Construction and characterization of a functional variant hFGF7 with enhanced properties by circular permutation

Hye-Ji Choi[1], Hanui Lee[2], Dae-Eun Cheong[1], Su-Kyoung Yoo[1], Dong-Eun Lee[2], and Geun-Joong Kim[1]

[1]Chonnam National University College of Natural Science
[2]Korea Atomic Energy Research Institute

March 30, 2024

Abstract

Human fibroblast growth factor 7 (hFGF7) is a member of the paracrine-acting FGF family and mediates various reactions such as wound healing, tissue homeostasis, and liver regeneration. These activities make it a plausible candidate for pharmaceutical applications as a drug. However, the low expression level and stability of the recombinant hFGF7 were known to be major hurdles for further applications. Here, we attempted to improve the expression level and stability of hFGF7 by changing the order of amino acids through circular permutation (CP), thereby expecting an alternative fate according to N-end rules. CP-hFGF7 variants were constructed systematically by using putative amino acid residues in the loop region that avoided the disruption of the structural integrity especially in the functional motif. Among them, cp-hFGF7$^{115\text{-}114}$ revealed a relatively higher expression level in the soluble fraction than the wild-type hFGF7 and was efficiently purified to apparent homogeneity by three consecutive steps. Approximately, 7 mg of the purified protein was obtained from a 1 L culture and showed a comparable activity to that of the wild type hFGF7. In addition, the spectral properties related to the apparent structure of cp-hFGF7$^{115\text{-}114}$ were quite similar to those of the wild type hFGF7. Therefore, CP could be an alternative tool for the functional expression of hFGF7 in *Escherichia coli*.

1. INTRODUCTION

Fibroblast growth factors (FGFs) are cytokine family proteins comprised of 22 members that function in cell survival, proliferation and differentiation.[1] The fibroblast growth factor 7 (FGF7), also known as keratinocyte growth factor 1 (KGF1), is a member of the paracrine-acting FGF family. These proteins are secreted from various types of mesenchymal cells and activate on epithelial cells.[2, 3] The activity of human FGF7 (hFGF7) elicits a local signaling response by primarily binding to the high-affinity FGF receptor 2b (FGFR2b) with heparin sulfate/heparin as a cofactor.[4] According to previous reports, hFGF7 have functioned well in various human diseases for repairing cells/tissues, diabetic wound healing, cartilage diseases and liver regeneration.[2, 5-7] Owing to the increasing demand for the production of hFGF7 with versatile function, the attempts to produce recombinant hFGF7 in various expression hosts including *E. coli*has been proliferating.[8-12]

The expression systems in *E. coli* are widely used for recombinant protein production because of the easy and diverse genetic tools and the availability of high cell density culture techniques. However, many proteins are yet expressed in the soluble fraction due to innate hurdles, such as the limited landscape of protein folding and incorrect pairing of disulfide bonds, leading to unsatisfactory quality and yields of recombinant proteins especially as inclusion bodies.[13-15] In this context, hFGF7 is also categorized as a difficult-to-express protein in *E. coli* . As an existing method to overcome these problems, solubility enhancers, such as Halo tag[9] and small peptide with 6×His tag (6HFh8),[11] are fused to the recombinant hFGF7. The fusion with 6HFh8

showed high solubility, but the removal of the fusion tag may negatively influence the structure and stability of hFGF7. The codon optimization and/or N-terminal truncation (commercially available Palifermin®) of hFGF7 has also been attempted to enhance the solubility and protein yield in *E. coli* .[12, 16] However, the high yield production of hFGF7 in the soluble fraction is still required especially as an intact protein containing the whole sequence. Therefore, a method for the high-yield production of a recombinant FGF7 without a fusion tag or deletion in *E. coli* is needed.

Systematic circular permutations (CP) of genes have emerged as a useful tool for conducting studies of polypeptide folding and stability.[17] Since the CP-applied protein is simply rearranged without altering the amino acid sequence, the tertiary structure is usually retained.[18, 19] Nevertheless, CP can significantly affect stability and activity.[19-21]Therefore, CP has been adopted by protein engineers to manipulate protein structure and function,[22, 23] and has been representatively used as a technique to expand the usefulness of fluorescent proteins.[19, 24, 25] According to the current Circular Permutation Database (CPDB), more than 4000 naturally occurring or artificially generated permutation proteins have been identified.[26] Nevertheless, there is no report on the generation of CP variants from FGF family proteins.

In this paper, we report the structural and functional characteristics of a CP variant of hFGF7 screened by a typical approach that generated a series of CP mutants by selective polymerase chain reaction (PCR) amplification using the duplicated gene as the template[27] and analyzed in terms of expression level and solubility in *E. coli* . The screened variant cp-hFGF7$^{115\text{-}114}$ showed more stability and productivity than the wild type hFGF7, and was also determined to have comparable activity to the wild type hFGF7.

2. MATERIALS AND METHODS

2.1. Bacterial strains, plasmids, and reagents

E. coli XL1-Blue (*ενδA1 γψρA96 [ναλP] τηι-1 ρεςA1 ρελA1 λαςγλν˝44 Φ'[::Tν10 προAB+ λαςIχΔ(λαςZ)M15] ησδP17[ρK-μK+]*) was used as a host for the cloning of the recombinant gene, and protein expression was induced in the host *E. coli* BL21(DE3) (*ομπT γαλ δςμ λον ησδΣB[ρB⁻μB⁻] λ(ΔE3 [λαςI λαςϒ˝5-T7π07 ινδ1 σαμ7 νιν5]) [μαλB+]K-12[λΣ]*). pET24a_hFGF7 was provided by the Korea Institute of Marine Science and Technology (Busan, Korea) and used as the template for PCR with a set of primers specially designed for the duplication of the gene encoding hFGF7 (Table S1). The vector pSCT5[28] was used to insert the duplicated hFGF7 gene for the generation of CP variants by PCR, and the pCold I vector (TaKaRa, Japan) was used to construct a series of CP hFGF7 variants (cp-hFGF7). n*Pfu* special polymerase (Enzynomics, Daejeon, Korea) was used for the amplification of the target gene by PCR. The restriction enzymes used were purchased from New England Biolabs (Ipswich, MA, USA) and Takara Bio Inc. (Shiga, Japan). T4 DNA ligase (Thermo Fisher, MA, USA) or the Infusion HD Cloning Kit (Clontech, Shiga, Japan) was used for the subcloning of the target gene into the plasmids according to the supplier's recommended protocol. Recombinant hFGF7 (indicated as wild type in this work) was purchased from PeproTech (Cranbury, NJ, USA) and used as a positive control for the structure and activity analyses of the purified cp-hFGF7.

2.2. Prediction of CP cleavage sites in hFGF7

CP cleavage sites were predicted by a web-based tool, CPred ().[26] Among residues with a probability score of greater than 0.8, the candidate positions for creating new termini were rationally selected from surface loops that predicted also without significant loss of structural traits by the secondary structure prediction using PSIPRED.[29]

2.3. Construction and analyses of the expression patterns of circularly permuted hFGF7 variants

For the construction of diverse CP variants, the gene encoding hFGF7 without an innate signal sequence (31 amino acids) was primarily duplicated by artificial fusion using PCR as follows. The DNA fragment encoding 163 amino acid residues of hFGF7 was amplified by PCR using pET24a_hFGF7 as the template and two sets of primers (pSCT5_hFGF7-Infu-F and hFGF7-Infu-R, hFGF7(×2)-Infu-F and pSCT5_hFGF7(×2)-Infu-R) under the typical conditions. Then, the resulting DNA fragment was cloned into a linearized pSCT5 vector to prepare the construct pSCT5_hFGF7(×2) containing the duplicate gene (Figure 1A). Using the

pSCT5_hFGF7(×2) vector as the template, CP variants were amplified by PCR using nine sets of primers (pSCold_CP1-hFGF7-F/R ˜ pSCold_CP9-hFGF7-F/R), and subcloned into the same restriction enzyme site (*Spe* I and*Hin* dIII) to prepare pSCold_cp-hFGF7 constructs expressing each of the nine CP variants (Figure 1B, C). Prior to the subcloning of cp-hFGF7, the removal of TEE-6×His-fXa cleavage sequence from the expression vector pCold I and incorporation of the*Spe* I-recognizing sequence ACTAGT (between 5'UTR and start codon) were simultaneously carried out by PCR using a pair of primer (pSCold_Vec-Infu-R and pSCold_Vec-Infu-F).[28] The resulting vector pSCold could express cp-hFGF7 without any additional amino acids at N-terminal region. All primer sequences used in this study are provided in Table S1.

To analyze the expression level and solubility of the cp-hFGF7 variants, each of the recombinant plasmid pSCold_cp-hFGF7 variants (CP1–CP9) was transformed into*E. coli* BL21(DE3) cells. Subsequently, a single colony was inoculated into 3.5 mL of LB medium containing ampicillin (100 μg/mL) and cultured at 37 under constant shaking (200 rpm). When the absorbance (OD_{600}) of the culture reached 2.0, an aliquot of culture broth was reseeded (2%, v/v) into the same LB medium. The resulting cells were cultured to an OD_{600} of 0.6 and treated with isopropyl-β-D-thiogalactoside (IPTG, 0.2 mM) to induce protein expression at 16 and 200 rpm for 36 h. The induced cells were then harvested by centrifugation at 12,000 × g for 10 min and resuspended in 10 mM sodium phosphate buffer (pH 6.5), then disrupted by irradiation with ultrasonic waves three times for 2 seconds. The resulting cell lysate (total fraction, T) was centrifuged at 4 and 12,000 x g for 25 min to obtain a supernatant (soluble fraction, S) from which insoluble aggregates had been removed. Both total and soluble fractions were loaded onto a Tricine-SDS-PAGE (10%) gel, and the expression level and solubility were analyzed under the same conditions previously reported.[30] The resulting gels were also subjected to western blot analyses as follows. The transfer of resolved proteins from gels onto nitrocellulose membrane (GenDEPOT, Texas, USA) was conducted using a Power Blotter-Semi-dry transfer system (Thermo-Fisher Scientific, MA, USA). The membrane was then blocked using 5% skim milk for 1 h at room temperature (RT), followed by incubating overnight at 4degC with anti-human FGF7 monoclonal antibody (1:5000, Abcam, US). Subsequently, the membrane was incubated with horseradish peroxidase-linked goat anti-mouse immunoglobulin G (1:5000, Enzo Life Sciences, US) at RT for 1 h. The proteins on the membrane were visualized with ECL detection kit system (Bio-Rad, US).

2.4. Purification of cp-hFGF7

The recombinant plasmid, pSCold_cp-hFGF7$^{115\text{-}114}$, was transformed into *E. coli* BL21(DE3) cells, and cultured at 37 until the cell density (OD_{600}) reached 2.0 in 4 mL of LB medium containing ampicillin (100 μg/mL). Then, 1 mL of the cultured cells was reseeded into fresh LB medium (100 mL) and cultured to an OD_{600} of 0.6–0.8, then induced with 0.2 mM IPTG at 16 and 200 rpm for 36 h. The induced cells were harvested by centrifugation, resuspended in 20 mM sodium phosphate (pH 6.5) buffer, lysed by sonication, and centrifuged at 4 and 10,000 × g for 60 min to remove the cell debris. Using the resulting supernatant, cp-hFGF7$^{115\text{-}114}$ was purified via a successive step consisting of heparin affinity, cation exchange, and size exclusion chromatography (SEC). Considering the functional structure for cofactor binding, the heparin HP column (1 mL, GE Healthcare, IL, USA) was selected for the primary step. The supernatant was then loaded onto a heparin column equilibrated with buffer A (20 mM sodium phosphate, pH 6.5). After binding, the column was thoroughly washed with the same buffer containing 0.2 M NaCl. Thereafter, a linear gradient was induced with 30 column volumes (CV) of buffer A and B (20 mM sodium phosphate, 1 M NaCl, pH 6.5). Fractions containing cp-hFGF7$^{115\text{-}114}$were collected and diluted 2-fold with a buffer (20 mM sodium phosphate, pH 7.3). Next, the eluted fractions from the heparin column were loaded onto a HiTrap SP HP column (5 mL, GE Hewlett, IL, USA) equilibrated with the same buffer. After complete washing with the same buffer, the bound proteins were eluted with 20 CV of the linear salt (NaCl) gradient buffer from 0 to 1.0 M. The final step of the purification was conducted using a Superdex 200 increase 10/300 GL column (GE Healthcare, Chicago, USA) with 20 mM sodium phosphate buffer (pH 7.3) containing 0.4 M NaCl. The purity and yield of the protein in the eluted fraction were determine by 10% Tricine-SDS-PAGE and western blot.

2.5. Spectroscopic property analyses of cp-hFGF7

To analyze the structural property of cp-hFGF7$^{115\text{-}114}$, UV-vis absorption scanning was performed by using a buffer (20 mM sodium phosphate, 0.4 M NaCl, pH 7.3) under the specified conditions. The protein solution (100 μL) was placed into a 1.0 cm quartz cuvette and the absorbance spectrum was measured by changing the wavelength from 260 to 600 nm at 5 nm intervals.[31] Fluorescence emission scanning was also performed by using the same concentration (150 μg/mL) of proteins. The change in fluorescence wavelength emitted from 280 to 500 nm was measured at 3 nm intervals by an excitation wavelength of 250 nm.[32]

The secondary structure analysis of cp-hFGF7$^{115\text{-}114}$was performed using a circular dichroism (CD) spectropolarimeter (Model J-1500, Jasco, Tokyo, Japan). Prior to CD measurement, the purified cp-hFGF7$^{115\text{-}114}$ from the size exclusion column was completely desalted in a 10 mM sodium phosphate (pH 7.3) buffer by using a PD-10 (GE Healthcare, IL, USA) column. The far-UV CD spectra of cp-hFGF7$^{115\text{-}114}$ (300 μg/mL) was recorded from 190 to 260 nm using a 0.1 cm path length cell at room temperature (25). Each spectrum was obtained three times at a scan rate of 100 nm/min, and then corrected by subtracting the spectral contribution of the buffer. The commercially available hFGF7 was used as a control for all spectral experiments.

2.6. Biological activity analyses of cp-hFGF7

Extracellular signal-regulated kinase (ERK) phosphorylation assay . The embryonic mouse fibroblast cell line NIH3T3 was obtained from the lab of Professor Tae-Hoon Lee at Chonnam National University and routinely maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% bovine calf serum (BCS) and penicillin-streptomycin (100 U/mL and 100 μg/mL, respectively) in an incubator (5% CO_2) at 37. For immunoblotting, 2×10^5 NIH3T3 cells were seeded onto 6-well plates in the same medium and cultured overnight. The cells were then serum-starved for 24 h prior to treatment with cp-hFGF7$^{115\text{-}114}$. After time- and dose-dependent treatment with cp-hFGF7$^{115\text{-}114}$, the treated cells were harvested and lysed using RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, and 0.1% SDS) containing protease and phosphatase inhibitor cocktails. The lysed cells were centrifuged at 4 and 12,000 x g for 20 min to obtain the supernatant. The protein concentration in the supernatant was measured using the BCA protein assay kit (Thermo Fisher, MA, USA). Then, the same amount of protein from each cell was separated by 10% Tricine-SDS-PAGE and transferred to a nitrocellulose membrane. The membrane was blocked in TBS-Tween 20 (0.1%) containing 5% skim milk at room temperature. Specific proteins on the membranes were detected by probing with specific primary antibodies, anti-phospho-specific ERK-1/2 (Thr202/Tyr204) and anti-ERK-1/2 antibodies from Cell Signaling Technology Inc. (Beverly, MA, USA) and the ∝-tubulin antibody from Santa Cruz Biotechnology Inc. (Santa Cruz, CA, USA), followed by incubation with the secondary antibodies conjugated to HRP (Enzo Life Sciences, MI, USA). The resulting specific binding was visualized by the ChemiDoc image analyzer (Bio-Rad, CA, USA) using an ECL chemiluminescence substrate (Bio-Rad, CA, USA).

Cytotoxicity and cell proliferation assay . Cytotoxicity analysis was carried out by live cell counting using NIH3T3 cells and an assay kit (Abcam, Cambridge, UK). Cells grown in the same medium described above were inoculated into 24-well plates and incubated for 24–36 h in a humidified incubator (37) containing 5% CO_2. When the cell confluence reached 70–80%, a new medium containing commercial hFGF7 and cp-hFGF7$^{115\text{-}114}$was added. After incubation for 24 to 72 h, 10 μL of WST-8 dye was added and an additional incubation was performed for 3 h. Afterward, the degree of color change was measured at 450 nm using a spectral microplate reader (SpectraMax ABS, Molecular devices, CA, USA). The results were expressed as a percentage of the control where the absorbance value of the untreated cells was normalized to 100%. After the cell proliferation proceeded in the same manner, the degree of color change of the WST-8 dye was measured at 460 nm. All assays were performed in triplicate.

Scratch wound healing assay . NIH3T3 cells used in the wound healing assay were inoculated into 24-well plates and cultured until 90% to 100% cell confluency was reached. A scratch wound was introduced with a 10 μL pipette tip. After washing with serum-free DMEM medium for cell debris removal after scratch formation, the cells were treated with rhFGF7 and cp-hFGF7$^{115\text{-}114}$ and incubated for 72 h. During the incubation period, sutures for wound closing were monitored and imaged with an optical microscope (Eclipse TE2000-E, Nikon, Tokyo, Japan).[33]

3. RESULTS AND DISCUSSION

3.1. Construction and expression analyses of circularly permuted variants cp-hFGF7s

Conventional approaches that have attempted to produce the recombinant hFGF7 in *E. coli* have limitations in improving soluble expression and stability.[8, 9, 11] As an alternative, we attempted to use a typical technique for inducing changes in the N- and C-terminal ends through the generation of new termini by using CP. CP, a phenomenon of rearrangement of such protein structures, can alter functional properties such as solubility and/or stability by generating new terminal regions from interior residues. Since a variant of the cytokine interleukin 4 had been successfully reported,[34] the CP technique has been presumed to be applicable for producing variants with novel properties in small single-domain proteins.

Practically, plausible CP sites had been selected through the combination of prediction results from two web-based tools (CPred[26] and PSIPRED[29]) because a deposited 3D structure (PDB: 1QQK) of hFGF7 was generated from a variant with N-terminal deletion and thus had different sequence from the wild type hFGF7 (Accession No. P21781) (Figure S1). Two high-scoring sites in a long loop (K100 to N107) and heparin binding region (K149 to K153)[35, 36] were arbitrary excluded to retain structural and functional integrity (Table S2). Using the selected sequences, specific PCR was performed using the tandemly duplicated hFGF7 gene as a template and a set of primers spanning the permutable loop residues (Figure 2A). The resulting DNAs were subcloned into a redesigned vector pSCold that removed the TEE-6×His-fXa cleavage sequence from the original pCold I vector, enabling direct comparison of the expression level and solubility of cp-hFGF7 to those of the wild type hFGF7 produced without any tag at the N-terminus. As shown in Figure 2B, when the protein expression was induced with 0.2 mM IPTG at 16 for 36 h, the expression pattern of each clone harboring a pSCold_cp-hFGF7 variant was quite different from each other. Five CP variants were expressed distinctly and three variants were hardly detected in the soluble and/or insoluble fractions. Changes in the cultivation conditions under different media and temperatures did not significantly improve the expression properties. Intriguingly, the CP variant cp-hFGF7$^{115\text{-}114}$ revealed an enhanced expression level and solubility when compared with those of the wild type hFGF7, thus it was selected to be a candidate for further analyses.

3.2. Purification of cp-hFGF7$^{115\text{-}114}$

The variant cp-hFGF7$^{115\text{-}114}$ with an enhanced expression level and solubility was purified to apparent homogeneity using the three successive steps consisting of a heparin affinity, cation-exchange, and size exclusion chromatography. To test whether the innate binding ability of hFGF7 to heparin was preserved in the purified variant, an affinity chromatography was primarily performed under the same conditions used for the wild type hFGF7.[37] As shown in Figure S2A, cp-hFGF7$^{115\text{-}114}$ was eluted in the fractions corresponding to the concentrations of 0.75 to 0.85 M NaCl after binding to the heparin column (Figure S2B). This result was comparable to those of the wild type hFGF7 that eluted with 0.5 to 0.6 M NaCl,[37] and thus provide a possibility that cp-hFGF7$^{115\text{-}114}$ was able to bind heparin more tightly by a subtle change of the heparin binding site[38] via CP. The structurally solid conformation of cp-hFGF7$^{115\text{-}114}$ was also partly proven by reproducible binding patterns to the heparin column. In the subsequent purification step using a cation-exchange column, cp-hFGF7$^{115\text{-}114}$was reproducibly eluted from the column with a salt concentration of approximately 0.7 M NaCl (Figure S2C). To remove the minor impurities remaining in the fractions from the ion-exchange column, we finally performed SEC and obtained a symmetrical peak corresponding to the target protein cp-hFGF7$^{115\text{-}114}$ (Figure S2D). The retention time of the eluted peak was further compared to those of the protein molecular weight markers and calculated to be a monomeric protein with a molecular mass of 19 kDa. This value was perfectly consistent with that of the wild type protein. As shown in Figure 2C, the purified protein was to apparent homogeneity (>95%) in Tricine SDS-PAGE and was reproducibly obtained via the purification procedures. The final yield of the purified cp-hFGF7$^{115\text{-}114}$ from the flask culture of the recombinant BL21(DE3) was approximately 7 mg/L. This was 2.5-fold higher than that of the previously reported result using *E. coli* as a host.[8]

Prior to the structural and functional characterizations, the storage stability of cp-hFGF7$^{115\text{-}114}$ was com-

pared to the commercially available wild type rhFGF7 under the same storage conditions at 4 for 35 days. As shown in Figure S3, the SDS-PAGE results revealed that cp-hFGF7$^{115\text{-}114}$ had comparable or greater stability than the wild-type hFGF7. The wild type hFGF7 was innately unstable and thus susceptible to proteolytic attack*in-vivo* and/or *in-vitro* even after purification to apparent homogeneity. Taken together, the screened variant FGF7$^{115\text{-}114}$ could have promising properties in terms of the expression level and stability, and thus was subjected to further analyses in structure and activity for comparison to the wild type protein.

3.3. Structural properties ofcp-hFGF7$^{115\text{-}114}$

To analyze the effect of CP on the structure of FGF with a single domain, UV-vis absorption, fluorescence spectroscopy, and the circular dichroism spectra of cp-hFGF7$^{115\text{-}114}$ were used and the results were compared to that of the wild type protein when necessary. These spectral properties were closely linked with the apparent structure (secondary and/or tertiary structure) of proteins. UV-vis absorption and fluorescence emission wavelengths are known as simple methods capable of confirming structural changes by monitoring the overall spectrum using photometric and fluorometric spectral scanning.[39] As shown in Figure 3A, there is also no distinct difference between cp-hFGF7$^{115\text{-}114}$ and hFGF7 in the adsorption spectra although the spectra were not overlapped due to the different concentration of proteins. Additionally, the fluorescence spectrum of cp-hFGF7$^{115\text{-}114}$perfectly overlapped with that of the wild type hFGF7 when excited at 250 nm and then emission scanned from 280 to 500 nm under the same protein concentration (200 μg/mL, Figure 3B). These results partly suggested that the difference in the apparent structure between the two proteins were negligible. We further analyzed the secondary structure of the purified cp-hFGF7$^{115\text{-}114}$ by using far-UV CD and compared it to the previously reported result. As shown in Figure 3C, cp-hFGF7$^{115\text{-}114}$ showed typical absorption bands at 193 to 223 nm, which are characteristic bands of the secondary structures of proteins. The contents of the calculated secondary structure based on the CD spectra are shown in Table S3. When compared to those of the reported results,[8] there is a slight difference in the contents of the secondary structure between the reported hFGF7 and cp-hFGF7$^{115\text{-}114}$ structures. These differences may be due to the remaining 6×His tag in the reported hFGF7 after affinity purification using Ni-Sepharose column.[8] There is also a possibility that CP marginally affected the secondary structure. Additionally, the equipment and buffer composition used between the two experiments were not the same.

3.4. ERK1/2 kinase activation activity of cp-hFGF7$^{115\text{-}114}$

hFGF7 can trigger Ras-Raf-ERK1/2 MARK pathway signaling by direct interaction with the FGFR2b receptor, thereby further stimulating reactions such as epithelial cell proliferation and migration.[40] Based on the previously reported works for mouse and human cell-based functional analysis of hFGF7,[41, 42] we also evaluated whether the MARK signaling pathway was activated by recombinant cp-hFGF7$^{115\text{-}114}$ using a mouse embryonic fibroblast NIH3T3 cell. NIH3T3 cells were treated with the purified cp-hFGF7$^{115\text{-}114}$, and ERK1/2 phosphorylation in these cells was evaluated in a time- and dose-dependent manner.[43] As shown in Figure 4A, cp-hFGF7$^{115\text{-}114}$ (50 ng/mL) induced clear phosphorylation of ERK1/2, and its effect decreased in a time-dependent manner. In addition, cp-hFGF7$^{115\text{-}114}$ also activated ERK1/2 phosphorylation in a dose-dependent manner that was similar to the commercially available wild type protein (Figure 4B). These results provided strong evidence that the CP did not profoundly change the biological activity of cp-hFGF7$^{115\text{-}114}$ and thus retained a capability for phosphorylation of ERK1/2 in experimental cell lines.

3.5. Cell proliferation activity of cp-hFGF7$^{115\text{-}114}$

The biological function of hFGF7 can also be proven by the controlling process for repairing epithelial cells *in-vitro* ; therefore, we further tested whether cp-hFGF7$^{115\text{-}114}$ could induce the proliferation of NIH3T3 cells. A cell proliferation assay for time course analyses was conducted at 1, 10, and 50 ng/mL cp-hFGF7$^{115\text{-}114}$ concentrations by using purified proteins. As shown in Figure 4C, cp-hFGF7$^{115\text{-}114}$retained a cell proliferation activity similar to that of the wild type hFGF7 under the same concentrations and incubation times. We also measured the cellular toxicity of cp-hFGF7$^{115\text{-}114}$ by using NIH3T3 cells. To do so, the cells were treated with cp-hFGF7$^{115\text{-}114}$ for 24 and 72 h at concentrations of 0.5, 1, 10, 50, and 100 ng/mL. The commercially

available hFGF7 was also used as the control. As a result, a survival rate of 98 to 100% was shown in the case of the purified cp-hFGF7$^{115\text{-}114}$ when the cells were treated for 24 h at a concentration of 50 ng/mL. At the 72 h culture time point, the cell viability did not decrease to 95% or less even with the purified protein at a concentration of 100 ng/mL (Figure S4). These results confirmed that the purified cp-hFGF7$^{115\text{-}114}$ did not cause more cytotoxicity than the wild type, and thus we were unable to observe a distinct difference between the two proteins.

3.6. Migration and would healing assay of cp-hFGF7$^{115\text{-}114}$

In-vitro scratch wound healing analysis was performed using NIH3T3 cells to evaluate the association between ERK activation and cell proliferation in FGF7-stimulated cells. As shown in Figure 4D, untreated NIH3T3 cells did not show migration after 72 h, whereas cells treated with cp-hFGF7$^{115\text{-}114}$ and hFGF7 (50 ng/mL) revealed cell proliferation and wound repair by migration after 48 h. In this result, cp-hFGF7$^{115\text{-}114}$ displayed comparable or better cell proliferation and wound healing ability than that of the wild type hFGF7. Additionally, it was found that cell proliferation was difficult to induce as that of the wild type hFGF7 using a low concentration of cp-hFGF7 (1 ng/mL). These results confirmed that cp-hFGF7$^{115\text{-}114}$ retained the same or higher activity than the commercially available rhFGF7.

Since the potential activities for cell proliferation and migration were discovered, FGF7 has been considered a practical drug candidate. In fact, although it is not an intact form, the variant Palifermin with a 24-amino acid deletion from the N-terminal region is commercially available as a drug for the treatment of oral mucositis.[38] Therefore, the wild type hFGF7 and its derivatives with suitable activity and stability would have potential for practical applications.

4. CONCLUSION

We herein developed the circularly permuted variant cp-hFGF7$^{115\text{-}114}$ without any deletion of the wild type hFGF7. This protein proved to be more soluble and stabile than the wild type, and had comparable or greater biological activity than the wild type protein. Taken together, although more sophisticated analyses and preclinical trials are needed for practical application, we expect that the variant cp-hFGF7$^{115\text{-}114}$ could be a potential alternative drug candidate for related symptoms. In addition, the CP technique was also broadly applicable for the generation of functionally soluble variants of the FGF family proteins.

CRediT authorship contribution statement

Hye-Ji Choi: Investigation, Visualization, Writing – original draft. **Hanui Lee:** Investigation, Formal analysis. **Dae-Eun Cheong:** Methodology, Validation, Writing – original draft. **Su-Kyoung Yoo** : Visualization, Formal analysis.**Dong-Eun Lee:** Methodology, Resources. **Geun-Joong Kim:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing, Project administration.

Declaration of Competing Interest

Geun-Joong Kim, Hye-Ji Choi, Hanui Lee, Dae-Eun Cheong and Su-Kyoung Yoo declare that they have conflict of interest as the method for preparing recombinant FGF7 through circular permutation has been filed as patents (Korean Patent Application No. 10-2022-0051241, United States of America Patent Application No. 18/061,836). Dong-Eun Lee declare no competing interests.

Data availability statement

All data generated or analyzed during this study were included in this published article (and its supplementary information files).

Acknowledgement

This work was supported by a grant from the Marine Biotechnology Program (Grant No. 20170305) funded by the Ministry of Oceans and Fisheries, Korea. This work was also partly supported by the Commercializations Promotion Agency for R&D Outcomes (COMPA) grant (No. 2021C300) funded by the Korea government

(MSIT) and Basic Science Research programs (Grant No. 2021R1A2C1006734) of the National Research Foundation (NRF) of Korea funded by the Ministry of Education, Science and Technology of Korea (MEST).

Reference

1. Ornitz, D. M. & Itoh, N. (2015). The fibroblast growth factor signaling pathway. *Wiley Interdisciplinary Reviews: Developmental Biology,* 4(3), 215-266.

2. Zhang, X., Pu, X., Pi, C. & Xie, J. (2023). The role of fibroblast growth factor 7 in cartilage development and diseases. *Life Sciences,* 326, 121804.

3. Katoh, M. (2008). Cancer genomics and genetics of FGFR2 (Review).*International Journal of Oncology,* 33(2), 233-237.

4. Zinkle, A. & Mohammadi, M. (2019). Structural biology of the FGF7 subfamily. *Frontiers in Genetics,* 10, 102.

5. Xue, P., Zhu, X., Shi, J., Fu, H., Zhang, J., Liu, M., Jiang, C. & Li, X. (2014). Highly efficient expression of functional recombinant human keratinocyte growth factor 1 and its protective effects on hepatocytes. *Applied Microbiology and Biotechnology,* 98(9), 3933-3945.

6. Movahednia, M. M., Kidwai, F. K., Jokhun, D. S., Squier, C. A., Toh, W. S. & Cao, T. (2016). Potential applications of keratinocytes derived from human embryonic stem cells. *Biotechnology Journal,* 11(1), 58-70.

7. Peng, C., Chen, B., Kao, H. K., Murphy, G., Orgill, D. P. & Guo, L. (2011). Lack of FGF-7 further delays cutaneous wound healing in diabetic mice. *Plastic Reconstructive Surgery,* 128(6), 673e-684e.

8. Jafari, B., Hamzeh-Mivehroud, M., Moosavi-Movahedi, A. A. & Dastmalchi, S. (2018). Identification of novel single-domain antibodies against FGF7 using phage display technology. *SLAS DISCOVERY: Advancing Life Sciences R&D,* 23(2), 193-201.

9. Liu, M., Chen, Z. Y., Huang, S. S., Chu, S. H., Issaro, N., Tian, H. S., Hu, H. Y. & Jiang, C. (2018). Effective protection on acute liver injury by Halo tag-flanked recombinant fibroblast growth factor 7.*Biotechnology Journal,* 13(7), e1700411.

10. Luo, Y., Cho, H. H., Jones, R. B., Jin, C. & McKeehan, W. L. (2004). Improved production of recombinant fibroblast growth factor 7 (FGF7/KGF) from bacteria in high magnesium chloride. *Protein Expression and Purification,* 33(2), 326-331.

11. Kim, Y. S., Lee, H. J., Han, M. H., Yoon, N. K., Kim, Y. C. & Ahn, J. (2021). Effective production of human growth factors in*Escherichia coli* by fusing with small protein 6HFh8.*Microbial Cell Factories,* 20(1), 1-16.

12. An, Y. J., Lee, K. W., Jung, Y. E., Jeong, Y. E., Kim, S. J., Woo, J., Jin, J., Lee, W. K., Cha, K., Cha, S. S., Lee, J. H. & Yim, H. S. (2022). Improvement of FGF7 thermal stability by introduction of mutations in close vicinity to disulfide bond and surface salt bridge.*International Journal of Peptide Research and Therapeutics,*28(3), 85.

13. Cheong, D. E., Choi, H. J., Yoo, S. K., Lee, H. D. & Kim, G. J. (2021). A designed fusion tag for soluble expression and selective separation of extracellular domains of fibroblast growth factor receptors. *Scientific Reports,* 11(1), 21453.

14. Ke, N. & Berkmen, M. (2014). Production of disulfide-bonded proteins in *Escherichia coli* . *Current Protocols in Molecular Biology,* 108(1), 16.1B.1-16.1B.21.

15. Rosano, G. L. & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli* : advances and challenges.*Frontiers in Microbiology,* 5, 172.

16. Kim, Y. S., Lee, H. J., Handoko, G. A., Kim, J., Kim, S. B., Won, M., Park, J. H. & Ahn, J. (2023). Production of a 135-residue long N-truncated human keratinocyte growth factor 1 in *Escherichia coli* . *Microbial Cell Factories,* 22(1), 98.

17. Lindberg, M., Tangrot, J. & Oliveberg, M. (2002). Complete change of the protein folding transition state upon circular permutation.*Nature Structural Biology,* 9(11), 818-822.

18. Cheong, D. E., Park, S. Y., Shin, H. J. & Kim, G. J. (2009). A new cloning system using a mutant esterase containing MCS as an indicator for gene cloning. *Journal of Microbiological Methods,* 77(3), 302-7.

19. Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology,* 24(1), 79-88.

20. Yu, Y. & Lutz, S. (2011). Circular permutation: a different way to engineer enzyme structure and function. *Trends in Biotechnology,*29(1), 18-25.

21. Paes, G., Berrin, J.-G. & Beaugrand, J. (2012). GH11 xylanases: structure/function/properties relationships and applications.*Biotechnology Advances,* 30(3), 564-592.

22. Carey, J., Lindman, S., Bauer, M. & Linse, S. (2007). Protein reconstitution and three-dimensional domain swapping: benefits and constraints of covalency. *Protein Science,* 16(11), 2317-2333.

23. Iwakura, M., Nakamura, T., Yamane, C. & Maki, K. (2000). Systematic circular permutation of an entire protein reveals essential folding elements. *Nature Structural Biology,* 7(7), 580-585.

24. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins.*Proceedings of the National Academy of Sciences of the USA,*96(20), 11241-11246.

25. Nakai, J., Ohkura, M. & Imoto, K. (2001). A high signal-to-noise Ca^{2+} probe composed of a single green fluorescent protein. *Nature Biotechnology,* 19(2), 137-141.

26. Lo, W. C., Wang, L. F., Liu, Y. Y., Dai, T., Hwang, J. K. & Lyu, P. C. (2012). CPred: a web server for predicting viable circular permutations in proteins. *Nucleic Acids Research,* 40(W1), W232-W237.

27. Shui, B., Wang, Q., Lee, F., Byrnes, L. J., Chudakov, D. M., Lukyanov, S. A., Sondermann, H. & Kotlikoff, M. I. (2011). Circular permutation of red fluorescent proteins. *Plos One,* 6(5), e20505.

28. Cheong, D. E., Ko, K. C., Han, Y., Jeon, H. G., Sung, B. H., Kim, G. J., Choi, J. H. & Song, J. J. (2015). Enhancing functional expression of heterologous proteins through random substitution of genetic codes in the 5' coding region. *Biotechnology and Bioengineering,* 112(4), 822-6.

29. McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics,* 16(4), 404-5.

30. Schagger, H. (2006). Tricine–SDS-PAGE. *Nature Protocols,*1(1), 16-22.

31. Suryawanshi, V. D., Walekar, L. S., Gore, A. H., Anbhule, P. V. & Kolekar, G. B. (2016). Spectroscopic analysis on the binding interaction of biologically active pyrimidine derivative with bovine serum albumin.*Journal of Pharmaceutical Analysis,* 6(1), 56-63.

32. Wang, F., Wang, Y., Jiang, L., Wang, W., Sang, J., Wang, X., Lu, F. & Liu, F. (2021). The food additive fast green FCF inhibits α-synuclein aggregation, disassembles mature fibrils and protects against amyloid-induced neurotoxicity. *Food & Function,* 12(12), 5465-5477.

33. Maretzky, T., Reiss, K., Ludwig, A., Buchholz, J., Scholz, F., Proksch, E., De Strooper, B., Hartmann, D. & Saftig, P. (2005). ADAM10 mediates E-cadherin shedding and regulates epithelial cell-cell adhesion, migration, and β-catenin translocation. *Proceedings of the National Academy of Sciences of the USA,* 102(26), 9182-9187.

34. Kreitman, R. J., Puri, R. K. & Pastan, I. (1994). A circularly permuted recombinant interleukin 4 toxin with increased activity.*Proceedings of the National Academy of Sciences of the USA,*91(15), 6889-6893.

35. Raman, R., Venkataraman, G., Ernst, S., Sasisekharan, V. & Sasisekharan, R. (2003). Structural specificity of heparin binding in the fibroblast growth factor family of proteins. *Proceedings of the National Academy of Sciences of the USA,* 100(5), 2357-2362.

36. Vadija, R., Mustyala, K. K., Nambigari, N., Dulapalli, R., Dumpati, R. K., Ramatenki, V., Vellanki, S. P. & Vuruputuri, U. (2016). Homology modeling and virtual screening studies of FGF-7 protein-a structure-based approach to design new molecules against tumor angiogenesis. *Journal of Chemical Biology,* 9(3), 69-78.

37. Huang, Z., Zhu, G., Sun, C., Zhang, J., Zhang, Y., Zhang, Y., Ye, C., Wang, X., Ilghari, D. & Li, X. (2012). A novel solid-phase site-specific PEGylation enhances the *in vitro* and *in vivo*biostabilty of recombinant human keratinocyte growth factor 1.*Plos One,* 7(5), e36423.

38. Sadeghi, S., Kalhor, H., Panahi, M., Abolhasani, H., Rahimi, B., Kalhor, R., Mehrabi, A., Vahdatinia, M. & Rahimi, H. (2021). Keratinocyte growth factor in focus: A comprehensive review from structural and functional aspects to therapeutic applications of palifermin. *International Journal of Biological Macromolecules,*191, 1175-1190.

39. Pan, X., Qin, P., Liu, R. & Wang, J. (2011). Characterizing the interaction between tartrazine and two serum albumins by a hybrid spectroscopic approach. *Journal of Agricultural and Food Chemistry,* 59(12), 6650-6656.

40. Uzan, B., Figeac, F., Portha, B. & Movassat, J. (2009). Mechanisms of KGF mediated signaling in pancreatic duct cell proliferation and differentiation. *Plos One,* 4(3), e4734.

41. Maretzky, T., Evers, A., Zhou, W., Swendeman, S. L., Wong, P.-M., Rafii, S., Reiss, K. & Blobel, C. P. (2011). Migration of growth factor-stimulated epithelial and endothelial cells depends on EGFR transactivation by ADAM17. *Nature Communications,* 2(1), 1-11.

42. Sher, I., Weizman, A., Lubinsky-Mink, S., Lang, T., Adir, N., Schomburg, D. & Ron, D. (1999). Mutations uncouple human fibroblast growth factor (FGF)-7 biological activity and receptor binding and support broad specificity in the secondary receptor binding site of FGFs. *Journal of Biological Chemistry,* 274(49), 35016-35022.

43. Choi, H. J., Cheong, D. E., Yoo, S. K. & Kim, G. J. (2023). One-step metal affinity purification of recombinant hFGF19 without using tags. *Protein Expression and Purification,* 201, 106186.

Figure Legends

Figure 1. Schematic diagram for the construction and cloning of cp-hFGF7 variants. (A) Genetic organization of the vector including a template polynucleotide for PCR amplification of the circular permutation (CP) fibroblast growth factor (FGF7) variants, in which the template was prepared by gene duplication. (B) After a tandem gene was constructed without a linker, circular permutants with new start (n) and stop (n-1) residues of hFGF7 were generated using PCR, then subcloned into the expression vector pSCold. The residues of cp-hFGF7 are numbered according to hFGF7 sequence without the signal peptide. (C) Genetic organization of the expression vector containing a cp-hFGF7 variant for the inducible expression under a low temperature.

Figure 2. Analyses of the expression and purification patterns of cp-hFGF7$^{115\text{-}114}$. (A) To visualize the selection results of circular permutation (CP) sites from the two web-based tools, CP site-containing loops are marked as yellow lines in the deposited tertiary structure of hFGF7 in the data bank (PDB ID: 1QQK). (B) The expression patterns of the constructed variants of cp-hFGF7 in *E. coli* BL21 (DE3) were analyzed by Tricine-SDS-PAGE under denaturing conditions. Western blots were also analyzed using the same PAGE gels. (C) The eluted fractions contacting cp-hFGF7$^{115\text{-}114}$ from each chromatography

were analyzed by Tricine-SDS-PAGE and western blot under the same conditions described in the Materials and Methods section. M, molecular weight markers; T, total fraction; S, soluble fraction; FT, flow through. Arrows indicate the protein band corresponding to cp-hFGF7$^{115\text{-}114}$.

Figure 3. Spectral profile analyses for the comparison of structural properties between the wild type and cp-hFGF7$^{115\text{-}114}$. (A) UV-visible spectra comparison between the wild type hFGF7 (commercially available) and cp-hFGF7$^{115\text{-}114}$. (B) Fluorescence spectra of both proteins were scanned from 280 to 500 nm by excitation at 250 nm. (C) The circular dichroism spectrum of cp-hFGF7$^{115\text{-}114}$ was also analyzed by using the purified proteins (300 μg/mL) under the specified conditions. Black line, buffer; orange line, rhFGF7; blue line, cp-hFGF7$^{115\text{-}114}$. All experiments were conducted at least three times according to the described procedure in the Materials and Methods section and their average values were plotted in these figures.

Figure 4. Biological activity of cp-hFGF7$^{115\text{-}114}$. Dose- (A) and time-dependent (B) extracellular signal-regulated kinase (ERK) phosphorylation activity analyses in NIH3T3 cells. A time course of ERK1/2 phosphorylation is shown in NIH3T3 cells treated with 50 ng/mL of wild type hFGF7 and cp-hFGF7$^{115\text{-}114}$. N.C, negative control. Dose-dependent ERK1/2 kinase activation in NIH3T3 cells treated with the defined amount of hFGF7 and cp-FGF7$^{115\text{-}114}$ for 30 min. (C) Comparisons of cell proliferation activities between hFGF7 and cp-hFGF7$^{115\text{-}114}$. The same amounts (1, 10, and 50 ng/mL) of both proteins were treated under the identical conditions and their activities were monitored over time. (D) Comparison of scratch wound healing activity between hFGF7 and cp-hFGF7$^{115\text{-}114}$ in NIH3T3 cells. After scratching, the cells were treated with a defined amount of both proteins under the same conditions and their effects on would healing were monitored for 72 h. All experiments were conducted at least three times according to the described procedure in the methods section. The commercially available hFGF7 was designated as the wild type and used as the positive control.