Heterologous gene expression using self-assembled supra-molecules with high affinity for HSP70 chaperone

Ji-Young Ahn, Hyung Choi, Yang-Hoon Kim, Kyung-Yeon Han, Jin-Seung Park, Sung-Sik Han[1] and Jeewon Lee*

Department of Chemical and Biological Engineering and [1]School of Life Sciences and Biotechnology, Korea University, Anam-Dong 5-1, Sungbuk-Ku, Seoul 136-713, South Korea

ABSTRACT

Contrary to the results of direct expression, various human proteins (ferritin light-chain, epithermal growth factor, interleukin-2, prepro-ghrelin, deletion mutants of glutamate decarboxylase and arginine deiminase, and mini-proinsulin) were all soluble in *Escherichia coli* cytoplasm when expressed with the N-terminus fusion of ferritin heavy-chain (FTN-H). Through systematic investigations, we have found that a specific peptide motif within FTN-H has a high affinity to HSP70 chaperone DnaK, and that the peptide motif was composed of a hydrophobic core of three residues (Ile, Phe and Leu) and two flanking regions enriched with polar residues (Gly, Gln and Arg). It was also observed that all the recombinant proteins expressed with the fusion of FTN-H formed spherical nanoparticles with diameters of 10–15 nm, as confirmed by the transmission electron microscopy image. The protein nanoparticles are non-covalently cross-linked supra-molecules formed by the self-assembly function of FTN-H. Upon the formation of the supra-molecule, its size is likely to be limited by the assembly properties of FTN-H, thereby keeping the self-assembled particles soluble. This study reports on the dual function of FTN-H for fusion expression and solubility enhancement of heterologous proteins: (i) high-affinity interaction with DnaK and (ii) formation of self-assembled supra-molecules with limited and constant sizes, thereby avoiding the undesirable formation of insoluble macro-aggregates of heterologous proteins.

INTRODUCTION

Ferritin plays a central role in iron storage and in the maintenance of intracellular iron balance for the growth and survival of most organisms (1–3). The functions of ferritin suggest that it may serve as a cytoprotective protein, minimizing oxygen-free radical formation by sequestering intracellular iron (4). Ferritin molecules isolated from vertebrates are composed of two types of subunits (heavy chain and light chain), whereas those from plants and bacteria contain only heavy chains. Mammalian heavy chain and light chain show ∼54% sequence identity. In mixed subunit 24mer (hetero-polymers) with a spherical shell structure, heavy (H) and light (L) chain subunits have similar conformations of principally four α helix bundles (5–8). Four years ago, the Korean Food and Drug Administration prohibited the commercial manufacture of animal ferritin (from equine spleen extracts), which had been used to treat human iron deficiency, because of its probable complication of virus infection. Thereafter, the production of recombinant human ferritin in *Escherichia coli* has been used as a possible substitute for equine ferritin extracts.

It was reported that contrary to ferritin L-chain (FTN-L), ferritin H-chain (FTN-H) and H–L hybrid (FTN-H fused at the N-terminus of FTN-L) are soluble in *E.coli* cytoplasm, probably due to their enhanced folding efficiency. The recombinant ferritin H–L hybrid as well as the H-chain showed the same iron storage activity as natural copolymers of human ferritin (9). Various human proteins [epidermal growth factor (EGF), interleukin-2 (IL-2), prepro-ghrelin (ppGRN), a deletion mutant of glutamate decarboxylase ($GAD_{512–585}$), a deletion mutant of arginine deiminase ($ADI_{132–410}$) and minipro-insulin (mp-INS)], forming inclusion bodies upon direct expression, were all soluble in *E.coli* cytoplasm when expressed with the N-terminus fusion of ferritin H-chain. Considering the structural diversity of the above human proteins

*To whom correspondence should be addressed. Tel: +82 2 3290 3304; Fax: +82 2 926 6102; Email: leejw@korea.ac.kr

fused to ferritin H-chain, it seems reasonable to assume that the *de novo* folding of proteins with ferritin H-chain fusion is not determined by direct protein–protein interaction between FTN-H and the target protein, but instead, the assistance of a global folding enhancer, a molecular chaperone, is much more important. Misfolding and aggregation of proteins are major damaging consequences of stress situations, such as heat shock and pathophysiological states (10–12). The central cellular defense against such damage is a molecular chaperone, which prevents aggregation, assists in refolding and mediates the degradation of misfolded proteins (13,14). Chaperones can cooperate *in vitro* as part of a functional network, in which 'holder' chaperones prevent the aggregation of misfolded proteins, whereas 'folder' chaperones actively assist in refolding (15,16). The Hsp70 chaperone DnaK participates in the folding of a subset of newly synthesized proteins, prevents protein aggregation in a denatured condition and refolds misfolded proteins.

Recently, Vandenbroeck *et al.* (17) identified conserved DnaK-binding sites in the N- and C-terminal halves of helix B and C of diverse 4-helix-bundle proteins, respectively, that belong to the superfamily of interferon-γ/interleukin-10-related cytokines. These cytokines belong to a relatively small group of homodimeric proteins with highly interdigitated interfaces exhibiting the strongly hydrophobic character of the interior core of a single-chain folded domain. They proposed that the binding of DnaK may constitute the hallmark of a novel conserved regulatory mechanism, in which HSP70-like chaperones assist in the formation of a hydrophobic dimeric 'folding' interface. Hidaka *et al.* (18) also found that the propeptide of guanylyl cyclase, which activates peptide II (GCAP-II), an endogenous ligand of guanylyl cyclase C, has a dual function in the proper folding of a mature peptide. This means that the N-terminus propeptide of pro-GCAP-II is critical, not only for disulfide-coupled folding but also for the net stabilization (or dimerization) of pro-GCAP-II.

In this study, diverse recombinant proteins, commonly with FTN-H fusion, were observed to form stable homopolymeric supra-molecules without the formation of insoluble protein aggregates. Seemingly, the supra-molecule size is limited by assembly properties of FTN-H, thereby keeping the supra-molecules soluble. The target proteins localized inside the spherical supra-molecules might have been in a protective environment against undesirable protein aggregation. We also identified a specific DnaK-binding motif within FTN-H and demonstrated the efficacy of FTN-H as a potent solubility enhancer upon heterologous protein expression in *E.coli*.

MATERIALS AND METHODS

PCR-based site-directed mutagenesis

Site-directed mutagenesis for preparing various mutants of FTN-H and light chain and glutathione *S*-transferase (GST) was performed according to the standard protocols for overlap PCR (19). PCRs were performed in a final volume of 20 μl with the following reagents: 20 mM Tris–HCl, pH 8.3, 100 pmol of each primer, 200 μM dNTPs, 1.5 mM magnesium chloride and 2.5 U *Taq* polymerase (BRL). PCR cycling was performed as follows: initial denaturation at 94°C for 120 s and then 20 cycles of 94°C (45 s), 60°C (60 s) and 72°C (120 s).

After the completion of 20 cycles, the reaction mixture was incubated for 5 min at 72°C. The products were gel purified after agarose gel electrophoresis.

Recombinant strains and plasmids

After PCR amplification using appropriate primers, each of the recombinant genes encoding human ferritin H- and L-chain (FTN-H and FTN-L) and their mutants, human EGF, human interleukin-2 (hIL-2), GST::(prepro-ghrelin) and other various fusion mutants (FTN-H::FTN-L, FTN-H::EGF, FTN-H::hIL-2, FTN-H::ppGRN, FTN-H::GAD$_{512–585}$, FTN-H::ADI$_{132–410}$ and FTN-H::mp-INS) was inserted into the NdeI–HindIII site of plasmid pET28a (Novagen) to construct expression vector. Upon the fusion construction above, specific cleavage sequence for thrombin protease was placed between C-terminus of FTN-H and N-terminus of target protein. After complete DNA sequencing of all the gel-purified plasmid vectors, *E.coli* strain BL21(DE3) [F$^-$ *omp*T *hsd*S$_B$(rB$^-$mB$^-$)] was transformed with the plasmid expression vectors, and kanamycin-resistant transformants were selected. For the constitutive co-expression of DnaK chaperone with each of the recombinant proteins above, *E.coli* dnaK gene was cloned into a compatible ampicillin-resistant plasmid pHCE19 (2) (Bioleaders Inc., Korea) (20). The resulting DnaK expression vector transformed the same bacterial host, *E.coli* BL21 (DE3), simultaneously with the expression vector for the target recombinant gene. Both ampicillin- and kanamycin-resistant transformants were selected using Luria–Bertani (LB) agar plates supplemented with ampicillin (100 mg/l) and kanamycin (100 mg/l).

Gene expression, purification of recombinant proteins and bioactivity assay

For the shake flask experiments (250 ml Erlenmeyer flasks, 37°C, 200 r.p.m. Shaking Incubator IB-3100, New Power Eng., Korea), LB media containing 100 mg kanamycin (or both ampicillin and kanamycin where applicable) per liter of culture was used. When the culture turbidity (OD$_{600}$) reached 0.5, the gene expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) (0.5 mM). After further 3–4 h cultivation, the recombinant cells were harvested by centrifugation (6000 r.p.m. for 5 min) and the cell pellets were resuspended in 5 ml distilled water. Cell disruption was done by using Branson Sonifier (Branson Ultrasonics Corp., Danbury, CT). The cell-free supernatant and insoluble protein aggregates were separated at 13 000 r.p.m. for 10 min. The isolated inclusion bodies, if any, were then washed twice with 1% Triton X-100. The cell-free supernatant and the washed inclusion bodies were subject to PAGE analysis, using 14% Tris-glycine precast gel (Novex, San Diego, CA). Coomassie-stained protein bands were scanned and analyzed using a densitometer (Duoscan T1200, Bio-Rad, Hercules, CA).

The purification of recombinant proteins was accomplished using metal affinity chromatography. The corresponding recombinant gene was cloned into the pET28a (Novagen) plasmid vector as mentioned previously, and expressed so that the soluble recombinant protein was synthesized as a fusion protein containing N-terminus six histidine residues. The resulting recombinant fusion protein with polyhistidine

tag has a high affinity for ProBond resin (Invitrogen). Recombinant protein was purified through a Ni^{2+} column with ProBond resin according to the following procedures. The cell-free supernatant containing 1–2 mg of polyhistidine-tagged recombinant protein was loaded onto ProBond resin (Ni^{+2}) column. Prior to sample loading, the resin was washed twice with 10 column volumes of binding buffer (50 mM potassium phosphate, 300 mM KCl and 20 mM imidazole, pH 7.0). Binding buffer contains 20 mM imidazole to minimize non-specific binding of untagged protein contaminants, and binding was carried out in a batch mode at 4°C. Subsequently, the resin was washed twice with 5–8 ml Tris–HCl (10 mM Tris, pH 8.0) prior to thrombin digestion step. Thrombin solution 50 μl (1 U/μl, Amersham, Catalog no. 27-0846-01) was mixed with Tris buffer 4950 μl, and the resulting reaction buffer was added to the column. After overnight thrombin digestion (16 h) at room temperature (22–25°C), the reaction solution was collected and properly concentrated using microcon YM-10 (Millipore, Ireland) before SDS–PAGE analysis.

The enzyme activity of recombinant arginine deiminase (ADI) was assayed by colorimetric determination of the enzyme reaction product, citrulline. The reaction mixture (0.1 M potassium phosphate, pH 6.5, 10 mM L-arginine and 0.1 ml enzyme solution in a final volume of 1 ml) was incubated at 37°C for 5 min, and the amount of citrulline was determined with diacetyl-monoxime (21). The sample heated to 95°C for 10 min and the optical density (A540 nm) was then determined. The protein concentration was determined using the Protein Assay Kit (Bio-Rad) with BSA as a standard. One unit of ADI is the amount of enzyme catalyzing 1 μmol of arginine to citrulline per min at 37°C under the assay conditions.

Peptide synthesis and *in vitro* DnaK-binding assay

Organic synthesis and purification of two peptides (GRIFLQD and GRALFQD) were implemented by Peptron, Inc. (Taejon, Korea). Each peptide was dissolved in Tris buffer (40 mM Tris–HCl and 0.5 mM ATP, pH 7.8), and the peptide concentration in the solution was adjusted to 10 mg/l. An aliquot of 20 μl of each peptide solution was dropped on PVDF membrane that was already wetted by methanol, and the blotted PVDF membrane was completely dried. [Through Ponceu S staining (0.25% Ponceu S in 1% acetic acid solution), we confirmed that a comparable amount of each peptide or protein was stably immobilized on the PVDF membrane.] The dried PVDF membrane was put into skim milk solution and slowly stirred for 30 min. Subsequently, the PVDF membrane was transferred into Tris buffer solution (40 mM Tris–HCl and 0.5 mM ATP, pH 7.8) containing 100 nmol of DnaK (Stressgen). After 1 h, the membrane was washed twice with phosphate-buffered saline buffer, followed by immuno-blotting assay using anti-DnaK Ab (monoclonal IgG, mouse, Stressgen) and anti-mouse immunoglobulin G–horseradish peroxidase as primary and secondary antibodies, respectively.

Transmission electron microscopy (TEM)

Unstained samples of a purified protein solution were prepared for electron microscopy by air-drying small drops of a sample solution onto carbon-coated copper electron microscopy grids. To obtain stained images of protein nanoparticles, the electron microscopy grids containing air-dried samples were incubated with a 2% (w/v) aqueous uranyl acetate solution for 10 min at room temperature and washed 3–4 times with distilled water. Protein nanoparticle images were examined using the Philips Technai 120 kV electron microscope. Electron diffraction patterns were recorded from a selected area that is well occupied with protein nanoparticles in order to obtain high diffraction intensities. Particle size distributions were made by measuring 50 protein nanoparticles.

RESULTS

Solubility enhancer vector system using gene fusion of FTN-H

In a previous study, recombinant human ferritin H-, L-chain, and H–L hybrid were produced using the T7 expression system of *E.coli* (9). Interestingly, the recombinant human ferritin H-chain (FTN-H) and H–L hybrid were intracellular soluble and active proteins. However, despite high-sequence homology with FTN-H, recombinant human ferritin L-chain (FTN-L) was aggregated into inclusion bodies in *E.coli* cytoplasm (9,22). With the expectation that the N-terminus fusion of FTN-H could be a plausible solution for soluble recombinant protein synthesis in *E.coli*, we attempted to heterologously express the following proteins, all of which were produced as inclusion bodies upon direct expression in *E.coli*: EGF, hIL-2, human ppGRN, deletion mutants of human GAD ($GAD_{512-585}$), Mycoplasma ADI ($ADI_{132-410}$) and human mini-proinsulin (mp-INS) (23). Figure 1 shows that the intracellular solubility of all recombinant fusion proteins (i.e. FTN-H::EGF, FTN-H::IL-2, FTN-H::ppGRN, FTN-H::$GAD_{512-585}$, FTN-H::$ADI_{132-410}$ and FTN-H::mp-INS) was significantly enhanced. For instance, almost 80% of synthesized FTN-H::ppGRN was soluble in *E.coli* cytoplasm, although most ppGRN formed inclusion bodies even by GST fusion (data not shown). ADI catalyzes the hydrolysis of arginine into citrulline and ammonia, which have been reported to inhibit cell proliferation by arresting cells in the G_1 phase and also to demonstrate anti-cancer activity. As shown in Figures 1 and 2A, we confirmed that ∼30% of the expressed fusion protein, FTN-H::$ADI_{132-410}$, was soluble in bacterial cytoplasm and biologically active as well, clearly showing arginine-degrading activity.

For the purification of ppGRN, FTN-H::ppGRN in a cell-free supernatant from whole-cell lysate was loaded to an Ni^{+2} column and retained inside the column due to the affinity interaction between Ni^{+2} and the N-terminus polyhistidine tag of FTN-H::ppGRN. Thrombin was subsequently added to the column to separate it from ppGRN by cleaving the fusion construct. As shown in Figure 2B (lane 3), ppGRN was successfully recovered through thrombin digestion of the fusion construct, FTN-H::ppGRN. Complete separation and purification of ppGRN was achieved after the removal of thrombin from the eluted solution. The other target proteins in the FTN-H fusion constructs were successfully purified through the same procedure (data not shown).

Search for putative DnaK-binding motif(s) in FTN-H

When 2-HEDS [2-hydroxyethyl disulfide ($HOCH_2CH_2$-S-S-CH_2CH_2OH), 15 mM] was present in the growth culture of

(A) Cytoplasmic solubility (%)
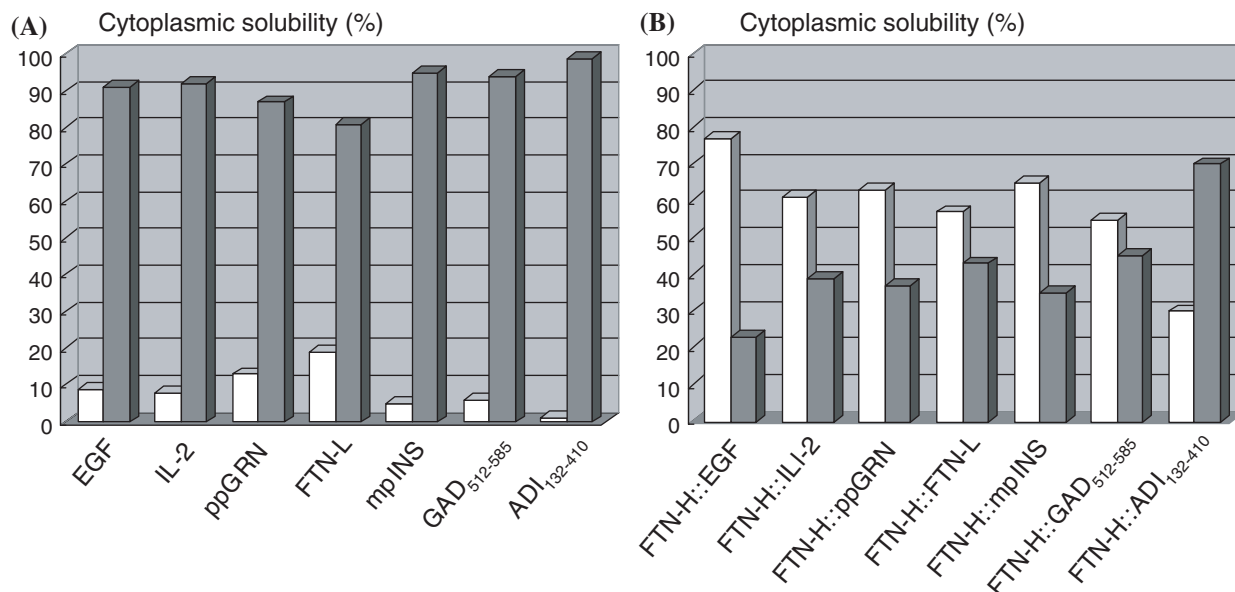

(B) Cytoplasmic solubility (%)

Figure 1. Cytoplasmic solubility (%) of various heterologous proteins in recombinant *E.coli*, either (**A**) directly expressed or (**B**) expressed with fusion of FTN-H at N-terminus of target heterologous proteins EGF, human IL-2, human ppGRN, human FTN-L, human mpINS, and C-terminal fragments of human GAD ($GAD_{512–585}$) and Mycoplasma ADI ($ADI_{132–410}$). White and gray bars represent percentages of soluble and insoluble proteins, respectively, in recombinant *E.coli* cytoplasm.
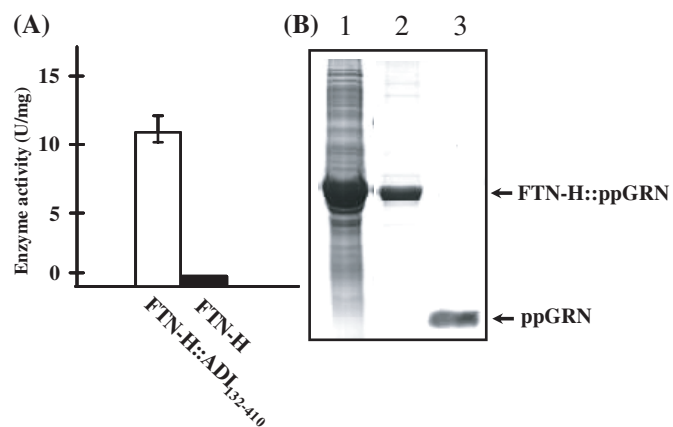
(A)

(B) 1 2 3

Figure 2. (**A**) Assay result of enzyme activity of $ADI_{132–410}$ expressed with the fusion of FTN-H (FTN-H::$ADI_{132–410}$). (**B**) Result of SDS–PAGE analysis of the fusion protein, FTN-H::ppGRN in cell-free supernatant from whole-cell lysate of recombinant *E.coli* (lane 1), in the solution eluted by imidazole after loading the cell-free supernatant to Ni^{+2} column for affinity purification of FTN-H::ppGRN (lane 2) and in the solution eluted after thrombin digestion to separate ppGRN from the Ni^{+2}-bound FTN-H::ppGRN inside the Ni^{+2} affinity column (lane 3).

(A) 2-HEDS: (−) (+) (+)
DnaK: (−) (−) (+)

(B) DnaK: (−) (+)

Figure 3. (**A**) Cytoplasmic solubility of recombinant FTN-H in the presence (+) or absence (−) of 2-HEDS with (+) or without (−) DnaK co-expression in the recombinant *E.coli* culture. (**B**) Cytoplasmic solubility of recombinant FTN-L with (+) or without (−) DnaK co-expression in the recombinant *E.coli* culture.

the recombinant *E.coli*, almost 100% of the recombinant FTN-H formed insoluble aggregates after gene expression. However, in the absence of 2-HEDS, >80% of the recombinant FTN-H was soluble in bacterial cytoplasm (Figure 3A). 2-HEDS is the oxidized form of DTT. When 2-HEDS is present in bacterial cytoplasm, its reduction is subject to the inducement of thiol oxidation of proteins, thereby promoting the formation of mis-linked disulfide bonds and protein aggregation. Comparing the results of reducing and non-reducing SDS–PAGE analyses, the aggregation of recombinant FTN-H by 2-HEDS seems to have occurred via intramolecular
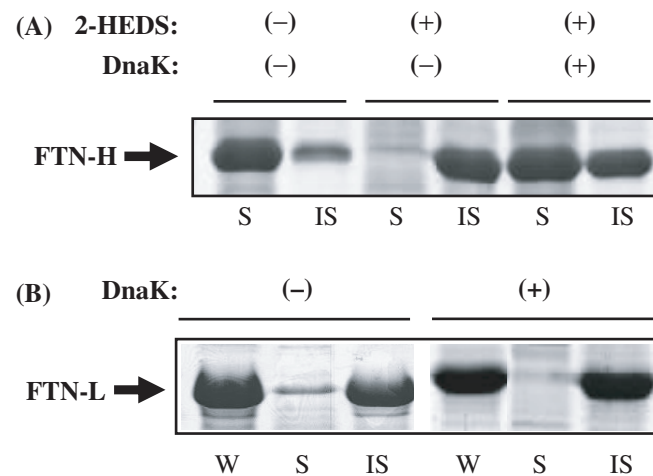
mislinkage of disulfide bridges (data not shown). From Figure 3A, even in the presence of 2-HEDS, the solubility of recombinant FTN-H was significantly increased when the HSP70 molecular chaperone, DnaK, was constitutively co-expressed. The co-expression of DnaK was carried out using a constitutive promoter localized in another expression vector, pHCE19 (2) (Bioleaders Inc., Korea) (20). Thus, it was possible to maintain the high cytoplasmic concentration of DnaK before the T7 expression of FTN-H was induced by IPTG. As shown in Figure 3B, the insoluble nature of recombinant FTN-L, however, never changed even with DnaK

co-expression using the same constitutive promoter system. The results shown in Figure 3 suggest that FTN-H must be capable of having a significant interaction with DnaK, presumably via special amino acid residues acting as high-affinity motif(s) for DnaK.

DnaK generally recognizes and binds to the hydrophobic region of the unfolded polypeptide chain. Thus, to identify the DnaK-binding motif(s), if any, the five hydrophobic clusters consisting of three amino acids (34VYL36, 81IFL83, 105LHL107, 133FIE35 and 161APE163) within the FTN-H sequence were first considered. They showed a relatively high hydrophobicity based on their theoretical hydropathy profile (Figure 4A). Using the site-directed mutagenesis technique, each of the five hydrophobic clusters was mutated

one at a time by replacing the hydrophobic amino acids in each cluster with arginine residue (34VYL36→34RRR36, 81IFL83→81RRR83, 105LHL107→105RHR107, 133FIE135→ 133RRE135 and 161APE163→161RRE163), thereby resulting in the construction of five FTN-H-derived mutants. Each mutant was expressed in the *E.coli* T7 expression system, and its intracellular solubility was analyzed. The purpose of mutating the wild-type sequence was to identify the hydrophobic patch that is recognized and bound by DnaK. If a certain hydrophobic cluster is responsible for DnaK binding, replacing the hydrophobic amino acids comprising the hydrophobic patch with strongly charged amino acids, such as arginine, will significantly affect the solubility of the polypeptide chain in bacterial cytoplasm. After gene expression, three
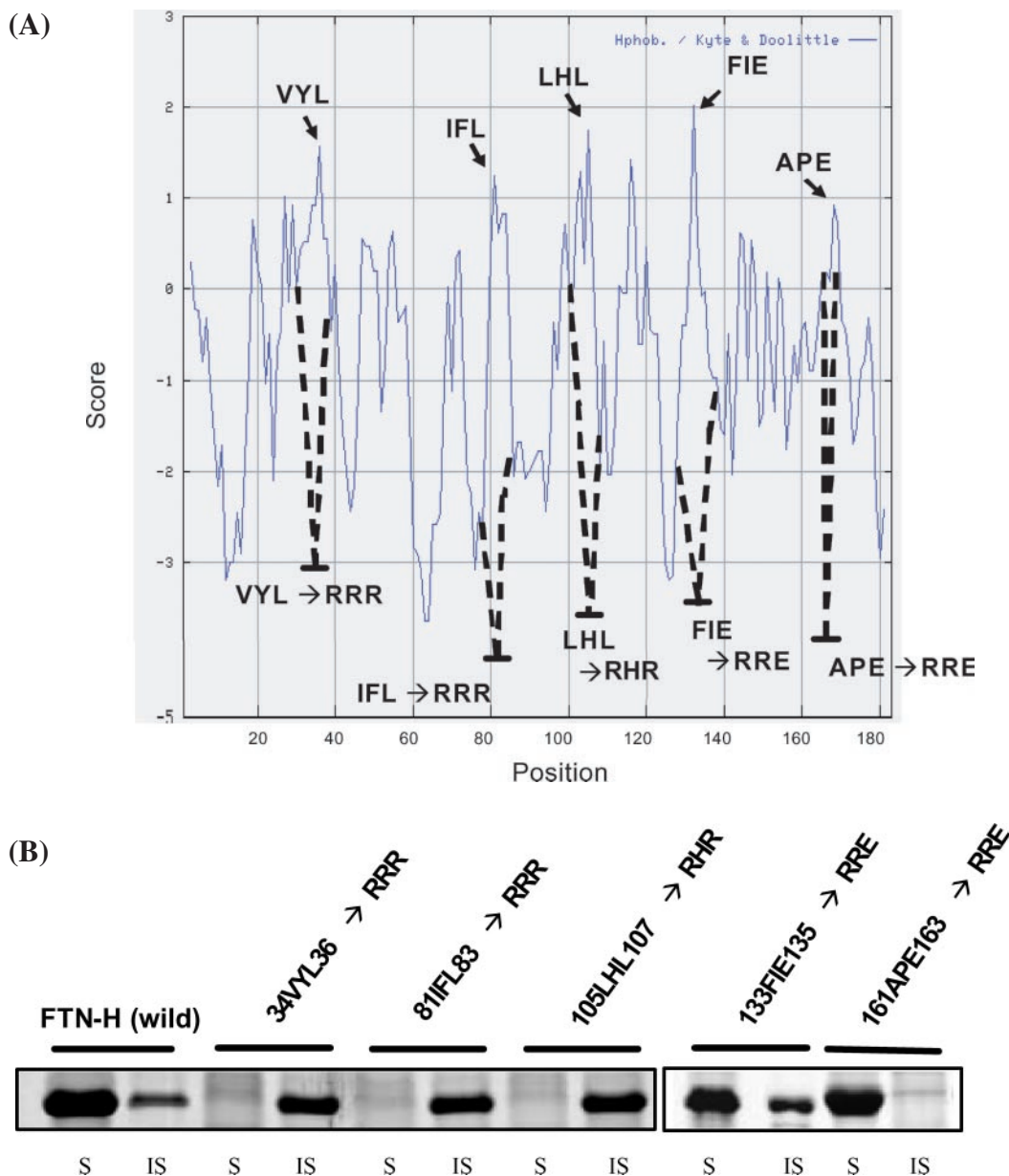

Figure 4. (**A**) Theoretical hydropathy profile of FTN-H and five hydrophobic tripeptide-residues to be replaced by arginine residues through site-directed mutagenesis. (**B**) *E.coli* cytoplasmic solubility of recombinant FTN-H mutants that were prepared by replacing the hydrophobic tripeptide residues [illustrated in (A)] with arginine residues.
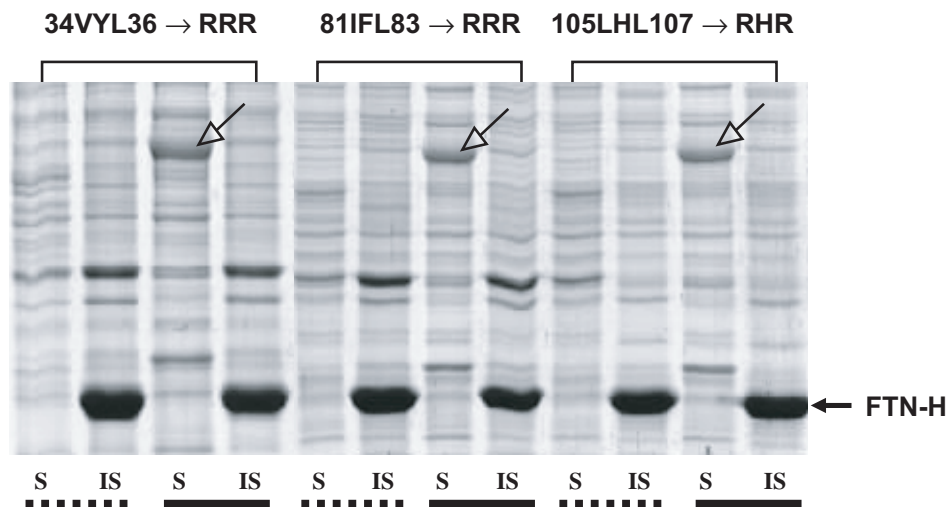
Figure 5. Solubility of three FTN-H mutants in recombinant *E.coli* cytoplasm, expressed with (represented by solid underlines) and without (represented by dotted underlines) DnaK co-expression. Arrows indicate co-expressed DnaK band.

FTN-H mutants (34RRR36, 81RRR83 and 105RHR107) were aggregated into inclusion bodies in *E.coli* cytoplasm (Figure 4B), implying that the three hydrophobic tripeptides (34VYL36, 81IFL83 and 105LHL107) could be putative motifs interacting with DnaK. Although DnaK was constitutively co-expressed with the three mutants (34RRR36, 81RRR83 and 105RHR107) using the same constitutive promoter system, the inclusion body formation of the three FTN-H mutants was never prevented (Figure 5). This is contrary to the case of wild-type FTN-H, in which the soluble FTN-H synthesis was resumed by DnaK co-expression even in the presence of 2-HEDS (Figure 3A). Therefore, it seems that there must be a special motif(s) in FTN-H with a high affinity to DnaK, leading to the syntheses of soluble FTN-H and other soluble heterologous proteins with FTN-H fusion.

Identification and validation of a specific DnaK-binding motif

Case of the ferritin-light chain mutant. The alignment of the amino acid sequences between FTN-H and FTN-L shows that the sequence of a putative DnaK-binding motif, 81IFL83, is completely different from the aligned cognate sequence (77ALF79) of FTN-L (Figure 6A). Through site-directed mutagenesis, the nucleotide sequence encoding 77ALF79 of FTN-L was replaced with the sequence coding for 77IFL79 so that FTN-L might have the putative DnaK-binding motif of FTN-H. The resulting substitution mutant of FTN-L (77IFL79) was expressed in *E.coli*. Surprisingly, the results of SDS–PAGE analysis showed that >50% of the synthesized FTN-L mutant was soluble in the bacterial cytoplasm (Figure 6B), whereas almost 100% of wild-type FTN-L was aggregated into inclusion bodies (Figures 1 and 3). Figure 6C shows that it is very similar to the case of FTN-H, in which most of the FTN-L mutant was aggregated by the addition of 2-HEDS, but the synthesis of the soluble FTN-L mutant was resumed by DnaK co-expression (Figure 6C). Consequently, it is evident that the tripeptide motif (IFL) from FTN-H played a critical role in increasing the solubility of the FTN-L mutant, probably by attracting the DnaK chaperone.

Case of the (GST mutant)::ppGRN. As described earlier in this study (Figure 1), the GST fusion protein, GST::ppGRN, was almost 100% aggregated upon expression in *E.coli*. The fusion expression of ppGRN was attempted using a GST mutant containing a DnaK-binding motif from FTN-H. First, several hydrophobic clusters in the N-terminus region of native GST were mutated by replacing the hydrophobic amino acids with arginine residues, and the cytoplasmic solubility of the resulting GST mutants was analyzed after their expression in the same recombinant *E.coli* system. As a result, a high degree of aggregation was observed in the synthesis of a GST mutant (85LAP87→85RRR87) (Figure 7A). This is very similar to the previous results obtained with FTN-H mutants (Figure 4). Thus, the hydrophobic cluster (85LAP87) and the sequence around 85LAP87 seem to be important sites in the interaction between GST and DnaK. To prepare GST mutants containing the putative DnaK-binding motif (81IFL83) of FTN-H, we first tried to replace the 85LAP87 of GST with the 81IFL83 of FTN-H, thereby constructing a substitution mutant of GST, 85LAP87→85IFL87 (GST-m1). We also constructed the following additional mutants of GST containing 80RIFLQ84 and 79GRIFLQD85 of FTN-H: 84LLAPV88→84RIFLQ88 (GST-m2) and 83QLLAPVN89→83GRIFLQD89 (GST-m3). Then, three fusion proteins of ppGRN (GST-m1::ppGRN, GST-m2::ppGRN and GST-m3::ppGRN) were synthesized using the same T7 gene expression system. From the results of the SDS–PAGE analysis, it is surprising that ∼50% of the synthesized GST-m3::ppGRN was in the form of soluble protein in *E.coli* cytoplasm (Figure 7B). Figure 7C shows that almost 100% re-aggregation of GST-m3::ppGRN was also observed in the presence of 2-HEDS, but that the soluble GST-m3::ppGRN appeared again after DnaK co-expression, which is very similar to the cases of FTN-H and the FTN-L mutant.

Apparently, the two mutants, GST-m3 and FTN-L (77IFL79), have a common foreign sequence, GRIFLQD, originating from FTN-H, which is therefore likely to be a peptide motif with a high affinity to DnaK. It is generally accepted that DnaK binds to the hydrophobic region in an almost extended polypeptide chain of unfolded protein at
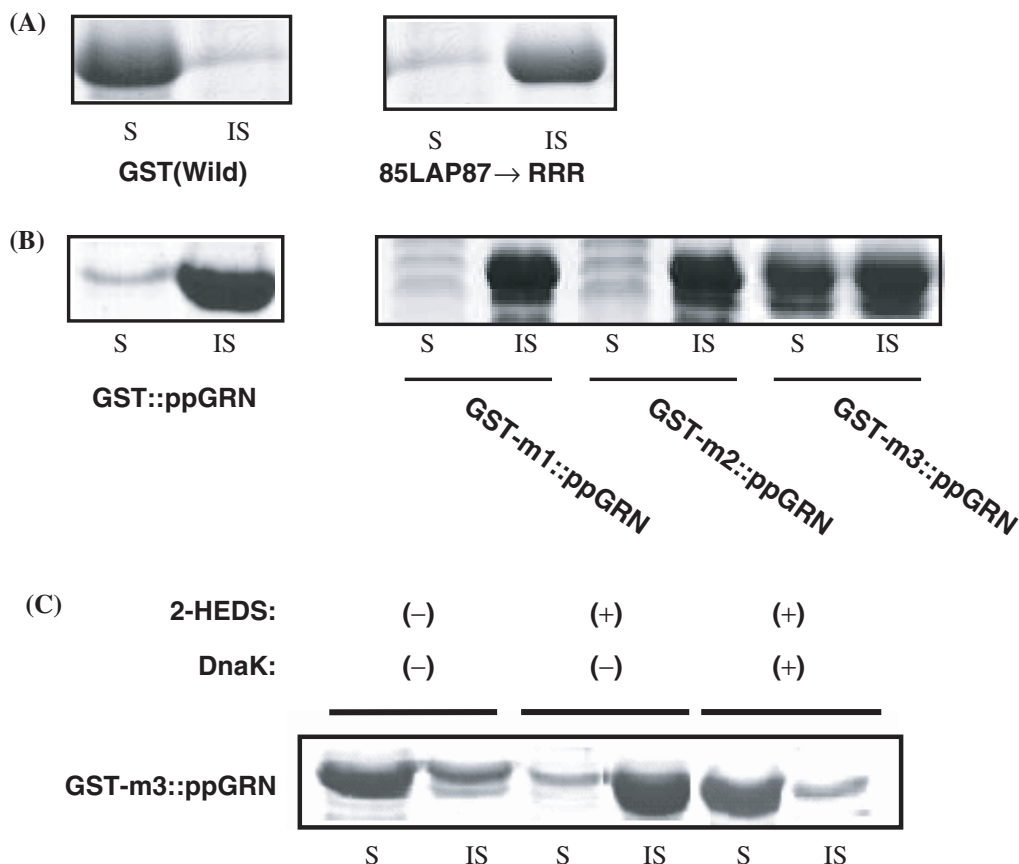
(A)

```
FTN-H     1  MTTASTSQVRQNYHQDSEAAINRQINLELYASYVYLSMSYYFDRDDVALK      50
             .:||:||||..|.|||:|..:||.|.|||.|| :..:||||||||||:
FTN-L     1       MSSQIRQNYSTDVEAAVNSLVNLYLQASYTYLSLGFYFDRDDVALE      46

FTN-H    51  NFAKYFLHQSHEEREHAEKLMKLQNQRGGRIFLQDIKKPDCDDWESGLNA     100
             ..:.:|...:.|:||..|:|:|:|||||||...||||..|:|....:|
FTN-L    47  GVSHFFRELAEEKREGYERLLKMQNQRGGRALFQDIKKPAEDEWGKTPDA      96

FTN-H   101  MECALHLCKNVNQSLLELHKLATDKNDPHLCDFIETHYLNEQVKAIKELG     150
             |:.|:.|  |.:||:||:||.|.:.:.|||||||:|||:|:|:||.||::|
FTN-L    97  MKAAMALCKKLNQALLDLHALGSARTDPHLCDFLETHFLDEEVKLIKKMG     146

FTN-H   151  DHVTNLRKMGAPESGLAEYLFDKHTLGDSDNES     183
             ||:|||.:||.||:||.||||:::.||...
FTN-L   147  DHLTNLHRLGGPEAGLGEYLFERLTLKHD     175
```

(B)

(C)

Figure 6. (**A**) Amino acid sequence alignment between FTN-H and FTN-L. (**B**) Schematic representation and *E.coli* cytoplasmic solubility of FTN-L (wild) and FTN-L mutant (77IFL79) containing a putative DnaK-binding motif from FTN-H. (**C**) Cytoplasmic solubility of recombinant FTN-L mutant in the presence (+) or absence (−) of 2-HEDS with (+) or without (−) DnaK co-expression in the recombinant *E.coli* culture.

the initial stage of protein folding. Therefore, the important factor determining the DnaK binding may not be protein conformation but the sequence structure of the binding motif. The sequence, GRIFLQD, consists of hydrophilic (polar) amino acid residues flanking the hydrophobic tripeptide. In the middle of the tripeptide, an amino acid with an aromatic side chain (Phe) is located (Figure 8). Presumably,

such a unique sequence in the extended polypeptide chain can be easily recognized by DnaK.

In vitro DnaK-binding assay. Through an *in vitro* assay experiment, we investigated the actual binding of DnaK to the specific peptide GRIFLQD (named 'peptide A') above. In addition to the peptide A, we prepared another peptide,

Figure 7. (**A**) Solubility of GST (wild) and GST mutant (85RRR87) in *E.coli* cytoplasm. (**B**) Cytoplasmic solubility of recombinant fusion proteins of ppGRN, prepared through N-terminus fusion of GST (wild) and three GST mutants (GST1-GST3) derived from the GST mutant (85RRR87). (**C**) Cytoplasmic solubility of recombinant GST-m3::ppGRN in presence (+) or absence (−) of 2-HEDS with (+) or without (−) DnaK co-expression in the recombinant *E.coli* culture.
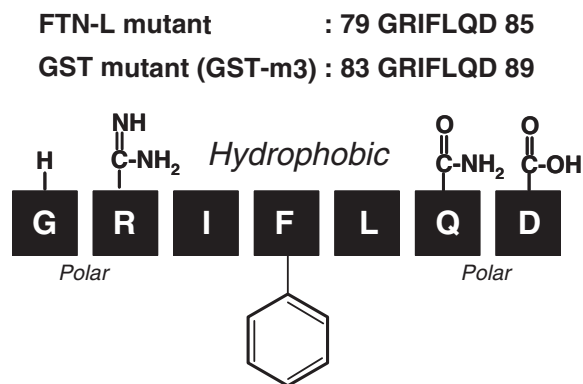
Figure 8. Amino acid sequence of putative DnaK-binding motif derived from FTN-H, which was commonly contained in the sequence of both FTN-L mutant (77IFL79) and GST3(83GRILFQD89)::ppGRN.

GRALFQD (named 'peptide B'), which exactly corresponds to the 79th–85th amino acids of FTN-L that were replaced by peptide A upon the preparation of the FTN-L mutant (Figure 6B). From the results of Ponceau S staining (Figure 9A), it seems obvious that both peptides A and B were well immobilized on the PVDF membrane. The addition of DnaK to the immobilized peptides and the subsequent

immunoblotting analysis using anti-DnaK monoclonal Ab showed that DnaK bound only to peptide A (Figure 9A), which clearly shows that the sequence 83GRILFQD89 within FTN-H has a high binding affinity for DnaK.

We also implemented an additional *in vitro* assay of DnaK binding to FTN-H and the FTN-L mutant containing GRIFLQD, and to wild type FTN-L that did not contain GRIFLQD (negative control). Since DnaK binds to an unfolded nascent chain at the early stage of folding, we performed the overlay assay using purified and unfolded FTN-H, FTN-L and FTN-L mutant. (The purified proteins were unfolded using a denaturation buffer containing 1 mM DTT and 8 M urea.) Evidently, from Figure 9B, DnaK bound to only the proteins containing GRIFLQD, i.e. FTN-H and the FTN-L mutant, whereas DnaK binding never happened to wild-type FTN-L. Through this additional *in vitro* assay, we confirmed that the peptide motif GRILFQD (i.e. the 83rd–89th amino acids of FTN-H) has a high binding affinity to DnaK and seems to be directly related to the synthesis of soluble heterologous proteins.

Supra-molecule formation by self-assembly function of N-terminus FTN-H

From TEM images of the purified FTN-H-fusion proteins (Figure 10A), it is very interesting that all the fusion proteins
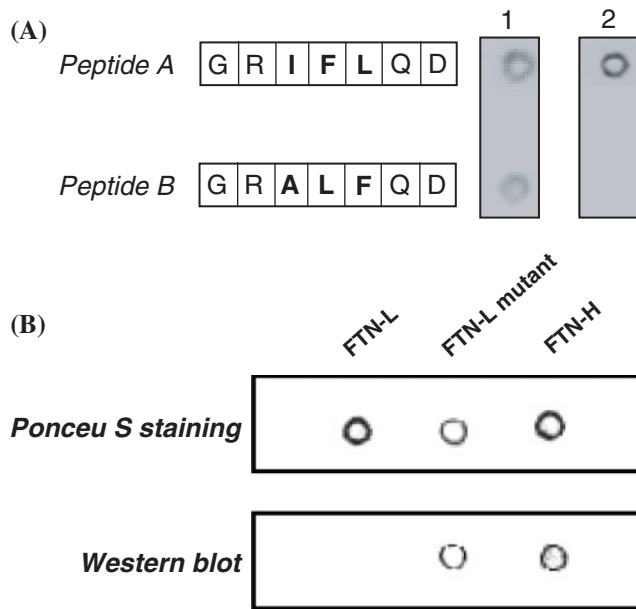
(A)

(B)

Figure 9. (**A**) Result of *in vitro* assay (Materials and Methods) to find binding affinity between DnaK and synthetic peptides (A and B). Lane 1, result of Ponceau S staining of immobilized peptides A and B prior to DnaK binding; lane 2, result of immunoblotting using anti-DnaK monoclonal Ab). (**B**) Result of *in vitro* assay (Materials and Methods) to find binding affinity between DnaK and recombinant proteins (FTN-H, FTN-L and FTN-L mutant).

formed spherical nanoparticles with diameters of 10–15 nm, i.e. non-covalently cross-linked supra-molecules by self-assembly. The particle size seems to be quite uniform for each fusion protein. Considering that all the heterologous proteins (EGF, IL-2, ppGRN, GAD$_{512-585}$ and mpINS) were aggregated into inclusion bodies when directly expressed in *E.coli*, the results imply that the supra-molecule formation in the shape of a spherical shell may provide significant advantage(s) in preventing protein aggregation. It seems that upon supra-molecule formation, there must be a limitation in the increase of the supra-molecule size due to the assembly properties of FTN-H. Therefore, self-assembled supra-molecules were produced with a constant size of 10–15 nm, thereby avoiding the undesirable formation of insoluble macroaggregates of heterologous proteins.

Immunoblotting analysis (western blot) following SDS–PAGE analysis of the purified FTN-H::ppGRN showed an apparent immunoreactive band (Figure 10B). When western blot was carried out after native PAGE, however, no immunoreactive bands were visualized (Figure 10C), which clearly indicates that a part of ppGRN, recognized by anti-ppGRN monoclonal Ab, was localized inside the supra-molecules. Thus, another plausible advantage of the self-assembled supra-molecule is its prevention of protein aggregation by non-specific hydrophobic interaction through the localization of critical hydrophobic patches of the target protein inside the supra-molecule.

DISCUSSION

As newly translated polypeptides emerge from the ribosome, they face a formidable task in the crowded environment of the cell: folding into their correct structure for proper functioning. Molecular chaperones are essential for the correct folding of a significant fraction of cellular proteins under both physiological and stress conditions (24), acting in a complex network as a main defense against protein aggregation. Some chaperones exhibit a holding activity, preventing polypeptides from aggregation (25–27). Others show a complementary folding role to assist in refolding and solubilization from aggregates (28,29). Chaperones also minimize aggregation by mediating the degradation of proteins that cannot be properly folded (30). The specific contribution of particular chaperones to this multifunctional folding network has only been partially identified. The Hsp70 chaperone DnaK serves as a central element of the chaperone network, assisting in *de novo* protein folding, protein translocation, oligomer dissociation and prevention of stress-induced protein aggregation (31–33). DnaK mediates protein folding through ATP-dependent interaction with short linear peptide segments that are exposed on unfolded proteins (34).

The heterologous expression of various human proteins (EGF, IL-2, ppGRN, GAD$_{512-585}$, ADI$_{132-410}$ and mp-INS) was attempted through the fusion of the target protein to the C-terminus of FTN-H, all of which were produced as inclusion bodies when directly expressed in *E.coli*. As a result, the intracellular solubility of recombinant fusion proteins (FTN-H::EGF, FTN-H::IL-2, FTN-H::ppGRN, FTN-H::GAD$_{512-585}$, FTN-H::ADI$_{132-410}$ and FTN-H::mp-INS) was significantly enhanced (Figure 1). As explained above, DnaK plays a central role in the initial folding of the nascent polypeptide chain, and thus it seems reasonable to assume that FTN-H has a significant interaction with DnaK and contains specific motif(s) that DnaK efficiently recognizes and binds to. The results shown in Figures 3 and 9 strongly suggest that this assumption is not irrational. Recently, Vandenbroeck *et al.* (17) identified conserved DnaK-binding sites in the N- and C-terminal halves of helix B and C of diverse 4-helix-bundle proteins, respectively, that belong to the superfamily of interferon-γ/interleukin-10-related cytokines. In this study, we also attempted to identify the specific amino acid residues of FTN-H, if any, with a high affinity to DnaK. Through systematic investigation using site-directed mutagenesis, the oxidative and forced protein aggregation technique and the DnaK co-expression vector system (Figure 4), several putative binding motifs were selected. It has been finally confirmed through an *in vitro* binding assay that DnaK recognizes an extended peptide strand composed of a hydrophobic core of three residues (Ile, Phe and Leu) and two flanking regions enriched in polar residues (Gly, Arg, Gln and Arg), such as the sequence GRIFLQD (Figure 8), which seems similar to the findings of Rudiger *et al.* (35). They reported that the DnaK-binding motif consists of a hydrophobic core of four or five residues enriched particularly not only in Leu, but also in Ile, Val, Phe and Tyr and two flanking regions enriched in basic residues. Although recombinant FTN-L was all aggregated into inclusion bodies in *E.coli* cytoplasm, the cytoplasmic solubility of the recombinant FTN-L mutant containing the putative DnaK-binding motif (GRIFLQD) was significantly increased (Figure 6B). The efficacy of the binding motif was also validated using a GST mutant (GST-m3) containing the same motif sequence, i.e. the fusion protein GST-m3::ppGRN was expressed as a cytoplasmic soluble protein. The sequence
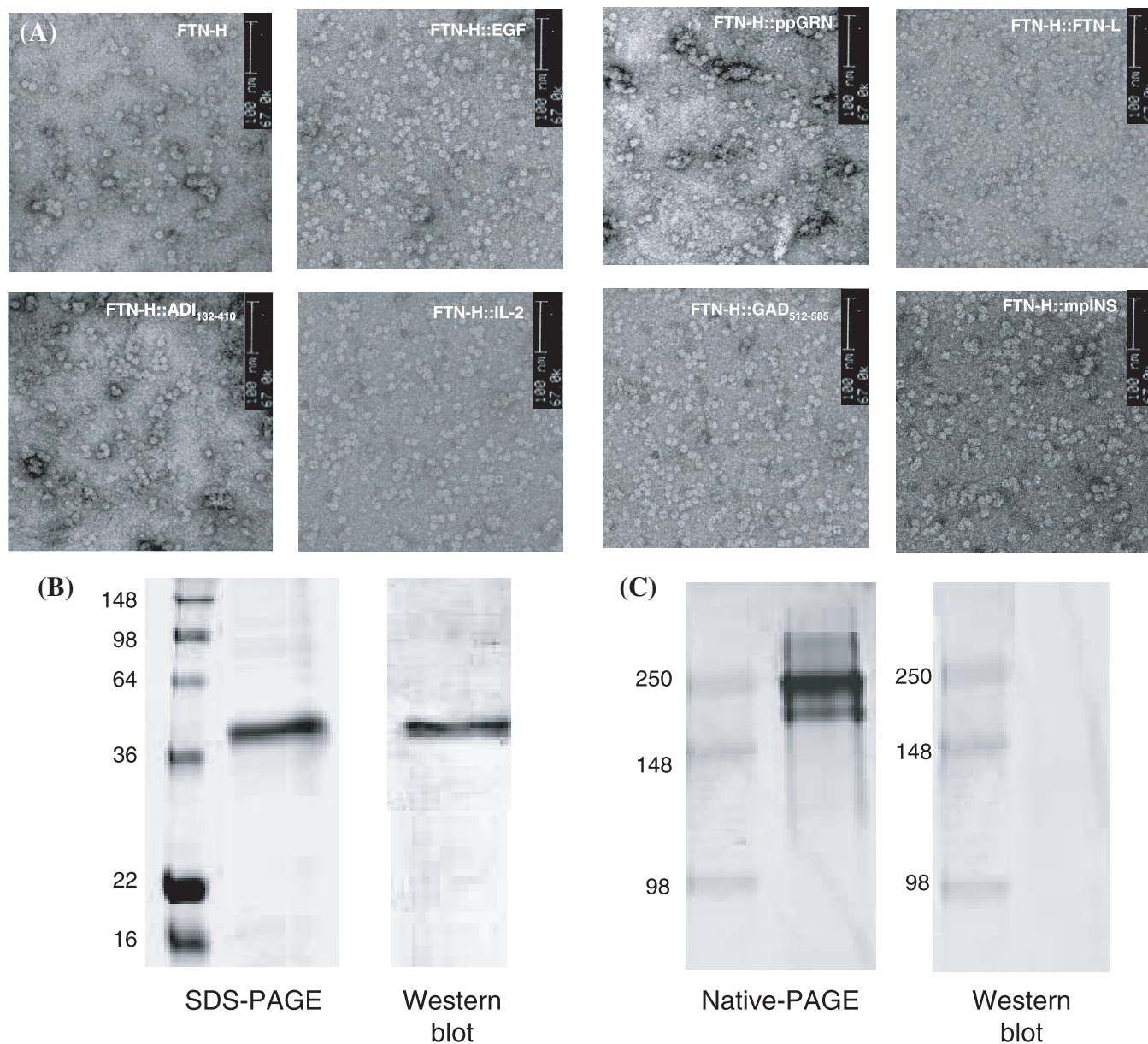
Figure 10. (**A**) TEM images of recombinant proteins (FTN-H, FTN-H::EGF, FTN-H::ADI$_{132–410}$, FTN-H::IL-2, FTN-H::ppGRN, FTN-H::FTN-L, FTN-H::GAD$_{512–585}$, FTN-H::ADI$_{132–410}$ and FTN-H::mpINS) purified after expressed in *E.coli* cytoplasm. (**B**) Result of western blot analysis following reducing SDS–PAGE of purified FTN-H::ppGRN. (**C**) Result of western blot analysis following native PAGE of purified FTN-H::ppGRN.

GRIFLQD of FTN-H is located in the long loop between the second and the third helix of the FTN-H, and the tertiary structure of the FTN-H shows that the flexible long loop containing the binding motif is localized in the protein exterior. It is therefore likely that association and dissociation of DnaK at the binding motif would be quite efficient at the early stage of protein folding.

Similar to native human ferritin that is heteropolymer (with the shape of a spherical shell) of 24 subunits of two types, heavy chain and light chain, all the recombinant proteins, including FTN-H, FTN-H::EGF, FTN-H::IL-2, FTN-H::ppGRN, FTN-H::GAD$_{512–585}$, FTN-H::ADI$_{132–410}$ and FTN-H::mp-INS, have formed stable homopolymers (i.e. supra-molecules) by non-covalent self-association. This

was confirmed by the TEM images showing nanometer-scale spherical particles of the recombinant proteins (Figure 10A). Recently, Louzada *et al.* (36) proposed the mechanism of a quaternary structure (dimer) formation in the folding pathway of glutathione reductase (GR). They proposed that the assembly of the native GR dimer proceed via the initial formation of molten-globule monomers that dimerize to form a partially folded, expanded dimer. Structural rearrangements then take place at the dimer interface, leading to a final, compact dimer structure. Based on this proposed mechanism, it is presumed that the intermolecular association of ferritin (or ferritin-fusion) monomers to form homopolymeric supra-molecules happens at the level of the partially folded (molten-globule) intermediate, and not at the level of the completely folded and
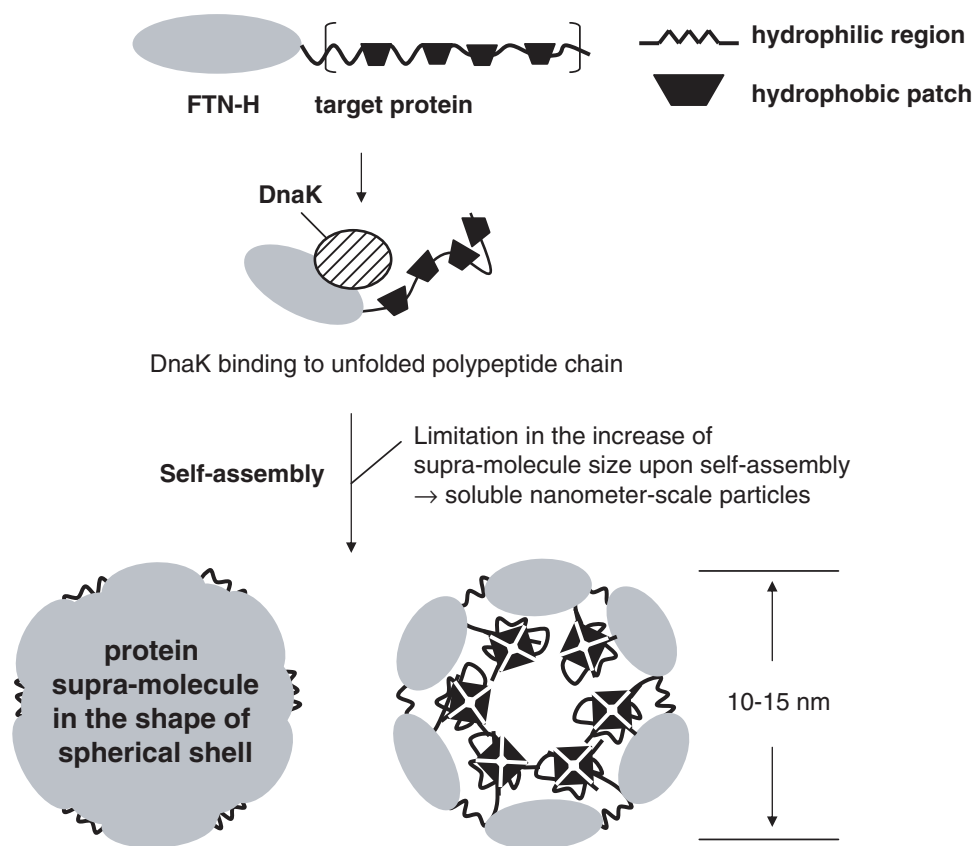
Figure 11. Schematic representation of dual function of FTN-H used for fusion expression of heterologous proteins.

thermodynamically optimal polypeptide chain. Protein aggregation, i.e. the off-pathway aggregation step, often initiates from partially folded intermediates upon intracellular chain folding and association (37). Within cells, nascent chains form sequentially on the ribosome, partially fold, and sometimes fail to reach their native state because of the intervention of various factors in a heterologous environment, including non-specific hydrophobic interactions with the intracellular cytoplasmic proteins (38), which sometimes cause the formation of large and insoluble protein aggregates, i.e. inclusion bodies. Supposing that as assumed earlier, the supramolecules of FTN-H or its fusion mutants are formed through the self-association of partially folded intermediates, they may have a distinct advantage in preventing the formation of insoluble protein aggregates; i.e. upon supra-molecule formation, the supra-molecule size would be limited by the assembly properties of FTN-H, and therefore, self-assembled supramolecules with constant sizes (10–15 nm) could be produced, thereby avoiding the undesirable formation of insoluble macro-aggregates of recombinant proteins (Figure 11).

As shown in Figure 11, a part of a target protein can be localized within the spherical shell of the supra-molecule upon the intermolecular association of a partially folded ferritin monomer, and thus, a protective environment against undesirable protein aggregation can be created. Evidently, from the results of the native PAGE and western blot analysis for FTN-H::ppGRN (Figure 10B and C), an anti-ppGRN mAb-specific epitope was shown to be localized inside the supra-molecule. Hydrophobic patches of ppGRN may be

hidden together inside, which, otherwise, may be involved in a non-specific hydrophobic interaction with the adjacent polypeptide chains, leading to protein aggregation.

In conclusion, this study proposes the dual function of FTN-H used for the fusion expression of heterologous proteins: (i) high-affinity interaction and association with DnaK and (ii) formation of self-assembled supra-molecules with limited and constant sizes, thereby preventing the undesirable formation of insoluble macro-aggregates of heterologous proteins and possibly creating a protective environment against protein aggregation. Similarly, Hidaka *et al.* (18) found that the propeptide of guanylyl cyclase activating peptide II (GCAP-II), an endogenous ligand of guanylyl cyclase C, has a dual function in the proper folding of the mature peptide. Thus, the N-terminus propeptide of pro-GCAP-II is critical not only for disulfide-coupled folding but also for the net stabilization (or dimerization) of pro-GCAP-II. The ferritin supra-molecular structure at the nanometer scale (or spherical protein nanoparticles) can contribute to the enhancement of the stability of synthesized recombinant proteins and can also be used as a potent 3D probe structure for an ultra-sensitive protein chip if a special expression system is designed so that the captured probe peptide can be properly displayed at the surface of the ferritin nanoparticles.

ACKNOWLEDGEMENTS

REFERENCES

1. Yang,X., Arosio,P. and Chasteen,D. (2000) Molecular diffusion into ferritin: pathways, temperature dependence, incubation time and concentration effects. *Biophys. J.*, **78**, 2049–2059.
2. Harrison,P.M. and Arosio,P. (1996) The ferritins: molecular properties, iron storage function and cellular regulation. *Biochim. Biophys. Acta*, **1275**, 161–203.
3. Treffry,A., Zhao,Z., Quail,M.A., Guest,J.R. and Harrison,P.M. (1998) How the presence of three iron binding sites affects the iron storage function of the ferritin (EcFtnA) of *Escherichia coli. FEBS Lett.*, **432**, 212–218.
4. Balla,G., Jacob,H.S., Balla,J., Rosenberg,M., Nath,K., Apple,F., Eaton,J.W. and Vercellotti,G.M. (1992) Ferritin: a cytoprotective antioxidant strategem of endothelium. *J. Biol. Chem.*, **267**, 18148–18153.
5. Lawson,D.M., Artymiuk,P.J., Yewdall,S.J., Smith,J.M., Livingstone,J.C., Treffry,A., Luzzago,A., Levi,S., Arosio,P. and Cesareni,G. (1991) Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. *Nature*, **349**, 541–544.
6. Liu,X., Jin,W. and Theil,E.C. (2003) Opening protein pores with chaotropes enhances Fe reduction and chelation of Fe from the ferritin biomineral. *Proc. Natl Acad. Sci. USA*, **100**, 3653–3658.
7. Santambrogio,P., Levi,S., Cozzi,A., Rovida,E., Albertini,A. and Arosio,P. (1993) Production and characterization of recombinant heteropolymers of human ferritin H and L chains. *J. Biol. Chem.*, **268**, 12744–12748.
8. Guo,J.H., Juan,S.H. and Aust,S.D. (1998) Mutational analysis of the four alpha-helix bundle iron-loading channel of rat liver ferritin. *Arch. Biochem. Biophys.*, **352**, 71–77.
9. Lee,J., Kim,S.W., Kim,Y.H. and Ahn,J.Y. (2002) Active human ferritin H/L-hybrid and sequence effect on folding efficiency in *Escherichia coli. Biochem. Biophys. Res. Commun.*, **298**, 225–229.
10. Allen,S.P., Polazzi,J.O., Gierse,J.K. and Easton,A.M. (1992) Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli. J. Bacteriol.*, **174**, 6938–6947.
11. Bardwell,J.C. and Craig,E.A. (1988) Ancient heat shock gene is dispensable. *J. Bacteriol.*, **170**, 2977–2983.
12. Bjellqvist,B., Sanchez,J.C., Pasquali,C., Ravier,F., Paquet,N., Frutiger,S., Hughes,G.J. and Hochstrasser,D. (1993) Micropreparative two-dimensional electrophoresis allowing the separation of samples containing milligram amounts of proteins. *Electrophoresis*, **14**, 1375–1378.
13. Buchberger,A., Valencia,A., McMacken,R., Sander,C. and Bukau,B. (1994) The chaperone function of DnaK requires the coupling of ATPase activity with substrate binding through residue E171. *EMBO J.*, **13**, 1687–1695.
14. Buchberger,A., Schröder,H., Hesterkamp,T., Schönfeld,H.-J. and Bukau,B. (1996) Substrate shuttling between the DnaK and GroEL systems indicates a chaperone network promoting protein folding. *J. Mol. Biol.*, **261**, 328–333.
15. Buchberger,A., Reinstein,J. and Bukau,B. (1999) The DnaK chaperone system: mechanism and comparison with other Hsp70 systems. In Bukau,B. (ed.) *Molecular Chaperones and Folding Catalysts: Regulation, Cellular Function and Mechanisms*. Harwood Academic Publishers, Amsterdam, The Netherlands, pp. 609–635.
16. Bukau,B. and Horwich,A.L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell*, **92**, 351–366.
17. Vandenbroeck,K., Alloza,I., Brehmer,D., Billiau,A., Proost,P., McFerran,N., Rudiger,S. and Walker,B. (2002) The conserved helix C region in the superfamily of interferon-gamma/interleukin-10-related cytokines corresponds to a high-affinity binding site for the HSP70 chaperone DnaK. *J. Biol. Chem.*, **277**, 25668–25676.
18. Hidaka,Y., Ohno,M., Hemmasi,B., Hill,O., Forssmann,W.G. and Shimonishi,Y. (1998) *In vitro* disulfide-coupled folding of guanylyl cyclase-activating peptide and its precursor protein. *Biochemistry*, **37**, 8498–8507.
19. Pogulis,R.J., Vallejo,A.N. and Pease,L.R. (1996) *In vitro* recombination and mutagenesis by overlap extension PCR. *Methods Mol. Biol.*, **57**, 167–76.
20. Lee,J., Kim,H.C., Kim,S.W., Kim,S.W., Hong,S.I. and Park,Y.H. (2002) Interplay of SOS induction, recombinant gene expression and multimerization of plasmid vectors in *Escherichia coli. Biotechnol. Bioeng.*, **80**, 84–90.
21. Boyde,T.R. and Rahmatullah,M. (1980) Optimization of conditions for the colorimetric determination of citrulline, using diacetyl monoxime. *Anal. Biochem.*, **15**, 424–31.
22. Kim,S.W., Kim,Y.H. and Lee,J. (2001) Thermal stability of human ferritin: concentration dependence and enhanced stability of an N-terminal fusion mutant. *Biochem. Biophys. Res. Commun.*, **289**, 125–129.
23. Shin,C.S., Hong,M.S., Bae,C.S. and Lee,J. (1997) Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of *Escherichia coli* BL21(DE3)[pET-3aT2M2]. *Biotechnol. Prog*, **13**, 249–257.
24. Tomoyasu,T., Mogk,A., Langen,H., Goloubinol,P. and Bukau,B. (2001) Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol. *Mol. Microbiol.*, **40**, 397–413.
25. Ehrnsperger,M., Graber,S., Gaestel,M. and Buchner,J. (1997) Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation. *EMBO J.*, **16**, 221–229.
26. Veinger,L., Diamant,S., Buchner,J. and Goloubinol,P. (1998) The small heat-shock protein IbpB from *Escherichia coli* stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *J. Biol. Chem.*, **273**, 11032–11037.
27. Mogk,A., Tomoyasu,T., Goloubinol,P., Rudiger,S., Roder,D., Langen,H. and Bukau,B. (1999) Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. *EMBO J.*, **18**, 6934–6949.
28. Parsell,D.A., Kowal,A.S., Singer,M.A. and Lindquist,S. (1994) Protein disaggregation mediated by heat-shock protein Hsp104. *Nature*, **372**, 475–478.
29. Ben-Zvi,A.P. and Goloubinol,P. (2001) Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. *J. Struct. Biol.*, **135**, 84–93.
30. Huang,H.C., Sherman,M.Y., Kandror,O. and Goldberg,A.L. (2001) The molecular chaperone DnaJ is required for the degradation of a soluble abnormal protein in *Escherichia coli. J. Biol. Chem.*, **276**, 3920–3928.
31. Gottesman,S., Wickner,S. and Maurizi,M.R. (1997) Protein quality control: triage by chaperones and proteases. *Genes Dev.*, **11**, 815–823.
32. Bukau,B., Schmid,F.X. and Buchner,J. (1999) Assisted protein folding. In Bukau,B. (ed.) *Molecular Chaperones and Folding Catalysts: Regulation, Cellular Function and Mechanisms*. Harwood Academic Publishers, Amsterdam, The Netherlands, pp. 3–10.
33. Hartl,F.U. and Hayer-Hartl,M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, **295**, 1852–1858.
34. Ewalt,K.L., Hendrick,J.P., Houry,W.A. and Hartl,F.U. (1997) *In vivo* observation of polypeptide flux through the bacterial chaperonin system. *Cell*, **90**, 491–500.
35. Rudiger,S., Germeroth,L., Schneider-Mergener,J. and Bukau,B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.*, **16**, 1501–1507.
36. Louzada,P.R., Sebollela,A., Scaramello,M.E. and Ferreira,S.T. (2003) Predissociated dimers and molten globule monomers in the equilibrium unfolding of yeast glutathione reductase. *Biophys. J.*, **85**, 3255–3261.
37. Hasse-Pettingell,C.A. and King,J. (1988) Formation of aggregates from a thermolabile *in vivo* folding intermediate in P22 tailspike maturation. A model for inclusion body formation. *J. Biol. Chem.*, **263**, 4977–4983.
38. Mitraki,A. and King,J. (1992) Amino acid substitutions influencing intracellular protein folding pathways. *FEBS Lett.*, **307**, 20–25.