

«مقاله پژوهشی»

آنالیز *in silico* ژن ۶- فروکتوز ترانسفراز در *Triticum aestivum* L.معصومه فلاح زيارانی^۱، مهدی صفایی‌زاده*

۱. دانشجوی دکتری گروه بیوتکنولوژی گیاهی، دانشکده علوم و فناوری زیستی، دانشگاه شهید بهشتی، تهران، ایران.
 ۲. استادیار گروه زیست‌شناسی سلولی- مولکولی، دانشکده علوم و فناوری زیستی، دانشگاه شهید بهشتی، تهران، ایران.
 (تاریخ دریافت: ۱۴۰۱/۰۲/۲۵- تاریخ پذیرش: ۱۴۰۱/۰۶/۲۷)

In Silico Analysis of Sucrose: Froctan 6- Fructose Transferase Gene in *Triticum aestivum* L.Masoumeh Fallah Ziarani¹, Mehdi Safaeizadeh^{2*}

1. Ph.D. Student in Plant Biotechnology, Faculty of Life Science and Biotechnology, Shahid Beheshti University, Tehran, Iran.

2. Assistant Professor, Faculty of Life Science and Biotechnology, Shahid Beheshti University, Tehran, Iran.

(Received: May 15, 2022- Accepted: Sep. 18, 2022)

Abstract

Sucrose Froctan 6 fructosyl transferase is an enzyme that catalyzes the transfer of the fructosyl group from sucrose to various receptors. The importance of this protein is due to its role in increasing plant resistance against stresses. Due to the importance of sucrose transport in the cell, this gene was studied with the help of various databases and various bioinformatics software and it was compared with genes involved in the hydrolysis of O-glycosyl in wheat as well as genes of other plant species. Phylogenetic analysis divided 21 haplotypes into eight distinct groups. An exon was detected on the positive strand of this gene by FgenesH program and it was found that its predicted protein is 619 amino acids long. MapViewer revealed that the sequence carrying the *Sucrose Froctan 6 fructosyl transferase (6-sft)* access gene (JQ728011) is located on chromosome 4 of wheat. The isoelectric point (pI) of the target protein is 5.19, its net charge is -24 and examination of the Instability index showed that this protein has a long life. The molecular weight of *Sucrose Froctan 6 fructosyl transferase (6-sft)* protein was 6877.38 Da and ontologically, the 6-sft protein remains inside the cell. This protein has two protected domains PF00251 (Glycohydrolase) and SSF75005 (Arabinana) at the N-terminus and a domain is SSF49899 (Cancanava) at the end of C. It was inferred that this protein was transported to the nucleus by the domain glyco-hydrolase after synthesis in the cytoplasm and regulates O-glycosyl hydrolysis in wheat. Comparison of the secondary structure of the protein confirmed the existence of alpha plates. Three-dimensional modeling of this protein in wheat was performed by homology modeling method using Swiss Model database after selecting the appropriate model with high similarity that was extracted from PDB database. To validation of modeled structure and esterochemical analysis, Ramachandran plot was drawn and dihydal angles were calculated. Structural quality evaluation results showed that the proposed models are good quality and stability. The study of protein structure may help to understand protein function and the details of its structure can be useful in studies of the active site of the protein and docking.

Keywords: Domain protected, O-glycosyle hydrolysis, Transfer of fructosyl group, Three-dimensional structure, Stress resistance.

چکیده

ساکارز: فروکتان ۶- فروکتوزیل ترانسفراز آنزیمی است که انتقال گروه فروکتوزیل از ساکارز را به گیرنده‌های مختلف کاتالیز می‌کند. اهمیت این پروتئین به دلیل نقش آن در افزایش مقاومت گیاه در برابر تنش‌ها می‌باشد. به دلیل اهمیت انتقال ساکارز در سلول، این ژن به کمک پایگاه‌های مختلف اطلاعاتی و نرم‌افزارهای گوناگون بیوانفورماتیک مطالعه گردید و با ژن‌های دخیل در هیدرولیز O-glycosyle در گندم و نیز ژن‌هایی از دیگر گونه‌های گیاهی مورد مقایسه قرار گرفت. آنالیز فیلوژنی ۲۱ گونه مورد بررسی را در هشت گروه مجزا قرار داد. با برنامه FgenesH بر روی رشته مثبت این ژن یک اگزون شناسایی گردید و مشخص گردید که پروتئین پیش‌بینی شده آن ۶۱۹ اسید آمینه طول دارد. با برنامه MapViewer مشخص شد که توالی حامل ژن ساکارز: فروکتان ۶- فروکتوزیل ترانسفراز (*6-SFT*) اکسشن (JQ728011) بر روی کروموزوم شماره ۴ گندم قرار دارد. نقطه ایزوالکتریک (pI) پروتئین مورد نظر ۵/۱۹، بار خالص آن -۲۴ می‌باشد و بررسی شاخص ناپایداری (Instability indx) نشان داد که این پروتئین عمر بالایی دارد. وزن مولکولی پروتئین ساکارز: فروکتان ۶- فروکتوزیل ترانسفراز معادل ۶۸ کیلو دالتون تعیین شد و از نظر پدیدایی (Ontology) این پروتئین در داخل سلول می‌ماند. این پروتئین دارای دو دمن حفاظت‌شده (PF00251 Glyco- hydrolase) و SSF75005 (Arabinana) در انتهای آمینی و یک دمن SSF49899 (Cancanava) در انتهای کربوکسیلی می‌باشد. استنباط گردید که این پروتئین به واسطه دمن Glyco-hydrolase پس از سنتز در سیتوپلاسم به هسته انتقال یافته و باعث تنظیم هیدرولیز O-glycosyle در گندم می‌گردد. مقایسه ساختار ثانویه پروتئین وجود صفحات آلفا را تأیید کرد. مدل‌سازی سه بعدی این پروتئین در گندم به روش همولوژی مدلیتینگ و با استفاده از پایگاه داده Swiss Model پس از انتخاب الگوی مناسب با میزان شباهت بالا که از پایگاه داده PDB استخراج شد، انجام گرفت. جهت اعتبارسنجی ساختاری مدل ترسیم شده سه‌بعدی و آنالیز استرنوشیمیایی، نمودار رامانچاندرا ترسیم و زوایای دی‌هیدرال محاسبه شدند. نتایج ارزیابی کیفیت ساختاری نشان داد که مدل‌های پیشنهادی دارای کیفیت و پایداری مناسبی می‌باشند. مطالعه ساختار پروتئین می‌تواند به درک عملکرد پروتئین کمک کند و بررسی جزئیات ساختار آن می‌تواند در مطالعات جایگاه فعال پروتئین و داکینگ سودمند باشد.

واژه‌های کلیدی: انتقال گروه فروکتوزیل، دمن حفاظت‌شده، ساختار سه‌بعدی، مقاومت به تنش، هیدرولیز O-glycosyle.

مقدمه

تحت شرایط تنش، طیفی از پاسخ‌ها در کل گیاه و در سطوح سلولی و مولکولی مشاهده می‌شود (Hasegawa *et al.*, 2000) و الگوی بیان بسیاری از ژن‌ها از جمله ژن‌های مربوط به پروتئین‌های درگیر در مسیرهای سیگنالی و پروتئین‌های عملکردی در پاسخ به تنش تغییر می‌کند. پاسخ‌های گیاهی در این شرایط به طور کلی منجر به حفظ ثبات، سم‌زدایی و تداوم رشد گیاه می‌گردد (Salekdeh *et al.*, 2002). در مناطق خشک و نیمه‌خشک با طی شدن رشد رویشی، گیاه با کاهش بارندگی مواجه شده، فراوانی آبیاری (کاهش دفعات آبیاری به دلیل کمبود آب) کاهش یافته و میزان تبخیر و تعرق گیاه با افزایش دما افزایش می‌یابد. این عوامل باعث تجمع نمک در محلول خاک (املاح و مواد معدنی موجود که در آب موجود در خاک حل شده‌اند) شده و می‌تواند تنش شوری را در مرحله رشد زایشی گیاه به وجود بیاورد (Karim *et al.*, 1993). به دلیل تنش اسمزی و یونی ناشی از شوری، گیاه از ساز و کارهای مختلفی نظیر تنظیم اسمزی و کده بندی سدیم^۱ و کلر در واکنش برای تداوم جذب آب و حفظ کارایی فتوسنتز استفاده می‌کند (Jenks *et al.*, 2007). تنظیم اسمزی تحت تأثیر عوامل محیطی بوده و باعث افزایش انواع اسمولیت‌ها مانند کربوهیدرات‌های محلول شامل فروکتان‌ها، ساکارز، گلوکز و فروکتوز طی شوری می‌گردد و در بین ارقام مختلف گندم تنظیم اسمزی یکسان نیست و تحت تأثیر عوامل ژنتیکی می‌باشد. قندهای محلول ساقه گندم به ویژه فروکتان در شرایط تنش اهمیت زیادی در پر شدن دانه و حفظ عملکرد گیاه دارند. ساکارز: فروکتان ۶- فروکتوزیل ترنسفرز (6-SFT) آنزیمی

است که انتقال گروه فروکتوزیل از ساکارز به گیرنده‌های مختلف را کاتالیز می‌کند. برخی از این آنزیم مشخص کرد که آب می‌تواند به‌عنوان پذیرنده فروکتوزیل^۲ عمل کند. به طور واضح این آنزیم برای سنتز فروکتان به‌عنوان یک گیرنده عمل می‌کند، اما به‌عنوان یک اهداکننده دارای تنوعی از الیگوفروکتان محتوی β (2-21) و β (2+6) متصل‌شونده به فروکتوزیل نمی‌باشد. بنابراین این نقش اصلی ساکارز: فروکتان ۶- فروکتوزیل ترنسفرز می‌باشد (Karim *et al.*, 1992). به دلیل کمبود آب و افزایش تنش خشکی اهمیت ژن‌هایی که در مقاومت گیاه به تنش نقش دارند مهم می‌باشد. با توجه به اهمیت و کارایی ژن ساکارز: فروکتان ۶- فروکتوزیل ترانسفرز که در ایجاد مقاومت در گیاه در تنش خشکی نقش دارد، این ژن به منظور آنالیزهای بیوانفورماتیکی انتخاب گردید تا از نتایج آن در تحقیقات آزمایشگاهی و مزرعه‌ای استفاده شود. تاکنون آنالیزهای بیوانفورماتیکی روی این ژن صورت نگرفته است.

مواد و روش‌ها

Blast P و BLASTn

Blast P و BLASTn به منظور جستجوی توالی مشابه با توالی ژنی و پروتئینی ژن ساکارز: فروکتان ۶- فروکتوزیل ترنسفرز (6-SFT) گیاه برنج به‌عنوان query انجام شد.

آنالیزهای فیلوژنتیکی

از برنامه Bioedit2007 برای هم‌ردیفی چندگانه توالی‌ها استفاده شد (Fallah Ziarani *et al.*, 2017). برای رسم درخت فیلوژنتیک نرم‌افزار MEGA5 به کار برده شد (Fallah Ziarani *et al.*, 2017).

دمین‌های حفاظت‌شده در پروتئین استفاده شد (Fallah Ziarani *et al.*, 2017).

بررسی ساختار دوم پروتئین

از ابزار آنالیز Phyre2 به منظور تعیین ساختار دوم پروتئین و مشخص شدن صفحات آلفا و بتا استفاده شد (Fallah Ziarani *et al.*, 2017).

بررسی ساختار سه‌بعدی پروتئین

مدل‌سازی سه بعدی این پروتئین به روش همولوژی مدلینگ و با استفاده از پایگاه داده Swiss Model پس از انتخاب الگوی مناسب با میزان شباهت بالا که از پایگاه داده PDB استخراج شد، انجام گرفت. جهت اعتبارسنجی ساختاری مدل ترسیم شده سه بعدی و آنالیز استرئوشیمیایی، نمودار رامانچاندرا ترسیم و زوایای دی‌هیدرال محاسبه شدند. مطالعه ساختار پروتئین می‌تواند به درک عملکرد پروتئین کمک کند و بررسی جزئیات ساختار آن می‌تواند در مطالعات جایگاه فعال پروتئین و Docking^۱ سودمند باشد.

آنالیز هضم و بررسی تغییرات پس از ترجمه

آنالیز هضم با تریپسین توسط PeptideCutter بررسی شد. تغییرات پس از ترجمه بوسیله ScanProsite مشخص شد.

نتایج و بحث

BLASTn

جستجوی توالی نوکلئوتیدی اکسشن FJ228688.1 برنج در مقابل پایگاه مجموعه نوکلئوتیدی برنج Non-redundant (n/r) توانست ۱۳۸ تطابق (hit) را نشان دهد که بالاترین امتیاز شاخص Max

آنالیز توالی کدکننده با برنامه FgenesH

بررسی خصوصیات ژن ساکارز: فروکتان ۶- فروکتوزیل ترانسفراز با نرم‌افزارهای FgenesH و Sequence Viewer (Fallah Ziarani *et al.*, 2017) انجام شدند. همچنین محتوی GC با استفاده از Genscan web server به دست آمد (Fallah Ziarani *et al.*, 2017).

آنالیز پروموتور، نقشه‌یابی فیزیکی ژن و هم‌ردیفی چندگانه توالی پروتئین مورد مطالعه

به منظور آنالیز ناحیه پروموتور ژن *SFT-6* از نرم‌افزار PromoterScan استفاده شد (Fallah Ziarani *et al.*, 2017). برنامه‌های SequenceViewer و MapViewer به منظور نقشه‌یابی فیزیکی ژن در ژنوم برنج (به دلیل اهمیت برنج در تغذیه) استفاده شدند. هم‌ردیفی چندگانه توالی پروتئینی پروتئین *SFT-6* با شش توالی پروتئینی مربوط به دیگر گونه‌های ژن *SFT-6* گندم با استفاده از برنامه MultiAlign انجام شد. همچنین هم‌ردیفی چندگانه توالی پروتئین *SFT-6* با چندین پروتئین درگیر در هیدرولیز O-glycosyle در گیاهان مختلف با برنامه MultiAlign (به‌منظور تأیید نتیجه توالی‌یابی با نرم‌افزار Bioedit) انجام شد.

بررسی خصوصیات ساختار اولیه پروتئین

به منظور بررسی خصوصیات ساختار اولیه پروتئین از برنامه PROTPARM استفاده شد. برای بررسی محل استقرار و نقش مولکولی پروتئین در سلول از برنامه UNIPROT استفاده شد. آنالیز وجود توالی‌های سیگنال توسط SignalP مورد بررسی قرار گرفت. محل استقرار نهایی پروتئین توسط برنامه MitoProt II مشخص شد. از برنامه جامع InterProScan به منظور شناسایی موتیف‌ها و

۱. داکینگ

نتایج Blast n (بر اساس توالی نوکلئوتیدی) و Blast P (بر اساس توالی پروتئینی) (در اینجا نتیجه همه هم‌ردیفی‌ها نشان داده نشده است) مشخص کرد که ژن 6-SFT در طول تکامل محافظت شده است.

آنالیزهای فیلوژنتیکی

برای این منظور، ۲۱ گونه اول که در آنالیز BLASTp با ژن 6-SFT همولوژی نشان دادند، انتخاب شدند. در گام اول این گونه‌ها با نرم‌افزار Bioedit2007 هم‌ردیف‌سازی شده و سپس برای انجام آنالیزهای فیلوژنتیکی از نرم‌افزار Mega5 استفاده شد.

آزمون شاخص نابرابری^۲ که با استفاده از نرم‌افزار Mega5 انجام شد، نشان داد که ژن 6-SFT همولوژی بالایی با همه گونه‌های مورد بررسی دارد (شکل ۱).

بر اساس آنالیز Substitution model selection، بهترین مدل جایگزینی برای آنالیز فیلوژنی، مدل WAG + G انتخاب شد که بر اساس آنالیز توسط نرم‌افزار Mega5 به دست آمد. این آنالیز به منظور تعیین بهترین مدل برای تشخیص نوع جهش انجام می‌شود تا مشخص شود که جهش‌های انجام شده در ژن در طول تکامل به چه صورتی می‌باشد (شکل ۲).

آنالیز فیلوژنی با روش UPGMA و انتخاب مدل جایگزینی WAG+G دندروگرامی ایجاد کرد که گونه‌های مورد مطالعه را در ۸ گروه کاملاً مجزا دسته‌بندی کرد (شکل ۳). مدل جایگزینی با آنالیز توسط نرم‌افزار Mega5 به دست آمد و با مقایسه تفاوت توالی‌ها و انتخاب بهترین مدل جایگزینی بهترین درخت فیلوژنتیکی رسم شد.

در ماتریس تشابه در مقایسه بین ۲۱ گونه مورد

score (۴۰۹۱) مربوط به تطابق با اکسشن EU981911.1 (ژن ساکارز: فروکتان ۶- فروکتوزیل ترنسفرز از *Triticum turgidum* subsp.) (Durum) با ۱۰۰ درصد پوشش و ۹۰ درصد شباهت بود؛ در حالی که کمترین امتیاز شاخص Max score (۶۰/۲) مربوط به تطابق با اکسشن AB190801.1 (ژن اینورتاز)^۱ در رقم *Musa Pyrus communis* PC-sAIV-3 با ۵ درصد پوشش و ۷۵ درصد شباهت بود. برای تشخیص بهترین انطباق از مقدار E-value هم تنها می‌توان استفاده کرد که ارزش کوچکتر نشان‌دهنده امتیاز بالاتر و معتبر بودن انطباق با احتمال بالاتر و ارزش زیر ۰/۰۱ نشان‌دهنده معنی‌داری است.

Blast P

هم‌ردیفی توالی پروتئینی با اکسشن نامبر JQ728011 در برنج مقابل پایگاه مجموعه نوکلئوتید Non-redundant (n/r) توانست ۱۰۱ تطابق (hit) را نشان دهد که بالاترین امتیاز شاخص Max score (۱۲۶۸) مربوط به هم‌ردیفی با اکسشن نامبر ACI43353.3 (*Triticum aestivum*) ژن 6-SFT- Max score با ۱۰۰ درصد پوشش و ۹۹ درصد شباهت بود؛ در حالی که کمترین امتیاز شاخص Max score (۷۲۴) مربوط به هم‌ردیفی با اکسشن نامبر AGO03823.1 (ژن اینورتاز در رقم *Sorghum*) با ۸۷ درصد پوشش و ۶۳ درصد شباهت بود. برای تشخیص بهترین انطباق از مقدار E-value هم تنها می‌توان استفاده کرد که ارزش کوچکتر نشان‌دهنده امتیاز بالاتر و معتبر بودن انطباق با احتمال بالاتر و ارزش زیر ۰/۰۱ نشان‌دهنده معنی‌داری است.

گونه می‌تواند نزدیک بودن درصد GC آن‌ها باشد
(Guex & Peitsch, 1997).

بررسی، کمترین عدم تشابه مربوط به عدد ۰/۰۱۲ می‌باشد که مربوط به *Agropyron cristotum* و *Aegilops tauschii* است که این دو شبیه‌ترین نسبت به هم هستند، یکی از دلایل شبیه‌بودن این دو

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. AFK29573.1 sucrose:fructan 6-fructosyltransferase Triticum aestivum		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000
2. ACI43225.3 sucrose:fructan 6-fructosyltransferase Triticum aestivum	1.000		0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.009	0.008	0.015	0.019	0.011
3. ACH73192.1 sucrose:fructan 6-fructosyltransferase Triticum urartu	1.000	1.000		0.002	0.000	0.000	0.000	0.000	0.000	0.009	0.013	0.002	0.004	0.008	0.000
4. XP020156722.1 sucrose:sucrose 1-fructosyltransferase-like Aegilops tauschii subsp. tauschii	1.000	1.000	0.426		0.002	0.000	0.002	0.000	0.000	0.013	0.015	0.038	0.032	0.036	0.000
5. ACH73194.1 sucrose:fructan 6-fructosyltransferase Aegilops tauschii	1.000	0.382	1.000	0.404		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6. ACH73193.1 sucrose:fructan 6-fructosyltransferase Aegilops searsii	1.000	1.000	1.000	1.000	1.000		0.002	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000
7. BAB2469.1 sucrose:fructan 6-fructosyltransferase Triticum aestivum	1.000	1.000	1.000	0.418	1.000	0.358		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
8. pirJJC7905 fructan 6-fructosyltransferase - wheat Triticum aestivum	1.000	1.000	1.000	1.000	1.000	1.000	1.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
9. ACH73195.1 sucrose:fructan 6-fructosyltransferase Triticum turgidum subsp. durum	1.000	1.000	1.000	1.000	1.000	0.408	1.000	1.000		0.000	0.000	0.000	0.000	0.000	0.000
10. AFP72240.1 sucrose:fructan 6-fructosyltransferase Hordeum vulgare subsp. vulgare	1.000	0.194	0.224	0.162	1.000	1.000	1.000	1.000	1.000		0.000	0.009	0.004	0.000	0.000
11. CAA58235.1 sucrose:fructan 6-fructosyltransferase Hordeum vulgare subsp. vulgare	1.000	0.298	0.196	0.156	1.000	1.000	1.000	1.000	1.000	1.000		0.004	0.000	0.000	0.000
12. AMM03226.1 sucrose:fructan 6-fructosyltransferase Leymus racemosus	1.000	0.252	0.432	0.064	1.000	1.000	1.000	1.000	1.000	0.268	0.406		0.000	0.008	0.000
13. CUM44319.1 sucrose:fructan 6-fructosyltransferase Desypium villosum	1.000	0.224	0.434	0.132	1.000	1.000	1.000	1.000	1.000	0.360	1.000	1.000		0.009	0.000
14. CUM44318.1 sucrose:fructan 6-fructosyltransferase Desypium villosum	0.408	0.274	0.398	0.096	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.034	0.054		0.006
15. ACN493835.1 putative sucrose:fructan 6-fructosyltransferase Eriomys pictus	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.360
16. AAK27319.1 sucrose:fructan 6-fructosyltransferase Agropyron cristatum	0.060	0.110	0.088	0.022	0.126	0.078	0.136	0.200	0.072	0.210	0.190	1.000	1.000	1.000	1.000
17. ANI24189.1 6-fructosyltransferase partial Leymus chinensis	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
18. EMT09010.1 6(O)-fructosyltransferase Aegilops tauschii	1.000	1.000	0.338	1.000	0.432	1.000	0.370	1.000	1.000	0.164	0.154	0.072	0.098	0.080	1.000
19. BAH130252.1 sucrose:fructan 6-fructosyltransferase Phleum pratense	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
20. BAO09621.1 fructosyltransferase Phleum pratense	0.208	0.278	0.282	0.278	0.442	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.276	0.180	1.000
21. ADP38050.1 putative sucrose:fructan fructosyl transferase 2 Poa pratensis	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

16	17	18	19	20	21
0.057	0.000	0.000	0.000	0.061	0.000
0.044	0.000	0.000	0.000	0.045	0.000
0.051	0.000	0.002	0.000	0.045	0.000
0.089	0.000	0.000	0.000	0.053	0.000
0.042	0.000	0.002	0.000	0.004	0.000
0.051	0.000	0.000	0.000	0.000	0.000
0.030	0.000	0.002	0.000	0.000	0.000
0.027	0.000	0.000	0.000	0.000	0.000
0.042	0.000	0.000	0.000	0.000	0.000
0.028	0.000	0.013	0.000	0.000	0.000
0.028	0.000	0.015	0.000	0.000	0.000
0.000	0.000	0.038	0.000	0.059	0.000
0.000	0.000	0.032	0.000	0.068	0.000
0.000	0.000	0.036	0.000	0.000	0.000
0.057	0.000	0.000	0.000	0.000	0.000
	0.000	0.089	0.000	0.023	0.000
1.000		0.000	0.000	0.000	0.000
0.012	1.000		0.000	0.053	0.000
1.000	1.000		1.000	0.000	0.000
0.332	1.000	0.222	1.000		0.000
1.000	1.000	1.000	1.000	1.000	

شکل ۱. نتیجه مقایسه الگوی توزیع جایگزینی‌ها و ماتریس تشابه در ۲۱ گونه مورد مطالعه

Model	#Param	BIC	AICc	InL	Invariant	Gamma	Freq A	Freq R
WAG+G	40	8613.4	8321.1	-4120.4	n/a	0.63522	0.0866	0.044
JTT+G	40	8618.3	8326.1	-4122.9	n/a	0.61741	0.0769	0.0511
WAG+G+I	41	8622.7	8323.1	-4120.4	0	0.63522	0.0866	0.044
JTT+G+I	41	8627.6	8328	-4122.9	0	0.61739	0.0769	0.0511
WAG+I	40	8649.2	8356.9	-4138.3	0.412172	n/a	0.0866	0.044
JTT+I	40	8654.1	8361.9	-4140.8	0.416599	n/a	0.0769	0.0511
cpREV+G	40	8714.1	8421.9	-4170.8	n/a	0.57165	0.0756	0.0621
cpREV+G+I	41	8723.4	8423.9	-4170.8	0.007141	0.57988	0.0756	0.0621
Dayhoff+G	40	8725.2	8433	-4176.4	n/a	0.59054	0.0871	0.0409
Dayhoff+G+I	41	8735	8435.5	-4176.6	0	0.59064	0.0871	0.0409
WAG+G+F	59	8739.1	8308.2	-4094.8	n/a	0.64382	0.0786	0.0556
WAG	39	8739.1	8454.1	-4187.9	n/a	n/a	0.0866	0.044
WAG+G+I+F	60	8748.4	8310.3	-4094.8	0	0.64383	0.0786	0.0556
JTT	39	8748.6	8463.6	-4192.7	n/a	n/a	0.0769	0.0511
JTT+G+F	59	8750.5	8319.6	-4100.5	n/a	0.62624	0.0786	0.0556
cpREV+I	40	8752	8459.7	-4189.7	0.441267	n/a	0.0756	0.0621
JTT+G+I+F	60	8759.6	8321.5	-4100.4	0	0.62619	0.0786	0.0556
Dayhoff+I	40	8762.8	8470.5	-4195.1	0.416645	n/a	0.0871	0.0409
rtREV+G	40	8767.8	8475.6	-4197.6	n/a	0.58262	0.0646	0.0453
WAG+I+F	59	8774.4	8343.5	-4112.5	0.409183	n/a	0.0786	0.0556
rtREV+G+I	41	8777.7	8478.1	-4197.9	0	0.58269	0.0646	0.0453
JTT+I+F	59	8786.6	8355.7	-4118.5	0.412017	n/a	0.0786	0.0556
Dayhoff+G+F	59	8799.2	8368.3	-4124.8	n/a	0.58753	0.0786	0.0556
Dayhoff+G+I+F	60	8808.5	8370.3	-4124.8	0	0.58753	0.0786	0.0556
cpREV+G+F	59	8809.5	8378.6	-4130	n/a	0.58062	0.0786	0.0556
rtREV+I	40	8814.4	8522.1	-4220.9	0.411207	n/a	0.0646	0.0453
cpREV+G+I+F	60	8818.8	8380.7	-4130	0.008692	0.59099	0.0786	0.0556
rtREV+G+F	59	8829.4	8398.6	-4140	n/a	0.58339	0.0786	0.0556

شکل ۲. انتخاب مدل جایگزینی



شکل ۳. دندروگرام به دست آمده از گونه‌های مورد مطالعه با روش UPGMA

آزمون تاجیما نشان داد که جمعیت مورد (۲۱ گونه مورد بررسی) مطالعه در تعادل دریفت- انتخاب قرار دارند، بدین معنی که این جمعیت در روند تکامل، دستخوش ناریبی تعادل آلی ناشی از عوامل بر هم زننده آن یعنی جهش و انتخاب قرار نگرفته است و در حالت تعادل می‌باشد. تست فرضیه خنثی اختلاف

آنالیز تنوع ژنتیکی توسط نرم‌افزار Mega5 نشان داد که در حدود ۴۴ درصد از طول توالی ژن 6-SFT جهش‌های نوع جایگزینی رخ داده که در این میان تنها حدود ۱۲ درصد از طول توالی ژن دارای چندشکلی‌های آگاهی‌بخش می‌باشد. تنوع نوکلوتیدی در جمعیت مورد مطالعه پایین و در حدود ۱۱ درصد می‌باشد (جدول ۱).

آنالیز توالی کد کننده با برنامه FGENESH

با استفاده از برنامه FgenesH مشخص شد که توالی کدکننده ژن *6-SFT* به طول ۳۱۴۶ جفت باز با اکسشن FJ228688.1 بر روی نوکلئوتید ۱ تا ۳۱۴۶ واقع است (شکل ۴). همان‌گونه که در شکل مشاهده می‌شود، این ژن ۴ اگزون و ۳ اینترون دارد که در رشته مثبت شناسایی شد. طول اگزون اول ۴۴۴ جفت باز (از نوکلئوتید ۱ تا ۴۴۴)، اگزون دوم به طول ۲۷ جفت باز (از نوکلئوتید ۵۶۲ تا ۵۸۸)، اگزون سوم به طول ۸۵۸ جفت باز (از نوکلئوتید ۱۰۹۳ تا ۱۹۵۰) و اگزون چهارم به طول ۷۰۲ جفت باز (از نوکلئوتید ۲۴۴۵ تا ۳۱۴۶) است و پروتئین پیش‌بینی شده آن ۶۷۷ اسید آمینه طول دارد. با برنامه Sequence Viewer نیز وجود یک اگزون در همان موقعیت گزارش شده به وسیله برنامه FgenesH تأیید گردید. درصد GC (درصد گوانین و سیتوزین در توالی مورد بررسی)، ۵۲/۸۹٪ به دست آمد. آنالیز محتوی GC مشخص کرد که ژن‌هایی با محتوی GC بالا پایداری بیشتری در برابر حرارت نسبت به ژن‌هایی با محتوی GC کمتر دارند (لول، ۲۰۱۳). تنوع در محتوای گوانین (G) و سیتوزین (C) مشاهده شده در بین گونه‌ها یکی از موضوعات مورد توجه در تکامل بیوانفورماتیکی است. چگونگی ارزش GC با محاسبه محتوای GC به کل ژنوم در موجودات مورد مطالعه می‌باشد فرگوسون و همکاران در سال ۲۰۱۱ و کائو در سال ۲۰۱۱ اعلام کردند که تنوع در محتوای درصد GC در یوکاریوت‌ها کمتر می‌باشد (Ferguson et al., 2011; Cao et al., 2011). علاوه بر این، آنالیز محتوی درصد GC به‌وسیله سرور ProtParam مشخص کرد که ژن‌هایی با درصد GC بالاتر در مقایسه با ژن‌هایی با درصد GC کمتر پایدارتر می‌باشند. این ممکن است ناشی از این حقیقت باشد که جفت‌های GC توسط سه باند هیدروژنی به هم متصل می‌شوند، اما جفت‌های AT

مشاهده شده برای یک جفت توالی که با الگوی جایگزینی مشابه تکامل یافته‌اند، را اندازه‌گیری می‌کند. اساس این کار مقایسه فرکانس نوکلئوتید یا آمینواسید در توالی جفتی و تعداد تفاوت مشاهده شده بین توالی‌ها است. در زیست‌شناسی مدل جایگزینی پروسه‌ای را توصیف می‌کند که در آن تغییر یک صفت را در یک پروسه زمانی بر اثر فشار انتخاب نشان می‌دهد. این مدل در تفسیر درخت تکاملی در فیلوژنتیک و شاخه‌بندی و شبیه‌سازی توالی با تست روش‌ها و الگوریتم‌های دیگر کاربرد دارد. یکی از آزمون‌هایی که برای بررسی اثر فشار انتخاب بر روی یک ژن به کار برده می‌شود تاجیما D است. این آزمون بر اساس مقایسه تنوع نوکلئوتیدی حاصل از فراوانی آلی جایگاه‌های چندشکلی است (Carlson et al., 2005).

با استفاده از تاجیما D که برای آزمون تئوری خنثی به کار برده می‌شود، مشخص شد ژن‌هایی که تنوع فراوانی آلی آنها بالا است، تاجیما D مثبت (بزرگتر از صفر) دارند و با انتخاب متعادل در جمعیت همراه هستند، در حالی که، ژن‌هایی که تنوع فراوانی آلی آنها پایین است D تاجیما منفی دارند و با فشار انتخابی همراه هستند که یک واریانت سودمند را در جمعیت جایگزین سایر واریانت‌ها کرده است. ژن‌هایی که به تازگی در معرض انتخاب قرار گرفته‌اند و آلل سودمند هنوز آلل اصلی در جمعیت نشده است، به سادگی با استفاده از تاجیما D قابل مطالعه نیستند. آزمون تاجیما منفی نشان داد که جمعیت مورد مطالعه در تعادل دریافت-انتخاب قرار دارند، بدین معنی که این جمعیت در روند تکامل، دستخوش ناریبی تعادل آلی ناشی از عوامل بر هم زننده آن یعنی جهش و انتخاب قرار نگرفته است.

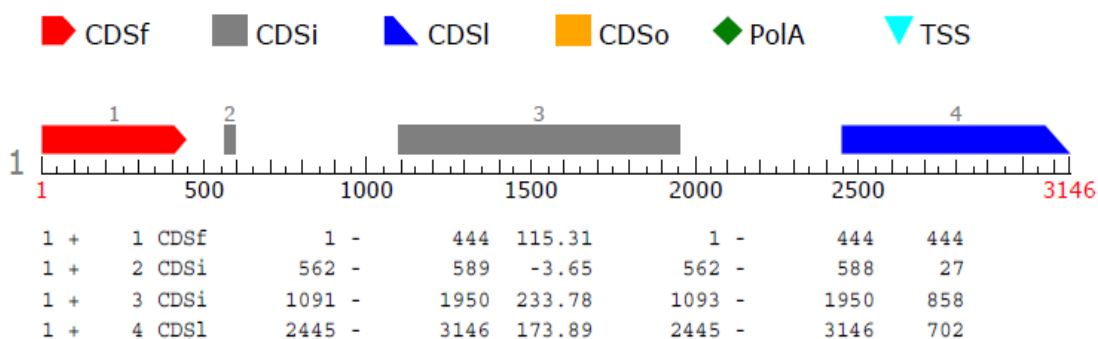
جدول ۱. نتیجه تست تاجیما برای ژن *6-SFT* در موجودات

مورد بررسی					
M	S	P _s	θ	π	D
21	235	0.445076	0.123710	0.116441	-0.242080

بوسیله دو باند هیدروژنی باند می‌شوند، باندهای هیدروژنی به پایداری بیشتر محصولات ژن کمک می‌کنند و تعداد باندهای هیدروژنی بیشتر نتیجه‌اش پایداری بیشتر پروتئین می‌باشد. نتایج نشان داد با توجه به بالابودن نسبی درصد GC پروتئین 6-SFT، این پروتئین تا حدودی قادر به حفظ ساختار فضایی خود در دماهای بالاتر می‌باشد (Burge & Karlin, 1997).

آنالیز ناحیه پرموتور ژن 6-SFT

برای شناسایی جایگاه شروع رونویسی (TSS) از برنامه PromoterScan استفاده شد. مقدار آستانه (cut off) معادل ۰/۵۳ در نظر گرفته شد. نتیجه این آنالیز براساس شکل ۵ نشان می‌دهد که ناحیه پرموتوری ژن 6-SFT از نوکلئوتید ۲۴۹۹ تا ۲۷۴۹ و محل شروع رونویسی نوکلئوتید ۲۵۱۶ می‌باشد (شکل ۵).



شکل ۴. نتیجه آنالیز اکشن FJ228688.1 برای جستجوی ژن 6-SFT با برنامه FgenesH.

Significant Signals:				
Name	TFD #	Strand	Location	Weight
NRE_Box2_CS	S01555	-	2516	1.229000
CREB	S00144	+	2550	1.912000
E4F1	S01249	+	2551	3.764000
ATF/CREB	S00534	+	2551	1.564000
E4F1	S01252	+	2551	3.824000
ATF	S01940	-	2556	3.721000
ATF	S01059	-	2556	1.591000
CREB	S00969	-	2556	4.589000
ATF	S00184	-	2557	12.909000
EivF	S00399	-	2557	3.227000
EivF/CREB	S00104	-	2557	1.564000
E4TF1	S00153	-	2557	1.750000
c-Fos_US5	S00676	-	2557	3.824000
AP-2	S00346	+	2591	1.355000
UCE.2	S00437	+	2663	1.278000
GCF	S01964	+	2710	2.361000
TTR_inverted_repeat	S01112	-	2715	3.442000
GCF	S01964	-	2716	2.284000
UCE.2	S00437	-	2733	1.216000

شکل ۵. آنالیز ناحیه پرموتور ژن 6-SFT برای یافتن ناحیه شروع رونویسی با برنامه PromoterScan.

سیس المنت‌ها مورد آنالیز قرار گرفت. تعداد ۳۰ نوع سیس المنت در ناحیه پرموتور مشخص شد که برخی جزو سیس المنت‌های عمومی و برخی سیس المنت‌های

همچنین ناحیه پرموتور (هزار نوکلئوتید بالادست کدون شروع ترجمه) که با برنامه PromoterScan تعیین شد با برنامه آنالیز PlantCare به منظور آنالیز

Motifs Found

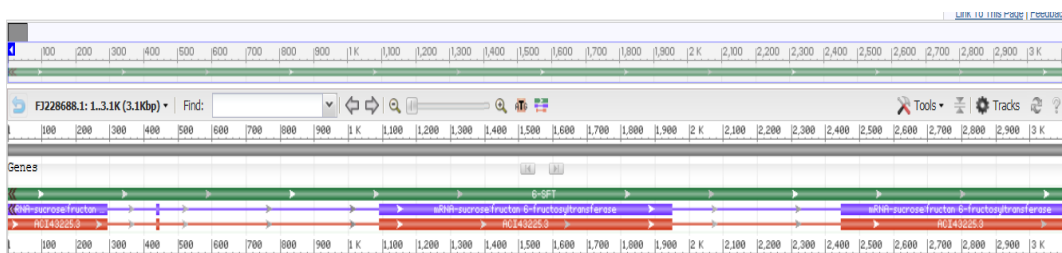
+	■	SUTR Py-rich stretch
+	■	ABRE
+	■	ACE
+	■	AE-box
+	■	Box 4
+	■	Box II
+	■	Box-W1
+	■	C-repeat/DRE
+	■	CAAT-box
+	■	CGTCA-motif
+	■	G-Box
+	■	G-box
+	■	GAG-motif
+	■	GC-motif
+	■	GT1-motif
+	■	GTGGC-motif
+	■	I-box
+	■	MBS
+	■	OBP-1 site
+	■	Skn-1_motif
+	■	Sp1
+	■	TATA-box
+	■	TCCACCT-motif
+	■	TGA-element
+	■	TGACG-motif
+	■	Unnamed_1
+	■	Unnamed_3
+	■	Unnamed_4
+	■	Unnamed_8
+	■	W box

شکل ۶ آنالیز سیس المنت‌های درگیر در ناحیه پرموتور ژن *6-SFT* با برنامه PlantCare.

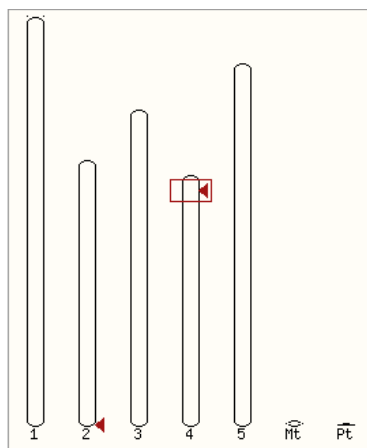
اختصاصی هستند (شکل ۶).

نقشه‌یابی فیزیکی ژن در ژنوم برنج

با برنامه SequenceViewer که در سایت NCBI قابل دسترس می‌باشد، مشخص شد که توالی حامل ژن *6-SFT* (اکسشن FJ228688.1) بر روی کروموزوم شماره ۴ برنج در موقعیت فیزیکی از نوکلئوتید ۱ تا نوکلئوتید ۳۱۴۶ قرار دارد (شکل ۷). به دلیل اهمیت برنج در تغذیه، این ژن در برنج مورد آنالیز قرار گرفت. نتیجه فوق از طریق برنامه MapViewer در سایت www.gramene.org نیز تأیید شد و نشان داده شد که ژن *6-SFT* بر روی نقشه کاربوتایی گندم بر روی بازوی کروموزوم ۴ قرار دارد (شکل ۸). همدیفی چندگانه توالی پروتئینی پروتئین *6-SFT* گندم با شش توالی پروتئینی مربوط به دیگر گونه‌های ژن *6-SFT* گندم انجام شد. نتیجه این آنالیز با برنامه MultiAlign نشان داد که همولوژی بالایی بین پروتئین *6-SFT* گندم و شش گونه دیگر از ژن *6-SFT* وجود دارد (شکل ۹).



شکل ۷. تعیین محل فیزیکی اکسشن حامل ژن *6-SFT* در ژنوم برنج با برنامه MapViewer در سایت NCBI



شکل ۸. تأیید محل فیزیکی اکسشن حامل ژن 6-SFT بر روی نقشه کاربوتایی کروموزوم‌های گندم با برنامه BLSATn در سایت www.gramene.org

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	
AFK29573.1	HGSHGKPLPYAYKPLPSGAAV-DADGERTGATRRVRCATVLTASAVVVVVGATLLAGFRVQAVDEEARAGFPMSNEHLQQRSGYHFQAKNYHSDPNGLMYNGYHMFQYNPVGTDADDDGHEG														
ACI43225.3	HGSHGKPLPYAYKPLPSGAAV-DADGERTGATRRVRCATVLTASAVVVVVGATLLAGFRVQAVDEEARAGFPMSNEHLQQRSGYHFQAKNYHSDPNGLMYNGYHMFQYNPVGTDADDDGHEG														
XP_020156722.1	HGSHGKPLPYAYKPLPSGAAV-DADGERTGATRRVRCATVLTASAVVVVVGATLLAGFRVQAVDEEARAGFPMSNEHLQQRSGYHFQAKNYHSDPNGLMYNGYHMFQYNPVGTDADDDGHEG														
ACH73192.1	HGSHGKPLPYAYKPLPS-----DADGERTGATRRVRCATVLTASAVVVVVGATLLAGFRVQAVDEEARAGFPMSNEHLQQRSGYHFQAKNYHSDPNGLMYNGYHMFQYNPVGTDADDDGHEG														
ACH73194.1	HGSHGKPLPYAYKPLPS-----DADGERTGATRRVRCATVLTASAVVVVVGATLLAGFRVQAVDEEARAGFPMSNEHLQQRSGYHFQAKNYHSDPNGLMYNGYHMFQYNPVGTDADDDGHEG														
ACH73193.1	HGSHGKPLPYAYKPLPS-----DADGERTGATRRVRCATVLTASAVVVVVGATLLAGFRVQAVDEEARAGFPMSNEHLQQRSGYHFQAKNYHSDPNGLMYNGYHMFQYNPVGTDADDDGHEG														
Consensus	HGSHGKPLPYAYKPLPS.....DADGERTGATRRVRCATVLTASAVVVVVGATLLAGFRVQAVDEEARAGFPMSNEHLQQRSGYHFQAKNYHSDPNGLMYNGYHMFQYNPVGTDADDDGHEG														
	131	140	150	160	170	180	190	200	210	220	230	240	250	260	
AFK29573.1	HAYSRNLVTRTLPTRAVADQHYDILGVLSSGHTVLPNGTYIHITGATNASAVEVQCIATPADPNDFLRRTKHPANPIHSPPGIGTKDFRDPHTAHYDESDDTARTLLGSKDDHGHGDIAMHYK														
ACI43225.3	HAYSRNLVTRTLPTRAVADQHYDILGVLSSGHTVLPNGTYIHITGATNASAVEVQCIATPADPNDFLRRTKHPANPIHSPPGIGTKDFRDPHTAHYDESDDTARTLLGSKDDHGHGDIAMHYK														
XP_020156722.1	HAYSRNLVTRTLPTRAVADQHYDILGVLSSGHTVLPNGTYIHITGATNASAVEVQCIATPADPNDFLRRTKHPANPIHSPPGIGTKDFRDPHTAHYDESDDTARTLLGSKDDHGHGDIAMHYK														
ACH73192.1	HAYSRNLVTRTLPTRAVADQHYDILGVLSSGHTVLPNGTYIHITGATNASAVEVQCIATPADPNDFLRRTKHPANPIHSPPGIGTKDFRDPHTAHYDESDDTARTLLGSKDDHGHGDIAMHYK														
ACH73194.1	HAYSRNLVTRTLPTRAVADQHYDILGVLSSGHTVLPNGTYIHITGATNASAVEVQCIATPADPNDFLRRTKHPANPIHSPPGIGTKDFRDPHTAHYDESDDTARTLLGSKDDHGHGDIAMHYK														
ACH73193.1	HAYSRNLVTRTLPTRAVADQHYDILGVLSSGHTVLPNGTYIHITGATNASAVEVQCIATPADPNDFLRRTKHPANPIHSPPGIGTKDFRDPHTAHYDESDDTARTLLGSKDDHGHGDIAMHYK														
Consensus	HAYSRNLVTRTLPTRAVADQHYDILGVLSSGHTVLPNGTYIHITGATNASAVEVQCIATPADPNDFLRRTKHPANPIHSPPGIGTKDFRDPHTAHYDESDDTARTLLGSKDDHGHGDIAMHYK														
	261	270	280	290	300	310	320	330	340	350	360	370	380	390	
AFK29573.1	TKDFLNVELIPGILHRVQRTGEWECIDFPYVGHRSNDSSEHLVHLKASHODERHYYSLGTYSARANTPTIDPELDLGTGLRYDHGKFYASTSFYDPAKRRVLYHGVEVDSKRDVYKGGASIQSV														
ACI43225.3	TKDFLNVELIPGILHRVQRTGEWECIDFPYVGHRSNDSSEHLVHLKASHODERHYYSLGTYSARANTPTIDPELDLGTGLRYDHGKFYASTSFYDPAKRRVLYHGVEVDSKRDVYKGGASIQSV														
XP_020156722.1	TKDFLNVELIPGILHRVQRTGEWECIDFPYVGHRSNDSSEHLVHLKASHODERHYYSLGTYSARANTPTIDPELDLGTGLRYDHGKFYASTSFYDPAKRRVLYHGVEVDSKRDVYKGGASIQSV														
ACH73192.1	TKDFLNVELIPGILHRVQRTGEWECIDFPYVGHRSNDSSEHLVHLKASHODERHYYSLGTYSARANTPTIDPELDLGTGLRYDHGKFYASTSFYDPAKRRVLYHGVEVDSKRDVYKGGASIQSV														
ACH73194.1	TKDFLNVELIPGILHRVQRTGEWECIDFPYVGHRSNDSSEHLVHLKASHODERHYYSLGTYSARANTPTIDPELDLGTGLRYDHGKFYASTSFYDPAKRRVLYHGVEVDSKRDVYKGGASIQSV														
ACH73193.1	TKDFLNVELIPGILHRVQRTGEWECIDFPYVGHRSNDSSEHLVHLKASHODERHYYSLGTYSARANTPTIDPELDLGTGLRYDHGKFYASTSFYDPAKRRVLYHGVEVDSKRDVYKGGASIQSV														
Consensus	TKDFLNVELIPGILHRVQRTGEWECIDFPYVGHRSNDSSEHLVHLKASHODERHYYSLGTYSARANTPTIDPELDLGTGLRYDHGKFYASTSFYDPAKRRVLYHGVEVDSKRDVYKGGASIQSV														
	391	400	410	420	430	440	450	460	470	480	490	500	510	520	
AFK29573.1	PRIATALDEKTRTNLLLPVEEIEIETLRNATELSDVTLNTGVSYTHIPLRQGTQDIETFHLDASAVAAALNEADVGYNCSSGGAVNRGALGPFGLLVLRAAGDRRGEQATRYFYVSRGLDGLLTSFCODE														
ACI43225.3	PRIATALDEKTRTNLLLPVEEIEIETLRNATELSDVTLNTGVSYTHIPLRQGTQDIETFHLDASAVAAALNEADVGYNCSSGGAVNRGALGPFGLLVLRAAGDRRGEQATRYFYVSRGLDGLLTSFCODE														
XP_020156722.1	PRIATALDEKTRTNLLLPVEEIEIETLRNATELSDVTLNTGVSYTHIPLRQGTQDIETFHLDASAVAAALNEADVGYNCSSGGAVNRGALGPFGLLVLRAAGDRRGEQATRYFYVSRGLDGLLTSFCODE														
ACH73192.1	PRIATALDEKTRTNLLLPVEEIEIETLRNATELSDVTLNTGVSYTHIPLRQGTQDIETFHLDASAVAAALNEADVGYNCSSGGAVNRGALGPFGLLVLRAAGDRRGEQATRYFYVSRGLDGLLTSFCODE														
ACH73194.1	PRIATALDEKTRTNLLLPVEEIEIETLRNATELSDVTLNTGVSYTHIPLRQGTQDIETFHLDASAVAAALNEADVGYNCSSGGAVNRGALGPFGLLVLRAAGDRRGEQATRYFYVSRGLDGLLTSFCODE														
ACH73193.1	PRIATALDEKTRTNLLLPVEEIEIETLRNATELSDVTLNTGVSYTHIPLRQGTQDIETFHLDASAVAAALNEADVGYNCSSGGAVNRGALGPFGLLVLRAAGDRRGEQATRYFYVSRGLDGLLTSFCODE														
Consensus	PRIATALDEKTRTNLLLPVEEIEIETLRNATELSDVTLNTGVSYTHIPLRQGTQDIETFHLDASAVAAALNEADVGYNCSSGGAVNRGALGPFGLLVLRAAGDRRGEQATRYFYVSRGLDGLLTSFCODE														
	521	530	540	550	560	570	580	590	600	610	621				
AFK29573.1	LRSRRAKDVTKRVIIGSTVPPVLDGEAFSHRVLVDHSIVQGFAMGGRTTHTSRVYPHERYQEAKEYLFNNAITGASVMAERLVVHEHDSAHNQLSNDDHYSYVQ														
ACI43225.3	LRSRRAKDVTKRVIIGSTVPPVLDGEAFSHRVLVDHSIVQGFAMGGRTTHTSRVYPHERYQEAKEYLFNNAITGASVMAERLVVHEHDSAHNQLSNDDHYSYVQ														
XP_020156722.1	LRSRRAKDVTKRVIIGSTVPPVLDGEAFSHRVLVDHSIVQGFAMGGRTTHTSRVYPHERYQEAKEYLFNNAITGASVMAERLVVHEHDSAHNQLSNDDHYSYVQ														
ACH73192.1	LRSRRAKDVTKRVIIGSTVPPVLDGEAFSHRVLVDHSIVQGFAMGGRTTHTSRVYPHERYQEAKEYLFNNAITGASVMAERLVVHEHDSAHNQLSNDDHYSYVQ														
ACH73194.1	LRSRRAKDVTKRVIIGSTVPPVLDGEAFSHRVLVDHSIVQGFAMGGRTTHTSRVYPHERYQEAKEYLFNNAITGASVMAERLVVHEHDSAHNQLSNDDHYSYVQ														
ACH73193.1	LRSRRAKDVTKRVIIGSTVPPVLDGEAFSHRVLVDHSIVQGFAMGGRTTHTSRVYPHERYQEAKEYLFNNAITGASVMAERLVVHEHDSAHNQLSNDDHYSYVQ														
Consensus	LRSRRAKDVTKRVIIGSTVPPVLDGEAFSHRVLVDHSIVQGFAMGGRTTHTSRVYPHERYQEAKEYLFNNAITGASVMAERLVVHEHDSAHNQLSNDDHYSYVQ														

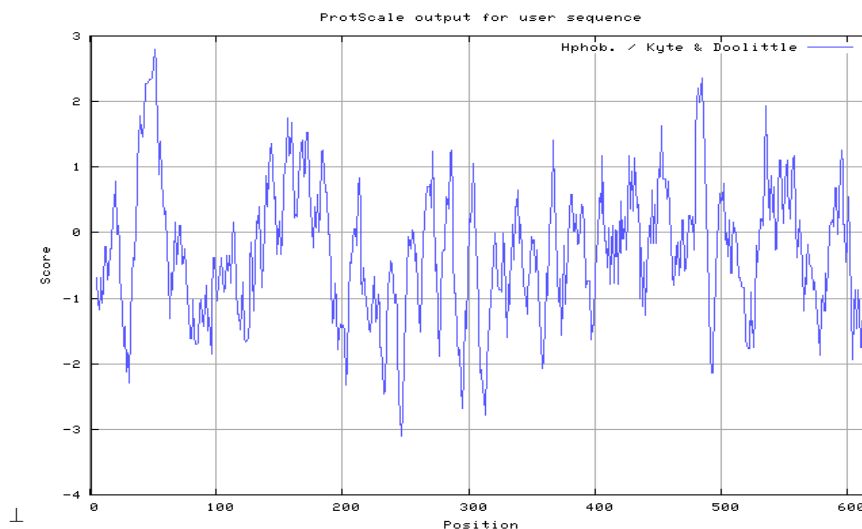
شکل ۹. هم‌ردیفی چندگانه با برنامه MultiAlign برای مقایسه توالی پروتئین 6-SFT گندم با شش توالی پروتئینی مربوط به دیگر گونه‌های ژن 6-SFT (ژن انتخاب شده در این مطالعه) می‌باشد.

تعیین شد (زیرا تعداد اسید آمینه‌های با بار مثبت ۵۳ و تعداد اسید آمینه‌های با بار منفی ۷۷ می‌باشد). از نظر ترکیب اسیدهای آمینه بالاترین فراوانی (۸/۷) درصد) متعلق به آلانین و کمترین فراوانی (صفر درصد) متعلق به دو اسید آمینه پیرولیزین و سلنوسیتئین می‌باشد. در نقطه ایزوالکتریک یک

خصوصیات ساختار اولیه پروتئین و بررسی خواص فیزیکوشیمیایی برای بررسی خواص فیزیکوشیمیایی پروتئین مورد بررسی از ساختار اولیه پروتئین و از برنامه PROTPARM استفاده شد. نقطه ایزوالکتریک (pI) پروتئین مورد نظر ۵/۱۹ و بار خالص آن ۲۴-

است. شاخص ناپایداری پروتئین‌ها بیانگر میزان پایداری آنها در لوله آزمایش در برابر حرارت می‌باشد و این به نوع اسیدآمینه‌های موجود بستگی دارد. شاخص ناپایداری محاسبه‌شده برای پروتئین‌های مورد بررسی نشان می‌دهند که پروتئین‌هایی که شاخص ناپایداری آنها کوچک‌تر از ۴۰ باشند با ثبات و در غیر این صورت بی‌ثبات خواهند بود. بنابراین پروتئین 6-SFT جزو پروتئین‌های پایدار می‌باشد. پروتئین‌های پایدار در شرایط تنش مثل تنش حرارتی می‌توانند ساختار سوم خود و در نتیجه عملکردشان را در شرایط تنش حفظ کنند (Pettersen *et al.*, 2004).
 با برنامه PROTSCALE نیز مشخص شد که پروتئین مورد مطالعه طبق روش کیت و دولیتل در نزدیکی به انتهای N و همچنین نزدیک به انتهای C خود دارای آبگریزی معنی‌داری می‌باشد (با توجه به خط برش برابر ۱ در شکل ۱۰).

آمینواسید و یا پروتئین بار خود را از دست می‌دهد و بنابراین در میدان الکتریکی با جریان مستقیم قادر به حرکت نمی‌باشد. اطلاع از نقطه ایزوالکتریک یک پروتئین می‌تواند برای انتخاب و بهینه‌سازی روش‌های مورد استفاده برای خالص‌سازی پروتئین اعم از کروماتوگرافی تبادل یونی و الکتروفورز ایزوالکتریک فوکوسینگ بسیار مفید باشد. به‌علاوه پروتئین‌هایی که در محدوده pH ایزوالکتریک قرار دارند و مجموع بار مثبت و منفی رشته پپتیدی با یکدیگر برابر بوده، پپتید توانایی نگهداری آب را ندارد و ناگزیر رسوب می‌کند. از طرفی میانگین آب‌گریزی کل آن (GRAVY) -0.298 می‌باشد. بررسی شاخص ناپایداری (Instability indx) نشان می‌دهد که این پروتئین عمر بالایی دارد زیرا شاخص ناپایداری آن بالای $32/96$ است. این موضوع حاکی از آن است که به این پروتئین معمولاً در سلول نیاز



شکل ۱۰. آنالیز آبگریزی پروتئین 6-SFT با برنامه PROTSCALE

مطالعه درنهایت در کجا استقرار خواهد یافت. متوسط GRAVY (خاصیت آبگریزی پروتئین) کل محاسبه‌شده برای پروتئین‌ها از تقسیم مجموع هیدروپاتی محاسبه‌شده برای تمام اسیدهای آمینه در پروتئین بر تعداد کل اسیدهای آمینه آن پروتئین به

این موضوع نشان می‌دهد که این پروتئین پس از سنتز در سیتوزول به‌واسطه خاصیت آبگریزی نمی‌تواند در فضای سیتوزولی بماند و باید به جای دیگری در سلول یا خارج از سلول نقل مکان کند. اما با این آنالیزها هنوز نمی‌توان گفت که پروتئین مورد

دست می‌آید.

در بررسی GRAVY پروتئین این نتیجه مشخص شد که پروتئین *6-SFT* آبگریز می‌باشد. متوسط هیدروپاتی کل (GRAVY) نشان‌دهنده شاخص آب‌گریزی پروتئین می‌باشد. در صورتی که GRAVY محاسبه شده برای یک پروتئین منفی باشد، بدین معنی است که آن پروتئین غیرقطبی است و در صورت مثبت بودن آن قطبی محسوب می‌شود. GRAVY محاسبه شده برای پروتئین *6-SFT* منفی می‌باشد و نشان‌دهنده غیرقطبی بودن آن می‌باشد. حالیت پروتئین‌های قطبی بیشتر از حالیت پروتئین‌های غیرقطبی در مواد حلال می‌باشد. در برابر حلال‌های ویژه هر پروتئین، پروتئین‌های غیرقطبی در وسط و پروتئین‌های قطبی در قسمت بیرونی تجمع پیدا می‌کنند. از پروتئین‌های غیرقطبی می‌توان به‌عنوان اهدافی برای انتقال دارو به نقطه مورد نظر استفاده کرد، با افزودن سیگنال خاص به پروتئین برای انتقال به قسمت مورد نظر و قرار دادن ماده هدف (دارو) در پروتئین این طراحی صورت می‌گیرد.

درخت فیلورنی براساس توالی پروتئین *6-SFT* نشان داد ۲۱ توالی در ۸ گروه قرار گرفتند. درخت فیلورنتیکی بر اساس شباهت توالی‌ها، زمان اشتقاق گونه‌ها را مشخص می‌کند و گونه‌ها را بر اساس شباهت و تفاوتشان در یک گروه یا گروه‌های مختلف قرار می‌دهد. همچنین شاخص آلفاتیک این پروتئین ۷۶/۲۵ می‌باشد که نشان‌دهنده عدم پایداری این پروتئین در برابر حرارت می‌باشد. شاخص آلفاتیک در واقع عبارت است از حجم نسبی پروتئین که با زنجیره‌های آلفاتیک اشغال شده‌اند (A, V, I, L) و به‌عنوان یک فاکتور مثبت در افزایش ثبات حرارتی پروتئین‌های گلوبول مانند محسوب می‌شود و به همین دلیل در باکتری‌های مقاوم به گرما، پروتئین‌ها دارای شاخص آلفاتیک بالایی می‌باشند. پروتئین‌هایی با شاخص آلفاتیک بسیار

بالا (بالتر از ۱۰۰) ممکن است در دامنه دمایی بسیار بالایی از خود ثبات نشان دهند. چون شاخص آلفاتیک پروتئین مور مطالعه کمتر از ۱۰۰ می‌باشد، در نتیجه این پروتئین در برابر حرارت پایداری بالایی ندارد و در دمای بالا قادر به حفظ ساختار سوم خود نمی‌باشد.

محل استقرار و نقش مولکولی پروتئین در سلول
به منظور تعیین محل قرارگیری پروتئین پس از تولید در سلول از برنامه UNIPROT استفاده شد. با این برنامه وزن مولکولی پروتئین *6-SFT* معادل ۶۸ کیلودالتون تعیین شد و از نظر هستی‌شناسی (Ontology) مشخص شد که پروتئین *6-SFT* در داخل سلول نمی‌ماند و به بیرون از سلول نقل مکان می‌کند. همچنین با جستجو در پایگاه InterPro نقش مولکولی پروتئین مورد مطالعه شناسایی شد و مشخص شد که این پروتئین قابلیت هیدرولازی (هیدرولیز ترکیبات O-glycosyl) دارد.

آنالیز وجود توالی‌های سیگنال

برای انجام آنالیز وجود توالی‌های سیگنال از برنامه SignalP استفاده شد. توالی‌های سیگنال مشخص می‌کند که پروتئین تولیدشده به کدام قسمت از سلول منتقل شود. نتایج این برنامه نشان می‌دهد که پروتئین *6-SFT* دارای توالی سیگنالی به منظور عبور از غشای سلولی است و در واقع پروتئین ترشعی می‌باشد. این نتیجه مشابه نتایج قبلی می‌باشد و آشکار کرد که این پروتئین بعد از سنتز در سیتوپلاسم باید به هسته منتقل شود. اما برای انتقال به هسته وجود توالی سیگنال استقرار هسته‌ای (NLS) در انتهای C آن ضرورت دارد (Murakami et al., 2002; Kojima et al., 2005). برای تأیید این موضوع نیاز به مطالعات بیشتری می‌باشد.

آنالیز رهگیری محل استقرار نهایی پروتئین

آنالیز توالی پروتئین با برنامه MitoProt II نشان می‌دهد که احتمال انتقال پروتئین *6-SFT* پس از تولید و تغییرات پس از ترجمه به میتوکندری بسیار پایین (۰/۰۰۴۵) است. همچنین آنالیز توالی پروتئین با برنامه TargetP نشان می‌دهد که احتمال انتقال پروتئین *6-SFT* به میتوکندری، کلروپلاست و مسیر ترشحی بسیار پایین است. بنابراین استنباط می‌شود که این پروتئین به میتوکندری منتقل نمی‌شود و این نتیجه یک بار دیگر تأییدکننده نتایج قبلی توسط برنامه SignalP مبنی بر داخل سلولی بودن پروتئین و نیز استقرار آن در هسته سلول می‌باشد.

شناسایی موتیف‌ها و دمین‌های حفاظت‌شده در پروتئین

به منظور شناسایی موتیف‌ها و دمین‌های حفاظت‌شده در پروتئین از برنامه جامع InterProScan استفاده شد. نتیجه آنالیز با این برنامه نشان داد که پروتئین *6-sft* دارای دو دمین حفاظت‌شده (PF00251 Glyco-) (hydrolysis) (پیوند گلیکوزیدی بین یک یا چند کربوهیدرات را هیدرولیز می‌کند) و SSF75005 (Arabinana) (در تخریب پکتین نقش دارد) در انتهای آمینی و یک دمین SSF49899 (Cancanava) (هیدرولیز گلوکز) در انتهای کربوکسیلی می‌باشد. همچنین آنالیز هستی‌شناسی ژن (GO) با این برنامه نشان داد که نقش مولکولی این پروتئین هیدرولیز ترکیبات o-glycosyl می‌باشد.

بررسی دقیق‌تر توالی پروتئین *6-fst* برای شناسایی موتیف‌های حفاظت‌شده در این توالی با برنامه CD-search نشان داد که پروتئین *6-SFT* به ابرخانواده Glyco-32 تعلق دارد. لازم به ذکر است که این سه موتیف ذکر شده در بالا معمولاً در پروتئین‌های درگیر

در هیدرولیز ترکیبات O-glycosyl یافت می‌شود. این دمین از ۱۳۰ اسید آمینه تشکیل شده است.

همردیفی چندگانه توالی پروتئین *6-SFT* با چندین پروتئین درگیر در هیدرولیز O-glycosyle در گیاهان مختلف

همردیفی چندگانه توالی پروتئین *6-fst* (اکسشن JQ728011) با توالی پروتئینی چندین ژن درگیر در هیدرولیز O-glycosyle در گیاهان مختلف با برنامه MultiAlign به منظور بررسی دمین‌های حفاظت‌شده آن‌ها انجام شد. نتیجه این آنالیز نشان داد که دمین حفاظت‌شده O-glycosyle در انتهای آمینی در ژن *6-sft* وجود دارد (شکل ۱۱). نتایج این آنالیز حاکی از آن است که ژن *6-sft* در هیدرولیز O-glycosyl نقش دارد و به واسطه داشتن سیگنال استقرار هسته‌ای (NLS) در پایانه آمینی خود، بعد از ترجمه و تغییرات پس از ترجمه به هسته منتقل می‌شود.

آنالیز هضم با تریپسین

برای این منظور از برنامه PeptideCutter استفاده شد که نشان می‌دهد پروتئین *6-sft* پس از هضم با آنزیم تریپسین به ۵۱ قطعه پپتیدی تقسیم می‌شود (شکل ۱۲). این نتیجه غیرحساسیت‌زا بودن این پروتئین را تأیید کرد. اگر پروتئین مورد مطالعه توسط آنزیم‌های برشی، برش نخورد آن پروتئین باعث حساسیت می‌شود.

تغییرات پس از ترجمه

برای پیدا کردن الگوهای شناخته‌شده و پیش‌بینی تغییرات پس از ترجمه از برنامه ScanProsite استفاده شد.

4EQV_E	12	HFTFNGKGMNDPGLWY.	[2]	.KDAKWHLYFYQYNP.	[9]	.FWGHATSDDLTNWEDQPIAI.	[1]	.PK.	[1]	.N.	[3]	.AF	82
gi 81744289	37	HYNTPKGFMDIQTIFK.	[1]	.SDGYHMYLLNS.	[9]	.EYHVHTKDEHFENLGVAI		PK		F.	[5]	.AV	106
gi 81668822	48	HFTFPDKWKNDPQKPIY		FDGKYHYFYLYNR.	[8]	.EWRHAVSEDLVHWTEGVAI.	[1]	.KY.	[3]	.D.	[2]	.IW	116
gi 81690869	46	HLTTPDKWKNDPQKPIY		FNGEYHYFYLYNR.	[8]	.EWRHAVSEDLVHWQDKGVAI.	[1]	.KY.	[3]	.N.	[2]	.PW	114
gi 81740874	150	HFSPLYGWMNDPGLWY		KDGEYHLFYQYNP.	[8]	.NWHGHAISKDLVNWHRFVAI.	[1]	.PD.	[3]	.T		.IF	216
gi 81314294	137	HHTFLYGWMNDANGLVY		KDGEYHLFYQYNP.	[8]	.HWGHSVSKDLVHWEHLEPAL.	[1]	.RD.	[3]	.H		.IF	203
gi 81365440	20	HFSTPVGNLADPGLVLF		YKGEYHLFHQKNG		TWAHAVSKDLLHWEHLEPAL.	[1]	.RD.	[3]	.Q		.AL	78
gi 657340881	33	HFTFDQYWMNDPGLIK		IGSTWHLFFQHNP.	[8]	.CWGHATSTDLMMHAKPTAI.	[1]	.DE.	[3]	.E		.AF	99
gi 81690116	55	HYTFRNWMNDPGLVY		ENGVYHLFYQYNP.	[8]	.SWGHAISRDLVHWDEQPVAM.	[1]	.AN.	[3]	.D		.IF	121
gi 209531372	64	HFSFAAYWMNDPGLIL		LDGVYHLFYQYAP.	[8]	.SWGHAISTDLLHWTEHGVAI.	[1]	.AT.	[3]	.E		.IF	130

شکل ۱۱. مقایسه توالی پروتئین ژن *6-sft* با توالی پروتئینی چندین ژن درگیر در هیدرولیز o-glycosyl در گیاهان مختلف. باکس‌ها (انتهای N در انتهای آمینی هر کدان از ژن‌های مورد بررسی) در پایانه آمینی مربوط به دمین موتیف Glycosyl hydrolyse

Name of enzyme	No. of cleavages	Positions of cleavage sites
Trypsin	51	28 33 35 59 83 92 133 139 199 200 203 219 222 237 243 258 260 274 277 292 305 312 333 342 347 359 360 361 362 374 375 380 390 397 399 414 436 475 491 492 504 520 523 525 529 530 547 563 569 580 596

شکل ۱۲. بررسی نحوه هضم پروتئین 6-fst با تریپسین بوسیله برنامه PeptideCutter (بخشی از نتایج نشان داده شده است).

یک جایگاه cAMP و cGMP وابسته به پروتئین کیناز فسفریلیشن^۳ و یک جایگاه پروتئین کیناز فسفریلیشن^۴ می‌باشد. اگر این تغییرات پس از ترجمه صورت نگیرد پروتئین کارایی خود را به دست نخواهد

با این برنامه جایگاه‌های تغییرات احتمالی بعد از ترجمه کلاً در چهار کلاس دسته‌بندی شده‌اند که عبارتند از ۱۰ جایگاه اتصال فسفوترئونین، ۸ جایگاه اتصال فسفوسرین، ۵ جایگاه اتصال -N- گلیکولیزیشن^۱ و ۹ جایگاه اتصال -N- مریستولیزیشن^۲.

- 2. N-myristoylation
- 3. cGMP-dependent protein kinase phosphorylation
- 4. Tyrosine kinase phosphorylation

- 1. N-glycosylation

به سختی داشته و انعطاف‌پذیری ناچیزی دارند. پیوندهای هیدروژنی بین‌رشته‌ای که میان گروه‌های CO یک رشته بتا و NH رشته بتای مجاور ایجاد می‌شوند، به صفحات بتا پایداری می‌بخشند و باعث می‌شوند که این صفحات ظاهری زیگزاگ داشته باشند. هرچه مارپیچ آلفا و بتا بیشتر باشد انعطاف‌پذیری پروتئین کمتر و پایداری آن بیشتر است. نتایج نشان داد پایداری پروتئین 6-SFT زیاد نیست.



شکل ۱۳. ساختار پیشنهادی برای پروتئین 6-sft با برنامه Phyre2

GOR4 :			
Alpha helix	(Hh) :	148 is	23.91%
3 ₁₀ helix	(Gg) :	0 is	0.00%
Pi helix	(Ii) :	0 is	0.00%
Beta bridge	(Bb) :	0 is	0.00%
Extended strand	(Ee) :	168 is	27.14%
Beta turn	(Tt) :	0 is	0.00%
Bend region	(Ss) :	0 is	0.00%
Random coil	(Cc) :	303 is	48.95%
Ambiguous states (?)	:	0 is	0.00%
Other states	:	0 is	0.00%

شکل ۱۴. آنالیز ساختار دوعدی پروتئین 6-sft با نرم افزار GORIV

ساختار سوم پروتئین

مدل‌سازی ساختار سه بعدی بر اساس انتخاب یک

آورد.

ساختار دوم پروتئین

با استفاده از نرم‌افزار آنالیز Phyre2 ساختار دوم پروتئین 6-sft پیش‌بینی شد. نتایج حاصل حاکی از آنست که ۶ ساختار مشابه در دیتابیس NCBI برای این پروتئین پیدا شد، یکی از این ساختارها، ساختار کریستالی *pachysandra terminalis* است که دارای شباهت ۱۰۰ با پروتئین 6-sft است. ساختار پیش‌بینی شده در شکل ۱۳ نشان داده شده است (شکل ۱۳). همچنین با نرم‌افزار GORIV مشخص شد که پروتئین 6-sft حاوی حدود ۲۴٪ مارپیچ آلفا هلیکس است و این اطلاعات به منظور بررسی پایداری پروتئین دارای اهمیت می‌باشد (شکل ۱۴). معمول‌ترین مکان برای مارپیچ آلفا سطح هسته‌های پروتئین است که یک رابط را برای تعامل با محیط آبی بیرون ایجاد می‌کند. قسمت داخلی مارپیچ تمایل به داشتن آمینواسیدهای آب‌گریز و قسمت بیرون تمایل به داشتن آمینواسیدهای آب‌دوست دارد، بنابراین در طول مارپیچ از هر چهار آمینواسید، سه آمینواسید آب‌گریز خواهند بود. سایر آمینواسیدهای موجود در هسته پروتئین و یا داخل غشای سلولی خاصیت آب‌گریزی دارند. به طور کلی مارپیچ‌های قرار گرفته در سطح، تعداد کمتری آمینواسید آب‌گریز دارند. از این ویژگی می‌توان در پیش‌بینی ساختار پروتئین‌ها کمک گرفت. برای مثال نواحی با مقادیر بیشتر از آلانین، گلوتامین اسید، لوسین و متیونین و مقادیر کمتر از پرولین، گلیسین، تیروزین و سرین تمایل به تشکیل مارپیچ آلفا دارند. ساختار صفحه‌های بتا، ساختار دوم بسیار کشیده و چین‌دار می‌باشد. یکی از تفاوت‌های مهم صفحه‌های بتا با مارپیچ آلفا این است که اسیدآمینو‌هایی که معمولاً در ساختار اول زنجیره پروتئینی با فاصله زیاد از هم قرار گرفته‌اند، برای تشکیل این ساختار در مجاورت یکدیگر قرار می‌گیرند، بنابراین صفحه‌های بتا تمایل

(Brown *et al.*, 2007) (شکل ۱۵).

الگوی با شباهت بالا با پروتئین هدف با استفاده از پایگاه اطلاعاتی Swiss Model انجام گرفت



شکل ۱۵. ساختار مدل‌سازی شده *6-fst* گندم در Swiss Model

جدول ۲. میزان همسانی اسیدآمینو پروتئین هدف با الگو

Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Sq Similarity	Range	Coverage	Description
4v7e	60%	Monomer	BLAST	ELECTRON MICROSCOPY	5.5 Å	0.33	1-389	81%	40S ribosomal protein S29

RNA در ریبوزوم در گیاه گندم می‌باشد. ساختار سه بعدی پروتئین‌های مدل‌سازی شده در شکل ۱۶ نشان داده شده است (شکل ۱۶).

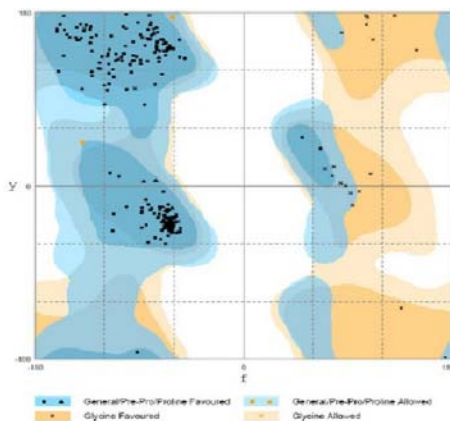
ارزیابی مدل‌های پیشنهادشده در Swiss Model براساس پارامترهای QMEAN و GMQE می‌باشد. تابع QMEAN4 یک تابع حسابی مرکب برای ارزیابی کیفیت global (کل ساختار) و local (هر اسید آمینه) ساختار مدل‌سازی شده می‌باشد. نمودار کیفیت Local برای هر اسید آمینه در مدل (محور X) شباهت مورد انتظار به ساختار بومی (محور Y) را نشان می‌دهد. به طور معمول، اسید آمینه‌هایی که امتیاز زیر ۰/۶ را نشان بدهند، از کیفیت پایینی برخوردارند. تابع QMEAN4 شامل چهار توصیف ساختاری پتانسیل زاویه چرخش، پتانسیل فاصله اتم‌ها، اثر متقابل کربن بتا و پتانسیل حلال‌پوشی می‌باشد. GMQE یک ارزیابی کیفی است که خصوصیات هم‌ترازی مدل الگو را ترکیب می‌کند. نمودارهای ارزیابی کیفی موضعی ساختار و میزان Z-

الگوی منتخب برای مدل‌سازی پروتئین *6-sft* در گندم (4v7e) دارای ۱۳۸ اسید آمینه بوده و در سال ۲۰۱۴ با قدرت تفکیک ۵/۵ آنگستروم از *Triticum aestivum* بوسیله میکروسکوپ الکترونی کشف شده است. میزان یکسانی الگوی 4v7e با پروتئین هدف ۸۰٪ و میزان شباهت آن ۳۳/۶۷٪ می‌باشد (جدول ۲). بر اساس خصوصیات مدل پیش‌بینی‌شده، می‌توان خصوصیات پروتئین مورد مطالعه را پیش‌بینی کرد.

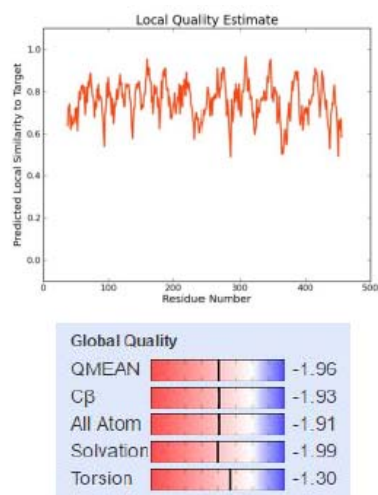
با توجه به میزان شباهت پایین الگوی 4v7e با پروتئین هدف این احتمال وجود دارد که مدل‌سازی به صورت ضعیف‌تری انجام شده باشد، اما باید خاطر نشان کرد که با توجه به اهمیت این پروتئین در کاتالیز انتقال گروه فروکتوزیل از ساکارز به گیرنده‌های مختلف که در تنش‌هایی مانند خشکی نقش اساسی دارد، اولین گام در انجام مطالعات بعدی جهت مطالعه جایگاه اتصال و یا لیگاندهای مهاری و یا فعال‌کننده، قطعاً مدل‌سازی این پروتئین می‌باشد. 4v7e بخشی از ساختار small

انجام شده است.

در روش مشابهت مدلینگ یا مدل‌سازی مقایسه‌ای، ساختمان پروتئین براساس مشابهت توالی با ساختمان‌های شناخته‌شده با روش‌های تجربی پیش‌بینی می‌شود. در واقع این روش بر این اصل استوار است که اگر دو پروتئین مشابهت توالی بالایی داشته باشند، احتمالاً ساختمان سه‌بعدی بسیار مشابهی دارند. ساختار سه‌بعدی پروتئین منبع مهم اطلاعاتی برای درک بهتر عملکرد پروتئین و اینترکشن آن با اجزای دیگر (لیگندها، پروتئین و ...) می‌باشد. به‌علت سختی و پرهزینه بودن فرآیند کریستالوگرافی به‌خصوص در مورد پروتئین‌های مهمی مانند *SFT-6* که نقش‌های حیاتی در مسیرهای شناخته‌شده دارند، پیشگویی ساختار آن از طریق ابزار *In silico* به منزله میانبری برای مطالعات بعدی و بررسی جزئیات ساختاری آن خواهد بود. به‌عبارت دیگر طراحی یک لیگاند مؤثر برای فعال‌سازی یا مهار پروتئین در یک مسیر خاص، با شبیه‌سازی ساختار سوم آن پروتئین مقرون‌به‌صرفه بوده و روند مطالعات را تسریع می‌کند. با توجه به نتایج ارزیابی کیفی مدل‌های ایجادشده، می‌توان با ضریب اطمینان بالایی از این مدل‌ها در آنالیزها و طراحی‌های بعدی استفاده نمود.



score در مورد تابع حسابی مرکب QMEAN4، تمامی اتم‌ها، کربن بتا، حلالیت و زوایای چرخشی در شکل ۱۶ نشان داده شده است (شکل ۱۶). ارزیابی مدل به منظور بررسی کیفیت مدل پیش‌بینی است و هر چه کیفیت مدل پیش‌بینی شده بالاتر باشد، خصوصیات آن به خصوصیت پروتئین مورد بررسی نزدیک‌تر است. ارزیابی کیفیت ساختار مدل شده پس از بهینه‌سازی انرژی به وسیله نمودار رامچاندران با استفاده از ابزار Rampage صورت گرفت (Schwede *et al.*, 2003). نمودار رامچاندران یک روش بسیار پرکاربرد جهت پلات زوایای چرخشی پروتئین و ارزیابی کیفیت استرئوشیمیایی ساختار مدل‌سازی شده می‌باشد.



شکل ۱۶. نمودار ارزیابی کیفی موضعی ساختار و نمودار نمودار رامچاندران با محاسبه درصد اسیدهای آمینه قرار گرفته در نواحی مطلوب، مجاز و یا غیرمجاز کیفیت مدل ورودی را تعیین می‌نماید. چنانچه در شکل ۱۷ دیده می‌شود ۹۹/۲٪، ۰/۸٪ و ۵۰ اسید آمینه برای ساختار مدل شده در گندم به ترتیب در مناطق مطلوب، مجاز و غیر مجاز قرار دارند و این بدان معنی است که هر مدل از کیفیت بالایی برخوردار است (شکل ۱۷). هرچه درصد اسیدهای آمینه در مناطق مطلوب بیشتر باشد، آن مدل کیفیت بهتری دارد و مدل‌سازی به‌درستی

با توجه به اهمیت نقش ساکارز ۶ فروکتوزیل ترنسفرز در انتقال گروه فروکتوزیل از ساکارز به گیرنده‌های مختلف و نقش آن‌ها در تنش‌هایی مانند خشکی در این تحقیق به بررسی بیوانفورماتیکی این ژن پرداخته شد. با بررسی جامعی که صورت گرفت، تاکنون هیچ منبع انگلیسی یا فارسی در رابطه با بررسی بیوانفورماتیکی این ژن منتشر نشده است و اطلاعات بیوانفورماتیکی در رابطه با این ژن وجود ندارد. از نتایج این تحقیق می‌توان در مطالعات بعدی مانند بررسی برهمکنش بین پروتئین‌ها و داکینگ مولکولی استفاده کرد.

شکل ۱۷. نمودار رامچاندرا ساختار 6-sft مدل‌سازی شده در گندم. نواحی آبی رنگ به معنی قرارگیری ایده آل اسیدهای آمینه در نواحی کرم رنگ به معنی نواحی مجاز، و نواحی زرد کم رنگ نواحی غیر مجاز می‌باشند.

نتایج این مقاله می‌تواند در مطالعه رفتار و کنش پروتئین 6-SFT در مسیر درگیر در هیدرولیز O-glycosyle در گیاهان مختلف بسیار کمک‌کننده باشد و به فهم چگونگی نحوه برهمکنش این پروتئین با سایر پروتئین‌ها کمک کرده و زمینه‌ساز روشن شدن نحوه فعالیت آن در سلول‌ها باشد (Labadorf et al., 2011; Misumi et al., 2010).

REFERENCES

- Bancal, P., Carpita, NC., & Gaudillère, JP. (1992). Differences in fructan accumulated in induced and field-grown wheat planis: an elongation-trimming pathway for their synthesis. *New Phytol*, 120, 450-457.
- Bancal, P., Gibeaut, DM., & Carpita NC. (1993). Analytical methods for the determination of fructan structure and biosynthesis. In M Suzuki, NJ Chatterton, eds, *Science and Technology of Fructans*. CRC Press, 32, 81-118.
- Brown, AK., Sridharan, S., Kremer, L., Lindenberg, S., Dover, LG., Sacchettini, JC., & Besra, GS. (2005). Probing the mechanism of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *J. Biol. Chem.*, 18, 54-63.
- Burge, C., & Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *J Mol. Biol*, 268, 78-94.
- Cao, H. (2011). Structure-function analysis of diacyl-glycerol acyltransferase sequences from 70 organism. *BMC Research Notes*, 4, 249-279.
- Carlson, CS., Thomas, DJ., Eberle, MA., Swanson, JE., Livingston, RJ., Rieder, MJ., & Nickerson, DA. (2005). Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Research*, 13, 120-131.
- Fallah Ziarani, M., Masoud, T., & Zahra, A. (2017). Bioinformatic analysis of Acyl Carrier Protein (ACP) in eukaryotes and prokaryotes. *Crop Biotech*, 17, 15-29.
- Ferguson, AA., Jiang, N., & Pack, M. (2011). Recycling and reshaping genes through GC-biased acquisition. *Mobile Genetic Elements*, 15, 2134-2141.
- Guex, N., & Peitsch, MC. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *lectrophoresis*, 18 (15), 2714-2723.
- Hasegawa, PM., Bressan, RA., Zhu, K., & Bohnert, HJ. (2000). Plant cellular and molecular responses to high salinity. *Annual Review of Plant Physiology and Plant Molecular Biology*, 51, 463-99.
- Jenks, MA., Hasegawa, PM., & Jain, SM. (2007). *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*. Springer. *The Netherlands*.

- Jun-Zhi, W., Jerry, C., & Steve, RL. (2001). Expression of sucrose: fructan 6 fructosyltransferase (6-SFT) and *myo*-inositol 1-phosphate synthase (MIPS) genes in barley (*Hordeum vulgare*) leaves. *Journal of Plant Physiology*, 158, 635- 643.
- Karim, MA., Nawata, E., & Shigenaga, S. (1993). Effect of salinity and temperature on yield, mineral ion concentrations and physiology in hexaploid triticale (X Triticosecale Wittmack). *Japenese Journal of Crop Science*, 62, 419-428.
- Kojima, S., Takahashi, Y., Kobayashi, Y. Monna, L., Sasaki, T., Araki, T., & Yano, M. (2002). *Hd3a*, a rice ortholog of the Arabidopsis *FT* gene, promotes transition to flowering downstream of *Hdl* under short-day conditions. *Plant Cell Physiol*, 43, 1096-1105.
- Labadorf, A., Link, A., & Rogers, MF. (2010). Genome-wide analysis of alternative splicing in *chlamydomonas reinhardtii*. *BMC Genomics*, 11, 114-124.
- Lovell, SC. Davis, IW. Arendall, WB. de Bakker, PI. Word, JM. Prisant, MG., & Richardson, DC. (2003). Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins: Structure, Function, and Bioinformatics*, 50 (3), 437-450.
- Misumi, C., Yoshida, Y., & Nishida, K. (2011). Genone analyses and its significance in four unicellular algae and *Thalassiosira pseudonana*. *J. Plant. Res*, 121, 3-17.
- Murakami, M., Matsushika, A., Ashikari, M., Yamashino, T., & Mizuno, T. (2005). Circadian-associated rice pseudo response regulators (*OsPRRs*): insight into the control of flowering time. *Biosci. Biotechnol. Biochem*, 69, 410- 414.
- Nathalie, D., Karlheinz, B., Urs, S., Andres, W., & Pierre, B. (1995). Sucrose: Fructan 6=Fructosyltransferase, a Key Enzyme for Diverting Carbon from Sucrose to Fructan in Barley Leaves. *Plant Physiol*, 107, 1249- 1255.
- Pettersen, EF., Goddard, TD., Huang, CC., Couch, GS., Greenblatt, DM., Meng, EC., & Ferrin, TE. (2004). UCSF Chimera-avisualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605-1612.
- Salekdeh, GH., Siopongco, J., Wade, LJ., Ghareyazie, B., & Bennett, J. (2002). A proteomics approach to analyzing drought- and salt- responsiveness in rice. *Field Crops Research*, 76, 199-219.
- Schwede, T., Kopp, J., Guex, N., & Peitsch, MC. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, 31, 3381- 3385.