Glucagon-Induced Self-Association of Recombinant Proteins in *Escherichia coli* and Affinity Purification Using a Fragment of Glucagon Receptor

Dae-Young Kim, Jeewon Lee, Vibhor Saraswat, Young-Hoon Park

Microbial and BioProcess Engineering Laboratory, Korea Research Institute of Bioscience and Biotechnology (KRIBB), P.O. Box 115, Yusong, Taejon 305–600, South Korea; telephone: + 82 42 860 4449; fax: + 82 42 860 4594; e-mail: jwlee@kribb4680.kribb.re.kr

Received 29 August 1999; accepted 20 March 2000

Abstract: The specific molecular interactions of α-helical peptide, human glucagon (i.e., intermolecular self-association and specific receptor-binding affinity) provided a rationale for using the glucagon as the fusion expression partner to achieve high productivity of foreign proteins both in vivo (in bacterial fusion-expression system) and in vitro (in affinity column chromatography). The fusion of glucagon peptide(s) effectively promoted homogeneous aggregate formation of recombinant proteins while avoiding intermolecular crosslinking by disulfide bridges. High sensitivity of the self-aggregation to sequence effects resulted from two distinct nonpolar domains of glucagon, determining specificity of molecular interaction and aggregate size of recombinant proteins. An N-terminal domain of glucagon molecule (Phe6-Tyr10-Tyr13) could be a certain hydrophobic moiety involved in intermolecular self-association (probably, via helix–helix docking), while a C-terminal domain (Phe22-Trp25-Leu26) seems to critically affect the oligomer size in the off-pathway aggregation of synthesized fusion proteins. An N-terminal extracellular domain of human glucagon receptor was recombinantly expressed in *Escherichia coli*, immobilized to a chromatography column, and efficiently renatured to a conformation that attains high specificity in interaction with N-terminus glucagon molecules of recombinant fusion proteins. Through column chromatography employing the receptor fragment as affinity ligand, the recombinant proteins were efficiently purified from total intracellular proteins, and the long-term ligand stability was evidently proven through multiple cyclic-purification experiments. Major scaffolds for using protein ligands are large-scale production in a low-cost expression system and long-term stable operation with selective-binding affinity. From this point of view, the extracellular fragment of human glucagon receptor used in this study seems to be a new potent ligand for fusion protein-based affinity chromatography. © 2000 John Wiley & Sons, Inc. *Biotechnol Bioeng* **69**: 418–428, 2000.

Keywords: glucagon; fusion protein; self-association; receptor; affinity purification

Correspondence to: Jeewon Lee

Contract grant sponsor: Ministry of Commerce, Industry, and Energy of South Korea

INTRODUCTION

In heterologous expression of many pharmaceutically important proteins, in vivo accumulation of polypeptide chains in the form of aggregated non-native states is a potential fate in bacterial cytoplasm (Hendrick and Hartl, 1993; Wetzel, 1994). Because isolation of the inclusion bodies is the most efficient initial step in the purification process, followed by refolding in vitro (Cleland, 1993), in vivo formation of homogeneous aggregates (which favors efficient conversion of non-native aggregates to native state) by a specific intermolecular interaction is of great advantage compared to covalently crosslinked heterogeneous multimers.

The inclusion body problem is one of intracellular polypeptide folding and association. During the in vitro refolding of certain proteins, aggregates often form from partially folded intermediates in an off-pathway step, and the properties of inclusion bodies are consistent with a similar origin: off-pathway steps from intermediates in intracellular chain folding and association pathways (Hasse-Pettingell and King, 1988) Within cells, nascent chains are forming sequentially on the ribosome, and must reach their native state without the intervention of a major change of heterologous environment. The ribosome, factors interacting with the nascent chains, cofactors and prosthetic groups, cytoplasmic membranes, chaperones, and the intracellular cytoplasmic environment including temperature and ionic composition, are all factors that can influence the outcome of in vivo folding processes (Mitraki and King, 1989). However, property inherent in the protein sequence during the folding process is most important as shown in the results (Chrunyk et al., 1993; Mitraki et al., 1993; Wetzel, 1992) of specific mutational effects on the folding-dependent aggregation of monomeric, as well as oligomeric, proteins. Hence, there would appear to be opportunities to manipulate intermolecular aggregation using a specific sequence in an engineered protein. Similar to the propeptides of some zymogens (e.g., bacterial subtilisin, α-lytic proteases, and carboxypeptidase) acting as namely cotranslational chaperones (i.e., mediating correct folding without being components of

the final three-dimensional structure) (Ikemura et al., 1987; Saen and Agard, 1989; Silen et al., 1989; Winther and Sorensen, 1991; Zhu et al., 1989), a specific propeptide fused to a mature protein sequence could have a dramatic effect on conformation of whole fusion protein and hence, an intrinsic tendency to aggregate as well.

Glucagon is a 29-residue peptide hormone that antagonizes insulin action and thereby leads to a stimulation of hepatic glucose production (Unger and Dobbs, 1978). X-ray analysis has demonstrated that in crystals the peptide adopts a mainly α-helical conformation, and in dilute solution, the glucagon molecules have a strong tendency to stabilize the α-helical conformation by hydrophobic interactions either as an oligomer by self-association between molecules or as a complex with the receptor (Sasaki et al., 1975). The binding of glucagon to receptors on liver cells initiates a signal-transduction cascade resulting in the production of glucose by glycogenolysis and gluconeogenesis (Unger and Orci, 1981). These dual properties of glucagon regarding intermolecular association provide a rationale for using glucagon and receptor as a fusion expression partner and affinity ligand in column chromatography, respectively, in the production and purification of recombinant proteins from *E. coli*. Protein–protein interactions can confer tremendous specificity in separation processes, and especially in fusion protein-based affinity purification, the exploitation of these specific interactions can translate into significant simplification and savings in downstream processing (Przybycien, 1998). Glucagon receptor belongs to a subfamily of peptide hormone receptors that are characterized by a seven transmembrane-domain structure and their coupling via GTP-binding proteins (G-proteins) to adenyl cyclase (Lok et al., 1994). The cDNA of human glucagon receptor was recently isolated, and the sequence and structure were determined (Lok et al., 1994). Also, the glucagon-binding site was presumably shown to be an amino-terminal extension which is also involved in antibody recognition (Buggy et al., 1995).

The present work describes: (1) the production of recombinant therapeutic protein [human interleukin-2 (rhIL-2) as a model protein] from *E. coli* using glucagon as N-terminus fusion peptide, and (2) the affinity purification of the recombinant fusion protein using a fragment of glucagon receptor (GR120H) as a glucagon-binding ligand. We found that by glucagon fusion, homogeneous aggregate formation was effectively induced in vivo while avoiding the intermolecular crosslinking by disulfide bridges. The sequence effects of glucagon molecule on interaction specificity and association kinetics of folding intermediates were examined through various immunoblotting analyses. Efficacy of the GR120H ligand used for affinity column chromatography was also demonstrated in detail in the purification of two recombinant fusion proteins, fusion rhIL-2, and fusion magainin-2.

MATERIALS AND METHODS

Strains and Plasmid Vectors

Recombinant *E. coli* strain KCTC (Korean Culture-Type Collection 8258P) contains plasmid pNKM21 which consists of a structural gene for human interleukin-2 mutant (Cys125Ser) (hIL-2, 15.5 kDa). The entire coding sequence of hIL-2 gene in plasmid pNKM21 was amplified by PCR (Bio-Rad, Hercules, CA) using primers, pNIL (5′ GCA CAT ATG GCA CCT ACT TCA AGT TCT 3′) and pHIL (5′ GCA AAG CTT CTA TTA AGT TAG TGT TGA GAT GAT 3′) and inserted into *Nde*I-*Hin*dIII site of plasmid pT7-7 (Tabor and Richardson, 1985) to construct recombinant plasmid pT7-IL2 (Fig. 1).

Two oligonucleotide fragments, GLU1 (5′ CAC TCT CAG GGT ACT TTC ACT TCT GAC TAC TCT AAA TAC CTG GAC TCT 3′) and GLU2 (5′ AGT GTT CAT CAG CCA CTG AAC GAA GTC CTG AGC ACG ACG AGA GTC CAG GTA TTT AGA GTA 3′), were synthesized using a DNA synthesizer (PerSeptive Biosystem, Inc., Forster City, CA). The coding sequence of human glucagon was then PCR-amplified using primers, Ga (5′ CAC TCT CAG GGT ACT 3′) and Gb (5′ AGT GTT CAT CAG CCA 3′) with the GLU1 and GLU2 fragments as templates. The sequences coding for three glucagon mutants, Gα (F6G, Y10G, Y13G), Gβ (F22G, W25G, L26G), and Gγ (R17F, Q20F, Q24F) were prepared by following the same amplification procedure using corresponding templates and primers. The coding sequences of 8th to 64th residue- (T_L) and 8th to 12th residue-peptide (T_S) of human TNF-α mutant (Cha et al., 1998) were also obtained as the products of PCR amplification using cDNA template of the human TNF-α mutant. Using the nucleotide sequences coding for hIL-2, glucagon (or mutants), and/or T_L and T_S peptides above as templates, such sequences encoding various fusion rhIL-2 (T_LG·, T_SG·, G·, G2·, G3·, G3Δ$(H)_6$·, Gα·, Gβ·, and Gγ· IL-2) (shown in detail in Fig. 1) that are to be inserted into *Nde*I-*Hin*dIII site of plasmid pT7-7, were PCR-amplified with appropriate primers, and finally nine corresponding plasmid vectors (pT7-T1GIL2, -T2GIL2, -GIL2, -G2IL2, -G3IL2, -G3ΔHIL2, -GαIL2, -GβIL2, -GγIL2, respectively) were constructed.

The entire coding sequence of magainin-2 (MSI-78, 2.6 kDa) (Jacob and Zasloff, 1994) was amplified with templates (5′ GCA CTC GAG ATG GGT ATC GGT AAA TTC CTT AAA AAG GCT AAG AAA TTC GGT 3′ and 5′ GCA AAG CTT CTA TTA CTT TTT AAG GAT TTT AAC GAA AGC TTT ACC GAA TTT 3′) and primers Ma and Mb, which introduced *Xho*I and *Cla*I restriction sites at positions encoding the N-terminal and C-terminal end of the peptide, respectively. Then, the *Xho*I-*Cla*I fragment in the plasmid vector pT7-G3ΔHIL2 including the coding region of hIL-2 was replaced with the amplified sequence of magainin-2 to construct the plasmid vector pT7-G3ΔHM for the synthesis of G3Δ$(H)_6$·*M*.

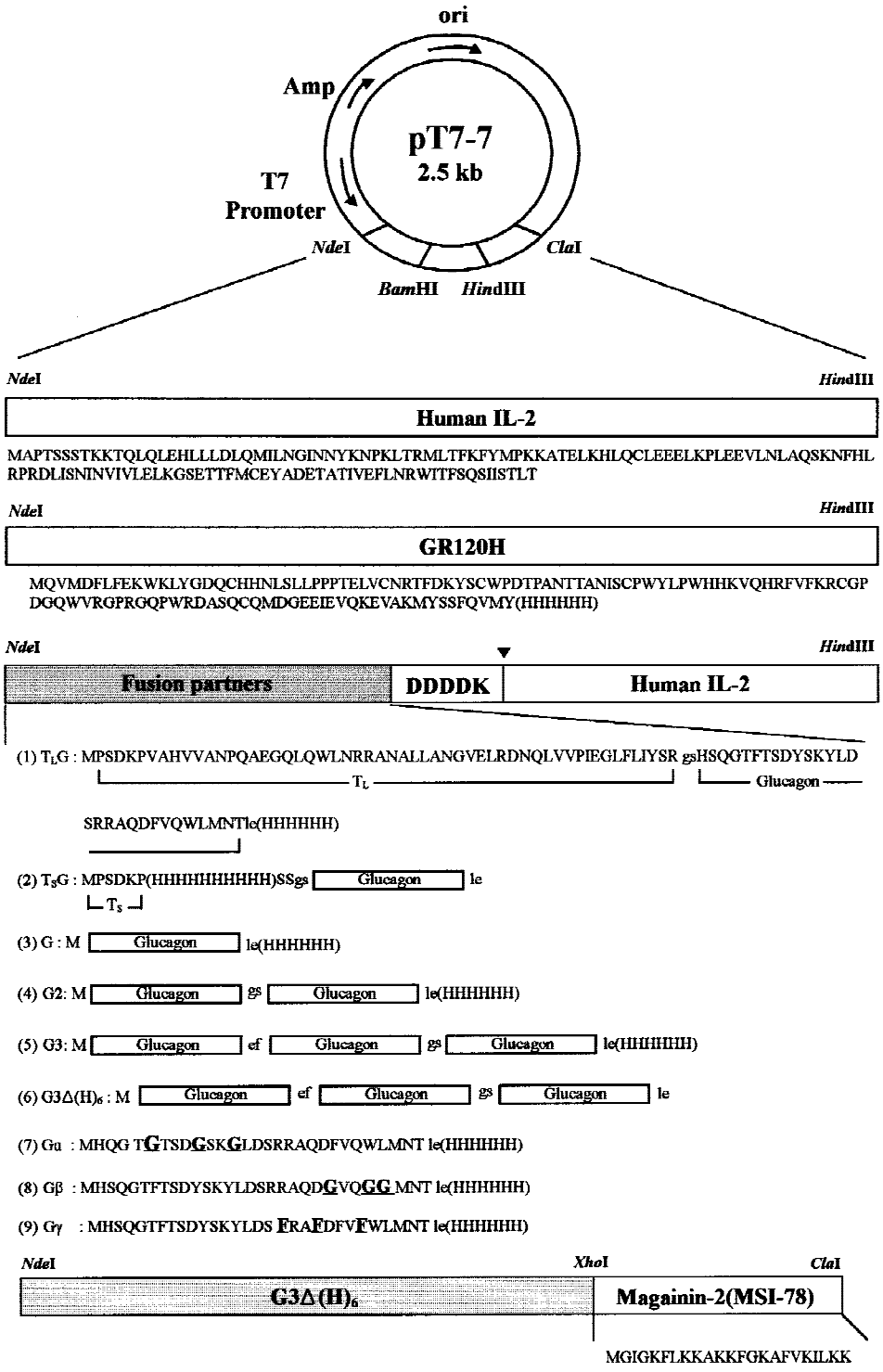

Figure 1. Schematic presentation of recombinant expression systems: Small letters in amino acid sequences represent restriction enzyme sites; polyhistidine tags were indicated in parentheses; new residues inserted by point mutation in the glucagon mutants (Gα, Gβ, and Gγ) were represented by bold and underlined letters; and the symbol (▼) indicates cleavage site for enterokinase. (Amp and ori: ampicillin resistance gene and origin of replication, respectively; *Nde*I, *Bam*HI, *Hin*dIII, and *Cla*I: restriction enzyme sites.)

The nucleotide sequence encoding extracellular 119 residues of human glucagon receptor (Fig. 1) was synthesized via four-stage PCR amplification where the products of 1st to 3rd PCR amplification (GR-1, GR-2, and GR-3 encoding 50, 73, and 96 residues, respectively) were used as template in next-stage PCR. As the product of 4th PCR amplification using primers introducing *Nde*I and *Hin*dIII sites at 5′ and 3′ end, respectively, the sequence coding for 126 residues

including N-terminal methionine and C-terminal 6 histidine residues was obtained and inserted into the *Nde*I-*Hin*dIII sites of the plasmid pT7-7 to construct the recombinant plasmid pT7-GR120HIS. After complete DNA sequencing (Sanger et al., 1977) of all plasmid vectors purified from transformant cells of *E. coli* XL1-blue, the *E. coli* strain BL21(DE3) [F^- *omp*T *hsd*S_B(rB^-mB^-)] was transformed with the constructed plasmid vectors above.

Culture Growth, Gene Expression, and Recovery of Insoluble Protein Aggregates

For the shake-flask experiments (250-mL Erlenmeyer flasks, 37°C, 200 rpm), LB media containing 100 mg ampicillin per liter of culture was used. When the culture $O.D._{600}$ reached 0.4, the gene expression was induced by adding IPTG (0.5 m*M*), and the induced cultures were harvested after further 3–4 h cultivation. The recombinant cells in 50-mL culture were spun down at 6000 rpm for 5 min, and the cell pellets were resuspended either in 5 mL distilled water or acidic solution (0.1*M* HCl) where applicable. After cell disruption by using Branson Sonifier (Branson Ultrasonics Corp., Danbury, CT), the insoluble protein aggregates were isolated at 5000 rpm for 10 min, washed twice with 1% Triton X-100, and subject to reducing SDS-PAGE analysis for estimating the purities of recombinant proteins. When necessary, the recovered inclusion bodies were subsequently dissolved (0.8 ~ 2 g/L) in 5 mL alkaline solution (pH 12) without adding any denaturing agents. The solution pH was quickly shifted back to 8 by adding 1*M* Tris-HCl buffer, pH 8, and the soluble protein aggregates in 50 m*M* Tris-HCl buffer were analyzed by various PAGE and subsequent Western blot.

In Vitro Coaggregation Experiment

Insoluble protein aggregates involving G·IL-2 or Gα·IL-2, isolated from the recombinant cells disrupted under acidic condition (0.1*M* HCl), were dissolved in 20 m*M* Tris-HCl buffer (pH 8.0) containing urea (2*M*). After adjusting concentration of the dissolved proteins to 0.3 g/L, the soluble proteins were reduced using 50 m*M* dithiothreitol (DTT) for 1 h at room temperature. The reduced protein solution was contained in dialysis membrane (Spectra/Por 3.1 with MW cut-off, 3.5 kDa, Spectrum Laboratories, Inc., Laguna Hills, CA) and incubated in the same Tris-HCl buffer for 3 d at 4°C. The buffer was exchanged with a fresh one every 18 h during the incubation, and the final coaggregation products were analyzed by reducing/non-reducing SDS-PAGE and Western blot.

Affinity Column Chromatography

Metal-chelating Sepharose (Pharmacia, Uppsala, Sweden) in 3 mL loading buffer (0.5*M* NaCl, 20 m*M* Tris-HCl, pH 8.0) was packed in a column, and nickel ion was chelated by loading 10 mL $NiSO_4$ solution (50 m*M*) at 2 mL/min using a peristaltic pump (Watson Marlow, Falmouth, Cornwall, England). Non-chelated nickel ion was washed out using the same Tris buffer (20 mL). Then, GR120H was bound to the Ni^{+2}-Sepharose column by loading the protein solution containing GR120H (0.2 g/L) at 2 mL/min for 30 min, and unbound GR120H and other impurities were removed from the column with additional 20 mL Tris buffer.

After the isolated intracellular aggregates involving $G3\Delta(H)_6 \cdot$ IL-2 or $G3\Delta(H)_6 \cdot M$ were dissolved by simple pH shift ($\rightarrow$12 $\rightarrow$8) as described earlier, urea, NaCl, and imidazole were supplemented. The recombinant protein [$G3\Delta(H)_6 \cdot$ IL-2 or $G3\Delta(H)_6 \cdot M$] in the resulting solution (0.16 g/L protein, 0.5 ~ 2*M* urea, 0.5*M* NaCl, 50 m*M* imidazole, and 20 *mM* Tris-HCl, pH 8) was purified through the column (GR120H-Ni^{+2}-Sepharose) chromatography using GR120H as an affinity ligand. The affinity column chromatography was operated by the following procedure: (1) loading 10 mL intracellular protein solution above (load-in) at 2 mL/min for 20 min with the effluent (load-out) recycle; (2) washing the column with 20 mL Tris-HCl buffer (20 m*M*, pH 8) containing 0.5*M* NaCl at 2 mL/min; (3) eluting the recombinant protein with 8*M* urea in 10 mL Tris-HCl buffer (20 m*M*, pH 8) at 1 mL/min; and (4) renaturing the immobilized GR120H with 10 mL Tris-HCl buffer (20 m*M*, pH 8) at 1 mL/min. When necessary, the operation of affinity column chromatography above was repeated through 50 cycles in purifying $G3\Delta(H)_6 \cdot$ IL-2, and in each cycle the operational procedure was the same as above except that the load of intracellular protein solution was made every 10 cycle. The purity of $G3\Delta(H)_6 \cdot$ IL-2 in the eluted solution was analyzed using C8 reverse-phase HPLC column (250 × 4.6 mm, YMC, Milford, MA, USA) with a binary buffer system (buffer A: 100% CH_3CN + 0.1% TFA; buffer B: 0.1% TFA). The samples were loaded onto the column and eluted with a linear gradient of 0–100% buffer A for 40 min at 1 mL/min. The elution profile was monitored at 220 nm.

Electrophoresis and Immunoblotting Analyses

Expression levels (%) of recombinant proteins were estimated by subjecting the total cell lysates to reducing (15 m*M* DTT) SDS-PAGE (14% Tris-glycine precast gel, Novex, San Diego, CA) with Coomassie staining. The resulting protein bands were scanned using a densitometer (Bio-Rad, Hercules, CA), and the percentage of recombinant protein in total cellular proteins was subsequently analyzed. Purities of recombinant proteins in the isolated inclusion bodies were measured through the same reducing SDS-PAGE and scanning analyses above. Non-reducing SDS-PAGE and native PAGE analysis were also conducted for appropriate analysis of synthesized recombinant proteins.

The multimer formation and intermolecular interactions of fusion rhIL-2 were examined by various Western blot analyses: the mouse monoclonal IgG1, anti-human interleukin-2 (BioSource International, Inc., Camarillo, CA) and goat anti-mouse IgG conjugated to alkaline phosphatase

(Santa Cruz Biotechnology, Inc., Santa Cruz, CA) were used as primary and secondary antibodies, respectively. The electrophoresed proteins were transferred on nitrocellulose using electrotransfer kit (Novex), and the filter was blocked for 1 h in PBS buffer containing 1% dried milk, followed by incubation with the primary antibody (1:2, 500) at room temperature for 1 h. The filter was washed extensively with PBS for 10 min and incubated in the same buffer containing the secondary antibody for 1 h at room temperature. The membrane was washed in PBS and developed using NBT/BCIP (Bio-Rad) as described (Ausubel et al., 1988).

RESULTS

Self-Association of Recombinant Proteins Induced by N-terminus Glucagon Fusion

For fusion expression of rhIL-2, various N-terminus fusion partners consisting of glucagon or glucagon-derived mutants were designed: T_LG (10.8 kDa), T_SG (6.1 kDa), G (4.7 kDa), G2 (8.3 kDa), G3 (12.0 kDa), G3$\Delta(H)_6$ (11.2 kDa), Gα (4.4 kDa), Gβ (4.4 kDa), and Gγ (4.7 kDa) (Fig. 1). The T_S and T_L peptides, derived from N-terminal human TNF-α (known as forming β-sheet structure) were previously reported as effective fusion partners in the high-level synthesis of other therapeutic proteins (Chang et al., 1998; Shin et al., 1997; 1998) and also employed in this study to clarify the role of glucagon fusion. After gene expression, all recombinant fusion proteins were accumulated in the form of insoluble aggregates (inclusion bodies) in bacterial cytoplasm, and all isolated aggregates involving recombinant fusion proteins were easily dissolved at room temperature by simple pH shift (→12 →8) only without using any denaturing agents. To examine the characteristics of the molecular interaction in recombinant aggregation and coaggregation, in vitro analysis conditions were appropriately chosen with the isolated protein aggregates.

Direct Expression Versus Fusion Expression of rhIL-2 Using Glucagon or Glucagon-Derived Peptides as N-Terminus Fusion Partner

The results of Western blot analysis following native PAGE of the dissolved aggregates (Fig. 2A) show that hIL-2 antibody-specific immunoreactive proteins in all types of isolated aggregates formed large multimers of high-molecular weight. Through immunoblotting analysis of the aggregates involving the directly expressed rhIL-2, a high degree of immunoreactivity was extensively detected at various multimer positions under the nonreducing condition of SDS-PAGE (lane 2 in Fig. 2B), while under the reducing condition, immunoreactive proteins appeared only at monomer and homo-dimer positions (lane 4 in Fig. 2B). (Some extent of immunoreactivity at homo-dimer position seems to result from strong hydrophobic molecular interaction of rhIL-2.) These results certainly imply that in the direct-expression mode, various heterogeneous multimers were significantly formed via nonspecific intermolecular crosslinking by disulfide bridges. However, the immunoblotting analysis of the aggregates involving fusion proteins, which were electrophoresed in the nonreducing SDS-PAGE gel (Fig. 2C) shows that most of the immunoreactive proteins appeared at monomer position. Hence, evidently with glucagon fusion, the intermolecular covalent linkage was effectively circumvented, and the multimerization took place via noncovalent (most probably hydrophobic) interaction.

Upon the fusion expression using the glucagon-derived mutants, Gα and Gγ, native PAGE-Western blot analysis of isolated aggregates shows similar mobilities of immunoreactive large multimers (lanes 5 and 6 in Fig. 2D), compared to the earlier results of fusion expression using wild-type glucagon molecule(s) (Fig. 2A). With the fusion of Gβ, however, the size of immunoreactive multimers was dramatically reduced to about 110 kDa (from the analysis results by gel filtration chromatography) (lane 7 in Fig. 2D), and the 110-kDa oligomer was found to be formed via noncovalent (hydrophobic) interaction, too (lane 3 in Fig. 2D).

For all recombinant proteins synthesized in various recombinant *E. coli* cultures, the fraction in the isolated aggregates (defined as aggregate purity of recombinant protein, %) and in total cellular proteins (defined as expression level, %) were estimated by reducing SDS-PAGE. Figure 3 shows that the aggregate purity of directly expressed rhIL-2 was very low (≈40%) due probably to nonspecific binding between rhIL-2 and other host proteins, while the purity of G3 · IL-2 increased to the highest level (≈80%). The aggregate purity of foreign proteins did not show a noticeable correlation with the expression level (Fig. 3). The purities of recombinant proteins, T_LG· and T_SG · IL-2 were apparently lower than that of G · IL-2, which suggests that the additional N-terminus fusion of human TNF-α fragments diminishes the glucagon effect on homogeneous aggregate formation. Also, with the N-terminus fusion of glucagon mutants (Gα, Gβ, and Gγ) the purities of the recombinant fusion proteins were lowered, and especially the fusion of the mutant Gα resulted in the most significant decrease of purity to < 40% level (Fig. 3).

In Vitro Coaggregation Study: Glucagon (G) Versus Glucagon Mutant (Gα) As a Fusion Partner

To elucidate the role of glucagon peptide in the molecular interaction of recombinant fusion proteins, an in vitro coaggregation experiment was conducted. Total intracellular aggregates (i.e., recombinant plus host-derived proteins), coprecipitated under the highly acidic condition (0.1*M* HCl) after cell disruption, were subsequently denatured using strong denaturing agents, DTT (50 m*M*) and urea (2*M*). Using this procedure, two different mixtures of denatured total proteins were prepared from the induced recombinant cultures, BL21(DE3)[pT7-GIL2] and BL21(DE3)[pT7-GαIL2] producing G · IL-2 and Gα · IL-2, respectively. In vitro coaggregation process was obliged to proceed inside a membrane tube (containing the denatured total protein mix-

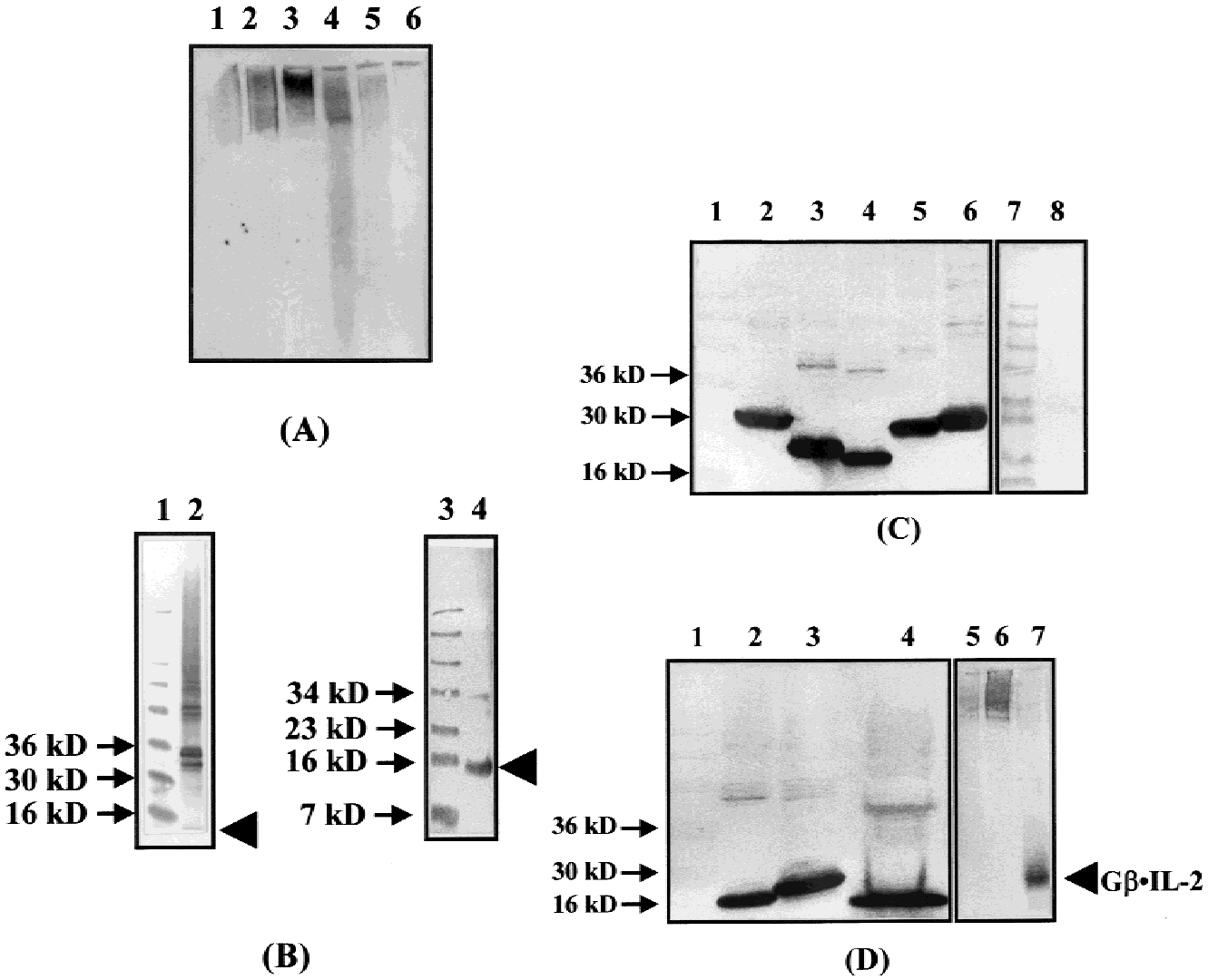

Figure 2. (A) Result of Western blot analysis after native PAGE of inclusion bodies involving directly expressed rhIL-2 (lane 1), $T_LG \cdot$ IL-2 (lane 2), $T_SG \cdot$ IL-2 (lane 3), G · IL-2 (lane 4), G2 · IL-2 (lane 5), and G3 · IL-2 (lane 6). (B) Result of Western blot analysis after nonreducing (lane 2) and reducing (lane 4) SDS-PAGE of inclusion bodies involving directly-expressed rhIL-2. (lanes 1,3, markers; "◀" indicates monomeric rhIL-2.) (C) Result of Western blot analysis after nonreducing SDS-PAGE of inclusion bodies involving $T_LG \cdot$ IL-2 (lane 2), $T_SG \cdot$ IL-2 (lane 3), G · IL-2 (lane 4), G2 · IL-2 (lane 5), and G3 · IL-2 (lane 6). [Lanes 1,7, markers; lane 8, total cell lysate sample from wild-type BL21(DE3)] (D) Result of Western blot analysis after nonreducing SDS-PAGE (lanes 1–4) and native PAGE (lanes 5–7) of inclusion bodies involving Gα · IL-2 (lanes 2, 5), Gβ · IL-2 (lanes 3,6), and Gγ · IL-2 (lanes 4,7). (Lane 1, markers).

ture above) with removing DTT and urea by dialysis (Experimental protocols). As DTT and urea are removed from the protein mixture, the coaggregation is subject to progressively take place. At the end of each coaggregation experiment, the insoluble aggregates were never formed in the solution, and the coaggregation products were assayed by reducing/non-reducing SDS-PAGE and subsequent Western blot analysis. Figure 4 shows that at the beginning of the experiment, monomeric G· and Gα · IL-2 appeared even under the nonreducing condition of SDS-PAGE (lanes 2 and 4), while at the end of the coaggregation experiment, the Gα · IL-2 monomer disappeared in the nonreducing SDS-PAGE (lane 8). Evidently, most of G · IL-2 was coaggregated via hydrophobic interaction, while in the coaggregation of Gα · IL-2, the multimer formation via intermolecular covalent (S-S) bridge(s) was much more significant.

Affinity Purification Using a Glucagon Receptor Fragment

Synthesis of Recombinant Receptor Fragment in Escherichia coli

The cDNA of human glucagon receptor encodes a 477-residue protein with 7 putative transmembrane (TM) domains (Lok et al., 1994). An N-terminus domain of 119 residues (Gln27 to Tyr145) between the predicted signal peptide and the first putative TM domain was presumed to

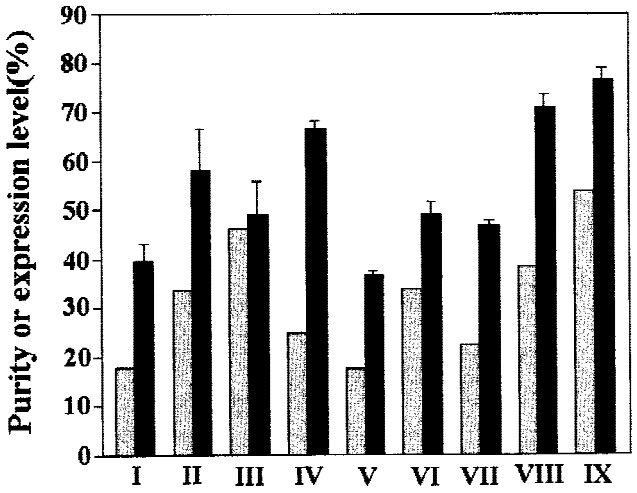

Figure 3. Aggregate purity, % (black bars) and expression level, % (grey bars) of various recombinant proteins: directly expressed rhIL-2 (I), $T_LG \cdot$ IL-2 (II), $T_SG \cdot$ IL-2 (III), G · IL-2 (IV), G-g3α · IL-2 (V), G-g3β · L-2 (VI), G-f3 · IL-2 (VII), G2 · IL-2 (VIII), G3 · IL-2 (IX).

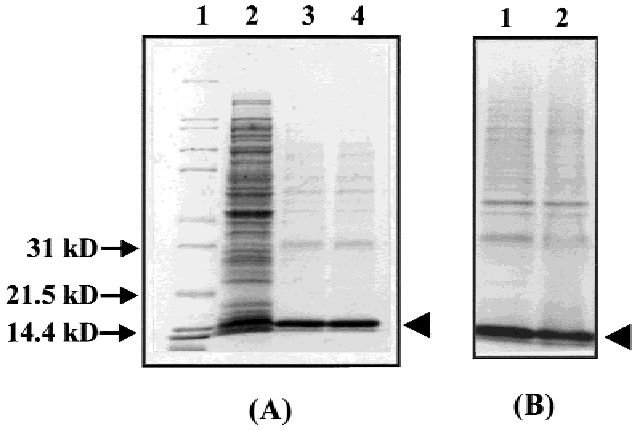

Figure 5. (A) Result of reducing SDS-PAGE analysis of total cell lysates (lane 2) and inclusion bodies involving GR120H, before (lane 3) and after (lane 4) dissolved by simple pH shift (→12 →8). (lane 1, markers) (B) Results of reducing (lane 1) and nonreducing (lane 2) SDS-PAGE analysis of inclusion bodies involving GR120H, after dissolved by simple pH shift (→12 →8). ("◀" indicates recombinant GR120H.)

be extracellular and contain a specific glucagon-binding site (Buggy et al., 1995; Lok et al., 1994). With C-terminus fusion of polyhistidine tag, $(His)_6$, the extracellular fragment of glucagon receptor was expressed in recombinant *E. coli* using the same T7 promoter system (Fig. 1), and the intracellular recombinant protein (i.e., GR120H, 14.8 kDa) was accumulated in the form of insoluble aggregates. The isolated aggregates were surprisingly dissolved by simple pH shift (→12 →8) like the earlier aggregates of recombinant fusion proteins expressed with glucagon fusion (Fig. 5A). Fig. 5B shows that a major portion of the dissolved GR120H appeared at monomer position even in the nonreducing SDS-PAGE and therefore, the intermolecular S-S bond formation was not significant. The dissolved GR120H in Tris-HCl buffer (20 m*M*, pH 8) was loaded onto Ni^{+2}-Sepharose column and then bound by virtue of their His-tags to the chelated metal ion (Ni^{+2}) on the column (Fig. 6A).

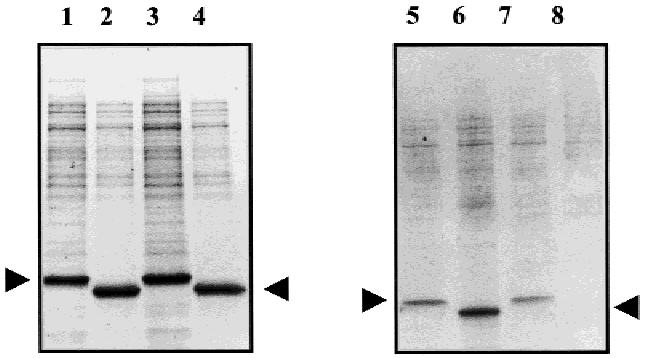

Figure 4. Results of reducing (lanes 1,2,5,6) and nonreducing (lanes 3,4,7,8) SDS-PAGE analysis of total cellular proteins involving G · IL-2 (lanes 1,3,5,7) and Gα · IL-2 (lanes 2,4,6,8) before (lanes 1–4) and after (lanes 5–8) in vitro coaggregation experiment. ("▶" and "◀" indicate recombinant G· and Gα · IL-2, respectively.)

Affinity Column Chromatography and Stability of GR120H Ligand

The affinity chromatography column (GR120H-Ni^{+2}-Sepharose) was used to purify $G3\Delta(H)_6 \cdot$ IL-2 from total intracellular protein mixtures. First, to effectively dissociate the hydrophobic multimers of $G3\Delta(H)_6 \cdot$ IL-2 without denaturing the α-helical conformation of glucagon molecules, a range of urea concentration (0.2 to 2.0*M*) was investigated in the load solution (0.16 g/L total intracellular proteins in 0.5*M* NaCl, 50 m*M* imidazole, 20 m*M* Tris-HCl buffer, pH 8) The lanes 2 to 5 in Figure 6B were subject to scanning analysis using a densitometer, and the percentage of $G3\Delta(H)_6 \cdot$ IL-2 was increased to >90% in the lanes 3 to 5 where urea concentration was $\leq$ 1*M* in the load. Also, Figure 6C shows that $G3\Delta(H)_6 \cdot$ IL-2 was not purified through GR120H-unbound Ni^{+2}-Sepharose column. These results indicate that the recombinant $G3\Delta(H)_6 \cdot$ IL-2 was purified by virtue of specific binding to the GR120H ligand with a highly enhanced purity. The high purity of $G3\Delta(H)_6 \cdot$ IL-2 in the eluted solution was evidently shown in the result of reverse-phase HPLC analysis (Fig. 6E).

To more fully elucidate the efficacy of GR120H as an affinity ligand, the same chromatography column was used again to purify $G3\Delta(H)_6 \cdot$ M [recombinant fusion magainin-2 (MSI-78)] from the total intracellular proteins. (The preparation of the expression vector pT7-G3ΔHM and the synthesis of recombinant $G3\Delta(H)_6 \cdot$ M in *E. coli* are well described in the Materials and Methods section.) Although the fraction of recombinant $G3\Delta(H)_6 \cdot$ M in the intracellular protein mixture was quite low due to a low expression level (< 5%), the $G3\Delta(H)_6 \cdot$ M was purified with high purity using GR120H as an affinity ligand (Fig. 6D). Consequently, the recombinantly synthesized GR120H ligand using a bacterial-expression system was immobilized on the column and finally attained such a conformation that

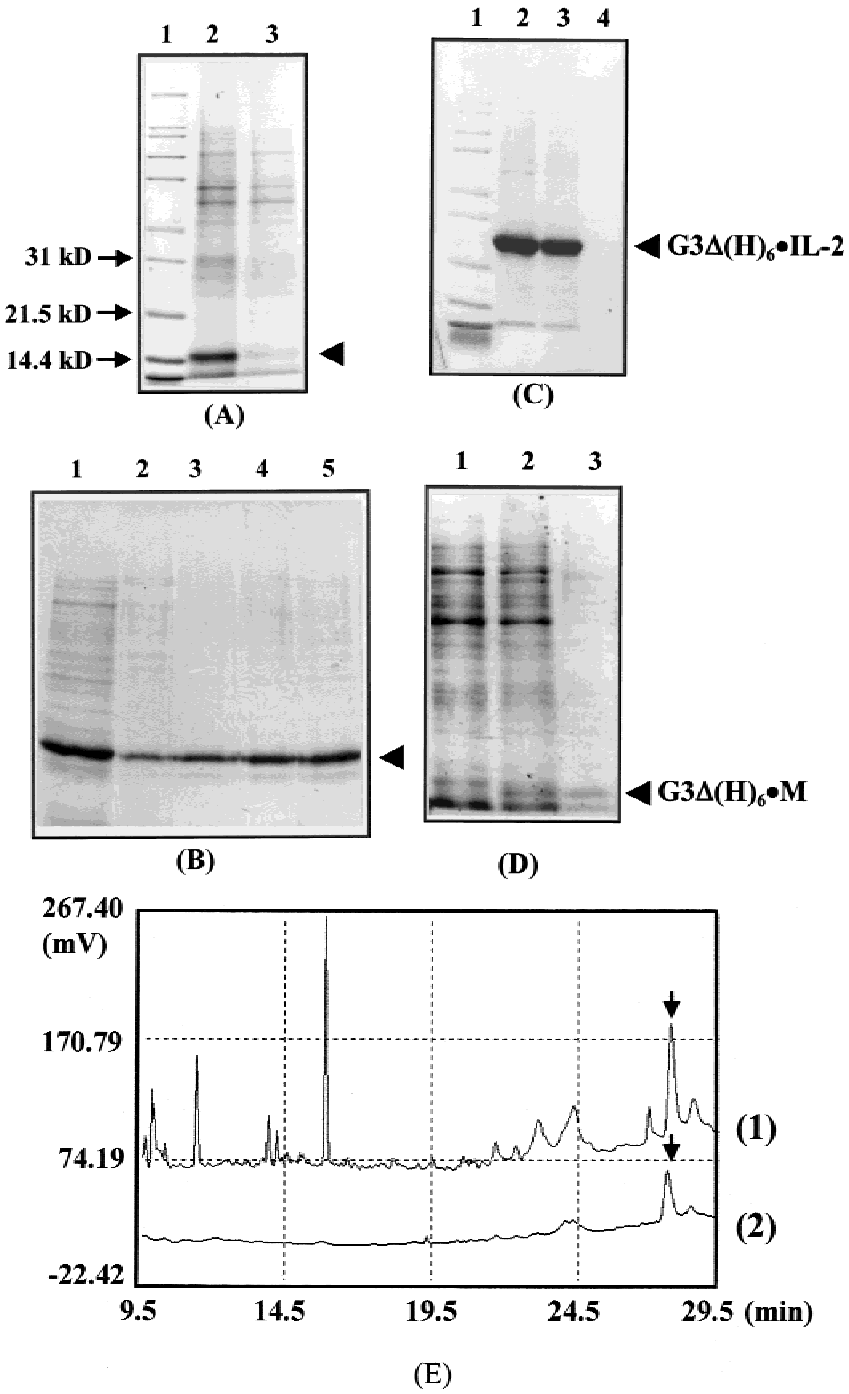

Figure 6. (A) GR120H binding to Ni^{+2}-Sepharose column, analyzed by reducing SDS-PAGE: Lane 1, markers; lane 2, inclusion bodies involving GR120H, dissolved by simple pH shift (→12 →8) ("load-in" solution); lane 3, effluent of unbound proteins from the column ("load-out" solution). (B) Purification of $G3\Delta(H)_6$ · IL-2 using GR120H-Ni^{+2}-Sepharose column, analyzed by reducing SDS-PAGE: Lane 1, total intracellular proteins containing $G3\Delta(H)_6$ · IL-2 ("load-in" solution); lanes 2–5, eluted solutions from the column loaded with 2*M* (lane 2), 1*M* (lane 3), 0.5*M* (lane 4), and 0.2*M* (lane 5) urea. (C) Purification of $G3\Delta(H)_6$ · IL-2 using Ni^{+2}-Sepharose column (GR120H-unbound), analyzed by reducing SDS-PAGE: Lane 2, total intracellular proteins containing $G3\Delta(H)_6$ · IL-2 ("load-in" solution); lane 3, effluent of unbound proteins from the column ("load-out" solution); lane 4, eluted solution from the column. (Lane 1, markers) (D) Purification of $G3\Delta(H)_6$ · M using GR120H-Ni^{+2}-Sepharose column, analyzed by reducing SDS-PAGE: Lane 1, total intracellular proteins from noninduced recombinant culture; lane 2, total intracellular proteins from induced recombinant culture ("load-in" solution); lane 3, eluted solution from the column. (E) Purification of $G3\Delta(H)_6$ · IL-2 using GR120H-Ni^{+2}-Sepharose column, analyzed by C8 reverse-phase HPLC: Chromatograms 1 and 2 were obtained by analyzing the same "load-in" and eluted solutions, shown in lanes 1 and 4, respectively, in Figure 6B.

allows N-terminus glucagon peptides of recombinant fusion proteins to selectively bind to it.

The stability of GR120H ligand was examined during the extended period of operation, that is, the affinity column (GR120H-Ni^{+2}-Sepharose) was used through 50 cycles to purify recombinant $G3\Delta(H)_6 \cdot$ IL-2 from total intracellular proteins (Materials and Methods). As seen in the results in Figure 7 (lanes 3–8), the affinity of GR120H ligand to $G3\Delta(H)_6 \cdot$ IL-2 was so stably maintained during the multiple cyclic operation that the purification was successful even at the 50th cycle. After the 50th cycle of operation, the GR120H ligand was eluted by passing 0.5*M* imidazole through the column (lane 10 of Fig. 7), which clearly indicates the longterm operational stability of affinity ligand, GR120H retained in the chromatography column.

DISCUSSION

In vivo inclusion body formation presumably represents a situation of overexpression in the heterologous environment, where nascent polypeptide chains which are synthesized continuously, may aggregate with any partially folded intermediates within the cells. In vivo aggregation is sometime so nonspecific that coaggregation of various host proteins significantly lowers the aggregate purity of target recombinant protein. The design of a special fusion peptide inducing, in general, the self-assembly of foreign protein is of great advantage in recovering relatively pure protein from the inclusion body state. The small α-helical peptide, glucagon which has special dual properties regarding intermolecular association: (1) strong self-association, and (2) stable complex formation with a receptor, is a very attractive candidate as a fusion expression partner in the production of recombinant proteins from *E. coli*. The specific glucagon-receptor binding could provide a convenient means, i.e., affinity column chromatography using the specific protein (glucagon)-ligand (receptor) interaction, for producing highly purified protein.

Mechanism of Glucagon-Induced Self-Association and Coaggregation

The various electrophoresis and immunoblotting analyses have shown that the directly expressed rhIL-2 interacted with various host-derived proteins via disulfide bridge(s) and formed large heterogeneous multimers. This suggests that in off-pathway aggregation of the partially folded rhIL-2 intermediate, the cysteine residues (Cys58 and Cys105) participated in the intermolecular disulfide linkage at a much higher rate before they were linked by intramolecular S-S bond. With the fusion of glucagon molecule(s), however, all recombinant proteins formed the hydrophobic aggregates where the intermolecular disulfide linkage was rarely found, and the fraction of recombinant protein was highly increased. Presumably, this drastic change in the pattern of intermolecular association arose from distinct conformational change of fusion proteins induced by N-terminus strong α-helical peptide(s) which effectively rescued the rhIL-2 molecules from being coagulated with various host proteins. Sasaki et al. (1975) reported that α-helical conformation of glucagon molecules favors intermolecular self-association to form trimers and possibly higher oligomers where the helical arrangement is more tightened up. Probably, the N-terminus glucagon in synthesized fusion protein was quickly transformed to an α-helical conformation and induced the self-association of folding intermediates. Because oligomeric proteins generally have a higher affinity for themselves than for other proteins (Jaenicke and Rudolph, 1986), the higher aggregate purity of $G3 \cdot$ IL-2 molecules might be due to the increased tendency

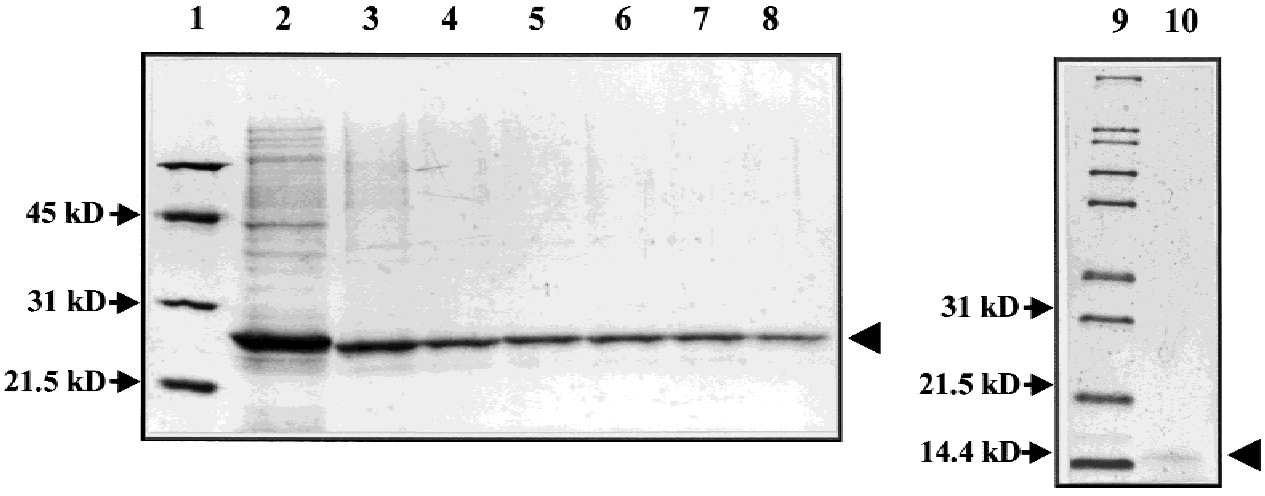

Figure 7. Purification of $G3\Delta(H)_6 \cdot$ IL-2 through multiple cyclic operation of affinity chromatography using GR120H-Ni^{+2}-Sepharose column, analyzed by reducing SDS-PAGE: Lane 2, "load-in" solution of total intracellular proteins containing $G3\Delta(H)_6 \cdot$ IL-2; lanes 3–8, eluted solutions at 1st, 10th, 20th, 30th, 40th, and 50th column chromatography, respectively; "◀" indicates $G3\Delta(H)_6 \cdot$ IL-2. (Lane 1,9, markers; lane 10, GR120H eluted by passing 0.5*M* imidazole after the 50th operation is finished, indicated by "◀.")

of self-association arising from more extended helical domain in the G3 trimer.

Compared to the fusion of G peptide, the use of Gα fusion partner caused a significant decrease in aggregate purity, and the tendency to form intermolecular S-S linkage was much more significant. It is presumed that the point mutation at three hydrophobic residues with aromatic sidechain (i.e., Phe6Gly, Tyr10Gly, Tyr13Gly) accompanied significant conformational change affecting the specificity of intermolecular association. Namely, an N-terminal domain of glucagon consisting of Phe6, Tyr10, and Tyr13 is responsible for such an α-helical conformation that promotes hydrophobic self-association via specific intermolecular packing. Probably, the domain of Phe6-Tyr10-Tyr13 is a specific site, directly involved in the intermolecular hydrophobic packing. In contrast to other fusion proteins, Gβ · IL-2 formed much lower oligomer (≈ 110 kDa) at the inclusion body state. If one protein folds or aggregates much faster than the other, then it effectively excludes the other from being further incorporated into a coaggregate complex (Mitraki et al., 1993; Speed et al., 1996). Probably switching the C-terminus residues (Phe22, Trp25, Leu26) with glycines increased the aggregation rate of the partially folded intermediate of fusion protein and hence, reduced the size of off-pathway aggregates. Glycine residues provide the flexibility of polypeptide conformation and can easily attain conformations which are energetically unfavorable to hydrophobic amino acids with sidechains (Sasaki et al., 1975). Thus, the C-terminus glycines of Gβ may change the conformation in a way that gives thermodynamic stability to oligomers at an early stage of aggregation and suppresses further multimerization. Consequently, it seems likely that there are two nonpolar domains on the glucagon molecule that are of crucial importance in determining intermolecular-binding specificity and final aggregate size of recombinant fusion proteins. The N-terminus domain (i.e., Phe6-Tyr10-Tyr13) could be a certain structural moiety involved in intermolecular self-association through a defined manner such as helix–helix docking, while the C-terminus domain (i.e., Phe22-Trp25-Leu26) seems to be related to the assembly size of folding intermediates.

Although a common assumption is that the off-pathway aggregation reactions are nonspecific, it has been recently reported that aggregation can occur by specific interaction of certain conformations of folding intermediates rather than by nonspecific coaggregation (Speed et al., 1996). Also, this study reports that the aggregation specificity of fusion protein was highly sensitive to sequence effects of N-terminus fusion peptide, glucagon. Therefore, there would appear to be opportunities to manipulate the in vivo aggregation by enforcing the specific aggregation of recombinant proteins, for instance, by using a special fusion peptide which efficiently induces the self-association of synthesized proteins via specific molecular interaction, as presented in this study.

Application of Active Recombinant Receptor Fragment to Affinity Purification

Fortunately, the analysis of isolated aggregates involving recombinant GR120H has shown that intermolecular S-S linkage formation by GR120H molecules was not significant, which is presumed to result from an inherent aggregation property of extracellular domain of glucagon receptor. Also, the GR120H immobilized to Ni^{+2}-Sepharose column demonstrated the high specificity in molecular interaction with both recombinant fusion proteins, $G3\Delta(H)_6 \cdot$ IL-2 and $G3\Delta(H)_6 \cdot$ M. It is strongly suggested that the immobilized recombinant GR120H (which contains six cysteine residues) was efficiently renatured and attained the specific-binding affinity to glucagon molecules like the extracellular domain of native receptor (though the structural authenticity was not confirmed by comparison with active glucagon receptor, which has been recently produced from animal cell cultures (Buggy et al., 1995; MacNeil et al., 1994)). Immobilizing recombinant proteins on a column has been recently reported to be of great advantage in protein refolding, because this approach can effectively prevent reaggregation as evidenced in the folding of two chloroplast membrane proteins (Toc75 and LHC2) (Rogl et al., 1998) and bacterial prosubtilisin (Volkov and Jordan, 1996). The reason for the anti-aggregation effect is that the hydrophobic faces exposed in folding intermediates are not free to associate with one another because the resin material may have a chaperone-like function in keeping folding proteins separated (Rogl et al., 1998). The six cysteine residues (Cys18, Cys33, Cys42, Cys56, Cys75, Cys96), 100% conserved in extracellular domains of other 13 GPCRs (G-protein-coupled receptors), are almost evenly distributed through the entire sequence of GR120H. Hence, once the intramolecular S-S bonds are correctly formed, they probably confer stable structural conformation to GR120H molecule. Therefore, the longterm operational stability, proven through the cyclic-purification experiments which required repeated denaturation and renaturation steps of the immobilized GR120H, may result from efficient renaturation process of the immobilized receptor fragment and/or inherent structural stability of GR120H molecule.

Adding a certain amount of urea to the load solution was shown effective in increasing the binding selectivity and affinity of $G3\Delta(H)_6 \cdot$ IL-2 to the GR120H and hence, the recovery purity. The lowered purity of $G3\Delta(H)_6 \cdot$ IL-2 at high-urea concentration (2*M*) seems to result from the reduced binding specificity to GR120H, which is due probably to significant denaturation of α-helical conformation of glucagon peptides at 2*M* urea. The helix uncoiling of $G3\Delta(H)_6$ peptide at high-urea concentration enabled urea to be used as an effective eluent dissociating the specific binding between GR120H and $G3\Delta(H)_6 \cdot$ IL-2.

The discovery of new affinity ligands which have high affinity to a certain peptide via specific molecular interaction is itself an important area in fusion protein-based affinity chromatography. In various approaches (Boder and Wittrup, 1997; Georgiou et al., 1997, Lima et al., 1997;

Nord et al., 1997; Tripet et al., 1996) for developing the new protein ligand, two major requirements for protein ligands are: (1) large-scale production in a low-cost expression system, and (2) longterm operational stability. From the viewpoint of this criteria, GR120H seems to provide a firm basis for being used in practice as a protein ligand for affinity-column chromatography.

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Struhl K, Chanda VB. 1988. Current protocols in molecular biology. New York: John Wiley & Sons.

Boder ET, Wittrup KD. 1997. Yeast surface display for screening combinatorial libraries. Nat Biotechnol 15:553–557.

Buggy J, Rossomando A, MacDougall M, Mierz D, Wunderlich D, Yoo-Warren H. 1995. Human glucagon receptor monoclonal antibodies: Antagonism of glucagon action and use in receptor characterization. Horm Metab Res 28:215–219.

Cha S-S, Kim J-S, Cho H-S, Shin N-K, Jeong W, Shin H-C. 1998. High resolution crystal structure of a human tumor necrosis factor-α mutant with low systemic toxicity. J Biol Chem 273:2153–2160.

Chang SG, Kim DY, Choi KD, Shin JM, Shin HC. 1998. Human insulin production from a novel mini-proinsulin which has high receptor-binding activity. Biochem J 329:631–635.

Chrunyk BA, Evans J, Lillquist J, Young P, Wetzel R. 1993. Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. J Biol Chem. 268:18053–18061.

Cleland JL. 1993. Impact of protein folding on biotechnology. In: Cleland JL, editor. Protein folding: In vivo and in vitro. American Chemical Society symposium series No. 526. Washington, DC: American Chemical Society. p. 1–21.

Georgiou G, Stathopoulos C, Daugherty PS, Nayak AR, Iverson BL, Curtiss R. 1997. Display of heterologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines. Nat Biotechnol 15:29–34.

Hendrick JP, Hartl F-U. 1993. Molecular chaperone functions of heat-shock proteins. Ann Rev Biochem 62:349–384.

Hasse-Pettingell CA, King J. 1988. Formation of aggregates from a thermolabile in vitro folding intermediate in P22 tailspike maturation. J Biol Chem 263:4977–4983.

Ikemura H, Tagaki H, Inouye M. 1987. Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli*. J Biol Chem 262:7859–7864.

Jacob L, Zasloff LM. 1994. Potential therapeutic applications of maganins and other antimicrobial agents of animal origin. In: Marsh J, and Goode JA, editors. Antimicrobial peptides, Ciba Foundation Symposium 186. Chichester: John Wiley & Sons. p. 197–223.

Jaenicke R, Rudolph R. 1986. Refolding and association of oligomeric proteins. Meth Enzymol 131:218–250.

Lima VLM, Correia MTS, Cechinel YMN, Sampaio CAM, Owen JS, Coelho LCBB. 1997. Immobilized *Cratylia mollis* lectin as a potential matrix to isolate plasma glycoproteins, including lecithin-cholesterol acyl transferase. Carbohyr Polym 33:27–32.

Lok S, Kuijper JL, Jelinek LJ, Kramer JM, Whitemore TE, Sprecher CA. 1994. The human glucagon receptor encoding gene: Structure, cDNA sequence and chromosomal localization. Gene 140:203–209.

MacNeil DJ, Occi JL, Hey PJ, Strader CD, Graziano MP. 1994. Cloning and expression of a human glucagon receptor. Biochem. Biophys Res Commun 198:328–334.

Mitraki A, King J. 1989. Protein folding intermediates and inclusion body formation. Bio/Technol 7:690–697.

Mitraki A, Danner M, King J, Seckler R. 1993. Temperature-sensitive mutations and second-site suppressor substitutions affect folding of the P22 tailspike protein in vitro. J Biol Chem 268:20071–20075.

Nord K, Gunneriusson G, Ringdahl J, Stahl S, Uhlen M, Nygren P-A. 1997. Binding proteins selected from combinatorial libraries of an α-helical bacterial receptor domain. Nat Biotechnol 15:772–777.

Przybycien TM. 1998. Protein-protein interactions as a means of purification. Curr Opin Biotechnol 9:164–170.

Rogl H, Kosemund K, Kuhlbrandt W, Collinson I. 1998. Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilized by nickel chelating chromatography. FEBS Lett 432:21–26.

Saen JL, Agard DA. 1989. The alpha-lytic protease pro-region does not require a physical linkage to activate the protease domain in vivo. Nature 341:462–464.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. PNAS USA 74:5463–5467.

Sasaki K, Dockerill S, Adamiak DA, Tickle IJ, Blundell T. 1975. X-ray analysis of glucagon and its relationship to receptor binding. Nature 257:751–757.

Shin CS, Hong MS, Bae CS, 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.

Shin CS, Hong MS, Bae CS, Kim DY, Shin HC, Lee J. 1998. Growth-associated synthesis of recombinant human glucagon and human growth hormone in high-cell-density cultures of *Escherichia coli*. Appl Microbiol Biotechnol 49:364–370.

Silen JL, Frank D, Fujishige A, Bone R, Agard DA. 1989. Analysis of prepro-alpha-lytic protease expression in *Escherichia coli* reveals that the pro region is required for activity. J Bacteriol 171:1320–1325.

Speed MA, Wang DIC, King J. 1996. Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. Nat Biotechnol 14:1283–1287.

Tabor S, Richardson CC. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. PNAS USA 82:1074–1078.

Tripet B, Yu L, Bautista DL, Wong WY, Irvin RT, Hodges RS. 1996. Engineering a *de novo*-designed coiled coil heterodimerization domain for the rapid detection, purification and characterization of recombinantly expressed peptides and proteins. Protein Eng 9:1029–1042.

Unger RH, Dobbs RE. 1978. Insulin, glucagon, and somatostatin in the regulation of metabolism. Ann Rev Physiol 40:307–343.

Unger RH, Orci L. 1981. Glucagon and the A cell: Physiology and pathophysiology. N Engl J Med 304:1518–1524.

Volkov A, Jordan F. 1996. Evidence for intramolecular processing of prosubtilisin sequestered on a solid support. J Mol Biol 262:595–599.

Wetzel R. 1992. Stability of protein pharmaceuticals. In: Ahern TJ, Manning MC, editors. *In vivo* pathways of degradation and strategies for protein stabilization. New York: Plenum Press. p. 43–88.

Wetzel R. 1994. Mutations and off-pathway aggregation of proteins. TIBTECH 12:193–198.

Winther JR, Sorensen P. 1991. Propeptide of carboxypeptidase Y provides a charperone-like function as well as inhibition of the enzymatic activity. PNAS USA 88:9330–9334.

Zhu X, Ohta Y, Jordan F, Inouye M. 1989. Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. Nature 339:483–484.