The micro domain responsible for ligand-binding of guanylyl cyclase C

Yuji Hidaka*, Yoshiko Matsumoto, Yasutsugu Shimoshini

Division of Organic Chemistry, Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565-0871, Japan

Received 21 May 2002; revised 12 July 2002; accepted 15 July 2002

First published online 30 July 2002

Edited by Gianni Cesareni

Abstract Guanylyl cyclase C (GC-C), a member of membrane-bound guanylyl cyclases, is a receptor protein for guanylin and uroguanylin. The binding of a ligand to the extracellular domain of GC-C (ECDGC-C) triggers signal transduction, resulting in the regulation of intestinal fluids and electrolytes. A previous study proposed that a ligand-binding site on GC-C is localized near the transmembrane region. To further investigate the mechanism by which GC-C is activated, the C-terminal ligand-binding region, was over-expressed in Escherichia coli and its ligand-binding ability was examined. The micro domain showed ligand-binding activity (IC₅₀ = 1×10⁻⁹ M). This result clearly indicates that a ligand-binding site is located in close proximity to the membrane-bound region, and that the micro domain is capable of independently binding the ligand, without assistance from other domains. The use of this micro binding domain in the study of interactions between GC-C and ligands could be a useful tool and could lead to a better understanding of GC-C signal transduction.

1. Introduction

Guanylyl cyclase C (GC-C) is a membrane-bound protein, which consists of an N-terminal extracellular domain (ECDGC-C), a transmembrane domain, and a C-terminal intracellular domain [1,2]. ECDGC-C is responsible for the ligand binding, resulting in the activation of the catalytic domain and, ultimately, the regulation of intestinal fluids and electrolytes, via the cystic fibrosis transmembrane conductance regulator protein [3]. GC-C is stimulated by several ligands, including the endogenous peptide hormones, guanylin and uroguanylin, and heat-stable enterotoxin (STα), the exogenous ligand [4-10]. Previously, we reported that a ligand-binding site is located at the region between Ser387 and Lys393, near the transmembrane domain [11]. This result poses the question of the minimal peptide region required to recognize the ligand.

Guanylyl cyclase A (GC-A), which is thought to have a tertiary structure similar to GC-C, is a member of the membrane-bound guanylyl cyclase family and is a receptor protein for atrial natriuretic peptide [12]. Recently an analysis of the crystal structure of the extracellular domain (ECDGC-A) of GC-A revealed that ECDGC-A consists of two subdomains, the N-terminal α-helical and the C-terminal β-sheet domains [12]. The identity of the region of the amino acid residues between GC-A and GC-C is about 10%, but that of the C-terminal regions of ECDGC-C and ECDGC-A, which corresponds to the β-sheet domain of ECDGC-A, is relatively high (23%), as shown in Fig. 1. In addition, a secondary structure prediction of ECDGC-C implies that the secondary structure of the C-terminal region of ECDGC-C consists of β-strands, which contain the ligand-binding site and is located in the same relative region as that of the C-terminal region of ECDGC-A, as shown in Fig. 1 [13]. Therefore, we speculate that the tertiary structure of the C-terminal region of ECDGC-C, referred to as the micro domain, is similar to that of the C-terminal region of ECDGC-A and that the micro domain might possess the ability to bind the ligand, without assistance from other domains.

Based on this hypothesis, the C-terminal polypeptide of ECDGC-C (the micro domain) was prepared and its ligand-binding ability examined. Based on the findings herein, the micro domain itself has a tertiary structure that is capable of binding the ligand. The data obtained permit a better understanding of the nature of signal transduction of GC-C.

2. Materials and methods

2.1. Materials

Restriction enzymes were purchased from Toyobo (Osaka, Japan) and New England BioLabs, Inc. (Beverly, MA, USA). Taq polymerase and T4 DNA ligase were obtained from Takara Shuzo Co. (Kyoto, Japan), STα (porcine STα with the amino acid sequence from positions 4 to 17) and [35S]-ANB-STα (4-17) were prepared according to a previously described procedure [14,15]. All other chemicals and solvents were of reagent grade. Polymerase chain reaction (PCR) was carried out using a Sanyo DNA amplifier MIR-D30 (Osaka, Japan).

2.2. Construction of the expression vector of the micro domain for the expression by Escherichia coli cells

The cDNA encoding the extracellular domain (Ser1-Gln407) of porcine GC-C(C349A), in which the Cys349 residue was mutated to an Ala residue, was prepared according to a previously reported method [14,15] using primers for the mutation (sense, ACTTGGAGAACAAGCCTCGGATATTGC; antisense, GTCATAATCCCGGGCGTGTGCACAGT). The cDNA was inserted into pVL1392

*Corresponding author. Fax: (81)-6-6879 8603.

E-mail address: yuji@protein.osaka-u.ac.jp (Y. Hidaka).

Abbreviations: PCR, polymerase chain reaction; TFA, trifluoracetic acid; HPLC, high performance liquid chromatography; Tris-HCl, tri(hydroxymethyl)aminomethane hydrochloride; IPTG, isopropyl β-thiogalactopyranoside; GC-C, guanylyl cyclase C; STα, heat-stable enterotoxin; STp(4-17), porcine STα with the amino acid sequence from positions 4 to 17; PAGE, polyacrylamide gel electrophoresis; ECDGC-C, extracellular domain of GC-C; ECDGC-A, extracellular domain of GC-A.
The protein was first purified by means of Cosmosil 140C18-OPN (Nacarai Tesque Inc., Kyoto, Japan), as described previously [7]. The mutant protein was finally purified by high performance liquid chromatography (HPLC), as described previously [9], and lyophilized. The protein thus prepared was characterized by mass spectrometry and amino acid analysis, as described previously [7].

2.3. Expression of the micro domain by E. coli cells

E. coli BL21(DE3) cells, transformed with the expression vector pY42, were cultured at 37°C in Luria broth supplemented with ampicillin (50 μg/ml). Expression of the mutant protein was induced by the addition of 1 mM isopropyl β-thiogalactoside (IPTG) at the mid-log phase of cell growth. After incubation at 37°C for 3 h, the cells were harvested and washed with PBS(−) containing 1 mM phenylmethylsulfonyl fluoride. The cells were resuspended in the same buffer, sonicated on ice, and centrifuged (15,000 × g, 20 min). The resulting precipitates were incubated with 10 mM dithiothreitol in 50 mM tri(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) (pH 8.0) containing 6 M guanidine-HCl at 50°C for 30 min. The protein was then purified using a Sephasil 400-CL column (Pharmacia). However, it was difficult to completely remove the non-specific binding by the baculovirus expression system was low, and the protein aggregated readily after purification. In addition, it is generally known that small peptide fragments of receptor proteins, such as the peptide fragment consisting of seven amino acid residues from Ser387 to Lys393 of ECDGC-C, cannot fold into the native conformation sufficient to bind to the ligand. Therefore, the minimum functional binding domain of ECDGC-C was explored in order to further study the interaction between ECDGC-C and ligands.

A recent X-ray crystallographic structural analysis of ECDGC-A indicated that the C-terminal region of ECDGC-A is comprised of a β-sheet structure [12]. The sequence alignment between ECDGC-A and ECDGC-C shows that these proteins share a 23% homology at the C-terminal region (from Met341 to Gln407 of GC-C), as shown in Fig. 1, implying that the tertiary structure of the C-terminal region of ECDGC-C, referred to as the micro domain, might be the same as that of ECDGC-A. The region between the Phe376 and the Ala403 residues of ECDGC-A, which corresponds to the region between the Met356 and Pro383 residues of ECDGC-C (Fig. 1), is comprised of a β-sheet domain in the crystal structure of ECDGC-A. In the case of ECDGC-C, this region contains the ligand-binding site, suggesting that this domain should be included as a ligand-binding unit of GC-C. In addition, our previous study indicated that the Asp347 residue of ECDGC-C has not yet been resolved.

Therefore, the micro domain (Met341 to Gln407) was prepared and used to understand the nature of the essential ligand-binding domain. The N-terminal Met341 residue is also involved in the ligand binding. The N-terminal Met341 to Gln407 residues of ECDGC-C, which corresponds to the region between the Met356 and Pro383 residues of ECDGC-C, was expressed by the baculovirus expression system [11]. Therefore, this region should supply the specific tertiary structure required for the ligand binding. However, the tertiary structure of the entire and the essential ligand-binding domain of ECDGC-C has not yet been reported.

Previously, ECDGC-C was prepared using a system consisting of baculovirus and insect cells in order to obtain structural information on the ligand binding of ECDGC-C [14]. However, the expression efficiency of the recombinant ECDGC-C by the baculovirus expression system was low, and the protein aggregated readily after purification. In addition, it is generally known that small peptide fragments of receptor proteins, such as the peptide fragment consisting of seven amino acid residues from Ser387 to Lys393 of ECDGC-C, cannot fold into the native conformation sufficient to bind to the ligand. Therefore, the minimum functional binding domain of ECDGC-C was explored in order to further study the interaction between ECDGC-C and ligands.

A recent X-ray crystallographic structural analysis of ECDGC-A indicated that the C-terminal region of ECDGC-A is comprised of a β-sheet structure [12]. The sequence alignment between ECDGC-A and ECDGC-C shows that these proteins share a 23% homology at the C-terminal region (from Met341 to Gln407 of GC-C), as shown in Fig. 1, implying that the tertiary structure of the C-terminal region of ECDGC-C, referred to as the micro domain, might be the same as that of ECDGC-A. The region between the Phe376 and the Ala403 residues of ECDGC-A, which corresponds to the region between the Met356 and Pro383 residues of ECDGC-C (Fig. 1), is comprised of a β-sheet domain in the crystal structure of ECDGC-A. In the case of ECDGC-C, this region contains the ligand-binding site, suggesting that this domain should be included as a ligand-binding unit of GC-C. In addition, our previous study indicated that the Asp347 residue of ECDGC-C has not yet been resolved. Moreover, the sequence between Gly342 and Leu346 of GC-C is completely conserved in GC-Cs from several species, also implying that it plays a significant role in ligand binding. The N-terminal Met341 residue is also included since the Met residue is generated by the starting codon in the E. coli expression system used in this study. Therefore, the micro domain (Met341 to Gln407) was prepared and its ligand-binding ability examined, in an attempt to better understand the nature of the essential ligand-binding domain.

The region between the Met341 and the Gln407 residues of ECDGC-C contains one cysteine residue at position 349. This cysteine residue at position 349 is conserved in all GC-Cs. Moreover, the sequence between Gly342 and Leu346 of GC-C is completely conserved in GC-Cs from several species, also implying that it plays a significant role in ligand binding. The N-terminal Met341 residue is also included since the Met residue is generated by the starting codon in the E. coli expression system used in this study. Therefore, the micro domain (Met341 to Gln407) was prepared and its ligand-binding ability examined, in an attempt to better understand the nature of the essential ligand-binding domain.
Cys349 residue is not found in other species of GC-C, suggesting that it may not play a role in ligand binding. To avoid the formation of an intermolecular disulfide bond at the Cys349 residue, it was mutated to an Ala residue. For this purpose, mutant C349A of the entire ECDGC/C0C was first prepared and its ligand-binding ability examined, using a previously reported method [14]. The mutant C349A of ECDGC/C0C showed a ligand-binding ability identical to that of wild-type ECDGC/C0C (Fig. 2), indicating that the Cys349 is not absolutely required for the binding of GC-C with ligands.

The cDNA encoding the micro domain, in which the Cys349 residue was mutated to an Ala residue, was generated by PCR using the cDNA encoding the mutant C349A of ECDGC/C0C as a template, and was inserted into the expression vector, pET17b (Novagen), resulting in the construction of the expression vector pY42. The micro domain was expressed as an inclusion body in E. coli BL21(DE3) cells transformed with pY42, as shown in Fig. 3A. Therefore, the micro domain was purified by HPLC in a denatured form (Fig. 3B). The purified micro domain showed a signal at m/z = 8498.4, consistent with its mass value (8498.1) calculated from the amino acid sequence, in a mass spectrometric analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.

The solubility of the purified micro domain, which is in a denatured form, was poor at neutral pH, but it was soluble in 0.05% TFA (pH 2.3). It has been reported that GC-C binds to the ligand at pH 4–8 [17]. Therefore, the purified micro domain was refolded by the dilution method under various conditions, followed by altering the solution pH. The purified micro domain was dissolved in 0.05% TFA and then mixed with nine volumes of 50 mM Tris–HCl (pH 7.4) or 50 mM AcONa (pH 4.7). The micro domain existed as a soluble form at neutral pH after this procedure as well as under acidic conditions, implying that the micro domain was folded into the native conformation.

The ligand-binding ability of the refolded micro domain was examined using a photoaffinity-labeled STa analog, 125I-ANB-STp(4–17), as shown in Fig. 4A,B. The negative control in this experiment was performed in the presence of 6 M guanidine–HCl (pH 7.4) (lanes 1 and 2 in Fig. 4A,B). The micro domain could not be photoaffinity-labeled with 125I-ANB-STp(4–17) under denaturing conditions, indicating that non-specific binding was not significant. However, the refolded micro domain, refolded with 50 mM Tris–HCl (pH 7.4), was labeled with 125I-ANB-STp(4–17) and the binding was specifically inhibited in the presence of unlabeled STp(4–17) (lanes 3 and 4 in Fig. 4A). Uroguanylin also inhibited the binding of 125I-ANB-STp(4–17) to the micro domain (lanes 7 and 8 in Fig. 4B). However, the binding ability of uroguanylin was lower than that of STp(4–17) as well as in the case of GC-C [18], and uroguanylin could not completely inhibit the 125I-ANB-STp(4–17) binding to the micro domain even in the presence of 10^{-5} M uroguanylin (lane 8 in Fig. 4B).
Since GC-C is able to bind the ligand under acidic conditions [17], as described above, the ligand-binding ability of the micro domain was examined under acidic conditions. The refolded micro domain, which had been refolded with 50 mM AcONa (pH 4.7), was also able to bind STp(4^-17) at pH 4.7, but non-specific binding was significant (lanes 5 and 6 in Fig. 4A,B). These data clearly show that the micro domain possesses the ability to bind to the ligand.

To estimate the binding capacity of the micro domain, a competitive binding assay was performed in the presence of several concentrations of cold STp(4^-17) and the ligand-binding ability was estimated by autoradiographic analysis. Fig. 5 shows the competitive binding curve between the micro domain and ^125^I-ANB-STp(4^-17) at equilibrium with an IC\textsubscript{50} value of 1 \times 10\textsuperscript{-8} M. This result clearly shows that the micro domain binds the ligand in the absence of other domains of GC-C. In order to determine the K\textsubscript{d} value of the micro domain for ligand binding, we attempted to use a gel filtration method to separate free and bound ligands, as described in Section 2. However, it was difficult to completely separate these peptides. Therefore, it would be necessary to construct a new assay system for the determination of the K\textsubscript{d} value of the micro domain. However, the binding affinity of the micro domain was lower than that of ECD\textsubscript{GC-C} (IC\textsubscript{50} value = 1 \times 10\textsuperscript{-8} M) in our assay system, as shown in Fig. 5 [14]. The K\textsubscript{d} value of ECD\textsubscript{GC-C} was 4 \times 10\textsuperscript{-10} M, as previously reported [14]. Therefore, the K\textsubscript{d} value of the micro domain might be one order lower than that of ECD\textsubscript{GC-C}. The lower affinity of the micro domain is not dependent on the lack of the sugar chain at the Asn379 residue, since the presence of glycosylation at the Asn residue had no effect on the ligand-binding affinity of ECD\textsubscript{GC-C} [15]. These results indicate that other domains of GC-C may cooperatively aid in ligand binding at the micro domain. Indeed, a preliminary structural analysis of the micro domain, based on circular dichroism (CD) measurements, indicates that the micro domain consists of small amounts of \alpha-helix and \beta-sheet (data not shown) although the predicted secondary structure is the \beta-sheet, suggesting that the micro domain may need assistance from other domains of GC-C to sufficiently fold into the native conformation so as to bind the ligand.

Previously Nandi et al. reported on the expression of the N-terminal domain of GC-C by E. coli [19]. They prepared the entire ECD\textsubscript{GC-C} including the transmembrane domain, and a truncated form of ECD\textsubscript{GC-C}, which corresponds to the region 1^-304. They reported that the entire extracellular domain possessed ligand-binding ability but that the truncated form did not. These results are consistent with the results reported here, since the micro domain corresponds to the region Met341^-Gln407. Therefore, it can be concluded that the critical ligand-binding site is located at the micro domain.

The micro domain of ECD\textsubscript{GC-C} corresponds, relatively, to the region 361^-428 of ECD\textsubscript{GC-A}. The putative ANP-binding sites on GC-A were assigned to regions 173^-188 and 191^-198 [20]. Although the tertiary structure of GC-C has not yet been elucidated, ECD\textsubscript{GC-C} might have a tertiary structure similar to that of ECD\textsubscript{GC-A}. Based on this hypothesis, the identified
STα-binding site (Ser387–Lys393) on the micro domain would be located at the opposite site of the ANP-binding site on the β-sheet domain of ECD\textsubscript{GC–A}, suggesting that the ligand-binding site of GC-C is located in a different region from that of GC-A or that GC-C may possess another ligand-binding site. Indeed, the mutant micro domain, in which the Trp392 in the STα-binding site of the micro domain was mutated to an Ala residue, showed a lower ligand-binding affinity than that of the micro domain (data not shown). Therefore, it is possible that the ligand-binding mode of GC-C may be different from that of GC-A.

In conclusion, the micro domain of ECD\textsubscript{GC–C} possesses the ability to bind ligands in the absence of other domains of GC-C. This micro domain should be a powerful tool to further investigate interactions between GC-C and ligands, and will provide new insights into the signal transduction of GC-C.

Acknowledgements: The authors wish to thank Professor Katsuhide Yutani (Division of Protein Physical Chemistry, Institute for Protein Research, Osaka University) for his help in the CD measurements. Use of the facility at the Radio Isotope Research Center of Osaka University is acknowledged.

References