Archives of Biochemistry and Biophysics 479 (2008) 153–157

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Mapping the putative binding site for uPA protein in Esophageal Cancer-Related Gene 2 by heteronuclear NMR method

Yong Geng [a], Yingang Feng [a], Tao Xie [a], Yuanyuan Dai [b], Jinfeng Wang [a,*], Shih-Hsin Lu [c]

[a] National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, China
[b] Department of Clinical Pharmacology, Cancer Institute, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100021, China
[c] Department of Etiology and Carcinogenesis, Cancer Institute, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100021, China

ARTICLE INFO

Article history:
Received 15 July 2008
and in revised form 25 August 2008
Available online 18 September 2008

Keywords:
ECRG2
Kazal-type domain
Serine proteinase inhibitor
Urokinase-type plasminogen activator
Reactive site loops

ABSTRACT

Esophageal Cancer-Related Gene 2 (ECRG2) is a novel member of the KAZAL-type serine proteinase inhibitor family and plays an important role in the inhibition of human esophageal cancer cell proliferation. The previous studies have shown that ECRG2 can bind the urokinase-type plasminogen activator (uPA)/plasmin system and inhibit its activity. In this study, the strategy of cloning, overexpression, and purification of ECRG2 for obtaining a properly folded ECRG2 with accurately formed disulfide bonds was established. The heteronuclear NMR experiments were performed with isotope labeled ECRG2 to investigate the binding interface of the protein with uPA. The sequence regions of ECRG2 for uPA binding were determined. Analysis indicates that the uPA-binding loops of ECRG2 are in correspondence with the reactive site loops for binding of serine proteinase in turkey ovomucoid third domain (OMTKY3). The structural similarity of ECRG2 to OMTKY3 was identified and a model for ECRG2 was proposed.

Esophageal Cancer-Related Gene 2 (ECRG2,[1] GenBank Accession No.: AF268198) is down-regulated in esophageal squamous cell carcinoma (ESCC) and involved in the induction of the apoptosis in esophageal cancer cell lines [1]. Bioinformatics analysis has indicated that ECRG2 is a novel member of the KAZAL-type serine proteinase inhibitor family showing 97% sequences homologous to a KAZAL-type serine proteinase inhibitor (http://www.smart.embl-heidelberg.de). ECRG2 gene encodes a protein consisting of 85 residues. The N-terminal 1–19 residues encode a signal peptide, and the C-terminal 30–85 residues consist of a conserved Kazal-type domain with the Kazal motif ($C–X_{various}–C–X_7–C–X_6–Y–X_3–C–X_{2,3}–C–X_{17}–C$), whereas the residues 20–29 link these two regions in the protein (Fig. 1). In esophageal cancer cells, ECRG2 acts as a bifunctional protein associated with the regulation of cell proliferation and induction of apoptosis [2]. ECRG2 can interact with metallothionein 2A (MT2A) which is involved in the regulation of cell proliferation and apoptosis as well as various other physiological processes [3,4], playing an important role in the carcinogenesis of esophageal cancer [2,5]. Recently, the growing data of evidence have indicated that some members of the KAZAL-type proteinase inhibitor family may be involved in the regulation of invasion and metastasis of tumor cells. RECK (reversion-inducing cysteine-rich protein with Kazal motif) is a membrane-anchored glycoprotein that negatively regulates matrix metalloproteinases (MMPs) and inhibits tumor metastasis and angiogenesis by negatively regulating the activity of MMPs which are essential for proper extracellular matrices (ECM) remodeling [6–8]. A tumor-associated trypsin inhibitor (TATI) was found to be identical to the earlier described pancreatic secretory trypsin inhibitor (PSTI) [9], which is also called the Kazal inhibitor. The ECRG2 sequence is highly homologous to RECK and TATI (Fig. 1), suggesting the possible ability of ECRG2 for degradation of the ECM. More recent research has shown that ECRG2 can inhibit the activity of urokinase-type plasminogen activator (uPA)/plasmin system, and plays an important role in the prevention of tumor cell migration and invasion by regulation of plasmin-mediated proteolysis of the ECM [10].

The present study is aimed to explore the binding interface between ECRG2 and uPA by heteronuclear NMR experiments. Thus, the recombinant isotope labeled ECRG2 is required. However, the Kazal-type domain of ECRG2 contains conserved cysteine residues, Cys32, Cys45, Cys53, Cys64, Cys67, and Cys85, which are expected to form three pairs of disulfide bonds (Fig. 1). Here, we report the cloning, overexpression, and purification of ECRG2 which provided a properly folded protein with accurately formed disulfide bonds. The model structure of ECRG2 was obtained, and the putative binding site of uPA protein on ECRG2 was determined.

* Corresponding author. Fax: +86 10 64872026.
E-mail address: jfw@sun5.ibp.ac.cn (J. Wang).

[1] *Abbreviations used:* ECRG2, Esophageal Cancer-Related Gene 2; ESCC, esophageal squamous cell carcinoma; MT2A, metallothionein 2A; MMPs, matrix metalloproteinases; ECM, extracellular matrices; TATI, tumor-associated trypsin inhibitor; PSTI, pancreatic secretory trypsin inhibitor; uPA, urokinase-type plasminogen activator.

0003-9861/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2008.08.023

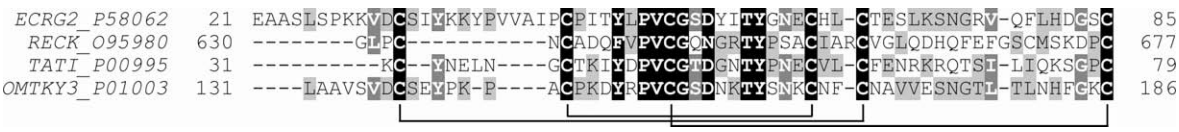

Fig. 1. Sequence alignment of human ECRG2 with RECK from human fibroblast, TATI from urine of a patient with ovarian cancer, and OMTKY3 from turkey egg white. The disulfide bridges between conserved cysteine residues in the Kazal domain of the sequences are indicated.

Materials and methods

Cloning, overexpression, and purification of ECRG2

Analysis of ECRG2 sequence has indicated that the segment of the N-terminal 19 residues behaviors as a trans-membrane peptide. For obtaining the recombinant ECRG2 protein, these 19 amino acids were truncated in the construction of the expression vector. Thus, peptide ECRG2(20–85) (called as ECRG2 hereafter), was expressed and purified for present study.

A DNA fragment encoding ECRG2 was amplified by PCR using the forward primer (cgcgggggaattcattgagggtcgctcagaagctgctagtctgtctcca) and backward primer (aaagaagtgctaccttcaacgattttcgagcggggc) which introduce an EcoRI site and a sequence encoding a factor Xa cleavage site at the 5′ end and a HindIII site at the 3′ end of the gene, respectively. The fragment was inserted into the T7 RNA polymerase-based expression vector pET-32a. Another DNA fragment encoding a linker peptide (SGSGSGSGSG) including EcoRI and EcoRV sites was synthesized and inserted into the vector pET-32a. The resulting expression vector (Supplementary material (SM) Fig. 1) for thioredoxin (Trx) fusion protein ECRG2, namely pET-32a-Trx-(linker-Xa)-ECRG2, was used for the expression and purification of ECRG2.

The pET-32a-Trx-(linker-Xa)-ECRG2 plasmid was transformed into the oxidative cytoplasm of *Escherichia coli* Origami (DE3) strain. The *E. coli* Origami (DE3)/pET-32a-Trx-(linker-Xa)-ECRG2 strain was grown overnight in 100 ml LB media (with 50 μg/ml ampicillin and 15 μg/ml kanamycin) at 37 °C. The overnight culture was used to inoculate another 1000 ml LB media. When the culture OD_{600} reached 0.6, the fusion protein expression was initiated by adding IPTG to a final concentration of 1 mM. After 20 h of additional cultivation at 16 °C, cells were harvested by centrifugation at 4000 rpm for 30 min. The cell was then resuspended in 20 ml of buffer A (50 mM Tris/HCl, pH 8.0, 300 mM NaCl, 10 mM imidazole) with 0.5 mM PMSF, 0.2 mg/ml lysozyme, 0.2 mg/ml DNaseI, and 2 mg/ml benzamidine. The cells were lysed by three 30 s sonication steps at 4 °C (Labsonic U sonicator at 150 W intensity) followed by overnight freezing at −80 °C. After thawing, the extract was centrifuged for 60 min at 40,000 rpm and 4 °C to separate the soluble and insoluble portions.

The soluble fusion protein Trx-(linker-Xa)-ECRG2 was purified using a metal-chelating affinity chromatography. For this, the supernatant was collected and loaded onto a Ni^{2+}-chelating column (10 ml) pre-equilibrated with buffer A. The column was extensively washed first with the buffer A, followed by buffer A containing 10 mM imidazole, and finally with factor Xa buffer (50 mM Tris–HCl, pH 8.0, 140 mM NaCl, 5 mM $CaCl_2$). Then the factor Xa (1 U/mg of Trx-Xa-ECRG2) was loaded on the column for cleavage of the Trx fusion protein, and the solution in the column was circulated by a pump for 16 h at room temperature. Afterwards, the column was washed again with 50 ml of the buffer A. The ECRG2 containing fractions were concentrated to 5 ml by an ultrafiltration. The collected solution was loaded onto a Superdex 50 column pre-equilibrated with 0.1 mM NH_4HCO_3 buffer. Then the column was washed with 0.1 mM NH_4HCO_3 buffer. The collected fraction containing ECRG2 was lyophilized. Eluted fraction of protein was >99% pure as verified by SDS–PAGE. To substantiate the identity of purified ECRG2, mass spectrum was recorded on an Applied Biosystems Voyager MALDI-TOF mass spectrometer.

Uniformly ^{15}N- and ^{13}C-single labeled and $^{15}N/^{13}C$-double labeled ECRG2 were obtained by growth in minimal media containing $^{15}NH_4Cl$ and/or [^{13}C]-glucose as the sole nitrogen and/or carbon sources. The purity of proteins was checked by SDS–PAGE to ensure a single band.

NMR spectroscopy

The NMR samples contained 1.0 mM ^{15}N-, ^{13}C-, or $^{15}N/^{13}C$-labeled ECRG2, 100 mM KCl, and 0.01% NaN_3 in 50 mM deuterated acetate buffer (pH 5.0). All NMR experiments were run on a Bruker DMX 600 MHz spectrometer equipped with a triple-resonance cryo-probe at 293 K. The 2D $^1H–^{15}N$ HSQC and 3D $^1H–^{15}N–^{13}C$ HNCO, HNCA, HN(CO)CA, HN(CA)CO, HNCACB, CBCA(CO)NH, and HBHA(CO)NH, and 3D $^1H–^{15}N$ TOCSY-HSQC and NOESY-HSQC experiments were adopted for the resonance assignments. All NMR data were processed and analyzed with FELIX software (Accelrys Inc.). ^1H chemical shifts were referenced to the internal DSS (2,2-dimethyl-2-silaentane-5-sulphonate). ^{15}N chemical shifts were referenced indirectly [11].

Results

Recombinant ECRG2 for NMR experiments

In this study, the stable-isotope-labeled ECRG2 is required for analysis the interaction of ECRG2 with uPA. Hence, the pET-32a-Trx-(linker-Xa)-ECRG2 expression plasmid was obtained, and the desired fusion protein Trx-(linker-Xa)-ECRG2 was expressed in the soluble form. As is well known, missing or wrong disulfide bridges often cause formation of insoluble aggregates when extracellular proteins containing cysteine residues are expressed in the conventional *E. coli* strains [12]. However, using *E. coli* Origami (DE3) strain which carries a trxB-/gor522-double mutation in the expression of the putative serine proteinase inhibitor allows disulfide formation in its oxidative cytoplasm [13,14]. Actually, expression of ECRG2 in the oxidative cytoplasm of *E. coli* Origami (DE3) provided a soluble form of the protein which can provide a properly folded ECRG2.

Moreover, the constructed plasmid and the expression and purification procedures can facilitate folding of ECRG2 with the correctly formed disulfide bonds. In the fusion protein Trx-(linker-Xa)-ECRG2, there is a linker peptide containing SGSGSGSGSG and a factor Xa cleavage site. This linker peptide makes a relatively large space between the two proteins, facilitating the proper folding of ECRG2 in the fusion protein and the proper cleavage of ECRG2 from the fusion protein using factor Xa. Factor Xa is a proteinase cleaving the fusion protein without the additional amino acid at the N-terminus of the target protein. Thus, the recombinant ECRG2 contains 66 residues (Ser20-Cys85) (Fig. 1). The mass of ECRG2 determined by MALDI-TOF MS analysis is 7192.77 (SM Fig. 2).

A recent report [15] has proposed that the ^{13}C chemical shifts can predict disulfide formation, providing simple rules for predict-