thomson, J. A. (1987) Measuring the conformational stability of a protein, in *Protein Structure: A Practical Approach* (Creighton, T., Ed.) pp 311−330, Oxford University Press, New York.
(31) Walters, J., Milam, S. L., and Clark, A. C. (2009) Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics. *Methods Enzymol. 455*, 1−39.
(32) Lavinder, J. J., Hari, S. B., Sullivan, B. J., and Magliery, T. J. (2009) High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. *J. Am. Chem. Soc. 131*, 3794−3795.
(33) Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global Kinetic Explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem. 387*, 20−29.
(34) Apetri, A. C., and Horwich, A. L. (2008) Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. *Proc. Natl. Acad. Sci. U. S. A. 105*, 17351−17355.
(35) Nieba, L., Honegger, A., Krebber, C., and Plückthun, A. (1997) Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved *in vivo* folding and physical characterization of an engineered scFv fragment. *Protein Eng. 10*, 435−444.
(36) Jung, S., and Plückthun, A. (1997) Improving *in vivo* folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng. 10*, 959−966.
(37) Frare, E., de Laureto, P. P., Zurdo, J., Dobson, C. M., and Fontana, A. (2004) A highly amyloidogenic region of hen lysozyme. *J. Mol. Biol. 340*, 1153−1165.
(38) Chen, L., Hodgson, K. O., and Doniach, S. (1996) A lysozyme folding intermediate revealed by solution X-ray scattering. *J. Mol. Biol. 261*, 658−671.
(39) Castillo, V., Graña-Montes, R., Sabate, R., and Ventura, S. (2011) Prediction of the aggregation propensity of proteins from the primary sequence: aggregation properties of proteomes. *Biotechnol. J. 6*, 674−685.
(40) Conchillo-Solé, O., de Groot, N. S., Avilés, F. X., Vendrell, J., Daura, X., and Ventura, S. (2007) AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinf. 8*, 65.
(41) Groot, N. S. d., Pallarés, I., Avilés, F. X., Vendrell, J., and Ventura, S. (2005) Prediction of "hot spots" of aggregation in disease-linked polypeptides. *BMC Struct. Biol. 5*, 18.
(42) Bryan, A. W., Jr., Menke, M., Cowen, L. J., Lindquist, S. L., and Berger, B. (2009) BETASCAN: probable β-amyloids identified by pairwise probabilistic analysis. *PLoS Comput. Biol. 5*, e1000333.
(43) Garbuzynskiy, S. O., Lobanov, M. Y., and Galzitskaya, O. V. (2010) FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics 26*, 326−332.
(44) Clark, E. D. B., Hevehan, D., Szela, S., and Maachupalli-Reddy, J. (1998) Oxidative renaturation of hen egg-white lysozyme. Folding vs aggregation. *Biotechnol. Prog. 14*, 47−54.
(45) Harms, N., Koningstein, G., Dontje, W., Muller, M., Oudega, B., Luirink, J., and de Cock, H. (2001) The early interaction of the outer membrane protein PhoE with the periplasmic chaperone skp occurs at the cytoplasmic membrane. *J. Biol. Chem. 276*, 18804−18811.
(46) Jaenicke, R. (1987) Folding and association of proteins. *Prog. Biophys. Mol. Biol. 49*, 117−237.
(47) Wetzel, R., and Chrunyk, B. A. (1994) Inclusion body formation by interleukin-1β depends on the thermal sensitivity of a folding intermediate. *FEBS Lett. 350*, 245−248.
(48) Wülfing, C., and Plückthun, A. (1994) Correctly folded T-cell receptor fragments in the periplasm of *Escherichia coli*. *J. Mol. Biol. 242*, 655−669.
(49) Hoyer, W., Ramm, K., and Plückthun, A. (2002) A kinetic trap is an intrinsic feature in the folding pathway of single-chain Fv fragments. *Biophys. Chem. 96*, 273−284.
(50) Fink, A. L. (1999) Chaperone-mediated protein folding. *Physiol. Rev. 79*, 425−449.
(51) Groot, N. S. d., Sabate, R., and Ventura, S. (2010) Amyloids in bacterial inclusion bodies. *Trends Biochem. Sci. 34*, 408−416.
(52) Jäger, M., and Plückthun, A. (1999) Domain interactions in antibody Fv and scFv fragments: effects on unfolding kinetics and equilibria. *FEBS Lett. 462*, 307−312.
(53) Wörn, A., and Plückthun, A. (2001) Stability engineering of antibody single-chain Fv fragments. *J. Mol. Biol. 305*, 989−1010.
(54) Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science 295*, 1852−1858.
(55) Zhou, H., Rivas, G., and Minton, A. P. (2008) Macromolecular crowding and confinement: biochemical, biophysical and potential physiological consequences. *Annu. Rev. Biophys. 37*, 375−397.
(56) Nakamoto, H., and Bardwell, J. C. A. (2004) Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim. Biophys. Acta 1694*, 111−119.
(57) Missiakas, D., Betton, J., and Raina, S. (1996) New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. *Mol. Microbiol. 21*, 871−884.
(58) Dartigalongue, C., Missiakas, D., and Raina, S. (2001) Characterization of the *Escherichia coli* σ^E regulon. *J. Biol. Chem. 276*, 20866−20875.
(59) Wu, S., Ge, X., Lv, Z., Zhi, Z., Chang, Z., and Zhao, X. S. (2011) Interaction between bacterial outer membrane proteins and periplasmic quality control factors: a kinetic partitioning mechanism. *Biochem. J. 438*, 505−511.
(60) Rowley, G., Skovierova, H., Stevenson, A., Rezuchova, B., Homerova, D., Lewis, C., Sherry, A., Kormanec, J., and Roberts, M. (2011) The periplasmic chaperone skp is required for successful *Salmonella* Typhimurium infection in a murine typhoid model. *Microbiology 157*, 848−858.
(61) Midelfort, K. S., Hernandez, H. H., Lippow, S. M., Tidor, B., Drennan, C. L., and Wittrup, K. D. (2004) Substantial energetic improvement with minimal structural perturbation in a high affinity mutant antibody. *J. Mol. Biol. 343*, 685−701.