

## **Molecular cloning of human basic fibroblast growth factor gene (*hbFgf*) and *in silico* analysis of physico-chemical characteristics of bFGF protein**

**Hojat Taheri Azizabadi**

PhD Student in Agricultural Biotechnology, Department of Biotechnology Engineering, Faculty of Agricultural and Natural Resources, Imam Khomeini International University, Qazvin, Iran.  
E-mail address: h\_taheri1986@yahoo.com

**Ramin Hosseini** 

\*Corresponding author. Associate Professor, Department of Biotechnology Engineering, Faculty of Agriculture and Natural Resources, Imam Khomeini International University, Qazvin, Iran. E-mail address: r.hosseini@ikiu.ac.ir

---

### ***Abstract***

#### **Objective**

This study aimed to clone the human basic fibroblast growth factor (hbFGF) gene into the pCAMBIA1304 binary vector for recombinant bFGF protein production in plant expression systems.

#### **Materials and Methods**

The full-length open reading frame (ORF) encoding the hbFGF gene was inserted into the pCAMBIA1304 binary vector and subsequently cloned into *Escherichia coli* strain DH5 $\alpha$ . The recombinant expression vector, designated pCAMBIA-hbFGF, was then introduced into competent *Agrobacterium tumefaciens* strains EHA105, GV3101, and LBA4404.

#### **Results**

The hbFGF gene, comprising 561 nucleotides, encoded a protein of 155 amino acid residues. The calculated molecular mass and predicted isoelectric point (pI) of the deduced polypeptide sequence were 17.25 kDa and 9.58, respectively. The instability index, aliphatic index, and hydrophobicity index of the bFGF protein were calculated as 36.60, 68.00, and -0.511, respectively, indicating high stability and a hydrophilic nature. Analysis using ProtScale and TMHMM programs confirmed the hydrophilic properties of bFGF and the absence of

transmembrane domains. Additionally, the MotifScan program identified a conserved fibroblast growth factor (FGF) family domain in the hbFGF sequence, while no potential N-glycosylation sites were detected. The predicted secondary structure of the mRNA sequence, calculated using the RNAfold program, exhibited a minimum free energy of approximately -136.90 kcal/mol, indicating relatively high stability with low entropy. The simulated three-dimensional structure of the hbFGF protein revealed a pyramidal arrangement of 10 non-parallel  $\beta$ -sheets. Ramachandran plot analysis confirmed that 86.3% of the amino acid residues were organized in regular structures, primarily non-parallel  $\beta$ -sheets, with only two residues (Thr8 and Gly24) exhibiting unusual  $\phi$  and  $\psi$  angles. Phylogenetic analysis demonstrated a high degree of similarity with *Pan troglodytes*, *Pongo abelii*, and *Camelus dromedarius*, showing 100%, 100%, and 98.71% sequence similarity, respectively.

### Conclusions

The bFGF protein was identified as a stable, hydrophilic protein without potential N-glycosylation sites. These characteristics suggest that the bFGF protein may be successfully produced heterologously in plant expression systems for potential commercialization.

**Keywords:** binary vector, cloning, fibroblast growth factor, hydrophobicity index, ramachandran plot

**Paper Type:** Research Paper.

**Citation:** Taheri Azizabadi H, Hosseini R (2025) Molecular cloning of human basic fibroblast growth factor gene (*hbFgf*) and *in silico* analysis of physico-chemical characteristics of bFGF protein. *Agricultural Biotechnology Journal* 17 (1), 1-26.

---

*Agricultural Biotechnology Journal* 17 (1), 1-26.

DOI: 10.22103/jab.2024.21823.1491

Received: September 24, 2024.

Received in revised form: December 03, 2024.

Accepted: December 04, 2024.

Published online: January 30, 2025.

Publisher: Faculty of Agriculture and Technology Institute of Plant



Production, Shahid Bahonar University of Kerman-Iranian  
Biotechnology Society.

© the authors

## همسانه‌سازی مولکولی ژن رمزگردان فاکتور رشد فیبروبلاستی بازی انسانی (*hbFgf*) و

### بررسی ویژگی‌های فیزیکی- شیمیایی پروتئین *hbFGF* در شرایط *in silico*

حجت طاهری عزیزآبادی

دانشجوی دکتری بیوتکنولوژی کشاورزی، گروه مهندسی بیوتکنولوژی، دانشکده کشاورزی و منابع طبیعی، دانشگاه بین‌المللی امام

خمینی (ره)، قزوین، ایران. رایانامه: h\_taheri1986@yahoo.com

رامین حسینی 

\*نویسنده مسئول: دانشیار گروه مهندسی بیوتکنولوژی، دانشکده کشاورزی و منابع طبیعی، دانشگاه بین‌المللی امام خمینی (ره)،

قزوین، ایران. رایانامه: r.hosseini@ikiu.ac.ir

تاریخ دریافت: ۱۴۰۳/۰۷/۰۳ تاریخ دریافت فایل اصلاح شده نهایی: ۱۴۰۳/۰۹/۱۳ تاریخ پذیرش: ۱۴۰۳/۰۹/۱۴

### چکیده

**هدف:** هدف از این پژوهش، همسانه‌سازی ژن *bFgf* انسانی (*hbFgf*) درون ناقل بیانی pCAMBIA1304 جهت بیان پروتئین نوترکیب bFGF در میزبان‌های گیاهی بود.

**مواد و روش‌ها:** توالی ORF رمزگردان ژن *hbFgf* درون ناقل بیانی pCAMBIA1304 درج و درون باکتری *Escherichia coli* سویه DH5 $\alpha$  همسانه‌سازی شد. ناقل بیانی نوترکیب pCAMBIA-hbFGF سپس به درون سلول‌های مستعد *Agrobacterium tumefaciens* (EHA105، GV3101 و LBA4404) منتقل شد.

**نتایج:** ژن *hbFgf* به طول ۴۶۸ bp، پروتئینی با ۱۵۵ اسیدآمینو را رمز می‌کند که دارای وزن مولکولی و نقطه ایزوالکتریک به ترتیب برابر با ۱۷/۲۵ و ۹/۵۸ است. شاخص‌های ناپایداری، آلیفاتیک و آب‌گریزی پروتئین hbFGF به ترتیب حدوداً ۳۶/۶۰، ۶۸/۰۰ و ۰/۵۱۱- محاسبه شد که نشان‌دهنده پایداری بالا و وضعیت آب‌دوستی آن است. به علاوه، بررسی‌ها با استفاده از برنامه‌های ProtScale و TMHMM تأیید کرد که پروتئین hbFGF آب‌دوست بوده و فاقد هر گونه دامنه‌های ترنس‌ممبران است. همچنین، برنامه MotifScan یک دامنه ویژه خانواده فاکتور رشد فیبروبلاستی (FGF) را در توالی اسیدآمینوای hbFGF شناسایی کرد و در حالی که جایگاه بالقوه جهت N-گلیکوزیلاسیون در توالی اسیدآمینوای hbFGF شناسایی نشد. حداقل انرژی آزاد محاسبه شده ساختار دوم توالی mRNA با استفاده از نرم‌افزار RNAfold برابر ۱۳۶/۹۰ kcal/mol- محاسبه شد که

نشان‌دهنده پایداری نسبتاً بالای ساختار دوم توالی mRNA با سطح آنتروپی پایین است. علاوه بر این، شبیه‌سازی ساختار سه‌بعدی پروتئین hbFGF نشان داد که پروتئین hbFGF شامل ۱۰ صفحه  $\beta$  غیرموازی است که در یک ساختار هرمی سازماندهی شده است. رسم نمودار رامچاندرا نیز نشان داد که پروتئین hbFGF عمدتاً از صفحات  $\beta$  غیرموازی تشکیل شده است. حدود ۸۶/۳٪ از اسیدهای آمینه در ساختارهای منظم شامل صفحات  $\beta$  غیرموازی سازمان‌دهی شده و تنها دو اسید آمینه Thr<sub>8</sub> و Gly<sub>24</sub> با زوایای  $\phi$  و  $\psi$  غیرمعمول در پیکربندی پروتئین hbFGF شناسایی شدند. بررسی میزان شباهت hbFGF با سایر توالی‌های bFGF، درجه بالایی از شباهت را با *Pan troglodytes*، *Pongo abelii* و *Camelus dromedaries* به ترتیب با ۱۰۰٪، ۱۰۰٪ و ۹۸/۷۱٪ شباهت نشان داد.

**نتیجه‌گیری:** پروتئین bFGF یک پروتئین پایدار و فاقد جایگاه بالقوه N-گلیکوزیلاسیون است که ممکن است به شکل نوترکیب درون سیستم‌های بیانی گیاهی در سطح تجاری تولید شود.

**کلیدواژه‌ها:** رشد فیبروبلاستی بازی، شاخص آب‌گریزی، فاکتور نمودار رامچاندرا، ناقل دوگانه، همسانه‌سازی  
**نوع مقاله:** پژوهشی.

**استناد:** طاهری عزیزآبادی حجت، حسینی رامین (۱۴۰۴) همسانه‌سازی مولکولی ژن رمزگردان فاکتور رشد فیبروبلاستی بازی انسانی (*hbFgf*) و بررسی ویژگی‌های فیزیکی-شیمیایی پروتئین hbFGF در شرایط *in silico*. مجله بیوتکنولوژی کشاورزی، ۱۱۷(۱)، ۲۶-۱.

Publisher: Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society.



© the authors

## مقدمه

توانایی بازسازی بافت‌ها و اندام‌های آسیب‌دیده و حتی اعضای قطع شده بدن، یک آرزوی دیرینه برای بشر و یک هدف اصلی و بسیار چالش برانگیز برای پزشکان، پژوهشگران و مهندسان است. متأسفانه، توان ترمیم در پستانداران بسیار محدود بوده و آسیب به اکثر بافت‌ها (بجز بافت برخی اندام‌ها نظیر کبد) نهایتاً در بیشتر موارد در فرآیند ترمیم منجر به تشکیل بافت غیراختصاصی از قبیل بافت جای زخم (اسکار) در آسیب به سامانه پوششی بدن می‌شود. برخلاف پستانداران، برخی جانداران توان ترمیم قابل توجهی دارند که عبارتند از کیسه‌تنان (Cnidarian) شامل *Polyp hydra* آب شیرین و پلاناریان‌ها (Planarian) به عنوان زیر مجموعه‌ای از کرم‌های پهن و برخی از گونه‌های مهره‌داران، مانند سمندر، قورباغه و ماهی زبرا. این جانداران قادر به رشد مجدد ساختارهای اصلی مانند سر یا دم، یا حتی موجود زنده کامل از قطعات بسیار کوچک بدن هستند (Farooq et al. 2021). فرآیندهای ترمیم و

بازسازی توسط طیف وسیعی از سیتوکین‌ها، عوامل رشد و تمایز کنترل می‌شوند. از جمله آن‌ها می‌توان به فاکتورهای رشد فیبروبلاستی (Fibroblast Growth Factor, FGF) اشاره کرد که تنظیم‌کننده اصلی اندام‌زایی و هم‌اوستازی بافت هستند. خانواده FGF متشکل از پروتئین‌های پیام‌دهنده ترشحی است که به گیرنده‌های تیروزین کیناز و FGF‌های غیر پیام‌دهنده داخل سلولی (Intracellular FGF, iFGF) که به عنوان کوفاکتور برای کانال‌های سدیم و سایر کانال‌های ولتاژی عمل می‌کنند، متصل می‌شوند. اعضای هر دو شاخه از خانواده FGF با حفظ توالی و ساختار مرتبط در مهره‌داران و بی‌مهرگان یافت می‌شوند. در پستانداران، خانواده FGF‌های ترشحی شامل ۲۲ پلی‌پپتید است که تقریباً در تمام بافت‌های بدن بیان شده و نقش‌های اساسی را در مراحل اولیه رشد جنینی، اندام‌زایی، پاسخ به آسیب، ترمیم و بازسازی بافت ایفا می‌کنند (Ornitz & Itoh 2015). FGF‌ها و گیرنده‌های آن‌ها در میان فرمانرو جانوران به شدت حفظ شده‌اند. بر اساس تجزیه و تحلیل فیلوژنتیکی، FGF‌ها را می‌توان به هفت زیرخانواده تقسیم‌بندی کرد که عبارتند از FGF1, FGF4, FGF7, FGF8, FGF9, FGF11, FGF15/19 و FGF2 (Ornitz & Itoh 2015). FGF2 اولین بار در سال ۱۹۷۳ از غده هیپوفیز گاو جداسازی شد و به دلیل دارا بودن اسیدهای آمینه بازی و نقطه ایزوالکتریک بالا به عنوان FGF بازی (bFGF) شناخته می‌شود (Armelin 1973). در میان اعضای خانواده FGF, FGF2 به فراوانی در سیستم عصبی مرکزی یافت می‌شود. به دلیل اثرات ترمیم و التیام زخم، FGF2 دارای خواص درمانی است و به طور گسترده در کاربردهای بالینی برای التیام زخم از طریق تحریک تکثیر سلولی استفاده می‌شود. این پروتئین همچنین برای حفظ سلول‌های بنیادی و مهار تمایز آن‌ها به محیط‌کشت اضافه می‌شود (Dolivo et al. 2022). در حال حاضر، تهیه bFGF که غالباً به منظور کاربردهای پزشکی و مراقبت از پوست استفاده می‌شود، بسیار پرهزینه است (Kwong et al. 2013). بنابراین، با توجه به اهمیت بالینی bFGF، توسعه یک سیستم بیانی با عملکرد بالا، کم هزینه و مقرون به صرفه جهت تولید تجاری به شکل نوترکیب از اهمیت بالایی برخوردار است. تاکنون از سیستم‌های بیانی مختلفی جهت تولید نوترکیب آن استفاده شده است که عبارتند از کرم ابریشم (*Bombyx mori*) (Wu et al. 2001)، سویا (*Glycine max*) (Ding et al. 2006)، باکتری *Bacillus subtilis* (Kwong et al. 2013)، برنج (*Oryza sativa*) (An et al. 2013; Poudel et al. 2019) و توتون (*Nicotiana tabacum*) (Wang et al. 2016) و باکتری *Escherichia coli* (Lingg et al. 2021). با توجه به عدم N-گلیکوزیلاسیون bFGF، به نظر می‌رسد که استفاده از انواع سیستم‌های کشت گیاهی شامل گیاه کامل، سوسپانسیون سلول گیاهی، ریشه موئین، خز، عدسک آبی و ریزجلیک‌ها به عنوان میزبان‌های بیانی جهت تولید تجاری bFGF نوترکیب می‌تواند بسیار کارآمد باشد. به‌علاوه، مطالعات و بررسی‌های به عمل آمده در اواخر دهه ۸۰ میلادی روشن نمود که سازوکارهای مولکولی در زمره مهم‌ترین فرایندهای ژنتیکی (مشتمل بر همانندسازی DNA، رونویسی، ترجمه و حتی نحوه تنظیم ژن‌ها) هستند (Barazandeh et al. 2016a; Safaei et al. 2022). ماده ژنتیکی یک سلول از تعداد زیادی ژن تشکیل شده است که هیچ‌گاه همه به طور هم‌زمان بیان نمی‌شوند و در هر لحظه فقط تعداد کمی از آن‌ها بیان شده و پروتئین یا آنزیم مورد نیاز سلول را تولید می‌کنند (Barazandeh et al. 2016b). محیطی که موجود در آن رشد

می‌کند مشخص می‌کند که آیا ژن بیان شود و یا این که نیازی به فرآورده آن نیست و باید غیرفعال و یا خاموش شود ( Bordbar et al. 2023; Jafari Ahmadabadi et al. 2022). ساز و کار بیان ژن اولین بار در باکتری *E. coli* کشف شد (Saadatabadi et al. 2023; Mohammadabadi et al. 2023). ژن‌های یوکاریوتی بیان‌شان تحت کنترل موقت و چندبعدی است. در هر یک از انواع بافت‌ها تنها یک مجموعه نسبتاً کوچک از تمام ژنوم بیان شده و نیز بیان ژن‌ها به مرحله نمو بستگی دارد (Masoudzadeh et al. 2020). بنابراین، بیان ژن در یوکاریوت‌ها برای هر بافت اختصاصی است. همچنین مقدار فرآورده‌های ژن که در همان بافت و نیز در سایر بافت‌هایی که آن محصول را می‌سازند، ساخته شده سبب تنظیم بیان آن ژن می‌شود (Mohammadabadi et al. 2021; Shokri et al. 2023). یکی از اقدامات اساسی در به‌نژادی ملکولی، مطالعه ژن‌ها و پروتئین‌های مرتبط با صفات مهم و مطالعه آن‌ها در سطح سلولی یا کروموزومی است ( Mohammadinejad et al. 2022; Shahsavari et al. 2022). لذا، در این پژوهش، توالی رمزگردان ژن *bFgf* انسانی (*hbFgf*) پس از درج درون ناقل بیانی pCAMBIA1304 درون باکتری *E. coli* سویه DH5 $\alpha$  همسانه‌سازی شده و سپس به درون سلول‌های مستعد سه سویه *Agrobacterium tumefaciens* (به نام‌های EHA105، GV3101 و LBA4404 منتقل می‌شود تا جهت بیان پروتئین bFGF نو ترکیب در میزبان‌های گیاهی مورد استفاده قرار بگیرد. علاوه بر این، ساختار دوم توالی mRNA ژن *hbFgf* توالی اسیدآمینو، ساختار دوم و مدل‌سازی مولکولی ساختار سوم پروتئین hbFGF و روابط فیلوژنتیکی و هم‌ریدیف‌سازی توالی چندگانه با سایر توالی‌های bFGF از جانداران مختلف با استفاده از نرم‌افزارهای بیوانفورماتیکی مورد مطالعه قرار می‌گیرد.

## مواد و روش‌ها

توالی رمزگردان ژن *hbFgf* (NCBI Genbank: AAA52448.1) از ناقل بیانی نو ترکیب pCAM-EXPCAS (هدا شده توسط Bastabi و Hosseini 2018) با استفاده از آغازگرهای رفت (-5' ataccatggATGGCTGCTGGGTCAATCAC-3') و برگشت (-5' tatggttaccTCACGACTTCGCCGACATAG-3') طی واکنش زنجیره‌ای پلیمرز<sup>۱</sup> (PCR) تکثیر شد. آغازگرهای رفت و برگشت به ترتیب حاوی جایگاه شناسایی برای آنزیم‌های محدودکننده *Nco* I و *Bst*E II بودند. چرخه‌های حرارتی شامل مرحله واسرشته‌سازی اولیه در دمای ۹۴ °C به مدت ۳ دقیقه، ۳۵ چرخه؛ مرحله واسرشته‌سازی در دمای ۹۴ °C به مدت ۳۰ ثانیه، مرحله اتصال در دمای ۵۸ °C به مدت ۶۰ ثانیه و مرحله بسط در دمای ۷۲ °C به مدت ۳۰ ثانیه بود. در انتها، مرحله بسط نهایی در دمای ۷۲ °C به مدت ۵ دقیقه انجام گرفت. از شاهد منفی بدون الگو جهت بررسی وجود یا عدم وجود آلودگی در PCR استفاده گردید.

<sup>1</sup> Polymerase chain reaction

فرآورده PCR با استفاده از کیت خالص‌سازی قطعات DNA (AccuPrep® Gel Purification Kit، شماره کاتالوگ K-3035-1، شرکت Bioneer، کره جنوبی) از روی ژل آگارز خالص‌سازی شد و با استفاده از آنزیم‌های محدودکننده *Nco I* و *BstE* II برش یافت. در مرحله بعد، قطعه برش یافته درون ناقل بیانی pCAMBIA1304 برش یافته با همان آنزیم‌ها با استفاده از آنزیم DNA لیگاز T4 جایگزین قطعه mgfp5-GUS شد. ناقل بیانی نوترکیب pCAMBIA-hbFGF با استفاده از روش شوک حرارتی به درون سلول‌های مستعد باکتری *E. coli* سویه DH5α همسانه‌سازی شد. سپس ناقل بیانی نوترکیب به روش لیز قلیایی با SDS استخراج شده (Birnboim & Doly 1979) و به روش PCR و برش آنزیمی تأیید شد. در مرحله بعد، ناقل بیانی نوترکیب به روش انجماد و ذوب به درون سلول‌های مستعد سه سویه آگروباکتریوم (*A. tumefaciens*) به نام‌های EHA105، GV3101 و LBA4404 همسانه‌سازی شده (Jyothishwaran et al. 2007) و پس از استخراج از سویه‌های آگروباکتریوم، به روش PCR تأیید شدند.

پیش‌بینی ساختار دوم توالی mRNA و حداقل انرژی آزاد (Minimum free energy MFE) با روش احتمال جفت‌بازی (Base-pair probabilities) به وسیله نرم‌افزار RNAfold ([www.rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi](http://www.rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi)) انجام شد. خصوصیات بیوشیمیایی توالی اسیدآمینه‌ای با استفاده از برنامه‌های ProtParam (Gasteiger et al. 2005)، TMHMM (<http://www.cbs.dtu.dk/services/>) و ProtScale (<http://arbl.cvmbs.colostate.edu/molkit/hydrophathy/index.html>) مورد بررسی قرار گرفته و شناسایی جایگاه‌های N-گلیکوزیلاسیون در توالی اسیدآمینه‌ای hbFGF توسط برنامه NetNGlyc انجام شد ([www.expasy.ch/tools/](http://www.expasy.ch/tools/)) همچنین ساختارهای دوم و سوم hbFGF به ترتیب با استفاده از برنامه‌های PSIPred (<http://bioinf.cs.ucl.ac.uk/psipred/>) و SWISS-MODEL (<https://swissmodel.expasy.org/>) تعیین شده و نمودار رامانچاندران نیز توسط برنامه Molprobity رسم شد (<http://molprobity.biochem.duke.edu>). دامنه‌ها یا دُمین‌های عملکردی و جایگاه‌های ویژه تغییرات پس از ترجمه به ترتیب با برنامه‌های MotifScan ([https://myhits.isb-sib.ch/cgi-bin/motif\\_scan](https://myhits.isb-sib.ch/cgi-bin/motif_scan)) و ScanProsite (<https://prosite.expasy.org/scanprosite/>) شناسایی شدند. پیش‌بینی نواحی غیرساختاری و جایگاه‌های اتصالی غیرساختاری در توالی اسیدآمینه‌ای پروتئین hbFGF از طریق برنامه IUPred2 انجام شد. بررسی شبیه‌سازی تطابق ساختارهای سوم اشکال پروتئینی بهینه‌سازی شده و نشده با برنامه 3d-SS (3 Dimensional Structural Superposition) (Russell et al. 1992) انجام شد. در نهایت، هم‌ردیف‌سازی چندگانه و ترسیم درخت فیلوژنتیکی فاقد ریشه با استفاده از الگوریتم Neighbour-Joining در نرم‌افزار ClustalW انجام شدند.

## نتایج و بحث

همسانه‌سازی ناقل بیانی نو ترکیب pCAMBIA-hbFGF: ناقل بیانی نو ترکیب pCAMBIA-hbFGF پس از انتقال به درون سلول‌های مستعد *E. coli*، به روش لیز قلیایی با SDS استخراج شده (Birnboim & Doly 1979) و با استفاده از PCR و برش آنزیمی مورد بررسی و تأیید قرار گرفت. بررسی ناقل بیانی نو ترکیب pCAMBIA-hbFGF با استفاده از PCR، یک نوار منفرد و واضح با اندازه مورد انتظار ۴۶۸ bp مطابق با ژن *hbFgf* را روی ژل نشان داد. ناقل بیانی نو ترکیب با استفاده از برش آنزیمی با آنزیم‌های برشی *Nco I* و *BstE II* نیز مورد بررسی قرار گرفت. برش آنزیمی با آنزیم‌های برشی *Nco I* و *BstE II* به شکل منفرد تنها یک تک نوار مربوط به ناقل بیانی نو ترکیب pCAMBIA-hbFGF خطی را روی ژل ایجاد کرد. در حالی که برش آنزیمی با جفت آنزیم‌های برشی *NcoI/BstE II* منجر به رهاسازی قطعه مربوط به ژن *hbFgf* با اندازه مورد انتظار ۴۶۸ bp از ناقل بیانی (با اندازه ۹۷۸۹ bp) شد (شکل ۱).

ناقل دوگانه نو ترکیب پس از استخراج از باکتری *E. coli*، با استفاده از روش انجماد و ذوب (Jyothishwaran et al. 2007) به درون سلول‌های مستعد اگروباکتریوم سویه‌های EHA105، GV3101 و LBA4404 انتقال داده شد. میزان بازده تراریختی با استفاده از شاهد مثبت حاوی ناقل دوگانه pCAMBIA1304 غیرنو ترکیب در حدود  $10^2 \times 3/8$  کلنی نو ترکیب به ازای یک میکروگرم ناقل فرایپچیده محاسبه شد. میزان بازده تراریختی در باکتری اگروباکتریوم در حدود  $10^3-10^2$  کلنی نو ترکیب به ازای یک میکروگرم ناقل فرایپچیده است (Hofgen & Willmitzer 1988). اما استفاده از سلول‌های باکتریایی در فاز لگاریتمی ( $OD_{600} = 0/3$ ) و استفاده از محلول  $MgCl_2$  ۱۰۰ mM جهت تهیه سلول‌های مستعد، میزان بازده تراریختی را تا  $10^5 \times 3/6$  افزایش می‌دهد (Jyothishwaran et al. 2007). ناقل دوگانه نو ترکیب pCAMBIA-hbFGF با روش لیز قلیایی با SDS از کلنی‌های نو ترکیب اگروباکتریوم استخراج شده (Birnboim & Doly 1979) و با استفاده از PCR مورد بررسی قرار گرفت. ناقل دوگانه نو ترکیب استخراج شده به شکل نوارهایی با اندازه‌های مختلف روی ژل مشاهده شدند که نشان‌دهنده اشکال ساختمانی متفاوت و در نتیجه کیفیت بالای آن‌ها است. بررسی ناقل‌های دوگانه نو ترکیب با استفاده از PCR و آغازگرهای اختصاصی ژن *hbFgf*، تکثیر یک قطعه DNA با اندازه مورد انتظار ۴۶۸ bp را روی ژل نشان داد. همچنین شاهد‌های منفی در PCR از قبیل شاهد منفی بدون الگو و شاهد منفی حاوی ناقل دوگانه pCAMBIA1304 فاقد محصول PCR تکثیر یافته بودند که نشان‌دهنده تراریختی موفقیت‌آمیز سویه‌های مختلف اگروباکتریوم با ناقل بیانی نو ترکیب بود (شکل ۲).

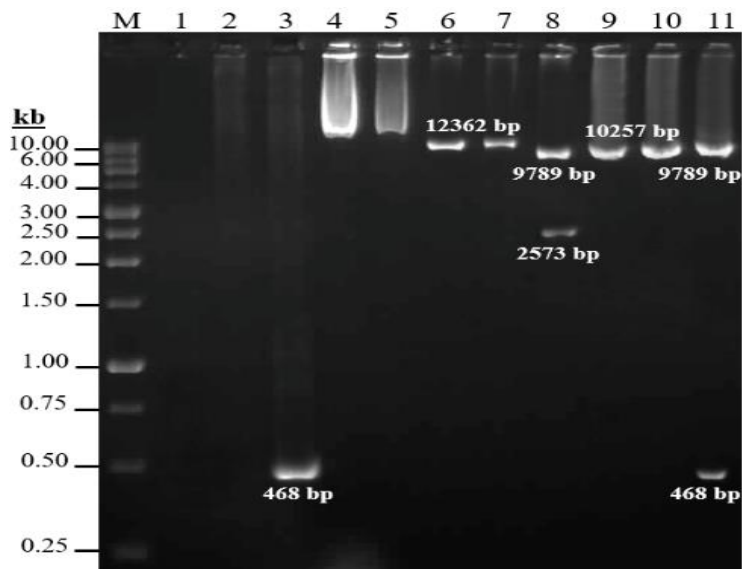
## مطالعات بیوانفورماتیکی پروتئین hbFGF- بررسی توالی اسید آمینه‌ای پروتئین hbFGF: ویژگی‌های

فیزیکی- شیمیایی توالی اسید آمینه‌ای پروتئین hbFGF با استفاده از برنامه ProtParam مورد بررسی قرار گرفت. پروتئین hbFGF با ۱۵۵ اسید آمینه دارای وزن مولکولی و نقطه ایزوالکتریک به ترتیب برابر با ۱۷/۲۵ kDa و ۹/۵۸ است.





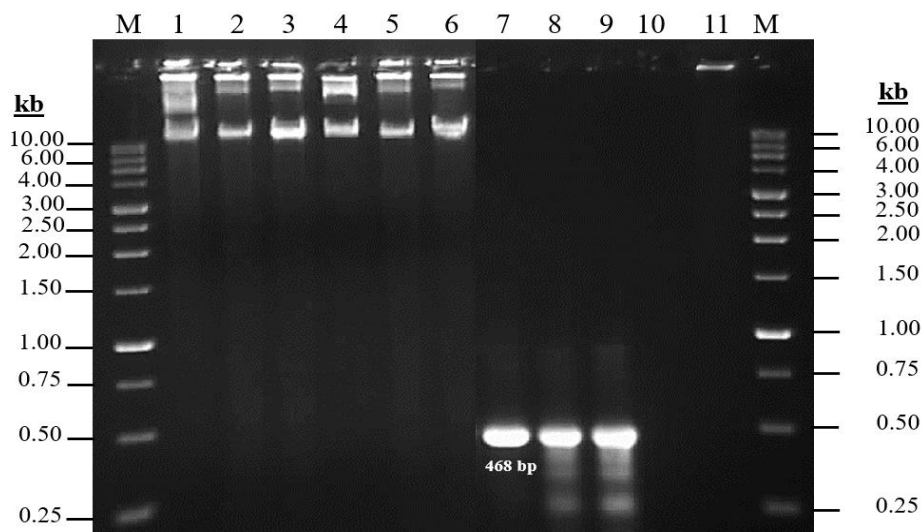
الف (A)



ب (B)

شکل ۱. الف) نقشه ناقل بیانی نوترکیب pCAMBIA-hbFGF و ب) تأیید آن با استفاده از PCR، برش آنزیمی و الکتروفورز ژل آگارز ۱.۲٪ (M) نشانگر وزن مولکولی DNA ۱ kb، ۱) شاهد منفی بدون الگو در PCR، ۲) شاهد منفی حاوی ناقل دوگانه pCAMBIA1304 در PCR، ۳) ژن *hbFgf* تکثیر شده از روی ناقل دوگانه نوترکیب pCAMBIA-hbFGF در PCR، ۴) ناقل دوگانه pCAMBIA1304، ۵) ناقل بیانی نوترکیب pCAMBIA-hbFGF، ۶-۸) ناقل دوگانه pCAMBIA1304 برش یافته به ترتیب با آنزیمهای برشی *Nco* I، *BstE* II و *Nco* I/*BstE* II، ۹-۱۱) ناقل بیانی نوترکیب pCAMBIA-hbFGF برش یافته به ترتیب با آنزیمهای برشی *Nco* I، *BstE* II و *Nco* I/*BstE* II

**Figure 1. A) The map of recombinant expression vector pCAMBIA-hbFGF and B) Confirmation it by PCR, enzyme digestion and 1.2% agarose gel electrophoresis. M) 1 kb DNA ladder, 1) PCR negative control without template, 2) PCR negative control with pCAMBIA1304 binary vector, 3) The *hbFgf* gene amplified in PCR containing the recombinant expression vector pCAMBIA-hbFGF as a template, 4) The pCAMBIA1304 binary vector, 5) The recombinant expression vector pCAMBIA-hbFGF, 6-8) The pCAMBIA1304 binary vector digested with *Nco* I, *BstE* II, and *Nco* I/*BstE* II, respectively, 9-11) The recombinant expression vector pCAMBIA-hbFGF digested with *Nco* I, *BstE* II, and *Nco* I/*BstE* II, respectively**



شکل ۲. تأیید کلنی‌های نوترکیب آگروباکتریوم با استفاده از PCR و الکتروفورز ژل آگارز ۱/۲٪ (M) نشانگر وزن مولکولی DNA ۱ kb (۱-۳) ناقل دوگانه pCAMBIA1304 استخراج شده به ترتیب از سویه‌های EHA105، GV3101، و LBA4404 (۴-۶) ناقل دوگانه نوترکیب pCAMBIA-hbFGF استخراج شده به ترتیب از سویه‌های EHA105، GV3101، و LBA4404 (۷-۹) ژن *hbFgf* تکثیر شده با PCR از روی ناقل دوگانه نوترکیب استخراج شده به ترتیب از سویه‌های EHA105، GV3101، و LBA4404 (۱۰) شاهد منفی بدون الگو در PCR (۱۱) شاهد منفی حاوی ناقل دوگانه pCAMBIA1304 در PCR

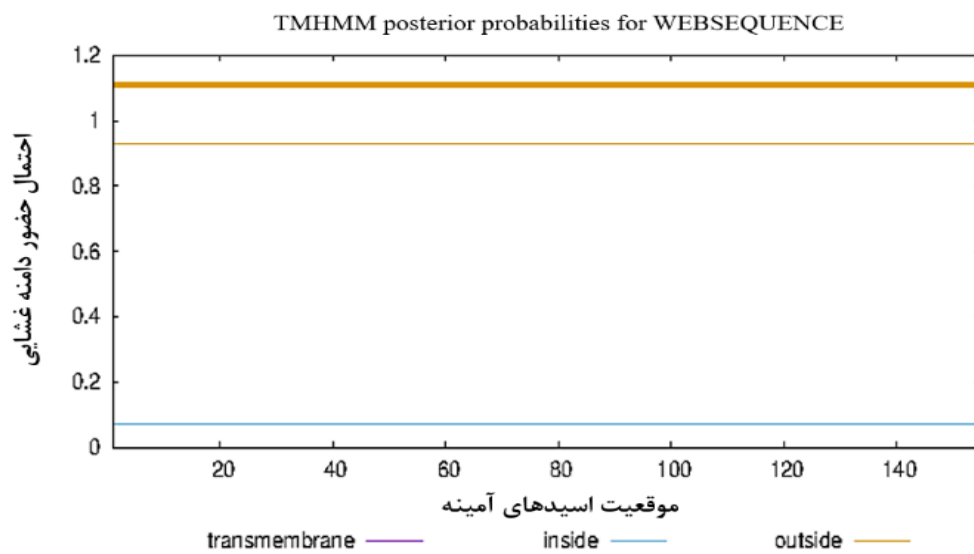
**Figure 2. Confirmation of the recombinant *Agrobacterium* colonies by PCR and 1.2% agarose gel electrophoresis. M) 1 kb DNA ladder, 1-3) The pCAMBIA1304 binary vector extracted from EHA105, GV3101, and LBA4404 strains, respectively, 4-6) The recombinant expression vector pCAMBIA-hbFGF extracted from EHA105, GV3101, and LBA4404 strains, respectively, 7-9) The *hbFgf* gene amplified in PCR containing the recombinant expression vector pCAMBIA-hbFGF extracted from EHA105, GV3101, and LBA4404 strains, respectively, 10) PCR negative control without template, 11) PCR negative control with pCAMBIA1304 binary vector**

شاخص ناپایداری (Instability index) به عنوان یک عامل مهم جهت شناسایی پروتئین‌های پایدار با نیمه‌عمر بالا، برای پروتئین *hbFGF* حدوداً ۳۶/۶۰ محاسبه شد که نشان‌دهنده پایداری بالای آن است. مقدار شاخص ناپایداری باید کمتر از ۴۰ باشد تا یک پروتئین به عنوان پروتئین پایدار شناخته شود (Claverie & Notredame 2007). شاخص آلیفاتیک (Aliphatic index) به عنوان یک عامل ضروری جهت شناسایی پروتئین‌های مقاوم در برابر حرارت (Kim et al. 2008)، در حدود ۶۸/۰۰

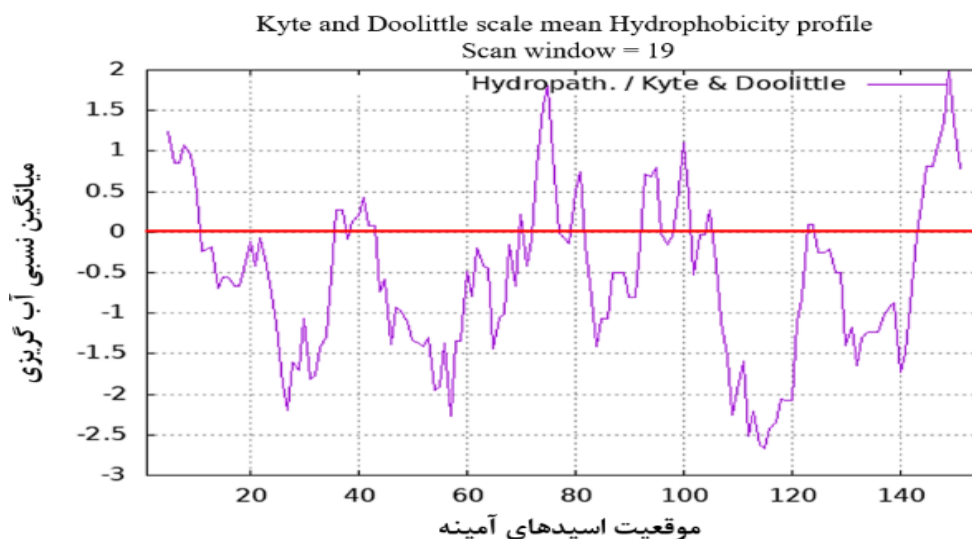
محاسبه گردید که بیانگر حساسیت بالای پروتئین hbFGF در برابر حرارت است. همچنین، شاخص آب‌گریزی (Hydrophobicity index) پروتئین hbFGF در حدود ۰/۵۱۱- برآورد گردید که نشانگر وضعیت آب‌دوستی آن است (Claverie & Notredame 2007). به علاوه، بررسی‌ها با استفاده از برنامه‌های ProtScale و TMHMM تأیید کرد که پروتئین hbFGF آب‌دوست بوده و فاقد هر گونه دامنه‌های غشای سلولی و ماریپچ‌های  $\alpha$  بین‌غشایی است (Kyte & Doolittle 1982) (شکل ۳).

دامنه‌های عملکردی در پروتئین hbFGF با استفاده از برنامه MotifScan مورد بررسی قرار گرفته و یک دامنه ویژه خانواده فاکتور رشد فیبروبلاستی (FGF) در موقعیت ۱۴۹-۲۸ اسیدآمین با شاخص‌های آماری  $\text{raw-score} = ۲۲۸/۹$  و  $\text{raw-score} = ۷۳/۸۶۴$  = N-score و  $E\text{-value} = ۲/۹e-۶۷$  شناسایی شد (شکل ۴ الف). بررسی‌ها با برنامه‌های ScanProsite و NetNGlyc هیچ‌گونه جایگاه بالقوه جهت تغییرات پس از ترجمه به ویژه جایگاه‌های N-گلیکوزیلایون را در توالی اسیدآمین‌های hbFGF شناسایی نکردند (شکل ۴ ب).

با استفاده از برنامه IUPred2، سه ناحیه غیرساختاری کوچک در توالی اسیدآمین‌های پروتئین hbFGF شناسایی شد (شکل ۵). هدف اصلی برنامه IUPred2 شناسایی نواحی پروتئینی ذاتاً غیرساختاری (Intrinsically Disordered Protein) (Regions, IDPRs)، یعنی نواحی‌ای که فاقد ساختار مونومر پایدار تحت شرایط بومی بر اساس یک مدل مبتنی بر بیوفیزیک است. این برنامه از دو الگوریتم مختلف IUPred2A و ANCHOR2 جهت پیش‌بینی IDPRها استفاده می‌کند. ماهیت غیرساختاری یک ناحیه پروتئینی می‌تواند وابسته به شرایط محیطی باشد، به این معنی که نواحی ویژه‌ای از پروتئین بسته به شرایط مختلف محیطی ممکن است بین وضعیت‌های ساختاری و غیرساختاری تغییر حالت بدهند. الگوریتم IUPred2A قادر است نواحی غیرساختاری وابسته به شرایط محیطی تحت تغییر عوامل محیطی، تغییر در وضعیت اکسیداسیون-احیاء (Redox) سلولی یا حضور یک عامل متصل‌شونده را پیش‌بینی کند. علاوه بر این، IDPRها ممکن است نتیجه تغییر بین حالت‌های تاخورده و تانخورده ناشی از تغییر وضعیت ردوکس سلولی باشند. این نواحی پروتئینی بسته به موقعیت‌شان در سلول ممکن است به حالت ساختاری یا غیرساختاری باشند. الگوریتم IUPred2A همچنین می‌تواند این نواحی غیرساختاری حساس به وضعیت ردوکس را شناسایی کند (Mészáros et al., 2018). IDPRها اغلب دارای جایگاه‌های متصل‌شونده جهت برهم‌کنش اختصاصی با یک دامنه کروی هستند. در اکثر موارد، در طی این برهم‌کنش، ناحیه غیرساختاری متصل‌شونده، ساختاری منظم در حالت اتصال شده به خود می‌گیرد. این حالت احتمالاً شایع‌ترین وضعیت غیرساختاری پروتئین وابسته به توالی است که در آن انتقال بین وضعیت‌های ساختاری و غیرساختاری در حضور یک پروتئین دیگر انجام می‌شود. این نواحی غیرساختاری متصل‌شونده با استفاده از الگوریتم ANCHOR2 شناسایی می‌شوند (Erdos et al., 2020).



الف (A)

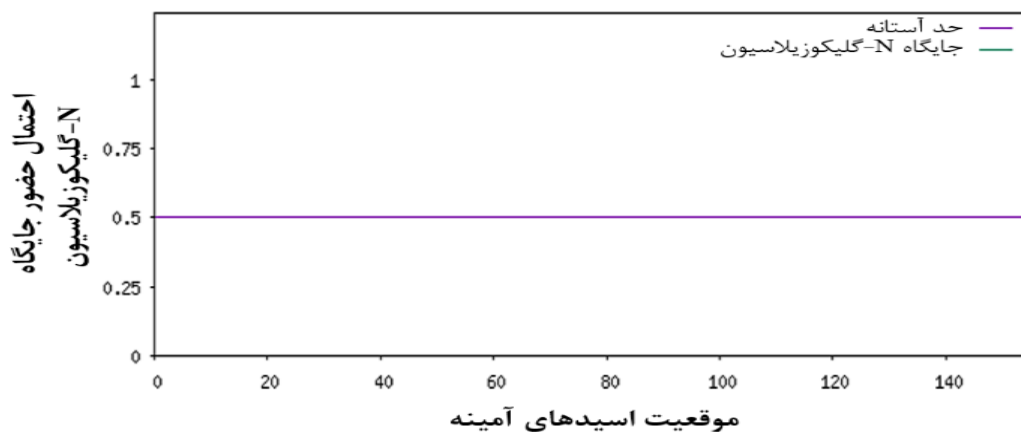


ب (B)

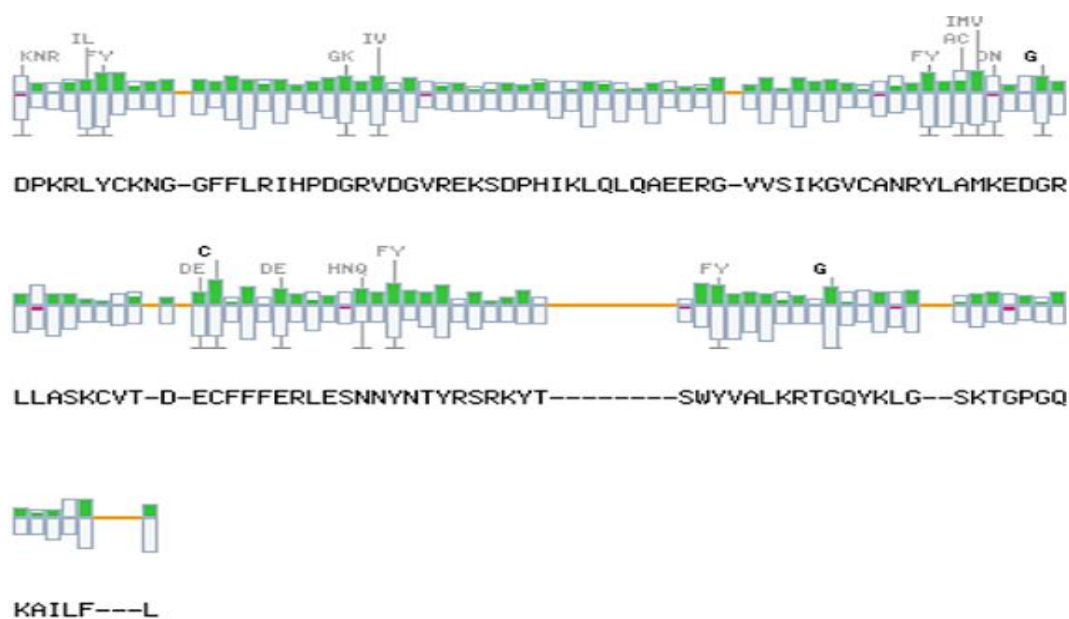
شکل ۳. ویژگی‌های آب‌گریزی پروتئین hbFGF. الف) بررسی شاخص آب‌گریزی در پروتئین hbFGF با استفاده از برنامه ProtScale. قله‌های بالاتر و پایین‌تر از خط آستانه به ترتیب نشان‌دهنده نواحی آب‌گریز و آب‌دوست هستند. ب) پیش‌بینی دامنه‌های غشای سلولی و مارپیچ‌های  $\alpha$  بین غشایی از طریق برنامه

TMHMM

**Figure 3. Hydrophobicity features of hbFGF deduced amino acid sequence. A) Hydrophobic index analysis of hbFGF deduced amino acid sequence by using ProtScale. Hydrophobic domains are above the line, and hydrophilic domains are below. B) Prediction of cell membrane domains and transmembrane helices via TMHMM program**



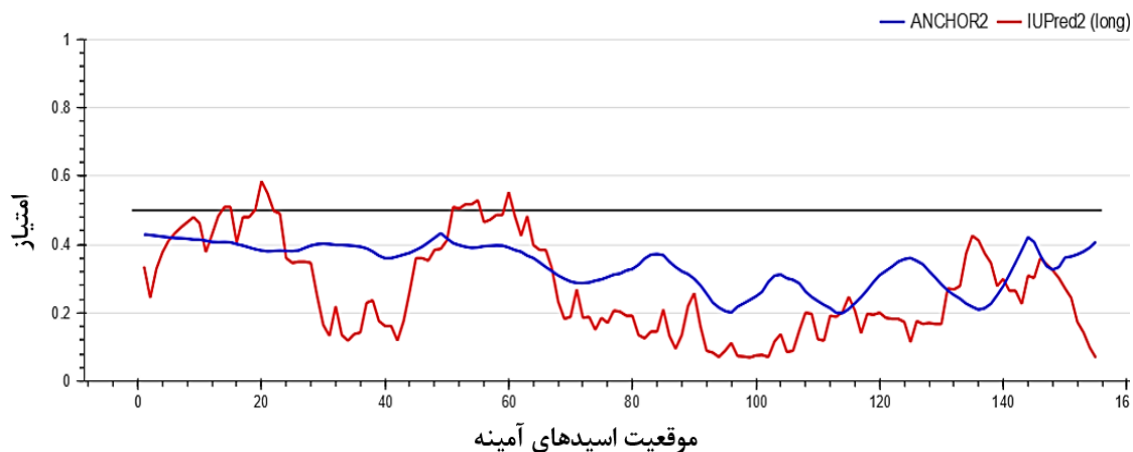
(A) الف



(B) ب

شکل ۴. پیش‌بینی دامنه‌های عملکردی و جایگاه‌های بالقوه N-گلیکوزیلاسیون در پروتئین hbFGF. الف) پیش‌بینی دامنه FGF با استفاده از برنامه MotifScan. جایگاه‌های حفاظت‌شده (نوار سبز) و غیرحفاظت‌شده (نوار قرمز) و همچنین میزان حفاظت‌شدگی (طول نوار) در دامنه FGF قابل مشاهده است. ب) پیش‌بینی جایگاه‌های بالقوه N-گلیکوزیلاسیون از طریق برنامه NetNGlyc

Figure 4. Prediction of functional domains and the potential sites for N-glycosylation in the hbFGF deduced amino acid sequence. A) Prediction of the FGF domain using MotifScan program. The conserved and non-conserved regions in the FGF domain are showed by green and red bands, respectively. The degree of conservation is also indicated via band length. B) Prediction of the potential sites for N-glycosylation through NetNGlyc program



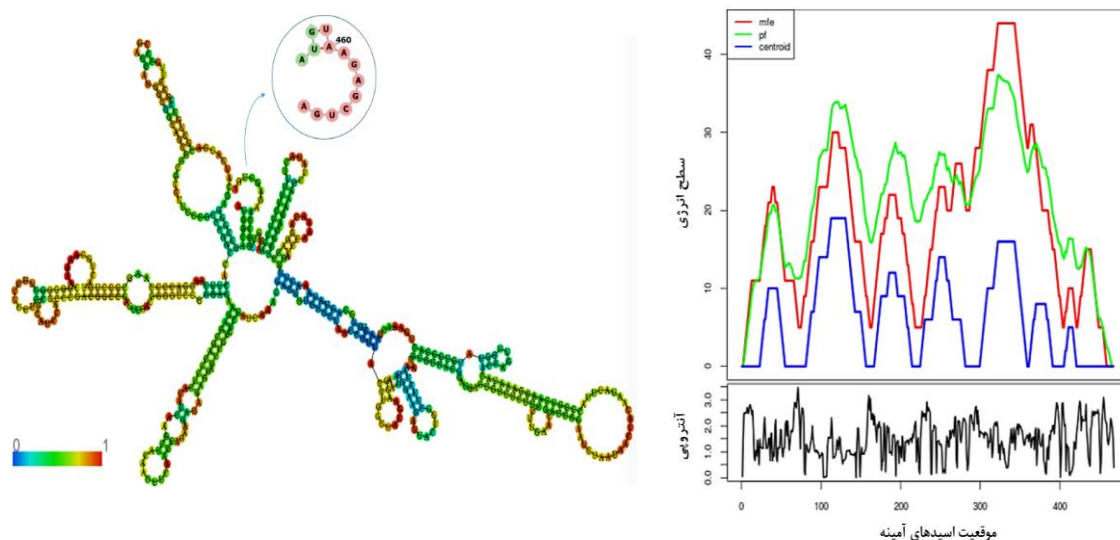
شکل ۵- شناسایی نواحی پروتئینی ذاتاً غیرساختاری (IDPRs) در توالی اسیدآمینه‌ای پروتئین hbFGF با استفاده از برنامه IUPred2

**Figure 5. Identification of IDPRs in the hbFGF deduced amino acid sequence using IUPred2 program**

**بررسی ساختار دوم توالی mRNA** در یک مولکول mRNA، ساختار دوم از طریق تشکیل پیوندهای هیدروژنی بین نوکلئوتیدهای مکمل تشکیل می‌شود. ساختار دوم یک مولکول mRNA سطح بیان بسیاری از ژن‌ها را در مرحله رونویسی تحت تأثیر قرار داده و نقش مهمی را در تنظیم فرآیندهای رونویسی، پیرایش، تخریب و ترجمه ایفا می‌کند. بنابراین، بررسی ساختار دوم توالی‌های mRNA به منظور تعیین فعالیت عملکردی یک رونوشت از اهمیت بالایی برخوردار است (Proctor & Meyer 2013). ساختار دوم توالی mRNA به منظور ارزیابی پایداری آن و تعیین موقعیت رمزهای آغازین و پایان با استفاده از نرم‌افزار RNAfold مورد بررسی قرار گرفت. حداقل انرژی آزاد به دست آمده از ساختار دوم توالی mRNA مربوط به *hbFgf* برابر  $-136/90$  kcal/mol محاسبه شد که نشان‌دهنده پایداری نسبتاً بالای ساختار دوم توالی mRNA با سطح آنتروپی پایین است. در توالی mRNA حاصل از *hbFgf* رمز آغازین در یک ساختار سنجاق سری چندگانه قرار می‌گیرد به طوری که نوکلئوتید A در ناحیه حلقه اما دو نوکلئوتید دیگر رمز آغازین (G و U) از طریق پیوند هیدروژنی با نوکلئوتیدهای مکمل جفت شده و ناحیه ساقه را تشکیل می‌دهد. در مقابل، رمز پایان (UGA) کاملاً در ناحیه حلقه واقع شده است (شکل ۶). بررسی ساختار دوم توالی mRNA نشان می‌دهد که رمز آغازین کاملاً در معرض مستقیم فعالیت ریبوزوم قرار ندارد که احتمالاً می‌تواند نرخ فرآیند رونویسی و به دنبال آن میزان بیان پروتئین نوترکیب hbFGF در گیاهان تراریخت شده با بیان گذرا را تحت تأثیر قرار دهد.

حداقل انرژی آزاد ساختار دوم یک مولکول mRNA از محاسبه مجموع انرژی‌های آزاد ساختارهای ساقه، حلقه‌های درونی و بیرونی، حلقه‌های چندگانه، ساختارهای سنجاق سری و برآمدگی‌ها در توالی mRNA به دست آمده و عدد حاصله تنها

معرف میزان پایداری کل مولکول mRNA خواهد بود (Mathews et al. 2004). بنابراین، جهت بررسی میزان رونویسی از یک مولکول mRNA، تنها میزان حداقل انرژی آزاد ساختار دوم مربوط به توالی کزاک و رمز آغاز حائز اهمیت می‌باشد. مقدار انرژی آزاد در این ساختارها باید در یک حد متعادل باشد، یعنی به اندازه کافی بالا برای پایداری در برابر نوکلئازها و به اندازه کافی پایین جهت تسهیل در تخریب پیوندهای هیدروژنی و آغاز ترجمه (De Smit & Van Duin 1990). به عنوان مثال، برای یک فرآیند بهینه ترجمه، حداقل انرژی آزاد برای ساختار دوم حاوی رمز آغاز باید از  $-6 \text{ kcal/mol}$  بیشتر باشد (Zhang et al. 2006).

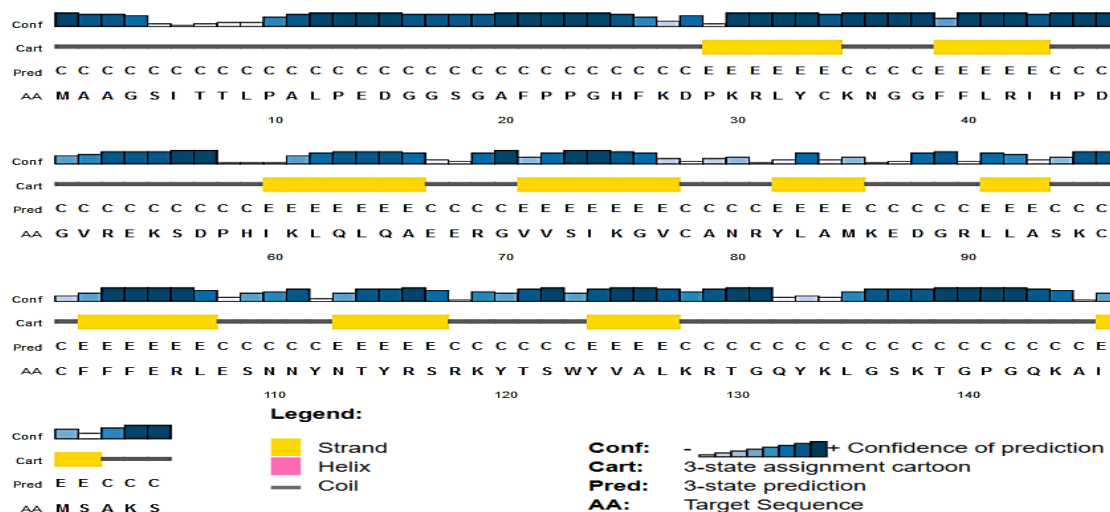


شکل ۶. پیش‌بینی ساختار دوم توالی mRNA مربوط به hbFgf به همراه نمودارهای سطوح حداقل انرژی آزاد و آنتروپی در هر موقعیت

**Figure 6. Prediction of the secondary structure of hbFgf mRNA with minimum free energy and the positional entropy for each position**

اثر ساختار دوم mRNA و حداقل انرژی آزاد روی سطوح بیان پروتئین نوترکیب hIFN- $\gamma$  در باکتری *E. coli* یک افزایش ۱۰ برابری در میزان بیان همراه با کاهش حداقل انرژی آزاد را نشان داد (Razaghi et al. 2016). به علاوه، اختلاف در ساختار دوم یک توالی mRNA می‌تواند کارایی ترجمه را به میزان ۱۰۰۰ برابر تحت تأثیر قرار دهد (Helke et al. 1993). در معرض قرار داشتن جایگاه اتصال ریبوزوم و رمز آغاز یک توالی mRNA جهت فعالیت ریبوزوم از عوامل موثر در فرآیند ترجمه به شمار می‌آیند (Mathews et al. 2004). اما همان طور که مشاهده شد، رمز آغازین در توالی mRNA مربوط به hbFgf از طریق پیوندهای هیدروژنی عمدتاً ساختار ساقه را تشکیل دادند. بنابراین، به نظر می‌رسد که احتمالاً میزان بیان پروتئین نوترکیب hbFGF در حد انتظار نباشد. با این وجود، Van Duin و De Smit (1990) نشان دادند که میزان ترجمه مستقیماً به حداقل انرژی آزاد ساختار دوم توالی کزاک و رمز آغازین بستگی دارد، به طوری که کاهش حداقل انرژی آزاد در این ساختارها به میزان  $1/4 \text{ kcal/mol}$ ، میزان بیان را تقریباً ۱۰ برابر کاهش می‌دهد.

**بررسی مدل‌سازی مولکولی پروتئین hbFGF:** بررسی ساختار دوم توسط برنامه PSIPred نشان داد که پروتئین hbFGF حاوی ۱۰ صفحه  $\beta$  و فاقد ماریج  $\alpha$  است (شکل ۷). صفحات  $\beta$  در مجموع ۳۲/۲۶٪ ساختار پروتئین را تشکیل داده و بزرگ‌ترین و کوچک‌ترین صفحات  $\beta$  به ترتیب ۷ و ۳ اسیدآمینه را شامل می‌شوند. با استفاده از توالی اسیدآمینه‌ای hbFGF به عنوان یک query، یک ساختار سوم برای پروتئین بومی bFGF از *Homo sapiens* (PDB ID: 1BLA.1) با بالاترین امتیاز (میزان شباهت: ۹۷/۴۲٪، E-value =  $1e-112$ ) از داده‌پایگاه مربوط به توالی‌های پروتئینی (Protein Data Bank, PDB) و از طریق الگوریتم BLASTp در NCBI به دست آمد. شبیه‌سازی ساختار سه‌بعدی پروتئین hbFGF با استفاده از برنامه SWISS-MODEL نشان داد که پروتئین hbFGF شامل ۱۰ صفحه  $\beta$  غیرموازی است که در یک ساختار هرمی سازماندهی شده است (شکل ۸). با استفاده از فناوری تولید پروتئین نو ترکیب، ساختار سه‌بعدی پروتئین FGF-2 با وزن مولکولی ۱۸ kDa توسط Thompson (1992) تعیین شد. چندین دامنه عملکردی در توالی FGF-2 شناسایی شده است. اعتقاد بر این است که اسیدآمینه‌های ۱۲۸-۱۴۴ ( $K_{128}RTGQYKLGSKTGPQK_{144}$ ) جایگاه اتصال به هپارین هستند. دو جایگاه بالقوه فسفوریلاسیون شامل  $Ser_{61}$  و  $Thr_{119}$  در توالی اسیدآمینه‌ای FGF-2 شناسایی شده است.  $Ser_{61}$  و  $Thr_{119}$  به ترتیب توسط پروتئین کیناز A و پروتئین کیناز C فسفریله می‌شوند. کینازهای سلولی مسئول فسفوریلاسیون FGF2 ممکن است هم در هسته و هم در سطح سلول قرار داشته باشند. پروتئین FGF-2 همچنین حاوی چهار سیستم است. با این وجود، هیچ پیوند دی‌سولفید درون مولکولی در این پروتئین تشکیل نمی‌شود (Thompson 1992).



شکل ۷. ساختار دوم پروتئین hbFGF با سطح حفاظت‌شدگی در طی تکامل توسط برنامه PSIPred

**Figure 7. The secondary structure of the hbFGF deduced amino acid sequence with the conservation degree through PSIPred program**



شبهه‌سازی تطابق ساختارهای سه‌بعدی با استفاده از برنامه 3d-SS نشان داد که ساختار سوم پروتئین hbFGF با ساختارهای سوم از پستانداران *Rattus norvegicus* (PDB ID: 2J3P, E-value = 9e-46) و *Oryctolagus cuniculus* (PDB ID: 1AFC, E-value = 3e-47) و *Bos taurus* (ID: 3HAL, E-value = 4e-47) به ترتیب با ۵۴/۸۹٪، ۵۴/۷۴٪ و ۵۴/۰۱٪ شباهت ساختاری کاملاً برهم‌دیگر منطبق هستند (شکل ۸).

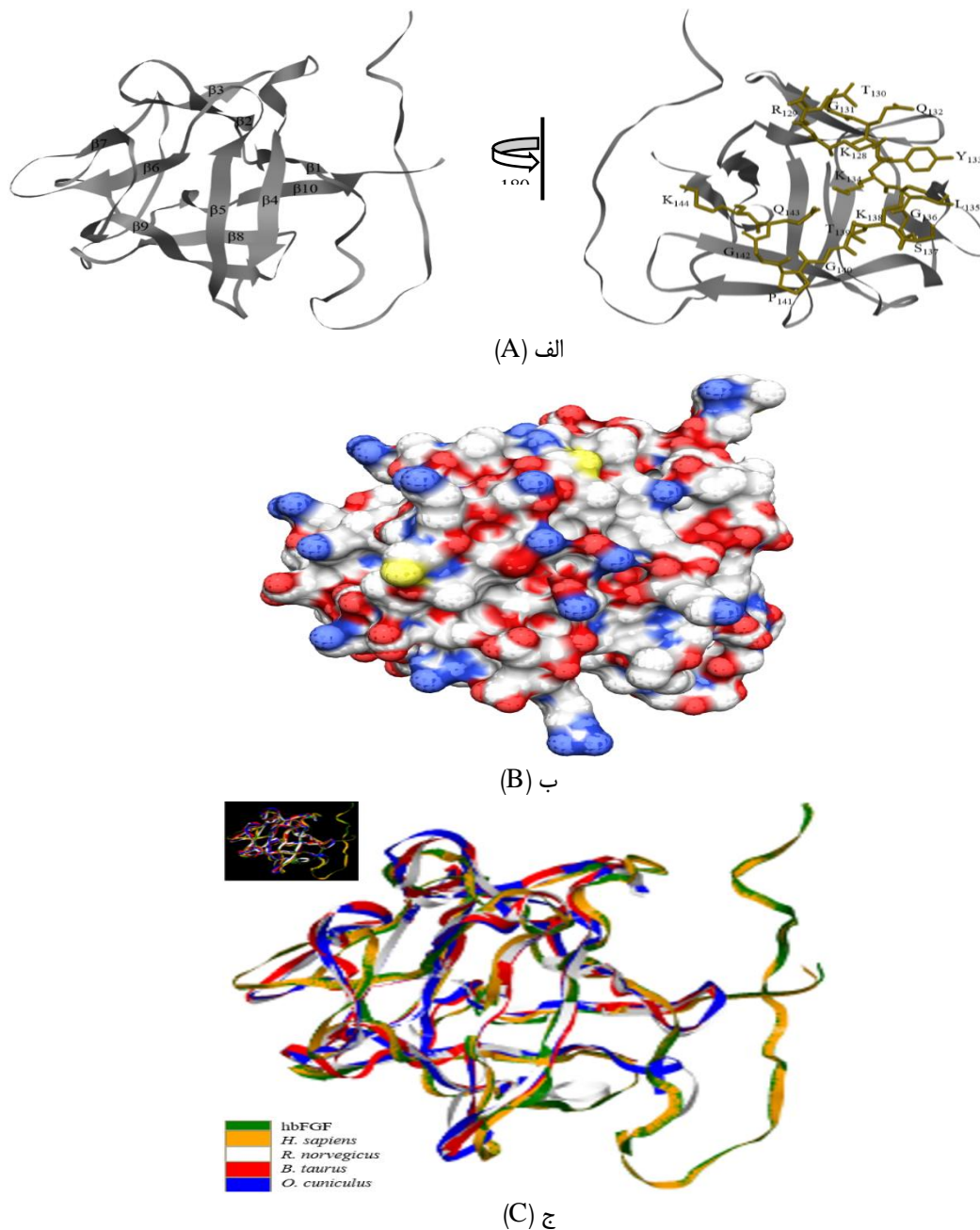
### پیش‌بینی نمودار رامچاندران (Ramachandran plot): نمودار رامچاندران مربوط به پیکربندی پروتئین hbFGF

با استفاده از برنامه MolProbity ترسیم شد. با توجه به نمودار، پروتئین hbFGF عمدتاً از صفحات  $\beta$  غیرموازی تشکیل شده است. تقریباً ۹۸/۷٪ از اسیدهای آمینه در ساختارهای منظم و نامنظم و حدود ۸۶/۳٪ از اسیدهای آمینه در ساختارهای منظم شامل صفحات  $\beta$  غیرموازی سازمان‌دهی شده است. تنها دو اسید آمینه Thr<sub>8</sub> ( $\psi = -۸۴/۸$ ) و Gly<sub>24</sub> ( $\psi = ۴/۶$ ) با  $\phi = -۱۵۹/۱$  و  $\psi = ۱۶۷/۵$  با زوایای  $\phi$  و  $\psi$  غیرمعمول در پیکربندی پروتئین hbFGF شناسایی شد. علاوه بر این، پروتئین hbFGF در مجموع حاوی ۱۲ اسید آمینه والین و ایزولوسین، ۹ اسید آمینه پرولین با آرایش ترانس و فاقد Pro با آرایش سیس است (شکل ۹).

نمودار گرافیکی مرتبط با انرژی‌های میان‌کنشی بین زنجیره‌های جانبی اسیدهای آمینه به عنوان تابعی از زوایای  $\phi$  و  $\psi$  اولین بار توسط Ramachandran و Sasiskharan (1968) برای نشان دادن پیکربندی‌های مجاز یک زنجیره پلی‌پپتیدی با در نظر گرفتن پدیده ممانعت فضایی ترسیم شد. در رسم این نمودار ابتدا اسید آمینه آلانین به عنوان مدل و نماینده بقیه اسیدهای آمینه انتخاب شد. با این وجود، آلانین مدل مناسبی برای گلیسین، پرولین و حتی اسیدهای آمینه حاوی زنجیره‌های جانبی حجیم نمی‌تواند باشد. زنجیره جانبی گلیسین فقط یک اتم هیدروژن دارد، پس در این اسید آمینه ممانعت فضایی ناشی از کربن  $\beta$  وجود ندارد. گلیسین با انعطاف‌پذیری بسیار در زنجیره پلی‌پپتیدی در نمودار رامچاندران فضای پیکربندی زیادی را به خود اختصاص می‌دهد. زنجیره جانبی ۳ کربنه پرولین یک حلقه ۵ ضلعی را از طریق کربن  $\alpha$  و نیتروژن آمینی خود ایجاد کرده که در زاویه  $\phi$  فقط در موقعیت ۶۰- و زاویه  $\psi$  فقط در موقعیت‌های ۵۵- و ۱۴۵+ دارای کم‌ترین سطح انرژی است. لذا در نمودار رامچاندران توالی اسید آمینه‌ای و موقعیت هر اسید آمینه در توالی را باید در نظر داشت (Hayward and Milner-White, 2021).

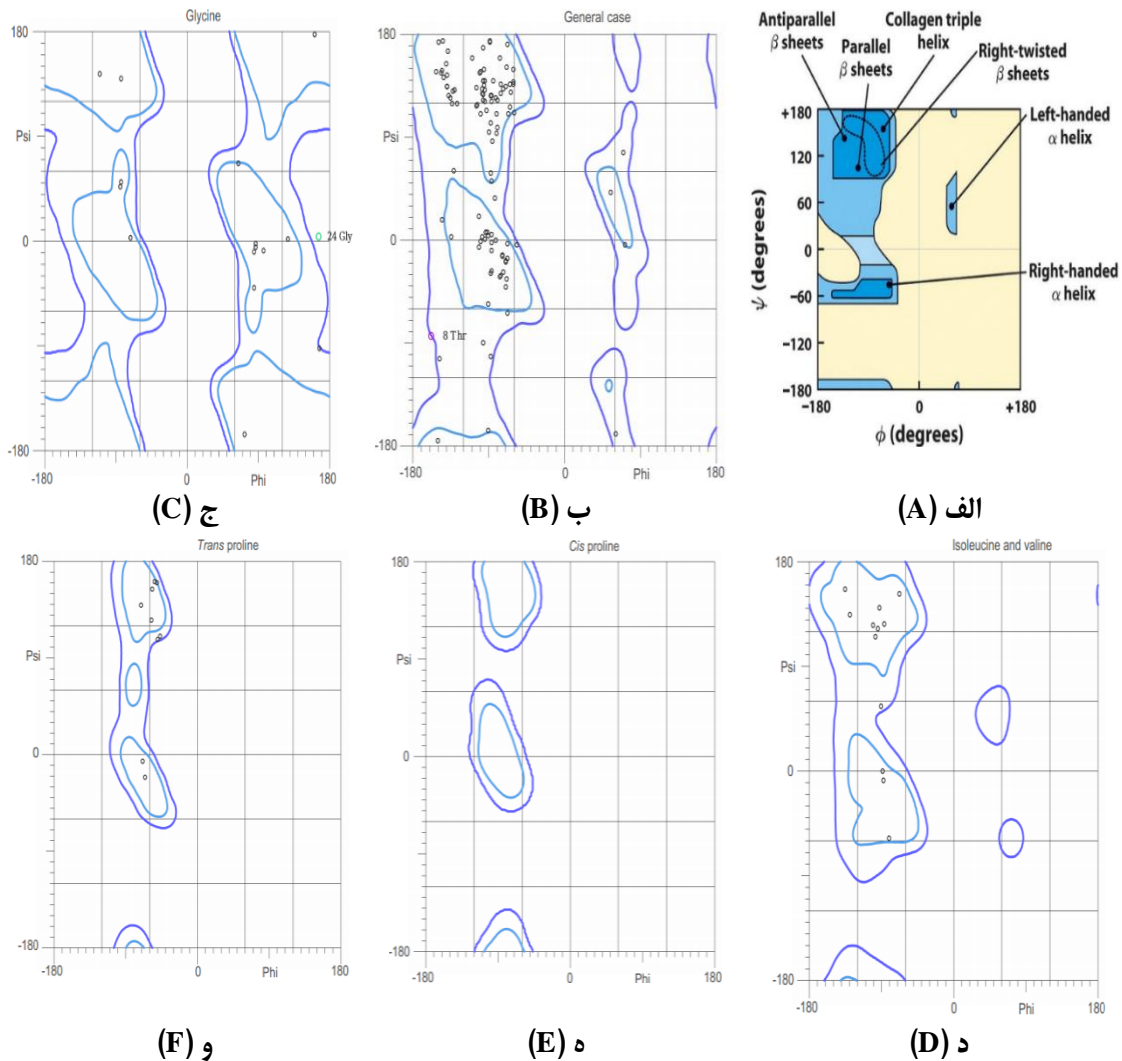
### مطالعه روابط فیلوژنتیکی و هم‌دیف‌سازی توالی چندگانه: یک درخت فیلوژنتیکی ریشه‌دار با توالی پروتئینی hbFGF و

سایر توالی‌های bFGF از پستانداران مختلف با استفاده از نرم‌افزار ClustalW2 با الگوریتم Neighbour-Joining ساخته شد. درخت فیلوژنتیکی به چهار شاخه اصلی شامل انسان و میمون‌ها، گیاهخواران، گوشتخواران و جوندگان تقسیم‌بندی می‌شود (شکل ۱۰ الف). بررسی میزان شباهت hbFGF با سایر توالی‌های bFGF از میمون‌ها، درجه بالایی از شباهت را با شامپانزه (*P. t. troglodytes*; PtbFGF با ۱۰۰٪ شباهت)، آرانگوتان (*P. abelii*; PabFGF با ۱۰۰٪ شباهت)، شتر (*C. dromedarius*; CdbFGF با ۹۸/۷۱٪ شباهت) و گوسفند (*O. aries*; OabFGF با ۹۸/۰۶٪ شباهت) نشان داد (شکل ۱۰ ب).

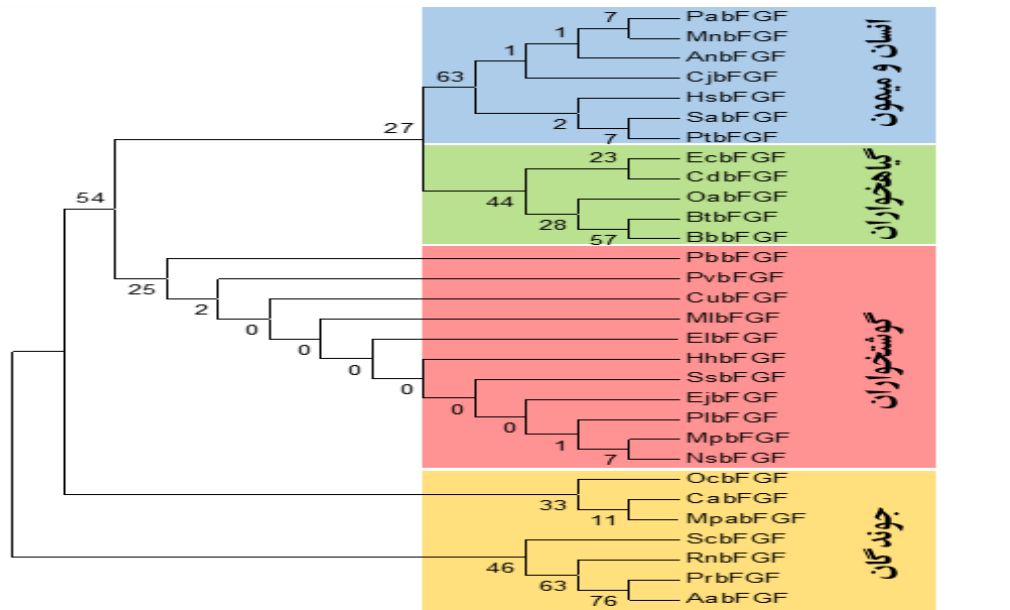


شکل ۸. ساختارهای سه بعدی پروتئین hbFGF پیش بینی شده با استفاده از برنامه SWISS-MODEL. الف) ساختار ریبون پروتئین hbFGF متشکل از ۱۰ صفحه  $\beta$  و آرایش اسیدهای آمینه درگیر در اتصال با هپارین ( $K_{128}RTGQYKLGSKTGPQK_{144}$ ). ب) ساختار سوم فضاپرکن پروتئین hbFGF. ج) انطباق ساختار سوم پروتئین hbFGF با ساختار سوم سایر پروتئین های bFGF از سایر پستانداران با استفاده از برنامه 3d-SS

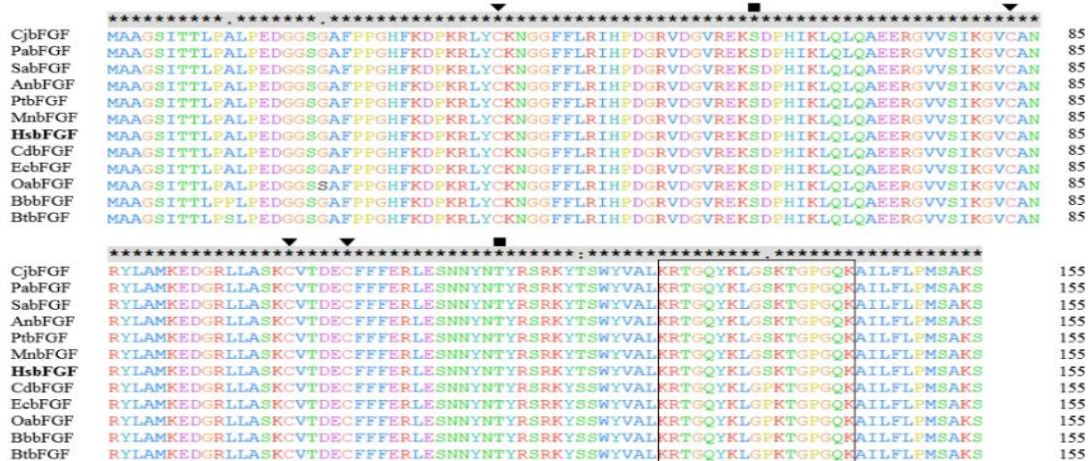
Figure 8. Three-dimensional models for hbFGF predicted by SWISS-MODEL program. A) Cartoon display of the three-dimensional structure of hbFGF including 10  $\beta$ -sheets and amino acid residues involving in interaction with heparin ( $K_{128}RTGQYKLGSKTGPQK_{144}$ ). B) Sphere display of the three-dimensional structure of hbFGF. C) Superimposition of 3D model of hbFGF (green) and the templates from other mammals using 3d-SS service



شکل ۹. الف) نمودار رامچاندرا نشان‌دهنده شاخص‌های مطلوب پیکربندی مرتبط با مقادیر مختلف زوایای φ و ψ برای تعدادی از ساختارهای منظم پروتئینی. ب) نمودار رامچاندرا مربوط به شناسایی ساختارهای منظم در پروتئین hbFGF. ج-و) توزیع اسیدهای آمینه Gly, Val, Ile و Pro در پیکربندی پروتئین hbFGF. Figure 9. A) Ramachandran plot showing the optimal configuration indices associated with different values of φ and ψ angles for a number of ordered protein structures. B) Ramachandran plot showing ordered protein structures in hbFGF. C-F) Distribution of Gly, Val, Ile, and Pro in the configuration of the hbFGF protein



(A) الف



(B) ب

شکل ۱۰. الف) ترسیم درخت فیلوژنتیکی و ب) انجام هم‌ردیف‌سازی توالی چندگانه با توالی پروتئینی hbFGF و سایر توالی‌های bFGF از پستانداران مختلف با استفاده از نرم‌افزار ClustalW2 با الگوریتم Neighbour-Joining. جایگاه متصل‌شونده به هیپارین، اسیدهای آمینه Cys و جایگاه‌های بالقوه فسفوریلاسیون به ترتیب درون کادر و با علامت‌های مثلث و مربع مشخص شده‌اند

Figure 10. A) Construction of phylogenetic tree and B) Multiple sequence alignment of hbFGF and selected bFGFs from different mammals using ClustalW2 with Neighbour-Joining algorithm. Heparin-binding site, Cys residues, and potential phosphorylation sites are indicated in the box and with triangle and square symbols, respectively

جدول ۱. شماره دستیابی NCBI توالی‌های bFGF از پستانداران مختلف مورد استفاده جهت ترسیم درخت فیلوژنتیکی و هم‌ردیف‌سازی چندگانه

**Table 1. NCBI accession numbers of bFGF sequences from different mammals used for phylogenetic tree drawing and multiple alignment**

شماره دستیابی Accession no.	پروتئین Protein	نام علمی Specific name	نام موجود زنده Organisms	ردیف No.
NP_001103711.1	PtbFGF	<i>Pan troglodytes</i>	شامپانزه Chimpanzee	1
PNJ83978.1	PabFGF	<i>Pongo abelii</i>	ارنگوتان Orangutan	2
XP_032148259.1	SabFGF	<i>Sapajus paella</i>	میمون کاپوچین Capuchin monkey	3
NP_001182150.1	EcbFGF	<i>Equus caballus</i>	اسب Horse	4
XP_035148890.2	CjbFGF	<i>Callithrix jacchus</i>	مارموست Marmoset monkey	5
XP_012328437.2	AnbFGF	<i>Aotus nancymaae</i>	میمون شب Night monkey	6
MBZ3874512.1	ScbFGF	<i>Sciurus carolinensis</i>	سنجاب Squirrel	7
NP_776481.2	BtbFGF	<i>Bos taurus</i>	گاو Cow	8
XP_034879159.1	MlbFGF	<i>Mirounga leonina</i>	فیل دریایی Walrus	9
XP_006866328.1	CabFGF	<i>Chrysochloris asiatica</i>	موش کور Mole	10
NP_001009769.1	OabFGF	<i>Ovis aries</i>	گوسفند Sheep	11
XP_010983858.2	CdbFGF	<i>Camelus dromedarius</i>	شتر Camel	12
XP_006079005.1	BbbFGF	<i>Bubalus bubalis</i>	بوفالو Buffalo	13
XP_044769019.1	NsbFGF	<i>Neomonachus schauinslandi</i>	خوک آبی Seal	14
XP_027958051.1	EjbFGF	<i>Eumetopias jubatus</i>	شیر دریایی Sea lion	15
XP_011747595.1	MnbFGF	<i>Macaca nemestrina</i>	میمون ماکاکا Macaca	16
XP_029807018.1	SsbFGF	<i>Suricata suricatta</i>	میرکت (تمس هندی) Meerkat	17
XP_042792627.1	PlbFGF	<i>Panthera leo</i>	شیر Lion	18
XP_022372242.1	EibFGF	<i>Enhydra lutris</i>	سمور آبی Otter	19
XP_047713548.1	PvbFGF	<i>Prionailurus viverrinus</i>	گره وحشی Wildcat	20
XP_040823319.1	OcbFGF	<i>Ochotona curzoniae</i>	موش پیکا Plateau pika	21
XP_043412008.1	PbbFGF	<i>Prionailurus bengalensis</i>	پلنگ Leopard	22
XP_004748022.1	MpbFGF	<i>Mustela putorius</i>	موش خرما Ferret	23
XP_039097538.1	HhbFGF	<i>Hyaena hyaena</i>	کفتار Hyena	24
XP_025729824.1	CubFGF	<i>Callorhinus ursinus</i>	فک Northern fur seal	25
NP_062178.1	RnbFGF	<i>Rattus norvegicus</i>	موش صحرایی Field mouse	26
XP_038203292.1	AabFGF	<i>Arvicola amphibius</i>	موش آبی اروپایی European water vole	27
XP_051041744.1	PrbFGF	<i>Phodopus roborovskii</i>	همستر Hamster	28
XP_021052229.1	MpabFGF	<i>Mus pahari</i>	موش Mouse	29

با توجه به هم‌دیف‌سازی توالی چندگانه، توالی اسیدآمینه‌ای پروتئین‌های bFGF به ویژه جایگاه متصل‌شونده به هپارین و انتهای آمینو در طی تکامل کاملاً حفظ شده‌اند. پژوهش‌ها نشان داده‌اند که انتهای آمینو به منظور حفظ ساختار پروتئین‌ها بسیار ضروری است، به طوری که حذف ۳-۷ اسیدآمینه ابتدایی انتهای آمینو منجر به ناپایداری ساختاری و تولید پروتئین نوترکیب نامحلول در باکتری *E. coli* می‌شود (Lundell & Narula 1994). در مقایسه با گیاهخواران که اسیدآمینه Pro<sub>137</sub> جایگزین Ser<sub>137</sub> شده است، اسیدهای آمینه واقع در جایگاه اتصال به هپارین (K<sub>128</sub>RTGQYKLGSKTGPQK<sub>144</sub>) در بین انسان و میمون‌ها به طور کامل در طی تکامل حفظ شده‌اند. سایر اسیدآمینه‌های محافظت شده مانند جایگاه‌های بالقوه فسفوریلاسیون (Ser<sub>61</sub>) و (Thr<sub>119</sub>) که به ترتیب توسط پروتئین کیناز A و پروتئین کیناز C فسفریله می‌شوند و همچنین چهار اسیدآمینه Cys که در تشکیل پیوند دی‌سولفید درون مولکولی نقشی ندارند (Thompson 1992) در شکل (۱۰ب) نیز نشان داده شده‌اند. شماره دستیابی توالی‌های bFGF مورد استفاده جهت ترسیم درخت فیلوژنتیکی و هم‌دیف‌سازی چندگانه در جدول ۱ آورده شده است.

**نتیجه‌گیری:** bFGF یا FGF-2 به عنوان یک عامل دارویی با اهمیت برای درمان بافت آسیب‌دیده، التیام زخم‌های پوستی و بازبانی بیماری‌های زوال سیستم عصبی مورد استفاده قرار می‌گیرد. بنابراین، تولید تجاری آن در سیستم‌های بیانی به ویژه گیاهان از اهمیت بالایی برخوردار است به ویژه اینکه این پروتئین پایدار و آب‌دوست، فاقد جایگاه N-گلیکوزیلاسیون است. به همین منظور، ژن *hbFgf* در ناقل دوگانه pCAMBIA1304 درج و درون سه سویه اگروباکتریوم جهت انتقال به میزبان‌های گیاهی همسانه-سازی شد.

**سپاسگزاری:** نگارندگان بر خود لازم می‌دانند از آقای دکتر رضا حیدری جاپلقلی، دانشگاه بین‌المللی امام خمینی (ره) قزوین، گروه مهندسی بیوتکنولوژی به خاطر مطالعه متن مقاله حاضر و ارائه نظرهای ارزشمند سپاسگزاری نمایند.

## منابع

- بسطامی میثم، حسینی رامین (۱۳۹۷) بهینه‌سازی توالی کدکننده فاکتور رشد فیبروبلاستی بازی انسانی بمنظور بیان در گیاه برنج و همسانه‌سازی آن تحت کنترل توالی‌های تنظیم‌کننده ژن *ramy3D* در ناقل دوتایی pCAMBIA1304. مجله بیوتکنولوژی کشاورزی. ۱۰، ۲۰-۱.
- جعفری احمدآبادی سید علی اصغر، عسکری همت حشمت‌اله، محمدآبادی محمدرضا (۱۴۰۲) تاثیر شاهدانه بر بیان ژن *DLK1* در بافت قلب بره‌های کرمانی. مجله بیوتکنولوژی کشاورزی، ۱۵(۱)، ۲۱۷-۲۳۴.
- شکری سمیرا، خضری امین، محمدآبادی محمدرضا، خیرالدین حمید (۱۴۰۲) بررسی بیان ژن *MYH7* در بافت‌های ران، دست و راسته بره‌های پرواری نژاد کرمانی. مجله بیوتکنولوژی کشاورزی، ۱۵(۲)، ۲۱۷-۲۳۶.

محمدآبادی محمدرضا، گلکار افروز، عسکری حسینی مجید (۱۴۰۲) اثر رازیانه (*Foeniculum vulgare*) بر بیان ژن فاکتور ۱ رشد شبه انسولین در بافت شکمبه گوسفند کرمانی. مجله بیوتکنولوژی کشاورزی، ۱۵(۴)، ۲۳۹-۲۵۶.

## References

- An N, Ou J, Jiang D, Zhang L, Liu J, Fu K, et al. (2013) Expression of a functional recombinant human basic fibroblast growth factor from transgenic rice seeds. *Int J Mol Sci* 14, 3556-3567.
- Armelin HA (1973) Pituitary extracts and steroid hormones in the control of 3T3 cell growth. *Proc Natl Acad Sci USA* 70, 2702-2706.
- Barazandeh A, Mohammadabadi MR, Ghaderi-Zefrehei M, Nezamabadipour H (2016a) Predicting CpG islands and their relationship with genomic feature in cattle by hidden markov model algorithm. *Iran J Appl Anim Sci* 6 (3), 571-579 .
- Barazandeh A, Mohammadabadi MR, Ghaderi-Zefrehei M, Nezamabadipour H (2016b) Genome-wide analysis of CpG islands in some livestock genomes and their relationship with genomic features. *Czech J Anim Sci* 61, 487.
- Birnboim HC, Doly J (1979) A rapid alkaline procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7, 1513-1525.
- Bordbar F, Mohammadabadi M, Jensen J, et al. (2022) Identification of candidate genes regulating carcass depth and hind leg circumference in simmental beef cattle using Illumina Bovine Beadchip and next-generation sequencing. *Anim* 12 (9), e1103.
- Claverie JM, Notredame C (2007) *Bioinformatics for dummies* (2nd edn), New York, John Wiley and Sons.
- De Smit MH, Van Duin J (1990) Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proc Natl Acad Sci USA*, 87: 7668-7672.
- Ding S-H, Huang L-Y, Wang Y-D, Sun H-C, Xiang Z-H (2006) High-level expression of basic fibroblast growth factor in transgenic soybean seeds and characterization of its biological activity. *Biotechnol Lett* 28, 869-875.
- Dolivo DM (2022) Anti-fibrotic effects of pharmacologic FGF-2: a review of recent literature. *J Mol Med* 100(6), 847-860.
- Erdős G, Dosztányi Z (2020) Analyzing protein disorder with IUPred2A. *Curr Protoc Bioinformatics* 70(1), e99.
- Farooq M, Khan AW, Kim MS, Choi S (2021) The role of fibroblast growth factor (FGF) signaling in tissue repair and regeneration. *Cell* 10(11), 3242.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed), *The proteomics protocols handbook*, Totowa, Humana Press, pp. 571-607.

- Hayward S, Milner-White EJ (2021) Determination of amino acids that favor the  $\alpha$ L region using Ramachandran propensity plots. Implications for  $\alpha$ -sheet as the possible amyloid intermediate. *J Struct Biol* 213(2), 107738.
- Helke A, Geisen RM, Vollmer M, Sprengart ML, Fuchs E (1993) An unstructured mRNA region and a 5' hairpin represent important elements of the *E. coli* translation initiation signal determined by using the bacteriophage T7 gene 1 translation start site. *Nucleic Acid Res* 21, 5705-5711.
- Hofgen R, Willmitzer L (1988) Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acid Res* 16(20), 9877.
- Jafari Ahmadabadi SAA, Askari-Hemmat H, Mohammadabadi M, et al. (2023) The effect of Cannabis seed on DLK1 gene expression in heart tissue of Kermani lambs. *Agric Biotechnol J* 15 (1), 217-234 (In Persian).
- Jyothishwaran G, Kotresha D, Selvaraj T, Srideshikan SM, Rajvanshi PK, Jayabaskaran C (2007) A modified freeze-thaw method for efficient transformation of *Agrobacterium tumefaciens*. *Curr Sci* 93, 770-772.
- Kim YJ, Shim JS, Krishna PR, Kim SY, In JG, Kim MK, et al. (2008) Isolation and characterization of a glutaredoxin gene from *Panax ginseng* CA Meyer. *Plant Mol Biol Reporter* 26(4), 335-349.
- Kwong KWY, Ng KL, Lam CC, Wang YY, Wong WKR (2013) Authentic human basic fibroblast growth factor produced by secretion in *Bacillus subtilis*. *Appl Microbi Biotechnol* 97, 6803-6811.
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157, 105-132.
- Lingg N, Cserjan-Puschmann M, Fischer A, Engele P, Kröb C, Schneider R, et al. (2021) Advanced purification platform using circularly permuted caspase-2 for affinity fusion-tag removal to produce native fibroblast growth factor 2. *J Chem Technol Biotechnol* 96(6), 1515-1522.
- Lundell DL, Narula SK (1994) Structural elements required for receptor recognition of human interferon-gamma. *Pharmacol Ther* 64, 1-21.
- Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci USA* 101(19), 7287-7292.



- Mészáros B, Erdős G, Dosztányi Z (2018) IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acid Res* 46(W1), W329-W337.
- Mohammadabadi M, Golkar A, Askari Hesni M (2023) The effect of fennel (*Foeniculum vulgare*) on insulin-like growth factor 1 gene expression in the rumen tissue of Kermani sheep. *Agric Biotechnol J* 15 (4), 239-256 (In Persian).
- Mohammadabadi M, Masoudzadeh SH, Khezri A, et al. (2021) Fennel (*Foeniculum vulgare*) seed powder increases Delta-Like Non-Canonical Notch Ligand 1 gene expression in testis, liver, and humeral muscle tissues of growing lambs. *Heliyon* 7 (12), e08542
- Mohammadinejad F, Mohammadabadi M, Roudbari Z, Sadkowski T (2022) Identification of key genes and biological pathways associated with skeletal muscle maturation and hypertrophy in *Bos taurus*, *Ovis aries*, and *Sus scrofa*. *Anim* 12 (24), 3471.
- Ornitz DM, Itoh N (2015) The fibroblast growth factor signaling pathway. *Wiley Interdiscip Rev: Dev Biol* 4, 215-266.
- Poudel SB, Min CK, Lee JH, Shin YJ, Kwon TH, Jeon YM, et al. (2019) Local supplementation with plant-derived recombinant human FGF2 protein enhances bone formation in critical-sized calvarial defects. *J Bone Miner Metab* 37, 900-912.
- Proctor JR, Meyer IM (2013) CoFold: an RNA secondary structure prediction method that takes co-transcriptional folding into account. *Nucleic Acid Res* 41, e102.
- Ramachandran GN, Sasisekharan V (1968) Conformation of polypeptides and proteins. *Adv Protein Chem* 23, 283-437.
- Razaghi A, Tan E, Lua L, Owens L, Karthikeyan OP, Heimann K (2016) Is *Pichia pastoris* a realistic platform for industrial production of recombinant human interferon gamma? *Biologic* 45, 52-60.
- Russell RB, Breed J, Barton GJ (1992) Conservation analysis and structure prediction of the SH<sub>2</sub> family of phosphotyrosine binding domains. *FEBS Lett* 304(1), 15-20.
- Saadatabadi LM, Mohammadabadi M, Ghanatsaman ZA, et al. (2023) Data of whole-genome sequencing of Karakul, Zel, and Kermani sheep breeds. *BMC Res Not* 16 (1), e353.
- Safaei SMH, Dadpasand M, Mohammadabadi M, et al. (2022) An *Origanum majorana* leaf diet influences myogenin gene expression, performance, and carcass characteristics in lambs. *Anim* 13 (1), e14.
- Shahsavari M, Mohammadabadi M, Khezri A, et al. (2022) Effect of fennel (*Foeniculum Vulgare*) seed powder consumption on insulin-like growth factor 1 gene expression in the liver tissue of growing lambs. *Gene Express* 21 (2), 21-26.

- Shokri S, Khezri A, Mohammadabadi M, Kheyroodin H (2023). The expression of MYH7 gene in femur, humeral muscle and back muscle tissues of fattening lambs of the Kermani breed. *Agric Biotechnol J* 15 (2), 217-236 (In Persian).
- Thompson SA (1992) The disulfide structure of bovine pituitary basic fibroblast growth factor. *J Biol Chem* 267, 2269-2273.
- Wang Y-P, Wei Z-Y, Zhong X-F, Lin C-J, Cai Y-H, Ma J (2016) Stable expression of basic fibroblast growth factor in chloroplasts of tobacco. *Int J Mol Sci* 17, 19.
- Wu X, Kamei K, Sato H, Sato S, Takano R, Ichida M, et al. (2001) High-level expression of human acidic fibroblast growth factor and basic fibroblast growth factor in silkworm (*Bombyx mori* L.) using recombinant baculovirus. *Protein Expr Purif* 21, 192-200.
- Zhang W, Xiao W, Wei H, Zhang J, Tian Z (2006) mRNA secondary structure at start AUG codon is a key limiting factor for human protein expression in *Escherichia coli*. *Biochem Biophys Res Commun* 349, 69-78.