Dual Function of the Prepropeptide of Prouroguanylin in the Folding of the Mature Peptide

DISULFIDE-COUPLED FOLDING AND DIMERIZATION*

Received for publication, January 18, 2000, and in revised form, May 15, 2000
Published, JBC Papers in Press, May 25, 2000, DOI 10.1074/jbc.M000543200

Yuji Hidaka‡§, Chisei Shimono‡, Megumu Ohno‡, Nobuaki Okumura‡, Knut Adermann‡, Wolf-Georg Forssmann‡, and Yasutugu Shimonishi†

From the ‡Division of Organic Chemistry and §Division of Protein Metabolism, Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565-0871, Japan and the ¶Lower Saxony Institute for Peptide Research, Fedor-Lyzen-Strasse 31, D-30625 Hannover, Germany

Guanylyl cyclase activating peptide II (GCAP-II), an endogenous ligand of guanylyl cyclase C, is produced via the processing of the precursor protein (prepro-GCAP-II). We have previously shown that the prepropeptide in pro-GCAP-II functions as an intramolecular chaperone in the proper folding of the mature peptide, GCAP-II (Hidaka, Y., Ohno, M., Hemmasi, B., Hill, O., Forssmann, W.-G., and Shimonishi, Y. (1998) Biochemistry 37, 8498–8507). Here, we report an essential region in pro-GCAP-II for the correct disulfide pairing of the mature peptide, GCAP-II. Five mutant proteins, in which amino acid residues were sequentially deleted from the N terminus, and three mutant proteins of pro-GCAP-II, in which N-terminal 6, 11, or 17 amino acid residues were deleted, were overproduced using Escherichia coli or human kidney 293T cells, respectively. Detailed analysis of in vivo or in vitro folding of these mutant proteins revealed that one or two amino acid residues at the N terminus of pro-GCAP-II are critical, not only for the chaperone function in the folding but also for the net stabilization of pro-GCAP-II. In addition, size exclusion chromatography revealed that pro-GCAP-II exists as a dimer in solution. These data indicate that the prepropeptide has two roles in proper folding: the disulfide-coupled folding of the mature region and the dimerization of pro-GCAP-II.

Endogenous peptide hormones are often synthesized in vivo in the form of precursor proteins with pre- (or signal) and prepro-leader sequences, which are subsequently processed into biologically active mature peptides after their release from the ribosome (1). Little is known, however, concerning the role of the prepropeptide in the pro-leader sequence in the processing of precursor proteins to the mature peptide hormones or their function in the folding process, which results in the mature hormones. Guanylin and uroguanylin (2–4), endogenous ligands of particulate guanylyl cyclase C (GC-C) (5), are thought to function in regulating the level of cGMP as a second messenger in intestinal and kidney cells, i.e. the regulation of chloride and water secretion from the inside of these cells to the outside (6, 7). Guanylin and uroguanylin are generated as precursor proteins (preproguanylin (prepro-GCAP-I) or preprouroguanylin (prepro-GCAP-II), respectively), which contain the prepro-leader sequences which precede the mature portion. After cleavage of the pre-sequence, pro-GCAP-I and/or pro-GCAP-II are further processed to give the mature peptides, guanylin or uroguanylin (Fig. 1). GCAP-II, a plasma form of uroguanylin, is one of the mature forms of pro-GCAP-II in vivo (4, 8, 12, 14–16). Recent studies in this laboratory have shown that spontaneous refolding to the native conformation is attained in pro-GCAP-II but not in GCAP-II (17), i.e. GCAP-II requires the prepropeptide, in order to efficiently fold into the bioactive form. There are a few examples, such as subtilisin, α-lytic protease, etc., in which the peptides of the pro-leader sequences in the precursor proteins aid the mature proteins in the proper assembly of the three dimensional structures in vitro and are referred to as “intramolecular chaperones” (18–20). The mature proteins are produced by enzymatic cleavage of the peptides in the pro-leader sequences from the folded precursor proteins. In these examples, the peptides in the pro-leader sequences function not only to diminish the activation energy but also to stabilize the rate-determining transition state(s) in the folding pathway (20, 21). Moreover, the N-terminal peptide in the pro-leader sequence of prosubtilisin, the precursor protein of subtilisin, mediates the folding of the protein intramolecularly. Prosousubtilisin exists as homodimer that is assembled during the folding of the protein (21, 22). However, the mechanism, at the molecular level, of the folding of these proteins via the peptides in the pro-leader sequence remains unclear.

In a recent study, it was demonstrated that guanylin, which is homologous to GCAP-II, requires the assistance of the prepropeptide of the precursor protein, pro-GCAP-I, not only to achieve correct folding but also for the formation of the native disulfide linkages (23). These findings, and our previous studies, led us to conclude that the prepropeptide in the pro-leader sequence of pro-GCAP-I and pro-GCAP-II plays a functional role as an intramolecular chaperone in the correct folding of the mature peptide and is also crucial for the disulfide-coupled folding of the reduced precursor (17, 23). Furthermore, these studies have led us to propose that the mature form of pro-GCAP-II, GCAP-II, is not at the thermodynamic ground state performance liquid chromatography; DTT, dithiothreitol; IGD, insulin-like growth factor.

This paper is available on line at http://www.jbc.org

25155
but, rather, is kinetically trapped in the precursor protein (17). Consequently, these studies raised questions as to which region(s) of the protein, or the manner in which the pro-leader sequence of pro-GCAP-II and, in turn, of GCAP-II. These residues function to stabilize the bioactive form of the mature portion during the folding of the entire protein. Moreover, we provide evidence that supports the existence of a homodimer, which is stabilized by intramolecular and non-covalent interactions between the region in the pro-leader sequence and, possibly, in the intermediate as well as the final steps of the folding process. The data obtained provide basic information, which is critical for our understanding of the role of the pro-leader sequence of the precursor proteins during the maturation of peptide hormones, such as GCAP-II and guanylin.

**EXPERIMENTAL PROCEDURES**

Materials—Restriction enzymes were purchased from Toyobo (Osaka, Japan) and New England Biolabs (Beverly, MA). Taq polymerase, T4 DNA ligase, and endoproteinasine Arg-C were obtained from Takara Shuzo Co. (Kyoto, Japan). Dulbecco’s modified Eagle’s medium (DMEM) and fetal bovine serum (FBS) were purchased from Nissui Pharmaceutical Co. (Tokyo, Japan) and Dainippon Pharmaceutical Co. (Osaka, Japan), respectively. Thr-Ile-Ala-uroguanylin and its disulfide isomers were synthesized according to a previously described procedure (17). All other chemicals and solvents were reagent grade. PCR was carried out using a Sanyo DNA amplifier MIR-D30 (Osaka, Japan).

**Construction of Expression Vectors of Deletion Mutants for and Their Expression by E. coli Cells**—The cDNAs encoding the deletion mutant proteins were subcloned into a pET17b expression vector (Novagen), following the introduction, by means of PCR, of an NdeI site at its 5′ end and an XhoI site at its 3′ end using pEX2 as a template. The cDNA sequences of the vectors were confirmed as described above. *E. coli* BL21(DE3) cells, which were transformed with the expression vector, were cultured at 37 °C in Luria broth medium supplemented with ampicillin (50 mg/liter). The production of the mutant proteins was induced by the addition of 1 mM isopropyl-1-thio-*b* lactam (IPTG) at 37 °C for 3 h; the cells were harvested and washed with phosphate-buffered saline without magnesium and calcium ions, containing 1% Triton and 1 mM phenylmethylsulfonyl fluoride. The cells were resuspended in the same buffer, sonicated on ice, and centrifuged at 5000 × g, 20 min. The mutant proteins, isolated as an inclusion body, possessed the Met residue at the N terminus derived from the NdeI site during subcloning. The proteins thus prepared were characterized by mass spectrometry and amino acid analysis.

**Construction of Expression Vectors of Deletion Mutants for Human Embryonic Kidney 293T Cells**—The pEX2 vector derived from the pcDNA3 vector (Invitrogen), which contains a strong cytomegalovirus enhancer-promoter sequence for a high level of protein expression in mammalian cells (17), was used in this experiment. The construction of the expression vectors of the N-terminal deletion mutants of pro-GCAP-II was carried out as follows. The pEX2 vector, which carries a cDNA encoding pre-pro-GCAP-II between a BamHI site at its 5′ end and a XhoI site at its 3′ end (8), was employed as a template for the construction of the expression vectors in carrying the cDNAs of the N-terminal deletion mutant proteins of pro-GCAP-II. To efficiently use the signal sequence (the pre-region of pre-pro-GCAP-II) for the expression of the mutant proteins, the cDNA fragment from a BamHI site to the end of the signal sequence was amplified by PCR using the pEX2 vector as a template and a sense (ATATAGGATCCAGGAGGCGC-GATG) as primer 2 and an antisense (TCTCTCTAGAAGTCTCTCCTC- GAGTGTACGTTGTCCTGTCG) as primer 2. The cDNA fragment encoding the signal sequence was inserted between the BamHI site and XhoI site in the construction of the cDNA3, which contains a unique XhoI site after the signal sequence between the BamHI site and the XhoI site. The cDNAs encoding the deletion mutant proteins were prepared by PCR and subcloned into the site between the XhoI site and the XhoI site in pcDNA3H. The expression vectors comprised the cDNA sequences, which encode the signal peptide of pre-pro-GCAP-II and each of the deletion mutant proteins, pro-GCAP-II(5–86), pro-GCAP-II(12–86), and pro-GCAP-II(18–86). The mutant proteins (pro-GCAP-II(7–86) and pro-GCAP-II(18–86)) contained two additional amino acid residues, which are derived from the XhoI site in the expression vector, at their N termini. The cDNA sequences of the vectors thus constructed were confirmed by analysis using an Applied Biosystems 373A sequencing system.

**Expression of Deletion Mutants in 293T Cells**—Human embryonic kidney 293T cells (24) were maintained in 10% FBS/DMEM and transferred in a 10-cm diameter plate at 60–80% confluence with 20 μg of each of the expression vectors and the SuperFect reagent (Qiagen, Hilden, Germany) according to the manufacturer’s specifications. After incubation for 16 h, the medium was replaced by DMEM (10 ml/plate) without FBS and the cells were incubated for an additional 2 days at 37 °C in a CO2 incubator.

**Purification of the Recombinant Proteins Expressed by E. coli Cells or Human Kidney 293T Cells**—The recombinant proteins, which were expressed as inclusion bodies in *E. coli* cells, were treated with 20 eq of DTT in 50 mM Tris/HCl (pH 8.0) (200 μl) containing 6 μg guanidine HCl under an N2 atmosphere at 50 °C for 1 h. The supernatant of the reaction mixture or the culture medium (20 ml) of the 293T cells were applied to a column of Cosmosil 140C3-OPN (1 ml) (Nacarai Tesque Inc., Kyoto, Japan) pre-equilibrated with and wash with 20 ml of solvent A (20% CH3CN in 0.05% trifluoroacetic acid). The adsorbed proteins, which were eluted with solvent B (80% CH3CN in 0.05% trifluoroacetic acid), were dialyzed and lyophilized. The lyophilized protein was purified by HPLC and analyzed by mass spectrometry, as described previously (17). The yield of the purified protein was 0.5–1 nmol/10 ml of the culture medium of 293T cells, as estimated by amino acid analyses.

**Endoproteasine Arg-C Digestion of the Recombinant Proteins**—The recombinant protein (1 nmol) was incubated with endoproteasine Arg-C (5 pmol) in 0.1 μl Tris/HCl (pH 8.0) (200 μl) at 37 °C for 18 h. The digest was treated with anhydrotrypsin agarose as described previously (17), and the supernatant was subjected to HPLC. The eluates were analyzed by mass spectrometry and amino acid analysis.

**Gel Filtration Chromatography**—The HPLC apparatus consisted of a Waters 600 multisolvent delivery system (Bedford, MA) equipped with a Hitachi L-3000 photodiode array detector and a D-2000 chromatographic integrator (Tokyo, Japan). The protein (1 nmol) was dissolved in 50 mM Tris/HCl (pH 7.4) (50 μl) containing 0.2 μ M NaCl and chromatographed on a TSK-Gel G3000SW column (7.8 × 300 mm; Tosoh, Tokyo, Japan). The protein was eluted with 50 mM Tris/HCl (pH 7.4) containing 0.2 μ M NaCl at a flow rate of 0.8 ml/min, and the eluate was monitored at 220 nm. The molecular mass of the protein was calibrated using a gel filtration calibration kit (Amersham Pharmacia Biotech) containing bovine serum albumin (67 kDa), ovalbumin (43 kDa), RNase A (13.7 kDa), and thioredoxin (20 kDa). Thioredoxin was prepared from *E. coli* cells transformed with pET32b (Novagen), which possesses the cDNA encoding thioredoxin, purified on a nickel-nitrilotriacetic acid resin (Qiagen), and identified by mass spectrometric analysis.

**In Vitro Complementary Refolding of Pro-GCAP-II(5–86) and Pro- GCAP-II(12–86)**—The recombinant pro-GCAP-II (12–86) was prepared as follows. The protein (2 nmol) was incubated with 20 eq of DTT in 50 mM Tris/HCl (pH 8.0) (200 μl) containing 6 μg guanidine HCl under an N2 atmosphere at 50 °C for 1 h. The reduced pro-GCAP-II (12–86) was purified by HPLC, as described above, and lyophilized. The reduced pro-GCAP-II (12–86) (1 nmol) was dissolved in 0.05% trifluoroacetic acid (20 μl) and mixed with 9 volumes of 50 mM Tris/HCl (pH 8.0) in the presence of 0.2 mM GSH and 1 μM GSSG as described previously, and incubated at room temperature for 2 days. The oxidative refolding experiment was also carried out in the redox buffer in the presence of the synthetic complementary N-terminal peptides (VYIQYQ or VYIQYQGRFVQ). The reaction mixture was analyzed by HPLC. All solutions used for the refolding experiment were flushed with N2, and the reaction was carried out in a sealed vial under an atmosphere of N2.

**RESULTS AND DISCUSSION**

**Mutational Analysis of the N-terminal Amino Acids for a Role in the in Vitro Folding of Pro-GCAP-II**—In a previous report, we demonstrated that the mature form of GCAP-II does not possess sufficient information to permit for its correct folding and that the propeptide in pro-GCAP-II aids in the folding process, yielding only the bioactive form of GCAP-II (17). This result provided confirmation that the function of the propeptide in the pro-leader sequence of pro-GCAP-II was to serve as an intramolecular chaperone in the folding of GCAP-II, and con-
acid residues are shaded. Disulfide linkages are between positions Cys\(^{41}\) and Cys\(^{54}\), Cys\(^{74}\) and Cys\(^{82}\), and Cys\(^{77}\) and Cys\(^{85}\) (17). Single-letter codes for amino acid residues are used. The \(H\), \(S\), and \(T\) in the secondary structure represent \(\alpha\)-helix, \(\beta\)-sheet, and \(\beta\)-turn, respectively.

FIG. 2. Schematic representation of the recombinant wild-type prouroguanylin and N-terminal deletion mutant proteins prepared in this study.

subsequently raised a number of questions, such as (i) which region(s) in the pro-leader sequence of pro-GCAP-II contribute to the correct folding of the mature peptide, and (ii) how does the propeptide play a role in the folding of GCAP-II in vivo and in vitro?

To address these problems, we first searched the sequence motif(s) in the pro-leader sequence of pro-GCAP-II in the primary structures of pro-GCAP-IIs, which have been determined thus far, and then deduced the secondary structure of pro-GCAP-II using the Chou-Fasman method (25), as shown in Fig. 1. The amino acid sequences of the N-terminal region (amino acid residues 1–23) and the C-terminal region (amino acid residues 38–65) in the pro-leader sequence of pro-GCAP-II are highly homologous in all species, whereas that in the central region (amino acids 24–37) is diverse. This raises the possibility that the N-terminal region (amino acid residues 1–23), along with the C-terminal region (amino acid residues 38–65), acts as an intramolecular chaperone for the correct folding of pro-GCAP-II to yield the bioactive conformation of the mature peptide, GCAP-II. Further, the secondary structure prediction implied that the N-terminal region (amino acids 1–6) and the C-terminal region (LCVNV, amino acid residues 76–80) in the mature region exist as \(\beta\)-strands, not only in pro-GCAP-II, but also in pro-GCAP-I. Schulz et al. (23) recently demonstrated that the N-terminal region (amino acids 1–5) is in close proximity to the C-terminal region (a portion of guanylin) in the solution structure of pro-GCAP-I, as evidenced by NMR measurement. This may be extended to the speculation that the N-terminal region shares a characteristic secondary structure of pro-GCAP-I with the C-terminal mature region. A similar molecular conformation may be imagined in the structure of pro-GCAP-II, since it is likely that pro-GCAP-II has a conformation similar to that of pro-GCAP-I, i.e. it is possible that the N-terminal region (amino acids 1–23) in the pro-leader sequence of pro-GCAP-II interacts with the C-terminal mature region. This interaction may lead to the proper folding of pro-GCAP-II and contribute to the stabilization of the three-dimensional structure of the mature portion of pro-GCAP-II, GCAP-II.

To examine the nature of the participation of the N-terminal region in the pro-leader sequence of pro-GCAP-II in terms of its correct folding, we prepared a series of mutant proteins of pro-GCAP-II, in which the N-terminal amino acid residues were sequentially deleted from the N terminus of pro-GCAP-II (Fig. 2). The recombinant proteins were generated using \(E. coli\) BL21(DE3) cells. All mutant proteins were expressed with an additional Met residue at their N termini, which originated from the starting codon. For example, the deletion of a Val residue at the N terminus of pro-GCAP-II resulted in the production of the mutant protein, Met\(^1\)-pro-GCAP-II (2–86) (Fig. 2). Since the mutant proteins were produced as inclusion bodies in the bacterial cells, as is the case for the expression of many eukaryotic proteins by \(E. coli\) cells, we were not able to define the conformational states of the mutant proteins immediately after expression in the cells. As a result, the recombinant proteins were purified by HPLC in the reduced form and then oxidatively refolded to the oxidized forms in the presence of 2 mM GSH and 1 mM GSSG following previously reported procedures (17). The refolded proteins were subjected to HPLC, which indicated the presence of a few isomers, which comprised different disulfide linkages (data not shown). It was not possible to completely separate these isomers from each other under the conditions used in this experiment. The refolded proteins were then directly digested by endoproteinase Arg-C and the resulting digests were separated by HPLC, as previously reported (17). The ratios of the disulfide isomers in the digests, which have different disulfide linkages and were clearly separated on HPLC as is the case in our paper (17), were estimated by measurement of their peak areas on HPLC and shown in Fig. 2. The mutant protein, Met\(^1\)-pro-GCAP-II (2–86), consisted predominantly of the native-type Thr-Ile-Ala-uroguanylin (the native-type Thr-Ile-Ala-uroguanylin contains two disulfide linkages between Cys\(^{74}\) and Cys\(^{82}\) and Cys\(^{87}\) and Cys\(^{88}\) (17).
(Ref. 17), along with small amounts of biologically inactive isomers 1 and 2 (the positions of the disulfide bonds of isomer 1 are between Cys74 and Cys85 and Cys77 and Cys82, and isomer 2 between Cys74 and Cys77 and Cys82 and Cys85 (Ref. 17)), as found in the folding of the wild-type pro-GCAP-II (17). This indicates that the mutation of the Val residue to Met at the N terminus had no significant effect on the folding of pro-GCAP-II. The mutant proteins, Met2-pro-GCAP-II-(3–86) and Met3-pro-GCAP-II-(4–86), were composed of the native-type Thr-Ile-Ala-uroguanylin, isomer 1 and isomer 2 in ratios of 1: 1.14: 0.82 and 1: 4.4: 2.10, respectively. These data suggest that the deletion of the amino acid residue at the N terminus greatly affects the construction of the native tertiary structure in the mature region of pro-GCAP-II, because the native-type disulfide pairing comprises only one-third of the mutant protein (Met2-pro-GCAP-II-(3–86)). Further, the mutant protein, Met3-pro-GCAP-II-(4–86), in which two amino acid residues were deleted from the N terminus, nearly completely lacked the ability to form the correct disulfide pairing in the mature region and, thus, was devoid of the chaperone function in the pro-leader sequence of pro-GCAP-II, because the ratio of the native type to isomers was comparable with that in the folding of the mature hormone, GCAP-II (17).

To further investigate whether the Tyr residue at position 2 from the N terminus is involved in the folding of pro-GCAP-II, the mutant protein, Met1,2-pro-GCAP-II-(3–86), was prepared in which the Tyr residue was replaced by Met. The result indicates that the distribution of the native-type isomers 1 and 2 in Met1,2-pro-GCAP-II-(3–86) were nearly the same as in the case of Met1-pro-GCAP-II-(2–86) and, therefore, that the replacement of the Tyr residue with Met had no effect on the function of the propeptide in the pro-leader sequence. Collectively, these results indicate that the two N-terminal amino acid residues in length, in particular the N-terminal residue, play an important role in the formation of the correct disulfide linkages of the mature portion of pro-GCAP-II in vitro and, thus, in the function of the intramolecular chaperone of the propeptide in the pro-leader sequence of pro-GCAP-II.

Expression of the N-terminal Deletion Mutants of Pro-GCAP-II in 293T Cells—Since the mutant proteins were expressed as inclusion bodies in the E. coli cells, we were not able to estimate the effect of the deletion of the N-terminal residue on the in vivo folding of pro-GCAP-II. Therefore, we prepared the wild-type pro-GCAP-II in human embryonic kidney 293T cells, as well as the N-terminal deletion mutants, in which the N-terminal amino acid residues were sequentially deleted from the N terminus, and three types of mutant proteins of pro-GCAP-II, which are devoid of the N-terminal region, as shown in Fig. 2: 1) pro-GCAP-II-(7–86), which is deprived of the 6 N-terminal amino acid residues; 2) pro-GCAP-II-(12–86), which lacks the 11 N-terminal residues; and 3) pro-GCAP-II-(18–86), which lacks the 17 N-terminal residues that comprise an invariant region (amino acid residue sequence 12–17) in both pro-GCAP-II and pro-GCAP-I. The N-terminal deletion mutant proteins, which lack the N-terminal Val or Val-Tyr residues, could not be isolated from 293T cells, although the reason for this is not clear at present. The mutant proteins, which were deleted in a portion of the peptide in the pro-leader sequence of pro-GCAP-II, might be due to its failure to fold in the endoplasmic reticulum, resulting in a protein that is susceptible to degradation by proteases and is not secreted from the endoplasmic reticulum (26). The other deletion mutants and the wild-type pro-GCAP-II were expressed in human kidney 293T cells via the expression vector, secreted into the culture media, and then purified by HPLC (Fig. 3) and analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Fig. 4). The wild-type pro-GCAP-II showed a signal at m/z 9487.0, which is consistent with the mass value (9487.9) calculated from the amino acid sequence. In contrast, pro-GCAP-II-(7–86), pro-GCAP-II-(12–86), and pro-GCAP-II-(18–86) exhibited mass spectral signals at m/z 17859.5, 16203.0, and 15248.0, respectively, which are twice the theoretical values, calculated from their amino acid sequences. No monomeric forms of the N-terminal deletion mutant proteins were detected on Fig. 3. These results indicate that the wild-type pro-GCAP-II was prepared as a monomer and, conversely, that the mutant proteins, which lack the N-terminal amino acid residues, were expressed as dimers. This dimer appears to be composed of two monomer units, which are connected via a covalent linkage(s), perhaps intermolecular disulfide linkage(s), because the dimer could be converted into a monomer by treatment with DTT, as described below.

To determine the location of the intermolecular disulfide linkage(s) found in the recombinant proteins, which are devoid of the N-terminal region, the wild-type pro-GCAP-II (as a control) and pro-GCAP-II-(12–86) were each digested with endoproteinase Arg-C and the hydrolysates examined by HPLC (Fig. 5). A comparison of the HPLC profiles of the digests of the wild-type pro-GCAP-II and pro-GCAP-II-(12–86) revealed that the wild-type pro-GCAP-II comprises the native-type Thr-Ile-Ala-uroguanylin covering the mature GCAP-II (observed mass value, 1953.0; theoretical value, 1953.2), whereas pro-GCAP-II-(12–86) contains a dimer (observed mass value, 3906.9; theoretical value, 3906.5) of Thr-Ile-Ala-uroguanylin. Two Cys residues at positions 41 and 54 in the pro-leader sequence of pro-GCAP-II were correctly bridged in both the wild-type pro-GCAP-II and pro-GCAP-II-(12–86). This clearly shows that the intermolecular disulfide linkage(s) in the recombinant pro-GCAP-II-(12–86) were connected in its mature region. The disulfide pairing in the dimer of Thr-Ile-Ala-uroguanylin in pro-GCAP-II-(12–86) could not be further defined, because the peptide was not soluble after purification by HPLC. The other N-terminal deletion proteins also gave the same results as found in pro-GCAP-II-(12–86) (data not shown). Consequently, these data indicate that the deletion of the N-terminal region in the pro-leader sequence of pro-GCAP-II greatly influenced the
linking of the disulfide bonds in the mature peptide of pro-GCAP-II, GCAP-II. In other words, the N-terminal region (amino acid residues 1 to 6) in the pro-leader sequence of pro-GCAP-II plays a critical role in the formation of the three-dimensional structure of pro-GCAP-II in vivo and in turn, the folding of the mature peptide, GCAP-II.

In Vitro Disulfide-coupled Folding of the N-terminal Deletion Mutants Expressed in 293T Cells—The S-protein (amino acid sequence 22–124) of RNase A folds complementarily with the S-peptide (amino acid sequence 1–21) to a stable conformation similar to that of intact RNase A (27–29). This led us to determine if pro-GCAP-II-(7–86) or pro-GCAP-II-(12–86) were able to adopt a tertiary structure similar to that of the intact protein by the aid of the complementary N-terminal peptides, which consist of 6 or 11 amino acid residues, respectively. The relative abundance of three disulfide isomers of the refolded proteins, pro-GCAP-II-(7–86) or pro-GCAP-II-(12–86) were able to adopt a tertiary structure similar to that of the intact protein by the aid of the complementary N-terminal peptides, which consist of 6 or 11 amino acid residues, respectively. The relative abundance of three disulfide isomers of the refolded proteins, pro-GCAP-II-(7–86) or pro-GCAP-II-(12–86) were nearly identical with those found in the refolding of the mature GCAP-II (17) and Met<sup>3</sup>-pro-GCAP-II-(4–86) (data not shown). Next, these N-terminal deletion proteins were reduced with DTT and, after removal of the reducing reagent, incubated with the corresponding peptides under the same conditions that were used for the refolding of the full-length protein. The N-terminal deletion proteins did not efficiently adopt a native conformation in the presence of the complementary peptides, as is the case for RNase A (data not shown). It should be noted, however, that the possibility that the complementary refolding of the N-terminal deletion proteins with the rest of the peptides could occur under conditions different from those used in this experiment cannot be excluded. Insulin and chymotrypsin, in contrast to their precursor proteins, have been observed to become thermodynamically unstable after the release of their propeptides. In this case, the mature proteins are required to fold to the proper structure in the original single-chained precursor proteins with the propeptides, which aid in guiding the proteins to the native conformation (30, 31). It is likely in the case of those proteins that the N-terminal portion in the pro-leader sequence of pro-GCAP-II functions to stabilize the three-dimensional structure of pro-GCAP-II.

In eukaryotic translation systems, a protein initiates cotranslational folding from the N-terminal region prior to the release of the nascent chain from the ribosome, and achieves a native conformation after the synthesis of the entire protein (32). As described above, the disulfide bond between two Cys residues at positions 41 and 54 in the pro-leader sequence of pro-GCAP-II was correctly formed in the N-terminal deletion proteins both in vitro and in vivo. This implies that the propeptide region in these proteins assemble the three-dimensional structure in a manner similar to that in the wild-type pro-GCAP-II. However, the N-terminal deletion proteins of pro-
GCAP-II gave rise to incorrect conformations of their mature regions, when expressed in vivo, and the native disulfide pairing was not predominant, when refolded in vitro. In addition, the N-terminal deletion proteins of pro-GCAP-II were refolded into only one type of disulfide isomer in vivo (Fig. 3), whereas these were obtained as a mixture of three disulfide isomers in vitro, as mentioned above. This may reflect a difference in the folding mechanism of pro-GCAP-II (or possibly proteins in general) under in vivo versus in vitro conditions. Considering these observations, we speculate that a propeptide region in the pro-leader sequence of pro-GCAP-II is folded at an earlier stage than the folding stage of the mature portion of pro-GCAP-II in vivo. The N-terminal deletion proteins did not efficiently achieve the proper tertiary structure but, rather, gave rise to, largely, misfolded protein in the folding pathway in vivo, because of the absence of the N-terminal region, which may interact with the mature region in the final stage of the folding of pro-GCAP-II. Indeed, our preliminary data for the CD analysis of the folding intermediate suggested that the misfolded GCAP-II was produced during the initial folding and gradually converted to the native type of pro-GCAP-II after the formation of the native conformation of the propeptide. In this stage we conclude that the propeptide region in the pro-leader sequence of pro-GCAP-II may contribute toward the thermodynamic control of the folding of pro-GCAP-II and regulates the proper folding of the mature region of pro-GCAP-II, GCAP-II.

**Dimerization of Pro-GCAP-II**—As is known in prosubtilisin, numerous proteases, such as subtilisin and α-lytic protease, are synthesized in the form of enzymatically inactive precursor proteins with pro-leader sequences, which then proceed to the active enzymes via the cleavage of the pro-sequences, which functions as an intramolecular chaperone. The precursor protein, prosubtilisin, forms a dimer, not only as a folding intermediate but also in the product just prior to processing to active enzymes (21, 22). Dimerization is an important step in the construction of the native tertiary structure of prosubtilisin. To understand the issue of whether pro-GCAP-II fits this case, i.e. if pro-GCAP-II dimerizes both as an intermediate(s) and as the final product in the folding process, we first investigated the behavior of pro-GCAP-II, which was expressed from 293T cells, and pro-GCAP-I (23) by size exclusion chromatography. Fig. 6 shows that pro-GCAP-II (9.5 kDa) and pro-GCAP-I (19 kDa) were eluted at retention times corresponding to 15.5 and 21 kDa, respectively. This observation demonstrates that these proteins behave as the homodimers in solution through non-covalent interactions, rather than as monomeric forms, although the observed molecular mass of pro-GCAP-II was smaller than the theoretical value for the dimer of pro-GCAP-II (19 kDa). Second, to examine whether dimerization is involved in the folding of pro-GCAP-II, we analyzed the efficiency of the refolding of pro-GCAP-II at various concentrations in vitro, since the dimerization of a protein is usually concentration dependent. pro-GCAP-II was refolded to the correct conformation in almost quantitative yield at 10⁻⁵ M protein concentration (17), although in a rather low yield (less than 10%) at concentrations of less than 10⁻⁶ M (data not shown). This result suggests that dimer formation is involved in the folding pathway of pro-GCAP-II.

In addition, we observed that the N-terminal deletion proteins do not form intermolecularly disulfide-linked dimers in the presence of protein disulfide isomerase in the in vitro refolding (data not shown). This suggests that the dimerization of the N-terminal deletion proteins is not the result of compliance with protein disulfide isomerase, which functions in protein folding in vivo (33, 34). Therefore, the dimeric form of the N-terminal deletion proteins may represent an intermediate in the folding pathway of pro-GCAP-II and could be kinetically trapped during the folding of the mutant proteins in vivo. Collectively, these results imply that pro-GCAP-II forms a dimer as an intermediate(s) in the folding pathway and suggests that pro-GCAP-II exists as a dimer in the final product. As described above, the deletion of the N-terminal region in the pro-leader sequence of pro-GCAP-II resulted in a failure to form the correct disulfide linkages in the mature portion. Consequently, the mature portion in the N-terminal deletion protein results in a different conformation from that of the wild-type pro-GCAP-II and in turn is unlikely to interact with each other to afford a dimer. This suggests that the propeptide portion, in particular, the region (amino acid sequence 18–62), in the pro-leader sequence of pro-GCAP-II might be involved in the creation of an interface in the interaction between two molecules of pro-GCAP-II for affording the dimer. Thus, it appears that the dimerization directly regulates the folding of pro-GCAP-II and contributes to the formation and stabilization of the three dimensional structure of both pro-GCAP-II and the mature portion of pro-GCAP-II, GCAP-II.

**Competitive Refolding of Pro-GCAP-II and Pro-GCAP-II-(12–86)**—The observation that pro-GCAP-II exists as a dimer in the intermediate and final states of the folding process raises the question of whether pro-GCAP-II is folded to a dimer intramolecularly (in cis) or intermolecularly (in trans). We tested the in vitro refolding of pro-GCAP-II in the presence of pro-GCAP-II-(12–86), which is devoid of the 11 N-terminal amino acid residues. If there are intermolecular interactions between the two proteins, pro-GCAP-II should yield both the correctly refolded protein and the misfolded protein. Mixing pro-GCAP-II with pro-GCAP-II-(12–86) in their unfolded states had no effect on the yield and the nature of their refolding (data not shown). This result demonstrates that pro-GCAP-II was refolded to the correct conformation without interference by pro-GCAP-II-(12–86), i.e. by an intramolecular interaction (in cis), and vice versa under the conditions used in this experiment. In order to determine whether pro-GCAP-II forms a dimer in vivo intramolecularly or intermolecularly, pro-GCAP-II was co-expressed with pro-GCAP-II-(12–86) in 293T cells and the culture media were directly analyzed by HPLC, as shown in Fig. 7 (A–D). pro-GCAP-II was expressed as a monomer (Fig. 7A, peak a) and pro-GCAP-II-(12–86) as a dimer (Fig. 7A, peak b). In addition, the formation of a hetero-dimer between pro-GCAP-II and pro-GCAP-II-(12–86) or other species related to these proteins have not been detected by HPLC. This indicates that pro-GCAP-II was expressed and folded to the native conformation without any influence by the co-existence of pro-GCAP-II-(12–86) in 293T cells. However, pro-GCAP-II-(12–86) was recovered in a low yield in the co-expression experiment, as shown in Fig. 7A. This might be due to the low expression efficiency of pro-GCAP-II-(12–86), but the precise reason for
Prosubtilisin, the precursor protein of subtilisin, forms a dimer through an intermolecular interaction (in cis) between two molecules, and has a “molten globule”-like structure, in the folding region, which are localized on the surface of each the propeptide region to be protected from exposure to solvent in an aqueous environment, because the propeptide in the pro-leader sequence of pro-GCAP-II is thermodynamically stabilized by the dimerization of a propeptide in the pro-leader sequence of pro-GCAP-II. Therefore, the role of the N-terminal amino acid residues in pro-GCAP-II may be to stabilize the native conformation of pro-GCAP-II, in order to kinetically trap the correct tertiary structure of the mature region, GCAP-II. The precise role of the N-terminal region in the folding of pro-GCAP-II might be explained by an analysis of the folding intermediates. This is currently under way in our laboratory.

Second, a propeptide in the pro-leader sequence of pro-GCAP-II is tightly associated with its dimerization via the intermolecular interaction between the region involving the propeptide. The mature form of pro-GCAP-II, GCAP-II, obeys “non-Anfinsen” folding (10). This is analogous to the folding behavior of insulin-like growth factor I (IGF-I) (10) but different from the propeptide-mediated folding of proteins such as subtilisin, in which the propeptide appears to act at the kinetic level by accelerating the folding rate (20). The dimerization of IGF-I itself is not involved in the folding pathway, but it has recently been reported that an IGF-binding protein can facilitate the correct folding of IGF-I and that the heterodimer between IGF-I (or a precursor form of IGF-I) and IGF-binding protein thermodynamically stabilizes the native conformation of IGF-I (13). pro-GCAP-II is not able to activate its receptor protein, guanylyl cyclase C, indicating that the mature portion of pro-GCAP-II, GCAP-II, is buried inside the structure of pro-GCAP-II and that the two disulfide bonds in the mature region are not exposed to the surface of the molecule or, if so, that they are concealed between the dimer of pro-GCAP-II. Therefore, one can speculate that the native conformation of pro-GCAP-II is thermodynamically stabilized by the dimerization of a propeptide in the pro-leader sequence of pro-GCAP-II. Possibly, the dimer might be formed at the intermediate(s) stage during the folding process. This might stabilize the rate-determining transition states, as well as the propeptide in subtilisin, and prevent the incorrect aggregation by protecting the hydrophobic surface of the protein molecule from environmental solvents. Interactions with other proteins may improve the kinetics of the appearance of the functional conformation or to prevent the formation of incorrectly folded structures, which might be lost to the folding pathway.

In conclusion, the present result shows that the propeptide of pro-GCAP-II has two roles, a chaperone function and a dimerization function.
ization function. The issue of whether these function independently or in a concerted manner will require further study. In any event, the data contribute to an explanation of the pathways by which pro-GCAP-I or pro-GCAP-II attain the proper folding to provide the biologically active mature peptides and generally, of the folding mechanism of propeptide hormones to the matured hormones.

Acknowledgments—We thank Professor Katsuya Nagai (Division of Protein Metabolism, Institute for Protein Research, Osaka University, Osaka, Japan) for help in performing the expression of recombinant proteins from human kidney 293T cells. Use of the facility at the Radio Isotope Research Center of Osaka University is acknowledged.

REFERENCES

1. Lazure, C., Seidah, N. G., Pelaprat, D. & Chretien, M. (1983) Can. J. Biochem. Cell Biol. 61, 501–515
2. Cuurie, M. G., Fok, K. F., Kato, J., Moore, R. J., Hamra, F. K., Duffin, K. L. & Smith, C. E. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 947–951
3. Hamra, F. K., Forte, L. R., Eber, S. L., Pidhorodeckyj, N. V., Krause, W. J., Freeman, R. H., Chin, D. T., Tempkings, J. A., Fok, K. F., Smith, C. E., Duffin, K. L., Siegel, N. R. & Currie, M. G. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10464–10468
4. Hess, R., Kuhn, M., Schultz-Knappe, P., Raida, M., Adermann, K., Kaever, V., Cetin, Y. & Forssmann, W.-G. (1995) FEBS Lett. 374, 34–38
5. Shultz, S., Green, C. K., Yuen, P. S. T. & Garbers, D. L. (1992) J. Biol. Chem. 267, 16019–16021
6. de Sauvage, F. J., Keshav, S., Kuang, W.-J., Gillett, N., Henzel, W. & Goeddel, D. V. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9089–9093
7. Hober, S., Hansson, A., Uhlen, M. & Nilsson, B. (1997) Biochemistry 36, 6758–6761
8. Hamra, F. K., Fan, X., Krause, W. J., Freeman, R. H., Chin, D. T., Smith, C. E., Currie, M. G. & Forte, L. E. (1996) Endocrinology 137, 257–265
9. Shultz, S., Chrisman, T. D. & Garbers, D. L. (1992) J. Biol. Chem. 267, 16019–16021
10. Whitaker, T. L., Witte, D. P., Scott, M. C. & Cohen, M. B. (1997) Gastroenterology 113, 1009–1006
11. Fan, X., Hamra, F. K., Freeman, R. H., Eber, S. L., Krause, W. J., Lim, R. W., Pace, V. M., Currie, M. G. & Forte, L. R. (1996) Biochem. Biophys. Res. Commun. 219, 457–462
12. Netzer, W. J. & Hartl, U. (1997) Nature 388, 343–349
13. Netzer, W. J. & Hartl, U. (1997) Nature 388, 343–349
14. Ruoppolo, M., Lundstrom-Ljung, J., Talano, F., Puci, P. & Marino, G. (1997) Biochemistry 36, 12259–12267
15. Netzer, W. J. & Hartl, U. (1997) Nature 388, 343–349
16. Ruoppolo, M., Lundstrom-Ljung, J., Talano, F., Puci, P. & Marino, G. (1997) Biochemistry 36, 12259–12267
17. Wilson, R., Lees, J. F. & Bulleid, N. J. (1998) J. Biol. Chem. 273, 9637–9643