


Research Paper

Bioinformatics Analysis of Flavanone-3'-hydroxylase (F3'H) in *Silybum marianum*

Hoda Alsadat Aghili¹, Azim Ghasemnezhad² , Hossein Moradi³ and Valiollah Ghasemi Omran⁴

1- Horticulture Science, Plant Production, Gorgan University of Agricultural Science and Natural Resources, Gorgan, Iran

2- Horticulture Science, Plant Production, Gorgan University of Agricultural Science and Natural Resources, Gorgan, Iran, (Corresponding author: ghasemnezhad@gau.ac.ir)

3- Horticulture Science, Crop Science, Sari University of Agricultural Science and Natural Resources, Sari, Iran

4- Crop Science Genetic & Agricultural Biotechnology Institute of Tabarestan, Sari University Agricultural Science and Natural Resources, Sari, Iran

Received: 04 June, 2025

Revised: 10 October, 2025

Accepted: 14 November, 2025

Extended Abstract

Background: Medicinal plants play a crucial role in human nutrition and the treatment of numerous diseases due to their ability to produce secondary metabolites with therapeutic properties. *Silybum marianum* L., a medicinal plant belonging to the Asteraceae family, contains silymarin as its principal bioactive compound, which is widely used in the treatment of liver-related disorders. A comprehensive understanding of the molecular pathways and the genes involved in the biosynthesis of this compound is essential for its improvement and enhanced production. Accordingly, functional analysis of the genes and proteins participating in the silymarin biosynthetic pathway holds significant importance. Flavonoid 3'-hydroxylase (F3'H) is one of the key enzymes involved in this pathway. Among the various approaches available for gene and protein function analysis, bioinformatics methods offer a powerful and efficient means of investigation. Therefore, a bioinformatics analysis of the nucleotide and protein sequences of the flavonoid 3'-hydroxylase (F3'H) gene was conducted in the present study.

Methods: A previously published article reporting the nucleotide sequence of this gene was utilized for the bioinformatics investigation of the nucleotide sequence of the flavonoid3-hydroxylase (F3'H) gene in the milk thistle plant (*S. marianum*). The mRNA sequence was retrieved from the NCBI database with the accession number KP861882.1. The nucleotide sequence analysis was performed using the ppugibo computational database. Additionally, the Plantcare website was employed to analyze the promoter of the gene of interest. The protein sequence of the gene of interest was obtained from the NCBI database under the accession number ALA39990.1, and information regarding the flavonoid3-hydroxylase protein from the milk thistle plant was examined through the Uniprot database with the accession number A0A0K2F2U. Furthermore, the protein domain was identified using the NCBI and InterPro databases. The physicochemical properties of the protein were analyzed using the ProtParam database. Motifs were identified from the meme database, and the functionality of these motifs was assessed through the elm database. The modeling of the secondary and tertiary structures of the F3'H protein was conducted using the phyre2 database, and the accessibility of the F3'H protein bases to solvents was predicted with the I-TASSER server. Finally, the quality of the stereochemical structure of the modeled F3'H protein was evaluated with the PROCHECK software using the Ramachandran plot.

Results: The investigations carried out in the Materials and Methods section indicate that this gene consists of 1557 nucleotides, with a C-G content of 51.87%. From the perspective of mRNA, it comprises 1317 base pairs that encode a protein with an amino acid length of 349. In the promoter sequence of the F3'H gene, 19 regulatory elements were predicted, with the TATA-box and CAAT-box being the most prominent regulatory elements in the promoter region of this gene. Additionally, this gene contains several light-responsive regulatory elements. The physicochemical properties obtained for this protein, which has 349 amino acids and a molecular weight of 48.320 kDa, indicated an aliphatic index of 32.99 and a predicted isoelectric point of 7.87, with a GRAVY index of -0.059. It was determined that there are 49 negatively charged amino acids (aspartic acid and glutamic acid) and 50 positively charged

amino acids (arginine and lysine). Furthermore, the extinction coefficient of this protein was found to be 80,330 at 280 nm in an aqueous environment, with an instability index of 34.65. Evaluations also revealed that this protein contains one domain from the P450 superfamily and four identifiable motifs. The structure of this protein consists of 45% alpha helices, 5% beta sheets, and 4% membrane helices, and the corresponding ligand-binding site was predicted. The analysis of the Ramachandran plot for the predicted model indicated that 87.5% of the amino acids fell within the favored region, 11.7% in the allowed region, 0.3% in the nearly allowed region, and 0.5% in the disallowed region. Moreover, the highest density of permissible points corresponds to right-handed alpha helices and beta sheets. The combined favored and allowed regions amounted to 99.2, thus rendering the model presented for the F3'H protein of milk thistle acceptable in this study.

Conclusion: In the bioinformatics study of the flavanone-3-hydroxylase gene, regulatory elements related to light response, abscisic acid response, cold temperature response, zein metabolism, and flavonoid biosynthesis pathways were identified in the promoter of this gene. Additionally, the bioinformatics analysis of the flavanone-3-hydroxylase protein demonstrates that this protein is positioned among relatively stable proteins. Furthermore, a domain of the P450 superfamily was identified in this protein, and the predicted model of the protein, with 45% alpha helices, 5% beta sheets, 4% membrane helices, and a corresponding ligand-binding site, also exhibits a favorable conformation. Considering the importance of silymarin in the treatment of liver diseases, improving and enhancing the production of this vital bioactive compound can lead to the advancement of pharmaceutical product quality and increased efficiency in medicinal plant agriculture. Therefore, the findings of this study have significant practical applications in biotechnology and the genetic improvement of *S. marianum* aimed at increasing silymarin content.

Keywords: Motif, Promoter regulatory elements, Protein, Ramachandran diagram, Silymarin

How to Cite This Article: Aghili, H. A., Ghasemnezhad, A., Moradi, H., & Ghasemi Omran, V. (2026). Bioinformatics Analysis of Flavanone-3'-hydroxylase (F3'H) in *Silybum marianum*. *J Crop Breed*, 18(1), 149-162. DOI: 10.61882/jcb.2026.1582



مقاله پژوهشی

بررسی بیوانفورماتیکی فلاونوئید ۳' هیدروکسیلاز (F3'H) در گیاه خارمریم (*Silybum marianum* L.)هدی السادات عقیلی^۱، عظیم قاسم نژاد^۲، حسین مرادی^۳ و ولی اله قاسمی عمران^۴

۱- گروه علوم باغبانی، دانشکده تولید گیاهی، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران
 ۲- گروه علوم باغبانی، دانشکده تولید گیاهی، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران، (نویسنده مسؤل: ghasemnezhad@gau.ac.ir)
 ۳- گروه علوم باغبانی، دانشکده علوم زراعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران
 ۴- پژوهشکده ژنتیک و زیست فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران

تاریخ پذیرش: ۱۴۰۴/۰۸/۲۳

تاریخ ویرایش: ۱۴۰۴/۰۷/۱۸
صفحه: ۱۴۹ تا ۱۶۲

تاریخ دریافت: ۱۴۰۴/۰۲/۱۴

چکیده مبسوط

مقدمه و هدف: گیاهان دارویی به دلیل تولید متابولیت‌های ثانویه با اثرات دارویی، اهمیت زیادی در تغذیه انسان و درمان بسیاری از بیماری‌ها دارند. گیاه دارویی خارمریم (*Silybum marianum* L.) متعلق به خانواده Asteraceae، دارای ماده موثره مهم سیلی‌مارین است که در درمان بیماری‌های کبدی مورد استفاده قرار می‌گیرد. شناخت دقیق مسیرهای مولکولی و ژن‌های دخیل در تولید این ترکیب برای بهبود و افزایش تولید آن اهمیت ویژه‌ای دارد. بر این اساس جهت آنالیز عملکرد ژن‌ها و پروتئین‌های مسیر تولید سیلی‌مارین حائز اهمیت است. فلاونوئید ۳' هیدروکسیلاز (F3'H) یکی از آنزیم‌های درگیر در مسیر بیوسنتز سیلی‌مارین است. یکی از روش‌های تحلیل عملکرد ژن‌ها و پروتئین‌ها مطالعه بیوانفورماتیکی آنها است. بنابراین، در این پژوهش به بررسی بیوانفورماتیکی توالی نوکلئوتیدی و پروتئینی ژن فلاونوئید ۳' هیدروکسیلاز (F3'H) پرداخته شد.

مواد و روش‌ها: برای بررسی بیوانفورماتیکی توالی نوکلئوتیدی ژن فلاونوئید ۳' هیدروکسیلاز (F3'H) در گیاه خارمریم (*S. marianum*)، از مقاله‌ای که توالی نوکلئوتیدی این ژن در آن گزارش شد، استفاده شد. به این منظور توالی mRNA آن از پایگاه داده NCBI به شماره دسترسی KP861882.1 دریافت شد. جهت آنالیز توالی نوکلئوتیدی از پایگاه محاسباتی ppuigbo استفاده شد. به منظور آنالیز پروموتور ژن مورد نظر از سایت Plantcare استفاده شد. از پایگاه داده NCBI توالی پروتئینی ژن مورد نظر به شماره دسترسی ALA39990.1 به دست آمد و از پایگاه داده uniprot اطلاعات مربوط به پروتئین فلاونوئید ۳' هیدروکسیلاز گیاه خارمریم به شماره دسترسی A0A0K2F2U بررسی شد. همچنین، دومین پروتئین از پایگاه داده NCBI و interpro شناسایی شد. با استفاده از پایگاه داده ProtParam خواص فیزیکی‌شیمیایی آن بررسی شد. موتیف‌ها از پایگاه داده meme و عملکرد موتیف‌های مورد بررسی از پایگاه داده elm ارزیابی شدند. مدل‌سازی ساختار دوم و سوم پروتئین F3'H با پایگاه داده phyre2 و پیش‌بینی میزان دسترسی بنیان‌های پروتئین F3'H به حلال با سرور I-TASSE انجام شد. در پایان، کیفیت ساختار استریوشیمی پروتئین F3'H مدل‌سازی شده با سرور phyre2، توسط نقشه رامچاندرا با نرم‌افزار PROCHECK ارزیابی شد.

یافته‌ها: طبق بررسی‌های انجام‌شده از مجموعه مراحل یادشده در بخش مواد و روش‌ها، این ژن دارای ۱۵۵۷ نوکلئوتید، با ۵۱/۸۷ درصد سیتوزین-گوانین است. از نقطه‌نظر mRNA، mRNA آن از ۱۳۱۷ جفت باز تشکیل شده است که پروتئینی را کد می‌کند که طول اسیدآمینه آن ۳۴۹ است. در توالی پروموتوری ژن F3'H، ۲۴ عنصر تنظیمی پیش‌بینی شد که TATA-box و CAAT-box بیشترین عنصر تنظیمی ناحیه پروموتوری این ژن بودند؛ همچنین، این ژن دارای چندین عنصر تنظیمی پاسخ به نور بود. خواص فیزیکی‌شیمیایی به دست آمده برای این پروتئین با ۳۴۹ اسیدآمینه و وزن مولکولی ۴۸/۳۲۰ کیلودالتون نشان داد که شاخص آلفاتیکی ۹۹/۳۲ و نقطه ایزوالکتریک پیش‌بینی شده ۷/۸۷، شاخص GRAVY ۰/۰۵۹- بود. تعداد ۴۹ اسیدآمینه (آسپاراتیک اسید و گلوتامیک اسید) دارای بار منفی و ۵۰ اسیدآمینه دارای بار مثبت (آرژینین و لیزین) تعیین شد. همچنین، در بررسی ضریب خاموشی این پروتئین مشخص شد که ضریب خاموشی پروتئین یادشده در ۲۸۰ nm در محیط آب برابر ۸۰۳۳۰ بود و شاخص ناپایداری آن ۳۴/۶۵ است. مطابق با ارزیابی‌های صورت گرفته، برای این پروتئین یک دومین از سوپر خانواده P450 و چهار موتیف قابل شناسایی هستند. در ساختار این پروتئین، ۴۵٪ ماریپیچ آلفا، ۵٪ صفحه بتا، و ۴٪ ماریپیچ‌های غشایی وجود دارند و محل اتصال لیگاند مربوط نیز پیش‌بینی شد. بررسی نمودار رامچاندرا برای مدل پیش‌بینی شده نشان داد که ۸۷٪ اسیدهای آمینه در ناحیه مطلوب، ۱۱٪ در ناحیه مجاز، ۳٪ در ناحیه تقریباً مجاز و ۵٪ در ناحیه غیر مجاز قرار داشتند. علاوه بر این، بیشترین تراکم نقاط مجاز با ماریپیچ آلفای راستگردان و صفحات بتا مطابقت دارد. مجموع ناحیه مطلوب و مجاز ۹۹/۲٪ می‌شود؛ بنابراین، مدل ارائه شده برای پروتئین F3'H گیاه خارمریم در این مطالعه قابل قبول است.

نتیجه‌گیری: در مطالعه بیوانفورماتیکی ژن فلاونوئید ۳' هیدروکسیلاز، عناصر تنظیمی مربوط به پاسخ به نور، پاسخ به آسبزیک اسید، پاسخ به دمای پایین، متابولیسم zein و مسیر بیوسنتز فلاونوئید در پروموتور این ژن شناسایی شدند. همچنین، بررسی بیوانفورماتیکی پروتئین فلاونوئید ۳' هیدروکسیلاز نشان داد که این پروتئین در ردیف پروتئین‌های غیر قطبی و نسبتاً پایدار قرار دارد. به علاوه، یک دومین از سوپر خانواده P450 و چهار موتیف در این پروتئین شناسایی شدند و مدل پیش‌بینی شده پروتئین با ۴۵٪ ماریپیچ آلفا، ۵٪ صفحه بتا، ۴٪ ماریپیچ‌های غشایی و محل اتصال لیگاند مربوط به "هم"، مدل نیز از وضعیت مطلوبی برخوردار بود. بنابراین، نتایج این مطالعه امکان درک بهتر مکانیزم‌های تنظیمی و عملکردی فلاونوئید ۳' هیدروکسیلاز را با بررسی ساختار ژنی، عناصر تنظیمی پروموتور، خصوصیات فیزیکی‌شیمیایی پروتئین F3'H، پیش‌بینی ساختار این پروتئین و محل اتصال لیگاند به F3'H را فراهم می‌آورد. با توجه به اهمیت سیلی‌مارین در درمان بیماری‌های کبدی، بهبود و افزایش تولید این ماده موثره مهم می‌تواند به ارتقای کیفیت محصولات دارویی و افزایش بهره‌وری در کشاورزی دارویی منجر شود. بنابراین، یافته‌های این مطالعه کاربردهای عملی مهمی در بیوتکنولوژی و اصلاح گیاه خارمریم به‌منظور افزایش سیلی‌مارین دارند.

واژه‌های کلیدی: سیلی‌مارین، عناصر تنظیمی پروموتور، پروتئین، موتیف، نمودار رامچاندرا

مقدمه

گیاهان حاوی منابع ارزشمندی از متابولیت‌های ثانویه دارای خواص دارویی هستند (Masoudi Jozchal et al., 2024). گیاه خارمریم با نام علمی *Silybum marianum* L. از

یکی از منابع بسیار ارزشمند منابع طبیعی ایران گیاهان دارویی آن است (Mohammadi Tabar et al., 2024). این

متابولیکی فراهم می‌آورد که می‌تواند در افزایش بهره‌وری و اصلاح ژنتیکی مورد استفاده قرار گیرد (Kholghi *et al.*, 2024).

هدف این مطالعه، بررسی بیوانفورماتیکی توالی نوکلئوتیدی و پروتئینی فلاونوئید^۳ هیدروکسیلاز (F3'H) در گیاه خارمریم، به منظور شناخت ساختار، عناصر تنظیمی و مدل‌سازی سه‌بعدی پروتئین آن است تا با فراهم آوردن اطلاعات بنیادی، زمینه را برای اصلاح ژنتیکی هدفمند و بهینه‌سازی تولید سیلی‌مارین در این گیاه دارویی فراهم سازد. بنابر این، در مطالعه حاضر بررسی بیوانفورماتیکی توالی نوکلئوتیدی و پروتئینی این ژن انجام شده است.

مواد و روش‌ها

برای بررسی بیوانفورماتیکی ژن فلاونوئید^۳ هیدروکسیلاز (F3'H) در گیاه خارمریم (*S. marianum* L.)، توالی نوکلئوتیدی آن، که قبلاً توسط گائو و همکاران (Gao *et al.*, 2020) گزارش شد، مورد استفاده قرار گرفت (Gao *et al.*, 2020). توالی mRNA آن از پایگاه داده NCBI به شماره دسترسی KP861882.1 دریافت شد (<https://www.ncbi.nlm.nih.gov/nucleotide/KP861882.1>). آنالیز توالی نوکلئوتیدی با پایگاه محاسباتی ppuigbo به نشانی <http://ppuigbo.me/programs/CAIcal> انجام شد. همچنین، به منظور آنالیز پروموتور از سایت Plantcare استفاده شد. در ادامه، توالی پروتئینی آن از پایگاه داده NCBI به شماره دسترسی ALA39990.1 به دست آمد و از پایگاه داده uniprot اطلاعات مربوط به پروتئین فلاونوئید^۳ هیدروکسیلاز گیاه خارمریم به شماره دسترسی A0A0K2F2U به نشانی <https://www.uniprot.org/s> بررسی شد. دومین پروتئین از پایگاه داده NCBI به نشانی <https://www.ncbi.nlm.nih.gov/interpro> و <https://www.ebi.ac.uk/interpro> شناسایی شد. با استفاده از پایگاه داده ProtParam، خواص فیزیوشیمیایی آن به نشانی <https://web.expasy.org/cgi-bin/protparam/protparam> بررسی شد. موتیف‌ها و عملکردشان از پایگاه داده meme به نشانی <https://meme-suite.org/meme> و <http://elm.eu.org> به نشانی <http://elm.eu.org> بررسی شد.

مدل‌سازی ساختار دوم و سوم پروتئین F3'H با پایگاه داده phyre2 به نشانی <http://www.sbg.bio.ic.ac.uk/phyre2> و پیش‌بینی میزان دسترسی بنیان‌های پروتئین F3'H به حلال با سرور I-TASSE به نشانی <https://zhanggroup.org/I-TASSE/> و TASSER انجام شد. در پایان، کیفیت ساختار استریوشیمی پروتئین F3'H مدل‌سازی شده با سرور phyre2 توسط نقشه رامانچاندرا با نرم‌افزار PROCHECK ارزیابی شد.

نتایج و بحث

بررسی‌ها نشان دادند که طول این ژن ۱۵۵۷ نوکلئوتید است (Gao *et al.*, 2020). درصد سیتوزین گوانین این ژن توسط پایگاه محاسباتی ppuigbo ۵۱/۸۷ درصد محاسبه شد (جدول ۱). همچنین، مطالعه توالی نوکلئوتیدی F3'H در

خانواده آفتابگردان (Asteraceae) است و ماده موثره مهم آن سیلی‌مارین (Silymarin) جهت مصرف مورد تأیید سازمان بهداشت جهانی بر اساس داده‌های بالینی این ماده، برای درمان حمایتی هپاتیت حاد و مزمن، سیروز ناشی از الکل، دارو و سم کاربرد دارد. مصرف مورد تأیید کمیسیون E (Commission E) آن در درمان مشکلات سوء هاضمه، سمیت کبدی، درمان حمایتی بیماری‌های التهابی کبد و سیروز کبدی است (Khalighi- Sigaroodi *et al.*, 2013). همچنین، طبق فارماکوپه گیاهی ایران در درمان مسمومیت‌های کبدی ناشی از سموم متابولیک مانند الکل و سموم قارچ آمانیتا، اختلالات کبدی پس از هپاتیت، نارسایی‌های مزمن کبدی مانند سیروز کبدی، کبد چرب و بیماری‌های خفیف استفاده می‌شود (Committee for Compilation of Herbal Pharmacopoeia of Iran., 2002).

سیلی‌مارین در پوشش بیرونی بذرها (آکن بذرها) تجمع پیدا می‌کند (Torres & Corchete, 2016). ساختار آن متشکل از فلاولیگنان‌های تاکسی‌فولین، سیلی‌بین، ایزوسیلی‌بین، سیلی‌دیانین، و سیلی‌کریستین است که ایزومر یکدیگر هستند (EISayed *et al.*, 2019). فلاونوئید^۳ هیدروکسیلاز، که فلاونول^۳ هیدروکسیلاز و فلاونون^۳ هیدروکسیلاز (F3'H) هم نامیده می‌شود، یکی از آنزیم‌های مهم در مسیر تولید سیلی‌مارین است (Drouet *et al.*, 2016; Torres & Corchete, 2016). بیوستز سیلی‌مارین از مسیر فنیل پروپانوئید شروع می‌شود. به این ترتیب که نخست فنیل‌آلانین تحت تاثیر آنزیم PAL (فنیل‌آلانین آمونیا لیاز) به ۴-کوماریل کوآنزیم A تبدیل می‌شود. در ادامه، ۴-کوماریل کوآنزیم A هم وارد مسیر فلاونوئیدی و هم مسیر مونولیگنان می‌گردد. در مسیر فلاونوئیدی، آنزیم‌های چالکون سینتاز (CHS) و چالکون ایزومراز (CHI) در دو مرحله، سه مولکول مالونیل کوآنزیم A را با یک ۴-کوماریل کوآنزیم A ترکیب و تولید فلاونون نارنجین می‌کنند. فلاونون نارنجین تحت فعالیت آنزیم فلاونون ۳ هیدروکسیلاز (F3'H) اکسید شده، هیدروکسی کامفرول تولید می‌شود. در پایان، این ترکیب با آنزیم فلاونوئید^۳ هیدروکسیلاز (F3'H) هیدروکسیله می‌شود و هیدروفلانول تاکسی‌فولین تولید می‌کند. این ترکیب به‌عنوان پیش‌ساز بخش فلاونوئیدی ترکیب سیلی‌مارین از اهمیت ویژه‌ای برخوردار است (Torres & Corchete, 2016). از آنجایی که فلاونوئید^۳ هیدروکسیلاز (F3'H) یکی از ژن‌های منتخب مهم در مسیر بیوستز سیلی‌مارین است، می‌توان با بررسی‌های بیوانفورماتیکی ساختار ژن‌ها و پروتئین‌ها و پیش‌بینی ساختارهای پروتئین‌ها در پیش‌بینی موقعیت سلولی پروتئین، بازسازی مسیرهای متابولیکی و تحلیل عملکردی آنها استفاده نمود (Khakpour *et al.*, 2019).

بیوانفورماتیک به‌عنوان رویکردی نوین، ابزاری برای مطالعه ژنی و پروتئینی در گیاهان، مانند مقایسه توالی، بررسی خانواده‌های ژنی و پروتئینی، همچنین مدل‌سازی پروتئین می‌باشد که نتایج آن درک عمیق‌تری را از مکانیسم‌های مولکولی، شناسایی ژن‌ها، عملکردهای احتمالی و مسیرهای

پایگاه داده NCBI نشان داد که این ژن، mRNA به طول ۱۳۱۷ جفت باز دارد که پروتئینی با ۳۴۹ اسید آمینه را کد می کند.

جدول ۱- توالی نوکلئوتیدی F3'H

Table 1. The nucleotide sequence of F3'H

%C+T	%A+C	%A+T	%G+T	%G+A	%G+C	%G	%T	%C	%A	G	T	C	A	طول (Length)	توالی ها / پارامترها (Sequences \ Parameters)
48.94	48.62	49.13	51.38	51.06	50.87	26.65	24.73	24.21	24.41	415	385	377	380	1557	

MYC, MBSI, ABRE4, MYc, STRE, MYB, Unnamed_1 و ABRE3 کارکردی شناسایی نشد. عمده این عناصر تنظیمی مانند MYB, STRE, LTR و ABRE در پاسخ به تنش‌ها نقش دارند (Arab et al., 2021; Karimi Goyjelo & Abdollahi Mandoulakani, 2021). عنصر تنظیمی Myb در پاسخ به دمای پایین درگیر است (Heidari et al., 2022). همچنین، عنصر تنظیمی W box جایگاه اتصال عامل رونویسی WRKY است که برای فعالیت ژن در برابر تنش زیستی ضروری است (Zinati et al., 2018). عنصر پاسخ به نور G-box, در پاسخ به نور قرمز در ژن‌های تنظیم‌شونده توسط فیتوکروم A است (Mir Hami dehsardi et al., Drikvand et al., 2023). (2015;

بررسی نواحی پروموتوری ژن F3'H

نتایج بررسی نواحی پروموتوری ژن F3'H (جدول ۲)، ۲۴ عنصر تنظیمی پیش‌بینی شده شامل عناصر تنظیمی مربوط به پاسخ به نور، پاسخ به آبسزیک اسید، پاسخ به دمای پایین، متابولیسم zein و مسیر بیوسنتز فلاونوئید را مشخص نمودند. بیشترین عنصر تنظیمی ناحیه پروموتوری این ژن، شامل TATA-box و CAAT-box بودند. TATA-box توالی شروع رونویسی و CAAT-box توالی فزاینده هستند (Zinati et al., 2018). همچنین، بجز SP1, Myb, LTR, MYc, W box, DRE1, O2-site, GA-motif و MBSI. باقی عناصر تنظیمی در موقعیت‌های مختلف تکرار شده بودند. با این وجود، برای عناصر تنظیمی Unnamed_4, Myb, MYc, AT~TATA-box, DRE1, W box, G-box.

جدول ۲- توالی پروموتوری ژن F3'H همراه با توالی موتیف و عملکرد عناصر سیس

Table 2. The promoter sequence of the gene F3'H along with the motif sequence and function of cis-acting elements

کارکرد (Function)	Matrix score.	تعداد (Number)	توالی (Sequence)	عناصر سیس (cis elements)	ردیف
عنصر هسته پروموتور در -۳۰ شروع رونویسی (core promoter element around -30 of transcription start)	6	81	ATATAT	TATA-box	1
بخشی از عنصر پاسخ به نور (part of a light responsive element)	6	2	GATAGGG	TC-motif	2
عنصر پاسخ به نور (light responsive element)	4	2	CTCC	Unnamed_4	3
	6	1	GGGCGG	SP1	4
	5	23	TATATA	AT~TATA-box	5
		1	TAACGT	Myb	6
بخشی از ماژول حفاظت شده DNA در پاسخ به نور (part of a conserved DNA module involved in light responsiveness)	6	2	ATTAAT	BOX4	7
عنصر مشترک تنظیمی در نواحی پروموتور و تقویت کننده (common cis-acting element in promoter and enhancer region)	8	37	(CC)CAAT(TT)	CAAT-box	8
عنصر پاسخ به دمای پایین (cis-acting element involved in low-temperature responsiveness)	6	1	CCGAAA	LTR	9
	5	2	AGGGG	STRE	10
	6	4	(C/T)AACCA	MYB	11
	7	1	TCTCTTA	MYc	12
	6	4	TACGTG	G-box	
	6	2	CACGAC		13
	8	1	ACACGTGT		
عنصر تنظیمی در تنظیم متابولیسم زین (cis-acting regulatory element involved in zein metabolism regulation)	9	1	GATGACATGG	O2-site	14
عنصر پاسخ به نور (cis-acting regulatory element involved in light responsiveness)	6	4	CACGTA	ABRE4	15
عنصر پاسخ به آبسزیک اسید (cis-acting element involved in the abscisic acid responsiveness)	6	2	CACGTT	G-Box	16
	5	7	ACGTG	ABRE	17
	6	1	TTGACC	W box	18
	5	10	CATTG	MYC	19
	5	4	CGTGG	Unnamed_1	20
	7	1	ACCGAGA	DRE1	21
بخشی از عنصر پاسخ به نور (part of a light responsive element)	8	1	ATAGATAA	GA-motif	22
سایت اتصال به MYB در تنظیم ژن‌های بیوسنتزی فلاونوئید (MYB binding site involved in flavonoid biosynthetic genes regulation)	10.5	1	aaaAaaC(G/C)GTTA	MBSI	23
	6	4	TACGTG	ABRE3	24

: اطلاعات عملکردی در پایگاه Plantcare موجود نبود.

دیگر، آبسزیک اسید در گیاه خارمریم (Drouet *et al.*, 2020) شناسایی شد که با نتایج به دست آمده در این پژوهش هم‌خوانی دارد.

خواص فیزیکوشیمیایی پروتئین F3'H

خواص فیزیکوشیمیایی به دست آمده از پایگاه داده ProtParam، برای این پروتئین با ۳۴۹ اسیدآمینه و وزن مولکولی ۴۸/۳۲۰ کیلودالتون نشان داد که شاخص آلفاتیکی ۹۹/۳۲ و نقطه ایزوالکتریک پیش‌بینی شده ۷/۸۷. شاخص GRAVY ۰/۰۵۹- بود. همچنین، تعداد کل آمینواسیدها دارای بار منفی ۴۹ آمینواسید (آسپاراتیک اسید و گلوتامیک اسید) و ۵۰ آمینواسید دارای بار مثبت (آرژنین و لیزین) تعیین شد. ضریب خاموشی این پروتئین در ۲۸۰ nm در محیط آب برابر ۴۲۱۹۰ و شاخص ناپایداری ۳۴/۶۵ محاسبه شدند (جدول ۴). همچنین، بیشترین تعداد اسیدآمینه در ترکیب این پروتئین به ترتیب مربوط به لوسین با ۵۷ عدد (۱۳ درصد) و آلین ۳۳ عدد (۷/۲ درصد) و کمترین تعداد اسیدآمینه مربوط به سیستئین ۴ عدد (۰/۹ درصد) بودند (جدول ۳).

بیشتر پروموتورهای ژن‌های یوکاریوتی شامل عناصر تنظیمی TATA-box و CAAT-box هستند؛ این عناصر تنظیمی نقش مهمی در شروع رونویسی دارند به طوری که توالی CAAT-box توسط فاکتور سیگمای RNA پلیمرز II شناسایی می‌شود و پس از آن در ناحیه TATA-box، دو رشته DNA از هم باز شده، کمپلکس باز رونویسی تشکیل می‌گردد (Karimi Goyjelo & Abdollahi, 2021). یکی از بیشترین عناصر تنظیمی شناسایی شده در در ناحیه پروموتوری ژن F3'H عناصر تنظیمی پاسخ به نور بودند. با توجه به عناصر تنظیمی شناسایی شده در ناحیه پروموتوری ژن F3'H، کارکرد عناصر تنظیمی آن بیشتر در پاسخ به تنش‌های زیستی و غیر زیستی به ویژه پاسخ به نور است. بنابر این، بیان آنزیم F3'H هم مانند سایر ژن‌های مسیرهای بیوسنتزی متابولیت‌های ثانویه تحت تاثیر تنش‌های زیستی و غیر زیستی است (Zare Mehrjerdi, 2015). در سایر پژوهش‌ها هم در پروموتور این ژن عناصر تنظیمی پاسخ به نور و تنش‌های غیر زیستی (Baba & Ashraf, 2019; Zheng *et al* 2025) در گیاهان

جدول ۳- ترکیب پروتئین فلاونوئید^۳ هیدروکسیلاز (F3'H) توسط پایگاه داده ProtParam

درصد اسیدآمینه در ترکیب			تعداد اسیدآمینه در ترکیب		
نام اسید آمینه	تعداد اسیدآمینه در ترکیب	درصد اسیدآمینه در ترکیب	نام اسید آمینه	تعداد اسیدآمینه در ترکیب	درصد اسیدآمینه در ترکیب
Name of amino acid	Number of amino acids in the protein composition	Amino acid percentage in the protein composition	Name of amino acid	Number of amino acids in the protein composition	Amino acid percentage in the protein composition
Ala (A)	40	9.1%	Leu (L)	57	13%
Arg (R)	28	6.4%	Lys (K)	22	5%
Asn (N)	19	4.3%	Met (M)	11	2.5%
Asp (D)	28	6.4%	Phe (F)	18	4.1%
Cys (C)	4	0.9%	Pro (P)	25	5.7%
Gln (Q)	9	2.1%	Ser (S)	26	5.9%
Glu (E)	21	4.8%	Thr (T)	21	4.8%
Gly (G)	32	7.3%	Trp (W)	6	1.4%
His (H)	13	3%	Tyr (Y)	6	1.4%
Ile (I)	20	4.6%	Val (V)	33	7.2%

جدول ۴- ویژگی‌های فیزیکوشیمیایی پروتئین فلاونوئید^۳ هیدروکسیلاز (F3'H) از پایگاه داده ProtParam

اندازه پروتئین (aa)	وزن مولکولی (kD)	شاخص آلفاتیکی (Aliphatic index_AI)	نقطه ایزوالکتریک (Isoelectric points_pI)	شاخص GRAVY	شاخص ناپایداری (Instability index (II)	ضریب خاموشی (Extinction (coefficient)	کل آمینو اسیدهای دارای بار مثبت (Total number of positively charged residues)	کل آمینو اسیدهای دارای بار منفی (Total number of negatively charged residues)
349	48.320	99.32	7.87	-0.059	34.65	42190	50	49

پروتئین‌هایی با شاخص آلفاتیکی بالاتر از ۱۰۰ در دمای بالا ثبات ساختاری دارند (Arab *et al.*, 2021; Mir Drikvand, 2023). از آنجایی که شاخص آلفاتیکی پروتئین F3'H نزدیک به ۱۰۰ برآورد شده است، لذا این پروتئین احتمالاً در دمای بالا مقاومت حرارتی نسبی خواهد داشت.

شاخص ناپایداری کمتر از ۴۰ (۳۴/۶۵) محاسبه شد که می‌توان این پروتئین را پایدار محسوب کرد (Fallahzadeh Mamaghani, 2016).

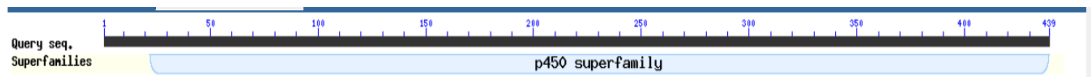
متوسط هیدروپاتی کلی (GRAVY) که نشان‌دهنده شاخص آبریزی پروتئین است اگر برای یک پروتئین منفی باشد، آن پروتئین غیر قطبی و در صورت مثبت بودن آن قطبی است (Fallah Ziarani & Safaezadeh, 2022). متوسط

نقطه ایزوالکتریک ۷/۸۷ برای این پروتئین بیانگر pH است که در آن پروتئین بار خود را از دست می‌دهد و در نتیجه در میدان الکتریکی با جریان مستقیم قادر به حرکت نیست. این شاخص برای انتخاب و بهینه‌سازی روش‌های خالص‌سازی پروتئین مانند کروماتوگرافی تبادل یونی و یا الکتروفورز ایزوالکتریک فوکوسینگ مفید است (Fallahzadeh Mamaghani, 2016).

شاخص آلفاتیکی برای پروتئین F3'H ۹۹/۳۲ محاسبه شد. این شاخص بیانگر حجم نسبی پروتئین است که با زنجیره‌های آلفاتیکی در آمینواسیدهای آلانین، والین، لوسین و ایزولوسین اشغال شده است. همچنین، این شاخص اثر مثبتی بر افزایش ثبات حرارتی پروتئین‌های کروی دارد، بنابر این، همبستگی مستقیم با پایداری حرارتی پروتئین دارد. لذا

غیر قطبی بودن این پروتئین است و با نتایج بدست آمده در این پژوهش همخوانی دارد (Ahmed, 2021). بررسی این پروتئین در پایگاه داده uniprot نشان داد که این پروتئین دارای یک محل اتصال "هم" یا یون آهن است. در پایگاه داده interpro و NCBI یک دومین حفاظت شده متعلق به سوپر فامیلی P450 در ناحیه ۲۲ تا ۴۳۹ برای این پروتئین شناسایی شد (شکل های ۱ و ۲).

هیدروپاتی کل پروتئین F3'H ۰/۰۵۹- محاسبه شد که بیانگر غیر قطبی بودن این پروتئین است. در بررسی مدل سازی همولوژی و آنالیز ساختاری ژن های فلاونون ۳-هیدروکسیلاز (F3H) و فلاونوئید ۳ هیدروکسیلاز (F3'H) در *Ginkgo biloba* (L.) نقطه ایزوالکتریک برای F3'H ۷/۷۱، شاخص آلیفاتیک ۸۸/۹۴ برآورد شد که نزدیک به برآوردهای F3'H در گیاه خارمریم بود و متوسط هیدروپاتی کل هم منفی محاسبه شد، که بیانگر



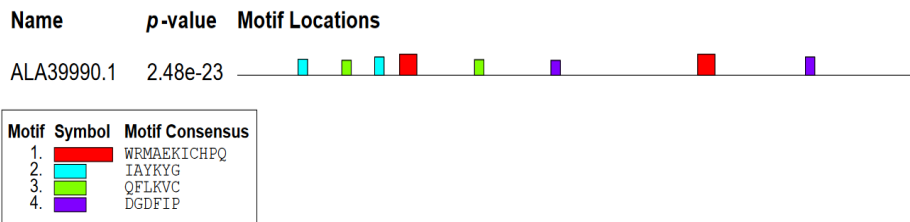
شکل ۱- دومین حفاظت شده پروتئین F3'H
Figure 1. Conserved domain of F3'H protein



شکل ۲- همترازی دومین حفاظت شده در پایگاه داده NCBI
Figure 2. The Sequence Alignment of the F3'H Protein Conserved Domain in the NCBI Database

۸۳_۶۸ سومین موتیف در موقعیت 4.5e+001 قرار دارد، در پایگاه داده meme برای این پروتئین چهار موتیف شناسایی کرد. یکی از این موتیف ها در موقعیت ۱۱۵_۱۰۵ (و ۱۵۸_۱۵۳) به طول ۶ اسیدآمین به E-value برابر ۱.1e+002، و چهارمین موتیف در موقعیت ۲۰۷_۲۰۲ (و ۳۷۰_۳۶۵) به طول شش اسیدآمین به E-value برابر 1.6e+002 شناسایی شده است (شکل های ۳ و ۴).

۳۰۶_۲۹۶ (و ۳۰۶_۲۹۶) به طول ۱۱ اسیدآمین به E-value برابر با 9.3e+000 قرار دارد. دومین موتیف در موقعیت ۴۵_۴۰ (و ۹۴_۸۹) به طول ۶ اسیدآمین به E-value برابر با



شکل ۳- موتیف های حفاظت شده پروتئین F3'H توسط پایگاه داده meme
Figure 3. The conserved protein motifs of the F3'H protein by the meme database



شکل ۴- موتیف لوگو توالی‌های پروتئین F3'H توسط پایگاه داده meme
Figure 4. The motifs logo of the F3'H protein sequences by the meme database

DEG_Nend_UBRbox_3 و DEG_Nend_UBRbox_2 دارای عملکرد یوبی کوئیتیناسیون و تخریب پروتئازومی هستند. به این ترتیب که موتیف N ترمینال، تخریب پروتئین را با اتصال به جعبه UBR از N-recognins آغاز می‌کند. این نوع عملکرد در هر چهار موتیف شناسایی شده وجود دارد.

پایگاه داده ELM برای موتیف‌های شناسایی شده پروتئین F3'H عملکردهایی را شناسایی کرده است که DEG_Nend_UBRbox_2، DEG_Nend_Nbox_1، DEG_Nend_UBRbox_3، DEG_CRBN_cyclicCter_1، LIG_14-3-، MOD_CAAXbox، 3_CanoR_1 و LIG_eIF4E_1 را شامل می‌شوند (جدول ۵).

جدول ۵- عملکردهای شناسایی شده و بخش سلولی موتیف‌های پروتئین F3'H توسط پایگاه داده ELM
Figure 5. The identified functions and cellular locations of protein motifs by the ELM database

شناسایی کلاس ELM ELM class identifier	بخش سلولی cellular location
DEG_Nend_Nbox_1	سیتوزول Cytosol
Nend_UBRbox_2	سیتوزول Cytosol
DEG_Nend_UBRbox_3	سیتوزول Cytosol
DEG_CRBN_cyclicCter_1	سیتوزول، غشای پلاسمایی، کمپلکس یوبی کوئیتین لیگاز Cul4A-RING E3 Cytosol, Plasma membrane, ubiquitin ligase complexes Cul4A-RING E3
LIG_14-3-3_CanoR_1	سیتوزول، غشای پلاسمایی Cytosol, Plasma membrane
MOD_CAAXbox	سیتوزول Cytosol
LIG_eIF4E_1	سیتوزول Cytosol

این پایگاه اطلاعاتی در موتیف دوم یک عملکرد DEG_CRBN_cyclicCter_1 را شناسایی کرد که موتیف دگرون امید حلقوی C ترمینال است که به دومین TBD در Cereblon متصل می‌شود. موتیف DEG_CRBN_cyclicCter_1 امیدهای حلقوی C ترمینال اسپار تیمید و بقایای آمینوگلو تار تیمید اولین موتیف‌های دگرون طبیعی هستند که توسط CRBN با ویژگی بالا شناخته می‌شوند. این اصلاح پس از ترجمه در نتیجه چرخه درون مولکولی باقیمانده‌های گلو تار تیمین یا اسپارژین و برش همزمان پپتید ایجاد می‌شود. نتیجه فعالیت آن یک پروتئین خرد شده است که اغلب فاقد عملکرد است. این خصوصیت به ویژه برای هر پروتئین چند دامنه‌ای که در آن ماژول‌ها با یکدیگر تعامل دارند قابل مشاهده است. در این نوع پروتئین‌ها، امیدهای

کاتابولیز پروتئین توسط سیستم یوبی کوئیتین - پروتئازوم انجام می‌شود که یک سیستم تخریب پروتئین در یوکاریوت‌ها است، به این ترتیب که پروتئین‌هایی که سیگنال‌های تخریب را نشان می‌دهند توسط لیگازهای یوبی کوئیتین شناسایی و پلی یوبی کوئیتیله می‌شوند که این فرایند منجر به تخریب پروتئازومی بعدی آنها می‌شود. زیرمجموعه‌ای از این سیگنال‌های تخریب را می‌توان در انتهای N پروتئین یا در پایانه N محصولات برش پروتئین مشاهده کرد. این موتیف‌های N ترمینال، که N-degron نامیده می‌شوند، توسط لیگازهای یوبی کوئیتین E3 شناخته شده به عنوان N-recognins شناخته می‌شوند، که برای تخریب پروتئازومی توسط پلی یوبی کوئیتیلایسون باقیمانده‌های Lys حائز اهمیت هستند (Bachmair et al., 1986).

یک یا دو سیستئین C ترمینال (CaaX) پروتئین های هدف است. این واکنش توسط سه آنزیم پروتئین فارنسیراز (PFT)، پروتئین ژرانیل ژرانیل ترانسفراز-۱ (PGGT) و Rab ژرانیل ژرانیل ترانسفراز (Rab/GGT/geranylgeranyltransferase II) کاتالیز می شود (Trinidad *et al.*, 2020). eIF4E یک پروتئین اتصال دهنده کلاهی است که در همکاری با پروتئین هایی مانند هلیکاز eIF4A و پروتئین داربست eIF4G به mRNA متصل می شود و امکان اتصال ریبوزومها و شروع ترجمه را فراهم می کند (Amorim *et al.*, 2018).

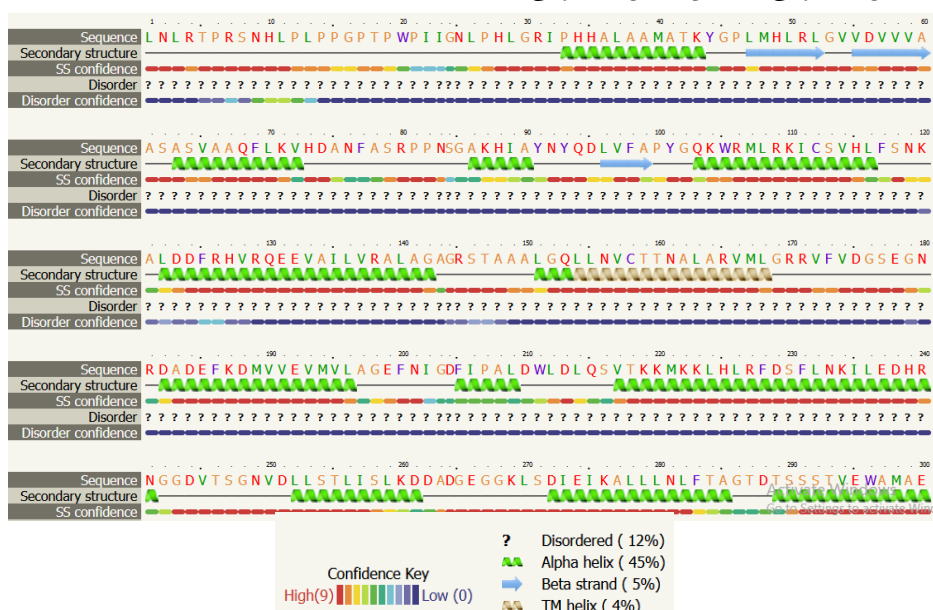
پیش بینی ساختار پروتئین

سرور phyre2 در ساختار دوم این پروتئین ۴۵ درصد ماریچ آلفا، ۵٪ صفحه بتا و ۴٪ ماریچ غشایی پیش بینی کرد (شکل ۵). در مدل سازی این پروتئین هم محل اتصال لیگاند مربوط به "هم" پیش بینی شد. این سرور با همتراز سازی مدل های مخفی مارکوف که دقت همتراز سازی را به صورت معنی داری افزایش می دهد، مدل سازی را انجام می دهد (Fallahzadeh Mamaghani, 2016).

حلقوی می توانند خود به خود ایجاد شوند. این اتفاق احتمالاً در پروتئین های پیر بیشتر خواهد بود. بنابر این، تشخیص آنها به عنوان مکانیزمی برای حذف موثر پروتئین های ناکارآمد و آسیب دیده اهمیت دارد (Friedrich *et al.*, 2019).

موتیف (14-3-3 binding) LIG_14-3-3_CanoR_1 phosphopeptide motif یک موتیف فسفره شده حاوی آرژنین استاندارد است که واسطه برهمکنش قوی با پروتئین های ۳-۳-۱۴ است. این موتیف متعارف دارای حداقل یک باقیمانده آرژنین است که در موقعیت ۳-، ۴- یا ۲- نسبت به باقیمانده فسفو قرار دارد و به قرارگیری پیتید در شیار اتصال کمک می کند. آرژنین با تشکیل پیوندهای هیدروژنی و همچنین یک پل نمک درون مولکولی با گروه فسفات، ترکیب فسفوپیتید را تثبیت می کند (Panni *et al.*, 2011).

موتیف MOD_CAAX box، موتیف پرنیلاسیون جعبه CAAX است که توسط فارنسیراز و ژرانیل ژرانیل ترانسفراز-۱ برای انتقال واحدهای چربی دوست به یک یا چند باقیمانده سیستئین C (Cys) ترمینال شناخته شده است. پرنیلاسیون یک اصلاح پس از ترجمه (PTM) در یوکاریوتی ها با حفظ شدگی بالا است. شامل اضافه شدن یک بخش ژرانیل ژرانیل ۱۵ کربنی یا ژرانیل ژرانیل ۲۰ کربنی به

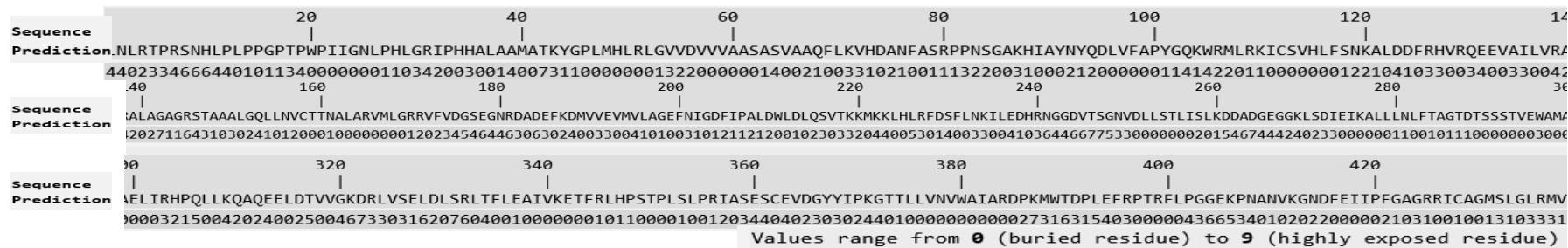


شکل ۵- پیش بینی ساختار دوم پروتئین F3'H با سرور phyre2

Figure 5. The secondary structure prediction of the F3'H protein by the phyre2 server

حلال قرار می گیرند، مقدار عددی ۹ و بنیان هایی که در درون پروتئین پنهان شده اند و در معرض حلال قرار نمی گیرند، مقدار عدد صفر داده می شود (Rezvannejad *et al.*, 2020).

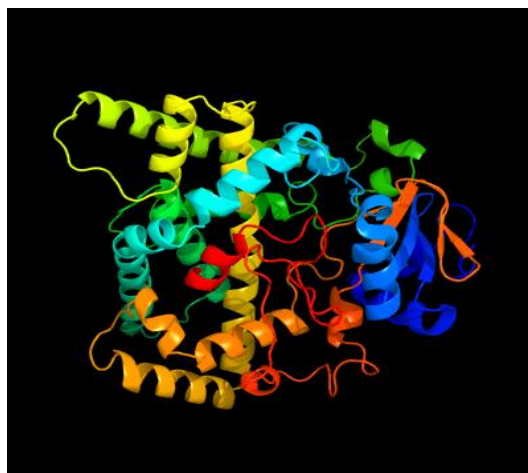
در این بررسی، میزان دسترسی بنیان های پروتئین F3'H به حلال با سرور I-TASSE پیش بینی شد. نتایج نشان می دهند که محدوده مقادیر بین صفر و ۹ است (شکل ۶). به بنیان هایی که کاملاً در سطح پروتئین هستند و در معرض



شکل ۶- پیش‌بینی میزان دسترسی بنیان‌های پروتئین F3'H به حلال با سرور I-TASSER
 Figure 6. The predicted solvent accessibility of the F3'H protein by the I-TASSER server

اطمینان و توسط الگوی با بالاترین امتیاز، مدل سازی کند (شکل ۷).

برای پیش بینی ساختار سوم، سرور phyre2 توانست ۴۱۹ باقیمانده اسید آمینه (۹۵٪ از دنباله مورد مطالعه) را با ۱۰۰٪



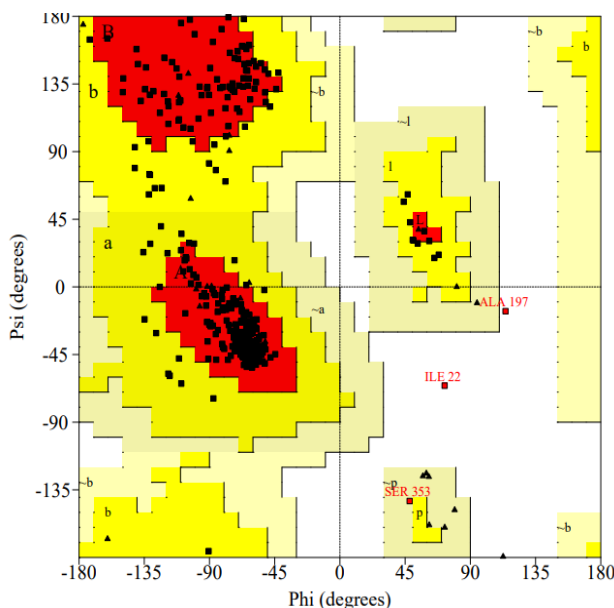
شکل ۷- پیش بینی ساختار سوم پروتئین F3'H با سرور phyre2
Figure 7. The 3D model prediction of the F3'H protein by the phyre2 server

چرخش داشته باشند که اتمها با یکدیگر تلاقی نکنند (Housaindokht & Chamani, 2022). پس اگر اتمها در فاصله ای نزدیکتر از آنچه شعاع واندروالسی آنها اجازه می دهد قرار گیرند، آنها در نواحی غیر مجاز نقشه مشخص می شوند. به جز گلاسیسین، به علت نداشتن زنجیره های جانبی، قرارگیری باقی اسید آمینه های از نظر برخورد فضایی در نواحی غیرمجاز ممکن نیست؛ از آنجایی که در گلاسیسین زنجیره جانبی هیدروژن است، پس مانند سایر اسید آمینه ها که محدودیت فضایی دارند، محدودیت ندارد و زوایای سای و فای می توانند مقادیر متعددی را داشته باشند (Housaindokht & Chamani, 2022; Ghorbanpour *et al.*, 2019). بنابر این، در تمام نقاط نقشه که برای سایر اسید آمینه ها محدودیت دارد، به خصوص نواحی مخصوص دورها در ساختار دوم، امکان استقرار دارد. مکان این اسید آمینه در نقشه حاصل از سرور PROCHECK با علامت مثلث مشخص شده است و در محاسبات آماری لحاظ نمی شود (Ghorbanpour *et al.*, 2019). در نمودار رامچاندران، محدوده مجاز زاویه چرخشی فای برای صفحات بتا (β _sheet) شامل $180^\circ < \phi < -45^\circ$ و $180^\circ < \psi < 360^\circ$ و برای زاویه چرخشی سای شامل $90^\circ < \psi < 180^\circ$ هستند؛ محدوده مجاز زاویه چرخشی فای برای مارپیچ آلفا (α _helix) $180^\circ < \phi < -45^\circ$ و $180^\circ < \psi < 360^\circ$ و برای زاویه چرخشی سای شامل $180^\circ < \phi < 180^\circ$ و $180^\circ < \psi < 360^\circ$ می شوند (Salehi Mojarrad *et al.*, 2018).

از نمودار رامچاندران برای ارزیابی کیفیت استرئوشیمیایی و مقادیر زوایای چرخشی پروتئین استفاده می شود (Bagheri *et al.*, 2019). نتایج بررسی نمودار رامچاندران مدل پیش بینی شده با سرور phyre 2 نشان می دهند که در مدل پیش بینی شده با سرور phyre 2 ۸۷/۵٪ اسیدهای آمینه در ناحیه مطلوب و ۱۱/۷٪ در ناحیه مجاز قرار گرفته اند. ۰/۳٪ در ناحیه تقریباً مجاز و ۰/۵٪ در ناحیه غیر مجاز قرار داشتند. همچنین، بیشترین تراکم نقاط مجاز با مارپیچ آلفای راستگردان و صفحات بتا مطابقت دارد (شکل ۸). مجموع ناحیه مطلوب و مجاز ۹۹/۲٪ می شود. بنابر این، می توان بیان داشت که مدل ارائه شده برای پروتئین F3'H گیاه خارمریم در این مطالعه قابل قبول است (Eftekharmanavi *et al.*, 2016).

زاویه چرخشی فای (ϕ) و زاویه چرخشی سای (ψ) زوایای چرخشی دو-وجهی هستند که برای تعیین وضعیت استقرار فضایی چارچوب یک رشته پلی پپتید به کار می روند (Miroliaei & Rastegari, 2009). در نمودار رامچاندران محور X زاویه چرخشی فای (ϕ) و محور Y زاویه چرخشی سای (ψ) را برای همه باقیمانده های ساختار پروتئین مورد بررسی نشان می دهد، که در آن زاویه فای نشان دهنده چرخش حول گروه N_{Ca} و زاویه سای چرخش حول گروه $C_{\alpha}C$ هستند. بنابر این، بسته به نوع زنجیره جانبی اسید آمینه، زوایای سای و فای می توانند در محدوده چرخش ۱۸۰ تا +۱۸۰ درجه قرار گیرند (Ghorbanpour *et al.*, 2019).

دو اتم می توانند فقط به اندازه شعاع واندروالسی به یکدیگر نزدیک شوند، بنابر این، زوایای سای و فای می توانند تا جایی



شکل ۸- نقشه رامچاندرا پروتئین F3'H مدل سازی شده با سرور phyre2 توسط سرور PROCHECK (رنگ قرمز: نواحی مطلوب (A, B, L)، رنگ زرد نواحی مجاز (a, b, l, p)، رنگ کرم نواحی تقریباً مجاز (~a, ~b, ~l, ~p) و رنگ سفید نواحی غیرمجاز) Figure 8. The Ramachandran plot of the F3'H protein modeled with the phyre2 server by the PROCHECK server (red color: regions favorable (A, B, L); Yellow color: additionally allowed regions (a, b, l, p); Cream color: generously allowed regions (~a, ~b, ~l, ~p) and white color: disallowed regions)

خاموشی این پروتئین در ۲۸۰ nm در محیط آب برابر ۸۰۳۳۰ و شاخص ناپایداری ۳۴/۶۵ محاسبه شدند. برای این پروتئین، یک دومین از سوپر خانواده P450 و چهار موتیف مشخص گردید. در ساختار این پروتئین، ۴۵٪ ماریپج آلفا، ۵٪ صفحه بتا، ۴٪ ماریپج های غشایی و محل اتصال لیگاند مربوط به هم پیش بینی شد. بررسی نمودار رامچاندرا برای مدل پیش بینی شده نشان داد که ۸۷/۵٪ اسیدهای آمینه در ناحیه مطلوب، ۱۱٪/۷ در ناحیه مجاز، ۳٪/۰ در ناحیه تقریباً مجاز و ۵٪/۰ در ناحیه غیر مجاز قرار داشتند. همچنین، بیشترین تراکم نقاط مجاز با ماریپج آلفای راستگردان و صفحات بتا مطابقت دارد. مجموع ناحیه مطلوب و مجاز ۹۹/۲٪ می شود. بنابر این، مدل ارائه شده برای پروتئین F3'H گیاه خارمریم در این مطالعه قابل قبول است. شناخت ساختار ژنی، عناصر تنظیمی پروموتور و خصوصیات فیزیوشیمیایی پروتئین F3'H، امکان درک بهتر مکانیزم های تنظیمی و عملکردی این آنزیم را فراهم می آورد. با توجه به کاربردهای گسترده سیلی مارین در درمان بیماری های کبدی، بهبود تولید آن می تواند به ارتقای کیفیت محصولات دارویی و افزایش بهره وری در کشاورزی دارویی منجر شود. بنابر این، یافته های این مطالعه کاربردهای عملی مهمی در حوزه بیوتکنولوژی و اصلاح نباتات دارند.

نتیجه گیری کلی

در گیاه خارمریم، ماده موثره مهم سیلی مارین، که در درمان بیماری های کبدی مورد استفاده قرار می گیرد، از مسیر فنیل پروپانوئید تولید می شود که فلاونوئید^۳ هیدروکسیلاز (F3'H) یکی از آنزیم های درگیر در مسیر بیوسنتز سیلی مارین است. این مطالعه نشان می دهد که این ژن دارای ۱۵۵۷ نوکلئوتید، با ۵۱/۸۷٪ سیتوزین-گوانین است. mRNA آن دارای ۱۳۱۷ جفت باز است که پروتئینی به طول ۳۴۹ اسیدآمینه را کد می کند. در توالی پروموتوری ژن F3'H، ۱۹ عنصر تنظیمی پیش بینی شد که شامل عناصر تنظیمی مربوط به پاسخ به نور، پاسخ به آسبیزیک اسید، پاسخ به دمای پایین، متابولیسم zein و مسیر بیوسنتز فلاونوئید است. همچنین، TATA-box و CAAT-box بیشترین عنصر تنظیمی ناحیه پروموتوری این ژن را شامل می شوند. خواص فیزیوشیمیایی به دست آمده از پایگاه داده ProtParam، برای این پروتئین با ۳۴۹ اسیدآمینه و وزن مولکولی ۴۸/۳۲۰ کیلودالتون نشان داد که شاخص آلفاتیگ ۹۹/۳۲ و نقطه ایزوالکتریک پیش بینی شده ۷/۸۷، شاخص GRAVY ۰/۰۵۹- است. همچنین، تعداد کل آمینواسیدها دارای بار منفی ۴۹ آمینواسید (آسپاراتیک اسید و گلوتامیک اسید) است و ۵۰ آمینواسید دارای بار مثبت (آرژنین و لیزین) ارزیابی شدند. ضریب

References

- Ahmed, M. Z. S. (2021). Homology Modeling and Structural Analysis of of the Flavanone 3-Hydroxylase (F3H) and Flavonoid 3hydroxylase (F3H) Genes from *Ginkgo Biloba* (L.). *Journal of Biotechnology and Biochemistr*, 7(1), 01-21. doi: 10.9790/264X-0701020121.
- Amorim, I. S., Lach, G., & Gkogkas, C. G. (2018). The role of the eukaryotic translation initiation factor 4E (eIF4E) in neuropsychiatric disorders. *Frontiers in Genetics*, 9, 561. doi: 10.3389/fgene.2018.00561.

- Arab, M., Najafi Zarrini, H., Nematzadeh, G., & Hashemi-petroudi, S. H. (2021). Bioinformatics analysis of calcium-dependent protein kinase (CPK) in *Aeluropus littoralis* L. *Crop Biotechnology*, 10(33), 1-19. doi: 10.30473/cb.2021.58258.1834. [In Persian]
- Baba, S. A., & Ashraf, N. (2019). Functional characterization of flavonoid 3'-hydroxylase, CsF3' H, from *Crocus sativus* L: Insights into substrate specificity and role in abiotic stress. *Archives of Biochemistry and Biophysics*, 667, 70-78. doi: 10.1016/j.abb.2019.04.012
- Bagheri, S., Fakheri, B. A., Ahadi, A. M., & Emamjomeh, A. (2019). Identification, cloning and bioinformatics analysis of the gene encoding NIa protein of wheat streak mosaic virus (WSMV). *Crop Biotechnology*, 9(27), 1-13. https://doi.org/10.30473/cb.2019.48896.1785. [In Persian]
- Committee for Compilation of Herbal Pharmacopoeia of Iran. (2002). Iranian Herbal Pharmacopoeia. *Ministry of health and medical education, Food and drug administration*. Tehran. [In Persian]
- Didelon, M., Khafif, M., Godiard, L., Barbacci, A., & Raffaele, S. (2020). Patterns of sequence and expression diversification associate members of the PADRE gene family with response to fungal pathogens. *Frontiers in Genetics*, 11, 491. doi: 10.3389/fgene.2020.00491
- Drouet, S., Tungmunthum, D., Lainé, É., & Hano, C. (2020). Gene expression analysis and metabolite profiling of silymarin biosynthesis during milk thistle (*Silybum marianum* (L.) Gaertn.) fruit ripening. *International Journal of Molecular Sciences*, 21(13), 4730. https://doi.org/10.3390/ijms21134730
- Eftekharnavi, S., Peyghan, R., Soltani, M., Ghorbanpour, M., & Ghalyanchilangeroudi, A. (2016). Multi-aspect molecular and structural bioinformatic study of outer membrane protein (ompts) of *aeromonas hydrophila* common cause of infectious septicemia in fish for immunization goals. *Journal of Animal Environment*, 8(1), 165_174. [In Persian]
- ElSayed, A. I., El-Hamahmy, M. A., Rafudeen, M. S., Mohamed, A. H., & Omar, A. A. (2019). The impact of drought stress on antioxidant responses and accumulation of flavonolignans in milk thistle (*Silybum marianum* (L.) Gaertn). *Plants*, 8(12), 611. doi: 10.3390/plants8120611.
- Fallah Ziarani, M., & Safaeizadeh, M. (2022). In Silico Analysis of Sucrose: Fructan 6-Fructose Transferase Gene in *Triticum aestivum* L. *Crop Biotechnology*, 11(38), 97-114. https://doi.org/10.30473/cb.2022.64450.1881. [In Persian]
- Fallah Ziarani, M., Tohidfar, M., Mirjalili, M. H., & Ahmadi Gavlighi, H. (2021). A bioinformatics analysis of antibacterial peptides in five species of prokaryote and eukaryote and the evaluation of antibacterial effects of nisin on gram-positive and gram-negative bacteria. *Journal of Microbial Biology*, 10(39), 29-50. 10.22108/bjm.2020.125232.1331. [In Persian]
- Fallahzadeh Mamaghani V. (2016). Characterization of Conserved Hypothetical Proteins from Proteome of *Xanthomonas citri* subsp. *citri*, with Ethylene Induction Activity on *Arabidopsis thaliana*. *Genetic Engineering and Biosafety Journal*, 5 (1), 31-39. [In Persian]
- Friedrich, M. G., Wang, Z., Schey, K. L., & Truscott, R. J. (2019). Mechanism of protein cleavage at asparagine leading to protein-protein cross-links. *Biochemical Journal*, 476(24), 3817-3834. doi: 10.1042/BCJ20190743.
- Gao, S., Zhou, J., & Chen, J. (2020). Identification of flavonoids 3'-hydroxylase from (*Silybum marianum* (L.) Gaertn) and its application in enhanced production of taxifolin. *Sheng wu gong cheng xue bao. Chinese Journal of Biotechnology*, 36(12), 2838-2849. doi: 10.13345/j.cjb.200178
- Ghorbanpour, R., Nikbakht, G., & Jalali, A. H. (2019). Immuno-Bioinformatics study of autotransporter protein, Antigen 43, in Enterotoxigenic *Escherichia coli* isolated from calves. *Journal of Veterinary Research*, 74(1), 128-141. doi: 10.22059/jvr.2018.211721.2504. [In Persian]
- Hami dehsardi m., Mirzaie-asl A., Deljou, A., & Ershadi, A. (2015). Bioinformatic analysis of CONSTANS gene sequence in beet plant. International Conference on Research in Science and Technology, Kuala Lumpur, Malaysia.
- Heidari, S., Heidari, B., & Heidari, P. (2022). The role of cis-acting elements of the promoter region in regulating the spatio-temporal expression of the salinity stress-responsive NHX1 gene in *Triticum aestivum* L. *Iranian Journal of Biological Sciences*, 16(4), 25-34. [In Persian]
- Housaindokht. M.R., & J. Chamani. (2022). Principles of biophysics. Ferdowsi University of Mashhad Press (FUM Press). Mashhad. [In Persian]
- karimi goyjelo R., & Abdollahi Mandoulakani, B. (2021). Isolation and characterization of cinnamate-4-hydroxylase gene promoter in *Ocimum basilicum* L. *Modern Genetics Journal*, 16(3), 211-218. [In Persian]
- Khakpour, A., Zolfaghari, M., & Sorkheh, K. (2019). Bioinformatics Study and Investigation of the Expression Pattern of Several Important Genes Involved in *Glycyrrhizin* Synthesis of *Glycyrrhiza glabra* L. in Autumn and Spring Seasons. *Plant Genetic Researches*, 6(1), 55-68.
- Khalighi- Sigaroodi, F., S. Jarvandi., & M. Taghizadeh. (2013). Therapeutic indications of medicinal plants. Arjamand. Tehran. [In Persian]
- Masoudi Jozchal, Z., Bagheri, N., Babaeian Jelodar, N., Ranjbar, G., & Farmani, J. (2024). Phytochemical analysis of the medicinal plant terrestrial orchid (*Orchis simia*) in the flowering stage. *Journal of Crop Breeding*, 16(2), 53-66. doi:10.61186/jcb.16.2.53. [In Persian]

- Mir Drikvand, R. M., Sohrabi, S. M., Sohrabi, S. S., & Samiei, K. (2023). Molecular identification and characterization of hevein antimicrobial peptide genes in barley (*Hordeum vulgare* L.). *Journal of Plant Genetic Research*, 8(2), 83-102. doi:10.61186/jcb.16.2.53. [In Persian]
- Miroliaei, M., & A. A. Rastegari. (2009). Fundamentals of biophysics. Aeeizh. Tehran. [In Persian]
- MOhammad Tabar, F., Fathi, S., Shameh, S., & Alirezaloo, A. (2024). Investigating Morphological Diversity, Essential Oil and Antioxidant Activity of some wild *Mentha longifolia* L. Genotypes in North-West of Iran. *Journal of Crop Breeding*, 16(49), 129-139. doi:10.61186/jcb.16.49.129. [In Persian]
- Panni, S., Montecchi-Palazzi, L., Kierner, L., Cabibbo, A., Paoluzi, S., Santonico, E., Landgraf., Ch. Volkmer-Enger., R. Bachi., A. Castagnoli., L. & Cesareni, G. (2011). Combining peptide recognition specificity and context information for the prediction of the 14-3-3-mediated interactome in *S. cerevisiae* and *H. sapiens*. *Proteomics*, 11(1), 128-143. doi: 10.1002/pmic.201000030.
- Rezvannejad, E., Lotfi, S., & Boustan, A. (2020). Modeling of heat stress protein 70 (HSP70) of honey bee by homology modeling and molecular simulation method and interaction to HSP40. *Cellular and Molecular Research (Iranian Journal of Biology)*, 32(4), 417-431. [In Persian]
- Salehi Mojarrad, M, H., Hasanzadeh Ghasemi, R., & Keramati, M. (2018). Nano-mechanical Study of a Bio Nano Actuator Under External Forces. *Modares Mechanical Engineering*, 17 (12), 457-462. <https://dor.net/dor/20.1001.1.10275940.1396.17.12.28.0> [In Persian]
- Sonenberg, N., Morgan, M. A., Merrick, W. C., & Shatkin, A. J. (1978). A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proceedings of the National Academy of Sciences*, 75(10), 4843-4847. doi: 10.1073/pnas.75.10.4843.
- Torres, M., & Corchete, P. (2016). Gene expression and flavonolignan production in fruits and cell cultures of *Silybum marianum*. *Journal of Plant Physiology*, 192, 111-117. doi: 10.1016/j.jplph.2016.02.004.
- Trinidad, J. L., Pabuayon, I. C. M., & Kohli, A. (2020). Harnessing protein posttranslational modifications for plant improvement. In *Advancement in Crop Improvement Techniques*. pp. 385-401. Woodhead Publishing. doi.org/10.1016/B978-0-12-818581-0.00023-1.
- Wang, A., & Krishnaswamy, S. (2012). Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. *Molecular Plant Pathology*, 13(7), 795-803. doi: 10.1111/j.1364-3703.2012.00791.x.
- Zare Mehrjerdi, M. (2015). Bioinformatical evaluation of alliinase promoter in Arabidopsis and rice. 2nd national conference of modern topic in agriculture, Tehran, Iran. [In Persian]
- Zheng, Y., Fan, Y., Ji, X., & Wu, X. (2025). Identification and Characterization of Flavonoid Biosynthetic Gene Families in Paeonia Species and Their Roles in Stamen Petalozation of *Paeonia lactiflora*. *Horticulturae*, 11(5), 463. <https://doi.org/10.3390/horticulturae11050463>.
- Zinati, Z., & Shamloo-Dashtpajardi, R. (2018). A Comparative Investigation of Catalase 2 and Thioredoxin H5 Genes Promoters in Arabidopsis to Determine of Their Functions in Response to Biotic and Abiotic Stresses. *Cellular and Molecular Research (Iranian Journal of Biology)*, 30(4), 349-360. doi:10.1001.1.23832738.1396.30.4.3.8. [In Persian]