



## Isolation, Cloning, and Bioinformatic Analysis of UGT Gene in Saffron (*Crocus sativus* L.) of Iran

Mohammad Javad Habibzadeh<sup>1</sup>, Seyed Mahdi Ziaratnia<sup>2\*</sup> and Ebrahim Dorani-Uliaie<sup>3</sup>

### Article type:

Research Article

### Article history:

Submitted 23 June 2023

Revised: 24 August 2023

Accepted: 27 September 2023

Available Online: 23 October 2023

### How to cite this article:

Habibzadeh, M.J., Ziaratnia, S.M. and Dorani-Uliaie, E. (2023). Isolation, Cloning, and Bioinformatic Analysis of UGT Gene in Saffron (*Crocus sativus* L.) of Iran. *Saffron Agronomy & Technology*, 11(3), 281-300.

DOI: 10.22048/jsat.2023.403704.1492

### Abstract

Apocarotenoids play a pivotal role in the economic value of saffron plants, underscoring the significance of isolating and studying the genes implicated in apocarotenoid metabolism. This research endeavor focused on the CsUGT gene, a key player in crocin biosynthesis. The gene was subjected to isolation, homogenization, and subsequent transfer into the *E. coli* strain DH5 $\alpha$ . The gene underwent cloning, sequencing, and registration within the NCBI database. A multi-step analysis was initiated to gain a deeper understanding of the attributes inherent to this gene. Firstly, the protein sequence of CsUGT was obtained. Subsequently, several servers and tools, including ProtParam, SOPMA, ProtScale, Pfam, ProtComp, SignalP, TMHMM, and ChloroP, were employed to delve into the protein's physical, chemical, and physiological characteristics. Moreover, the Swiss-Model server was leveraged to investigate the protein's three-dimensional structure. The outcome of this exploration included the construction of a Ramachandran diagram, which effectively validated the 3D model structure produced. According to the results, CsUGT protein with 462 amino acids has the conserved sequence of glycosyltransferase family proteins and was identified as a polar protein, stable at high temperatures and without a hydrophobic domain. CsUGT protein has no peptide signal or binding signals and has a cytoplasmic location. This research made it possible to isolate the saffron's CsGTS gene and optimized its transfer conditions into the vector. In addition, the results of CsUGT protein structure analysis provide the basis for future functional studies and can provide valuable information regarding the behavior and reaction of this enzyme in the synthesis of saffron apocarotenoids; also, these results can be useful in future programs of Iranian saffron genetic modification.

**Keywords:** *Crocus sativus*, CsUGT, Apocarotenoid, Phylogenetic analysis, Bioinformatics, Modeling.

1 -Graduated with a PhD in Agricultural Biotechnology, Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

2 -Assistant Professor, Department of Food Biotechnology, Research Institute of Food Science and Technology (RIFST), Mashhad, Iran

3 - Associate Professor, Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz.



Corresponding author: [m.ziaratnia@rifst.ac.ir](mailto:m.ziaratnia@rifst.ac.ir)

## مقاله پژوهشی

# همسانه‌سازی و بررسی بیوانفورماتیکی ژن *UGT* در زعفران (*Crocus sativus*) ایران

محمد جواد حبیب زاده<sup>۱</sup>، سید مهدی زیارت‌نیا<sup>۲\*</sup> و ابراهیم دورانی علیایی<sup>۳</sup>

تاریخ دریافت: ۲ تیر ۱۴۰۲

تاریخ بازنگری: ۲ شهریور ۱۴۰۲

تاریخ پذیرش: ۱۵ مهر ۱۴۰۲

حبیب زاده، م.ج.، زیارت‌نیا، س.م.، و دورانی علیایی، ا. (۱۴۰۲). همسانه‌سازی و بررسی بیوانفورماتیکی ژن *UGT* در زعفران (*Crocus sativus*) ایران. زراعت و فناوری زعفران، ۳(۱۱): ۲۸۱-۳۰۰.

## چکیده

آپوکاروتنوئیدها عامل ارزش اقتصادی در گیاه زعفران هستند، بنابراین جداسازی و مطالعه ژن‌های درگیر در متابولیسم آپوکاروتنوئیدها از اهمیت ویژه‌ای برخوردار است. در این تحقیق، با توجه به اهمیت ژن *CsUGT* در بیوسنتز کرو سین و به منظور شناخت بیشتر خصوصیات پروتئین *CsUGT*، اقدام به جداسازی، همسانه‌سازی و انتقال این ژن به باکتری *E. coli* سویه DH5 $\alpha$  شد. بدین منظور، ابتدا توالی پروتئینی ژن مورد نظر بدست آمد و سپس ویژگی‌های فیزیکی و شیمیایی و فیزیولوژی پروتئین *CsUGT* توسط سرورها و ابزارهای ProtParam، ProtScale، SOPMA، Pfam، ProtComp، SignalP، TMHMM و ChloroP بررسی شد. همچنین با استفاده از سرور Swiss-Model ساختار سه‌بعدی این پروتئین مورد بررسی قرار گرفت و جهت اعتبار سنجی ساختاری مدل ترسیم شده سه‌بعدی، نمودار رامانچاندرا ترسیم گردید. با توجه به نتایج حاصل، پروتئین *CsUGT* دارای ثبات در دماهای بالا، قطبی و فاقد دمین آب‌گریز می‌باشد. این پروتئین دارای ۴۶۲ اسید آمینه و حاوی توالی حفاظت شده پروتئین‌های خانواده گلیکوزیل ترانسفراز است. پروتئین *CsUGT* فاقد پپتیدهایی با نقش سیگنالی یا سیگنال‌های اتصال دهنده است و جایگاهی سیتوپلاسمی دارد. این تحقیق امکان جدا سازی ژن *CsGTS* زعفران را فراهم و شرایط انتقال آن را به داخل ناقل بهینه نمود. علاوه بر این، نتایج تجزیه و تحلیل ساختار پروتئین *CsUGT* زمینه را برای مطالعات عملکردی آتی فراهم می‌کند و می‌تواند اطلاعات با ارزشی در رابطه با رفتار و واکنش این آنزیم در مسیر سنتز آپوکاروتنوئیدهای زعفران فراهم کند.

**کلمات کلیدی:** آپوکاروتنوئید، کروسین، بررسی فیلوژنتیک، مدل‌سازی.

۱- دانش آموخته دکتری رشته بیوتکنولوژی کشاورزی، گروه به‌نژادی و بیوتکنولوژی گیاهی، دانشکده کشاورزی دانشگاه تبریز، تبریز، ایران.  
۲- استادیار، گروه زیست فناوری مواد غذایی، مؤسسه پژوهشی علوم و صنایع غذایی، مشهد، ایران.  
۳- دانشیار، گروه به‌نژادی و بیوتکنولوژی گیاهی، دانشکده کشاورزی دانشگاه تبریز، تبریز، ایران.  
\* - نویسنده مسئول: [m.ziaratnia@rifst.ac.ir](mailto:m.ziaratnia@rifst.ac.ir)

## مقدمه

کروسين امکان پذير نيست و لذا علاقه زيادي به توليد آن به روش هاي زيست فناوري وجود دارد (Demurtas et al., 2018). آناليز فيلوژنتيک GTهاي گياهي نشان داده است که اين آنزيمها را مي توان بر اساس خصوصيات پيش ماده در زير خانوادههاي متفاوتي تقسيم بندي نمود (Lim et al., 2003). GTها توسط خانواده بزرگ چند ژني کد مي شوند و مي توانند به وسيله شاخص امضا گونه اي که در توالي اوليه آنها وجود دارد شناسايي شوند (Hughes & Hughes, 1994)، که آنها را به زير خانواده شماره یک از GTها طبقه بندي مي کنند.

مسير گليکوزيلاسيون کروستين در سال ۱۹۹۷ توسط دوفرسن و همکاران (Dufresene et al., 1997) پيشنهاد شد. وي براي اولين بار نشان داد گليکوزيلاسيون کروستين منجر به تشکيل کروسين مي شود و ممکن است شامل فعاليت هاي دو آنزيم گليکوزيل ترانسفراز متفاوت باشد (Cote et al., 2000; 2001). اهرازم و همکاران (Ahrazem et al., 2015)، چهار ژن کدکننده آنزيم هاي گليکوزيل ترانسفراز که تحت عنوان *CsUGT2* (*UGT74AD1*) شناخته مي شوند و گليکوزيلاسيون کروستين در کلاله *C. ancyrensis* را کاتاليز مي کنند، جدا نمودند و مشاهده کردند که بيان *UGT74AD2* با تجمع مقادير بالاي کروسين در کلاله ارتباط دارد و نقش آن را در بيوسنتز کروسين پيشنهاد دادند. عليرغم تلاش هاي متعدد، گمز-گمز و همکاران (Gomez-Gomez et al., 2017) قادر به بازيابي توالي ژن *UGTcs2* از رونويس هاي خود نشدند و بجاي آن یک همولوگ نزديک به ژن *CsUGT74AD1* را شناسايي نمودند که بر اساس نتايج ايمنوبلات<sup>۴</sup> یک پروتئين بسيار فراوان در کلاله هاي زعفران است و در کروموپلاست کلاله ها قرار دارد.

ارزش اقتصادي زعفران به عنوان طلای سرخ به کلاله خشک شده، اندام جنسي ماده گل زعفران، مربوط مي باشد (Bathaie et al., 2006). کيفيت کلاله به دليل حضور آپوکاروتنوئيدهاي کروسين، پيکروکروسين و سافرانال است که به ترتيب مسؤل رنگ قرمز، طعم تلخ و رايحه تند زعفران هستند (Caballero-Ortega et al., Xi & Qian, 2006; 2007).

مسير بيوسنتز کروسين از زيزانتين شروع مي شود و شامل سه مرحله اصلي است که توسط سه آنزيم، ديوکسي ژناز جداکننده کاروتنوئيدها<sup>۱</sup> (*CsCCD2*)، آلدهيد دهيدروژناز<sup>۲</sup> (*ALD8*) و *UDP-glucuronosyltransferases* کاتاليز مي شوند (Bouvier et al., 2003; Mai et al., 2012). گليکوزيل تراز سفرازاها<sup>۳</sup> (*GTases*) آنزيم هاي هستند که یک مولکول قند را به یک گيرنده خاص متصل مي کنند، در نتيجه یک پيوند گليکوزيدي ايجاد مي کنند زعفران، گليکوزيلاسيون کروسين، منجر به تبديل کاروتنوئيد محلول در چربي، به کروسين که گليکوزيد آپوکاروتنوئيد پايدار محلول در آب است مي شود (Moraga et al., 2004). در گياه زعفران ايزوفرم هاي مختلفي از آنزيم گليکوزيل ترانسفراز شناسايي شده است و آناليز بيان اين ژن نشان مي دهد که تقريباً در کلاله تمامي گونه هاي زعفران نسخه هاي ژني گليکوزيل ترانسفراز يافت مي شود (Hosseinpour Azad et al., 2020). همانطور که اشاره شد، کروسين فرم گليکوزيله شده کروسين در کلاله هاي زعفران مي باشد (Hosseinpour Azad et al., 2017) ولي به دليل ساختار پيچيده و فراواني مراکز کايرال، از نظر شيميائي سنتز

1- Carotenoid Cleavage Dioxygenase (CCD)

2- Aldehyde Dehydrogenase

3 -Glycosyl Transferase

4 - Immunoblot

گیاهان هنوز ناشناخته است (Ahrazem et al., 2015). با توجه به نقش اساسی گلیکوزیل ترانسفراز (UGT) در مسیر بیوسنتز کروسین در کلاله‌های زعفران، هدف از انجام تحقیق حاضر، برای اولین بار جداسازی، هم‌سانه‌سازی و شناسایی توالی کامل ژن UGT از منابع DNA ژنومی استخراج شده از زعفران ایران و cDNA کدکننده این آنزیم تعیین شد همچنین مقرر گردید با استفاده از نرم‌افزارهای بیوانفورماتیکی، بررسی‌های ساختاری شامل پیش‌بینی جایگاه، عملکرد، ساختمان دوم و سوم پروتئین‌ها، تغییرات پس از ترجمه و نیز ارتباطات خویشاوندی این آنزیم بررسی و تحلیل گردد.

## مواد و روش‌ها

### استخراج DNA ژنومی

نمونه‌های پیاز و برگ زعفران از مزرعه تحقیقاتی مؤسسه پژوهشی علوم و صنایع غذایی جمع‌آوری گردید و در دمای  $70^{\circ}\text{C}$  تا زمان آزمایش نگهداری شدند. DNA ژنومی به روش CTAB تغییر یافته (Beiki et al., 2011) استخراج و با استفاده از نانودراپ (BOECO N-1C, Germany) و ژل آگارز یک درصد کمیت و کیفیت DNA استخراج شده (۵۰-۱۰۰ نانوگرم) مورد بررسی قرار گرفت. به منظور شناسایی توالی‌های ژنی UGT زعفران، اطلاعات مربوط به توالی‌های نوکلئوتیدی DNA و mRNA این ژن از بانک ژن NCBI<sup>۱</sup> استخراج و با استفاده از نرم‌افزار Oligo version 7.60 آغازگر برای ژن فوق‌الذکر طراحی شد (جدول ۱).

ژن UGT با واکنش زنجیره‌ای پلیمرز (PCR)<sup>۲</sup> شامل پنج دقیقه ۹۴ درجه سانتی‌گراد، ۱۰ چرخه دمایی شامل ۲۰ ثانیه ۹۴ درجه سانتی‌گراد، ۳۰ ثانیه ۵۳ درجه سانتی‌گراد (۰/۳) درجه افزایش در هر چرخه، یک دقیقه و ۴۵ ثانیه دقیقه در دمای ۷۲

برای تأیید بیشتر عملکرد UGTs2، دمور تاس و همکاران (Demurtas et al., 2018) رونوشت مورد نظر را در ترنسکرپتومی<sup>۱</sup> که از مطالعات خود و دیگران بدست آمده بود جستجو کردند و ژن جدیدی تحت عنوان *CsUGT74AD1* را پیدا کردند که نسخه دیگری از *UGTcs2* را کد می‌کند و پروتئین *CsUGT74AD1* مسئول گلیکوزیل‌اسیون اولیه کروسین تشخیص داده شد. براساس نتایج بدست آمده *CsUGT74AD1* تشکیل کرو سین را در سیتوپلاسم کاتالیز می‌کند و با توجه به فراوانی بسیار زیاد *CsUGT74AD1* و شباهت آن به *UGTcs2*، پیشنهاد شد که پیتیدهای شناسایی شده توسط گمز-گمز و همکاران (Gómez-Gómez et al., 2017) که به *UGTcs2* نسبت داده شد، متعلق به *CsUGT74AD1* می‌باشند (Demurtas et al., 2018). حسین‌پور آزاد و همکاران (Hosseinpour Azad et al., 2020) به شناسایی و بیان هترولوگ ایزوفرم ژنی کدکننده آنزیم گلیکوزیل ترانسفراز کلاله گیاه زعفران در مرحله گرده افشانی پرداخته و نتایج حاصل از توالی‌یابی و آنالیزهای صورت گرفته نشان داد این ژن به طول ۱۳۶۸ جفت باز و وزن مولکولی ۶۹/۵ کیلوالتون متعلق به خانواده گلیکوزیل ترانسفرازها بوده که بصورت آپوپلاستی در سلول ترشح می‌گردد.

در گیاهی مثل زعفران به علت با ارزش بودن کلاله و ترکیبات آن، که از ترکیبات آپوکاروتنوئیدی می‌باشند، شناخت ژن‌های شرکت کننده در مسیر سنتز آپوکاروتنوئیدها یکی از مسائل مهم در این گیاه است که با روشن شدن این مسیر می‌توان ابتکار عمل در کمیت و کیفیت محصول را در دست گرفت. اگرچه بسیاری از این آنزیم‌ها از چندین گونه گیاهی جدا سازی و در شرایط آزمایشگاهی مورد مطالعه قرار گرفته‌اند، ولی در بسیاری از موارد نقش آنها در متابولیسم ثانویه این

3- Polymerase Chain Reaction

1 - Transcriptom

2- www.ncbi.nih.gov/genbank

شرکت ( AccuZol™ Total RNA Extraction Solution ) از نمونه های کلا له گل زعفران جمع آوری شده در آبان ماه انجام شد و کمیت و کیفیت آن با استفاده از نانودراپ و ژل آگارز بررسی شد. پس از حذف DNA با آنزیم DNase، رشته اول cDNA با بکارگیری آنزیم نسخه بردار معکوس و استفاده از ۲ میکروگرم از RNA استخراج شده و آغازگر های ۱۸ نوکلئوتیدی الیگو dt با استفاده از کیت (Thermo Fisher Scientific) (RevertAID™ M-MuLV) (USA) طبق پروتکل ساخته و در دمای ۷۰°C- ذخیره شد.

درجه سانتی گراد و ۲۵ چرخه دمایی شامل ۳۰ ثانیه ۹۴ درجه سانتی گراد، ۳۵ ثانیه ۵۸ درجه سانتی گراد، یک دقیقه و ۴۵ ثانیه ۷۲ درجه سانتی گراد و ۱۰ دقیقه بسط نهایی در ۷۲ درجه سانتی گراد، از DNA ژنومی با آغازگرهای طراحی شده با استفاده از ترکیب آنزیم های *Taq polymerase* و *Pfu polymerase* تکثیر شدند و در ادامه با استفاده از کیت GeNet Bio بر اساس دستورالعمل شرکت سازنده، از ژل آگارز خالص سازی شد.

#### استخراج RNA کل و سنتز cDNA

استخراج RNA کل با استفاده از کیت استخراج RNA

جدول ۱- آغازگرهای مورد استفاده در بررسی ژن *CsUGT*

Table 1- Primers used for *CsUGT* gene analysis

ژن Gene	آغازگر Primer	توالی Sequence
<i>UGT</i>	GTS-F	5'-ATGTTGAACGGCAACAAATGCCACATCC-3'
	GTS-R	5'-TTAAACTAAGGAAATTTGGAGTCATTGC-3'

DNA sequencer به شرکت ماکروژن، کره جنوبی (Macrogen Inc., Seoul, Korea) ارسال شدند.

#### بررسی توالی نوکلئوتیدی

بررسی نتایج حاصل از توالی یابی با استفاده از نرم افزار Chromas version 2.6.2 انجام شد و برنامه های BLAST n و T-Coffee به منظور یافتن توالی های مشابه از پایگاه داده NCBI، استفاده شد. همچنین توالی اسیدهای آمینه با ترجمه توالی نوکلئوتیدی ژن *UGT* به وسیله نرم افزارهای ترجمه داده پایگاه مولکولی Expassy بدست آمد (Artimo et al., 2012). برای بررسی ویژگی های فیزیکی و شیمیایی پروتئین شامل وزن

#### همساز سازی و تعیین توالی *UGT*

قطعه استخراج شده از با استفاده از کیت pTG19-T PCR cloning vector (Malaysia، Vivantis) و آنزیم T4 DNA Ligase (USA، Thermo Fisher Scientific) در باکتری *Escherichia coli* سویه DH5α به کمک محیط کشت LB (Luria Bertani) حاوی آنتی بیوتیک آمپی سیلین، با استفاده از روش شوک حرارتی (Maniatis et al., 1995) هم سان شده. جهت انتخاب باکتری های تراریخت، دو روش ColonyPCR و هضم آنزیمی با آنزیم برشی *BamH I* (Jena Bioscience، Germany) انجام شد و پس از انتخاب کلون ها و کشت شبانه آنها، استخراج پلاسمیدهای نو ترکیب pTG19-UGT با استفاده از کیت استخراج TPG (پیشگامان انتقال ژن-ایران) انجام گرفت، به دنبال خالص سازی پلاسمیدهای استخراج شده، به منظور توالی یابی قطعه ژنی *UGT* با آغازگرهای یونیورسال M13 و همچنین آغازگرهای اختصاصی با روش automated

با gDNA و cDNA زعفران تکثیر و هم‌ردیفی توالی‌های نوکلئوتیدی استخراج شده از DNA ژنومی و cDNA نشان داد که این ژن فاقد اینترون می‌باشد که نتایج بدست آمده در سایر مطالعات (Hosseinpour Azad, 2020; bio-Moraga et al., 2004; Trapero et al., 2012; Mozes, 2013; Ahrazem et al., 2015) نیز فقدان اینترون در توالی ژن‌های گلیکوزیل ترانسفرازها را تأیید نموده است.

وزن مولکولی پروتئین CsUGT ۵۱/۳ کیلودالتون محاسبه شد. از مجموع کل اسید آمینه‌های CsUGT ۵۴ اسید آمینه با بار منفی (Asp+ Glu) و ۴۴ اسید آمینه با بار مثبت (Arg+ Lys) بود. نقطه ایزو الکتریک تئورتیکال ۵/۶۳ بود. بررسی و پیش‌بینی ویژگی‌های بیوشیمی‌یابی و بیوانفورماتیکی پروتئین CsUGT نشان می‌دهد که نقطه ایزوالکتریک آن در دامنه اسیدی (کمتر از ۷) بوده و خاصیت اسیدی دارد. اطلاع از نقطه ایزوالکتریک در ارزیابی حلالیت و نحوه حرکت پروتئین در میدان الکتریکی صفر مفید است و می‌تواند برای انتخاب و بهینه‌سازی روش‌های مورد استفاده برای خالص‌سازی پروتئین اعم از کروماتوگرافی تبادل یونی (IEC<sup>۲</sup>) و الکتروفورز ایزوالکتریک فوکوسینگ (IEF<sup>۳</sup>) بسیار مفید باشد. به علاوه پروتئین‌هایی که در محدوده pH ایزوالکتریک قرار داشته و مجموع بار مثبت و منفی رشته پپتیدی آنها با یکدیگر برابر می‌باشد، توانایی نگهداری آب را نداشته و ناگزیر رسوب می‌کنند. شاخص ناپایداری پروتئین CsUGT، ۴۸/۹۶ بوده که این شاخص بیشتر از ۴۰ برآورد شده است. شاخص ناپایداری پروتئین‌ها، تخمینی از ثبات پروتئین در شرایط لوله آزمایش می‌باشد (Mirhoseini et al., 2016) و این شاخص به نوع اسید آمینه‌های موجود بستگی دارد.

شاخص ناپایداری محاسبه شده برای پروتئین مورد بررسی

مولکولی، pI، شاخص آلفاتیکی، شاخص پایداری و نیمه عمر پروتئین از نرم‌افزار ProtParam (Gasteiger et al., 2005)، پیش‌بینی ساختار دوم پروتئین به بیان شوونده از SOPMA (Combet et al., 2000)، پیش‌بینی توالی حفاظت شده با نرم‌افزار Pfam (El-Gebali et al., 2019)، شناسایی جایگاه درون سلولی از نرم‌افزار ProtComp، الگوی آب‌گریزی-آبدوستی بر اساس روش کیت و دولیتل (Kite & Doolittle, 1982) از نرم‌افزار ProtScale (Gasteiger et al., 2005)، وجود سیگنال‌های پپتیدی در پروتئین از نرم‌افزار Signal P 5.0 (Almagro Armentro et al., 2019)، ناحیه درون‌غشایی و جهت‌گیری مرتبط با پروتئین از نرم‌افزار TMHMM (Krogh et al., 2001) و برای بررسی هدف‌گیری پروتئین از بانک‌های TargetP و ChloroP استفاده شد. با استفاده از پایگاه اطلاعاتی SWISS-MODEL (Waterhouse et al., 2017; Bienert et al., 2018) نیز مدل‌سازی به روش مدلینگ صورت گرفت و کیفیت مدل سه‌بعدی با استفاده از نمودار رامانچاندرا (Bertoni et al., 2017) مورد بررسی قرار گرفت.

## نتایج و بحث

نتایج توالی‌یابی صحت قطعه تکثیر یافته از مجموعه ژنی زعفران را که از ایزوفرم‌های ژنی خانواده گلیکوزیل ترانسفرازها است را تأیید کردند و ژن CsUGT بدست آمده شامل ۱۳۸۹ جفت باز بوده که ۴۶۲ اسید آمینه را کد می‌کند. ژن CsUGT در پایگاه NCBI با شماره دسترسی‌های MN688236 ثبت شد. بلاست توالی ژن CsUGT بدست آمده نشان داد که بیشترین تشابه (۹۹/۳۴٪) را با ژن UGT74AD1 (MF596166) (Demurtas et al., 2018) داشته است. همچنین ژن CsUGT

در پیش‌بینی ساختار دوم برای CsUGT، ۱۹۴ ساختار مارپیچ آلفا (۴۱/۵۴٪)، ۷۳ Extended strand (۱۵/۶۳٪)، ۲۷ Beta turn (۵/۷۸٪) و ۱۷۳ ساختار Random coil (۳۷/۰۴٪) مشاهده شد. ساختارهای ثانویه مارپیچ آلفا و صفحات بتا در دسته ساختارهای دوم منظم و پیچ‌های تصادفی، جزو ساختارهای ثانویه نامنظم تکرارپذیر و تصادفی طبقه‌بندی می‌شوند. حضور توالی آمینواسیدهای تیروزین، تریپتوفان، فنیل‌آلانین، لوسین، ایزولوسین، والین، سیستئین و اسپارژین دلیل اصلی منظم شدن ساختار این قسمت از پروتئین‌ها می‌باشد. البته عوامل دیگری شامل آبدوستی، انعطاف‌پذیری، وزن مولکولی و نقطه ایزوالکتریک نیز در منظم شدن این ساختارها نقش دارند (Yousefi Javan & Gharari, 2017).

در بین آنزیم‌های گلوکوزیل ترانسفراز در گیر در متابولیسم ثانویه گیاهان، دو دمین اصلی در ناحیه پایانه C بخوبی حفظ شده است (Hughes & Hughes, 1994). در این تحقیق نرم افزار Pfam جهت تعیین توالی حفاظت شده مورد استفاده قرار گرفت و یک توالی را به عنوان الگوی حفاظت شده پروتئین مشخص نمود که این توالی‌ها در ناحیه پایانه C، (۲۵۴-۴۳۸) برای CsUGT در خانواده گلیکوزیل‌ترانسفرازها قرار داشت. بررسی هدف‌گیری پروتئین با استفاده از بانک‌های ChloroP نشان داد که پروتئین CsUGT دارای سیگنال هدف‌گیری پلاستییدی هستند. چهار آنزیم UGT74A2، UGT74A3، UGT74A4 و UGT74A5 را در *C. ancyrensis* از نظر وجود سیگنال‌های هدف‌گیری در پایانه C و سیگنال‌های لنگر غشایی در پایانه N با استفاده از برنامه‌های تحت شبکه Target P و Chloro P مورد بررسی قرار گرفته‌اند. آنزیم‌های UGT74A2 و UGT74A3 حاوی یک سیگنال پپتیدی در پایانه N برای هدف قراردادن پلاستییدها همانند UGT74A1 بودند در حالی که آنزیم‌های UGT74A4 و UGT74A5 سیتوزولی پیش‌بینی شدند (Ahrazem et al., 2015). در مطالعه حاضر، پیش‌بینی

نشان می‌دهد که پروتئین‌هایی که شاخص ناپایداری آنها کمتر از ۴۰ باشند با ثبات و در غیر این صورت بی‌ثبات خواهند بود. پیش‌بینی ناپایدار بودن این پروتئین می‌تواند عدم فعالیت آنزیم نو ترکیب تولید شده در باکتری *E. coli* را توضیح دهد (Vogt & Jones, 2000). زیرا یکی از دلایل عدم توانایی تعیین فعالیت آزمایشگاهی آنزیم می‌تواند این باشد که CsUGT بسیار حساس است و فعالیت آنزیم در زمان خالص‌سازی از بین می‌رود. 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).

شاخص آلیفاتیک، به عنوان یک عامل مهم در ارزیابی مقاومت پروتئین در برابر حرارت، برابر ۸۶/۴۵ محاسبه گردید. شاخص آلیفاتیک عامل مهمی به منظور برآورد مقدار آب‌گریزی پروتئین و یک عامل مثبت برای پایداری حیاتی پروتئین و ساختار سه‌بعدی آن محسوب می‌شود که در واقع عبارت است از حجم نسبی پروتئین که با زنجیره‌های آلیفاتیک اشغال شده‌اند (A, I, L, V). پروتئین‌هایی با شاخص آلیفاتیک بسیار بالا (بیشتر از ۱۰۰) ممکن است در دامنه دمایی بسیار بالا از خود ثبات نشان دهند. محاسبه شاخص آلیفاتیک به عنوان یک فاکتور مثبت در مقاومت به حرارت نشان داد که پروتئین مورد مطالعه در این پروژه از ثبات بالایی در برابر دماهای بالا برخوردار است. تجزیه و تحلیل با نرم‌افزار ProtParam نشان داد که متوسط میزان آبدوستی - آب‌گریزی (GRAVY) محاسبه شده برای این آنزیم منفی 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، قطبی محسوب می‌شود. غیرقطبی بودن یک پروتئین باعث حالیت کمتر آن نسبت به پروتئین‌های قطبی در حلال می‌گردد و در برابر حلال‌های ویژه هر پروتئین، پروتئین‌های قطبی در قسمت بیرونی و پروتئین‌های غیرقطبی در وسط تجمع پیدا می‌کنند.

پروتئین همانطور که شکل ۱ مشاهده می شود، آنزیم CsUGT فاقد قطعه آب‌گریز و ناحیه درون غشایی می‌باشد و تنها دارای مارپیچ پوششی (اسیدهای آمینه ۱۴۳-۱۲۸) است با نتایج حاصل از نرم‌افزارهای تعیین سیگنال پپتیدی (شکل ۲) و سایر مطالعات نیز مطابقت داشت. آنزیم CsUGT فاقد قطعه آب‌گریز و درون غشایی بوده و توسط یک مارپیچ پوششی با غشا در ارتباط است. موزوس و همکاران (Mozos et al., 2013) با مطالعه بر روی پروتئین‌های UGT نیز بیان نمودند که این پروتئین‌ها فاقد دمین آب‌گریز و درون غشایی بوده که با نتایج این تحقیق مطابقت داشت.

شناسایی جایگاه درون سلول (شکل ۲) نشان می‌دهد که در پروتئین CsUGT اسیدهای آمینه پایانه N آن (۱۲۸-۱) جایگاه سیوپلاسمی دارند و اسیدهای آمینه ۱۴۴-۱۲۹ با غشا در ارتباط بوده و مابقی جایگاه خارج سلولی دارند. که این نتایج با نتایج نرم افزار ProtComp نیز که با امتیاز ۴/۵ جایگاه CsUGT (۲/۳) امتیاز) جایگاه خارج سلولی را پیش‌بینی کرده بود مشابهت داشت. پروتئین‌های جانبی معمولاً بصورت برگشت‌پذیر با غشاء دولایه لیپیدی ارتباط دارند و برهمکنش با غشاء ممکن است بر ساختار، فعالیت بیولوژیکی، پایداری ترمودینامیکی، اتصال آنها به سوبستراها و ماکرومولکول‌های محیطی تأثیر بگذارد (Lomize et al., 2007). در مطالعه‌ای که اخیراً توسط دمور تاس و همکاران (Demurtas et al., 2018) در خصوص تجزیه و تحلیل جایگاه پروتئین صورت گرفته است، نشان داد که CsCCD2، CsALDH31 و CsUGT74D1 به ترتیب در پلاستیدها، شبکه آندوپلاسمی و سیتوپلاسم قرار دارند و بر اساس این یافته‌ها پیشنهاد شد که شبکه آندوپلاسمی و سیتوپلاسم به عنوان مراکز انتقال متابولیت‌ها بوده و بیوسنتز در پلاستیدها آغاز و در واکوئل‌ها به عنوان مراکز ذخیره خاتمه

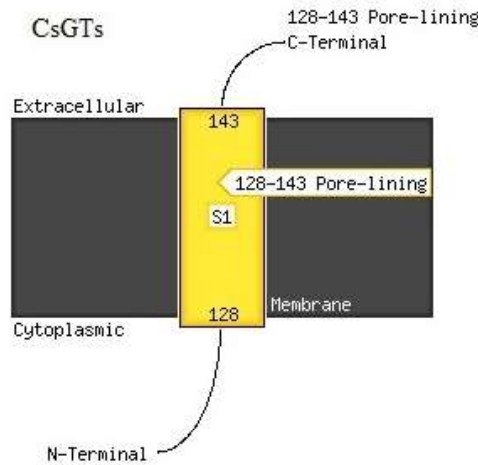
سیگنال‌های پپتیدی پروتئین CsUGT توسط برنامه‌های مبتنی بر شبکه Signal P و PSIPRED نشان داد که این پروتئین فاقد سیگنال پپتیدی می‌باشد و نواحی از پروتئین CsUGT که با غشا در تعامل می‌باشند بصورت مارپیچ پوششی<sup>۱</sup> نشان داده شده است که این نتیجه با سایر اطلاعات بدست آمده توسط موزوس و همکاران (Mozos et al., 2013) که بر روی سایر پروتئین‌های UGT انجام شده است مطابقت داشت. موراگا و همکاران (Rubio-Moraga et al., 2004) بر روی پروتئین UGTCS2 وجود یک سیگنال پپتیدی در توالی آمینو اسیدی UGTCS2 را پیش‌بینی نمودند ولی در سال ۲۰۰۹ نتایج تحقیقات آنها بر روی پروتئین CsGT45 نشان داد که این پروتئین فاقد سیگنال پپتیدی و سیگنال درون غشایی می‌باشد (Rubio-Moraga et al., 2009).

در شرایط محیط آزمایشگاهی، کاروتنوئیدها می‌توانند در بین غشا دولایه لیپیدی یا پلاستییدی ذخیره و انباشته شوند (Cunningham et al., 1998). لذا غشاهای لیپیدی محیطی مناسب برای آنزیم‌هایی هستند که سوبستراهای آب‌گریزی همانند کاروتنوئیدها را متابولیزه می‌کنند (Ahrazem et al., 2015). در پژوهش‌های انجام شده، حضور یک ناحیه آب‌گریز برای پروتئین UGTCS2 پیش‌بینی شده بود (Rubio-Moraga et al., 2004). ولی در این پژوهش نتایج پیش‌بینی الگوی آب‌گریزی-آبدوستی نتایج نشان داد که آنزیم CsUGT فاقد بخش‌های آب‌گریز است. وجود دمین آب‌گریز آنزیم‌ها امکان برقراری ارتباط با لیپیدهای غیرقطبی را فراهم می‌سازد (Rubio et al., 2008) و پیشنهاد می‌شود که آنزیم‌ها از طریق این ناحیه با غشا در ارتباط هستند (Rubio-Moraga et al., 2004).

در بررسی ناحیه درون غشایی و جهت‌گیری مرتبط با



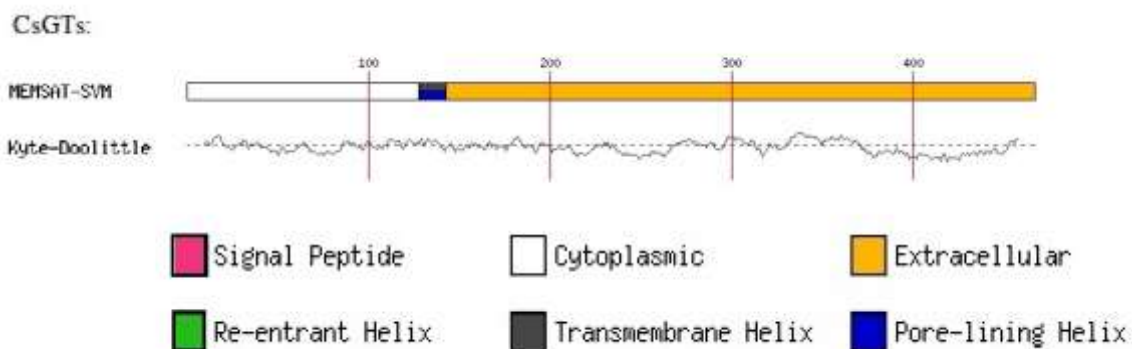
مى‌يابد. آنها گزارش نمودند كه آنزيم CsUGT74AD1 جايگاه سيتوپلاسمى دارد.



شکل ۱- دمین درون غشایی پروتئین CsUGT بدست آمده از سرور PSIPRED  
 Figure 1- Transmembrane domain of CsUGT protein obtained from PSIPRED server.  
 (<http://bioinf.cs.ucl.ac.uk/psipred>)

Trapero et al., 2012; Ahrazem et al., 2015; Gomez- در پژوهشى ديگر، گمز-گمز و همكاران (Gomez- et al., 2017) نیز بر اساس داده‌هاى پروبيوتيك نشان دادند كه UGTCs2 شناسايى شده در كلاله‌هاى زعفران در كروموپلاست‌هاى كلاله‌هاى زعفران مستقر مى‌باشند. هيچكدام از پروتئين‌هاى UGT شامل CsUGT74AD1 و ديگر CsUGTs شناسايى شده (Rubio-Moraga et al., 2004; )

در پژوهشى ديگر، گمز-گمز و همكاران (Gomez- et al., 2017) نیز بر اساس داده‌هاى پروبيوتيك نشان دادند كه UGTCs2 شناسايى شده در كلاله‌هاى زعفران در كروموپلاست‌هاى كلاله‌هاى زعفران مستقر مى‌باشند. هيچكدام از پروتئين‌هاى UGT شامل CsUGT74AD1 و ديگر CsUGTs شناسايى شده (Rubio-Moraga et al., 2004; )



شکل ۲- سيگنال‌هاى پپتيدى پيش‌بينى شده براى پروتئين CsUT با استفاده از سرور PSIPRED  
 نواحى نارنجى رنگ جايگاه خارج سلولى داشته و نواحى سفيد رنگ در سيتوپلاسم قرار دارند. عدم وجود ناحيه صورتى نشان دهنده فقدان سيگنال پپتيدى است.  
 Figure 2- Peptide signals predicted for the CsUT protein using the PSIPRED server.  
 The orange regions are extracellular, and the white regions are in the cytoplasm. The absence of a pink area indicates a lack of a signal peptide.

مدل سازی ساختار سه بعدی بر اساس انتخاب یک الگو با شباهت بالا با پروتئین هدف با استفاده از پایگاه اطلاعاتی Swiss model به انجام رسید (جدول ۲ و شکل ۳) و کیفیت پارامترهای ساختاری مدل‌های ساخته شده با نمودار رامچاندراان مورد ارزیابی قرار گرفت (شکل ۴).

پیش‌بینی توسط مدل سازی همولوژیکی بر اساس شباهت توالی پروتئین با یک یا چند پروتئین با ساختار شناخته شده می‌باشد. این روش، مدل سازی مقایسه‌ای (CA) نیز نامیده می‌شود. در این روش، ساختمان پروتئین با استفاده از روش‌های تجربی و بر اساس مشابهت توالی ساختمان با توالی یا ساختمان‌های شناخته شده، پیش‌بینی می‌شود. در واقع این روش بر این اصل استوار است که اگر دو پروتئین از نظر توالی تشابه بالایی داشته باشند، احتمالاً ساختمان سه بعدی بسیار مشابهی خواهند داشت (Ahmad Baba et al., 2015). در این تحقیق با توجه به قرار گرفتن توالی حفاظت شده پروتئین CsUGT در خانواده گلوکوزیل ترانسفرازها، معیارهایی نظیر وضوح در محدوده ۳ آنگسترم، شباهت بالای ۳۵٪ الگو با توالی مورد مطالعه و E-value پایین در نظر گرفته شد تا اعتبار و اطمینان‌پذیری مدل افزایش یابد. برای مدل سازی پروتئین CsUGT، پروتئین گلوکوزیل ترانسفراز OS79 در برنج (*Oryza sativa*) (5tmb.1.A) با ۴۶۷ اسید آمینه که در سال ۲۰۱۶ با قدرت

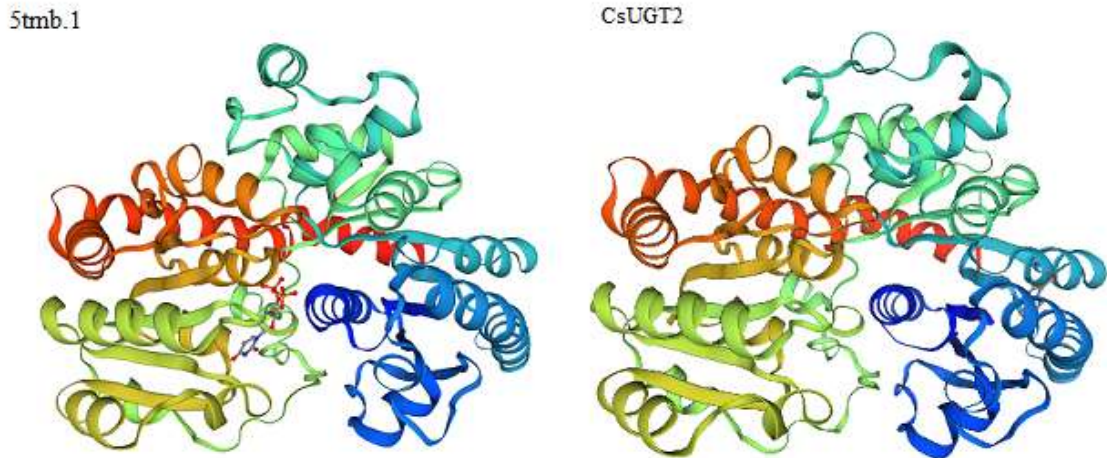
تفکیک ۲,۳۴ آنگسترم از برنج کریستالوگرافی شده است (Watterhorn et al., 2016)، به عنوان الگو انتخاب شد. ۴۷/۶۵٪ میزان تطابق الگو با پروتئین CsUGT و ۴۳٪ میزان شباهت آن تعیین شد (جدول ۲). بنابراین با توجه به میزان شباهت پایین الگوها با پروتئین مورد نظر، این احتمال وجود دارد که مدل سازی بصورت ضعیف انجام شده باشد. لذا این نکته که این آنزیم نقش مهم و حیاتی در مسیر سنتز کروسین در زعفران دارد، قطعاً اولین گام در راستای مطالعه جایگاه‌های اتصال و یا لیگاند‌های فعال کننده یا مهارکننده، مدل سازی این پروتئین می‌باشد.

پس از مدل سازی، ارزیابی کیفیت پارامترهای ساختاری مدل ساخته شده با نمودار رامچاندراان انجام شد، که در این نمودار با محاسبه درصد اسید آمینه‌های قرار گرفته در نواحی مطلوب، مجاز و یا غیرمجاز کیفیت مدل ورودی تعیین می‌شود. همانطور که در شکل ۴ مشاهده می‌شود، ۹۲/۸۱٪ اسید آمینه‌ها در CsUGT برای ساختارهای مدل شده در مناطق مطلوب و ۲/۲۵٪ اسید آمینه‌ها در مناطق مجاز قرار دارند. با جمع نواحی مطلوب و مجاز ۹۵/۰۶٪ بدست می‌آید و این بدان معنی است که این مدل از کیفیت بالایی برخوردار است و با توجه به نتایج ارزیابی کیفی مدل ایجاد شده می‌توان از این مدل در بررسی‌ها و طراحی‌های بعدی استفاده نمود.

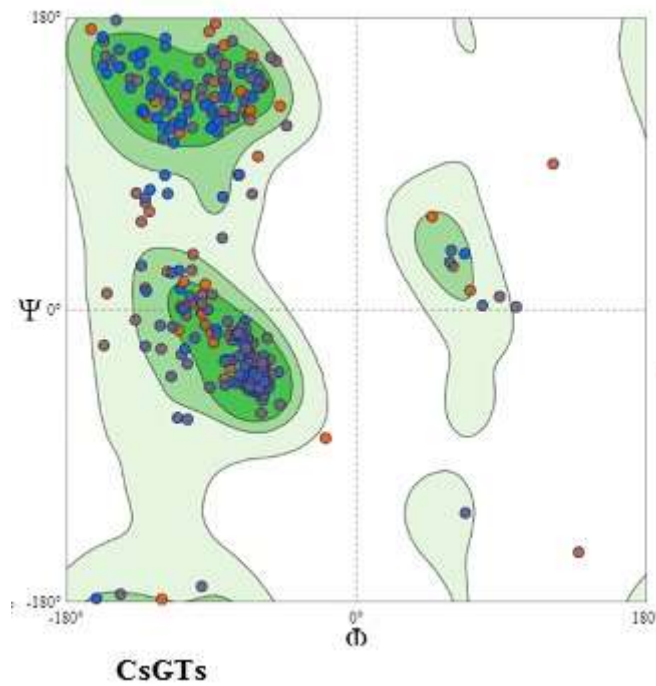
جدول ۲- میزان شباهت اسید آمینه‌های پروتئین CsUGT با الگو (پروتئین OS79 در برنج)

Table 2 - The similarity of the amino acids of CsUGT protein with pattern (OS79 protein in rice)

شرح Description	درصد پوشش Coverage	محدوده Range	میزان شباهت Seq similarity	وضوح Resolution	روش Method	یافت شده توسط Found By	ساختار Oligo-state	میزان تطابق Seq identity	الگو Template	پروتئین Protein
Glycosyltransferase, Os 79	97%	8- 454	0.43	2.34Å	X-ray	BLAST	Monomer	47.65%	5tmb.1.A	GTs



شکل ۳- مدل سازی ساختار سه بعدی بر اساس انتخاب یک الگو با شباهت بالا با پروتئین CsUGT با استفاده از پایگاه اطلاعاتی Swiss model  
 Figure 3- Modeling a 3D structure based on selecting a pattern with high similarity to the CsUGT protein using the Swiss model database (Waterhouse et al., 2018; Bienert et al., 2017).



شکل ۴- نمودار رامچاندران ساختار مدل سازی شده برای پروتئین CsUGT. مناطق سفید رنگ محدود غیر مجاز است و هر نقطه قرمز رنگ یک اسید آمینه با زاویه فی<sup>۱</sup> و سای<sup>۲</sup> غیر مجاز را نشان می دهد. مناطق رنگی به ترتیب از سبز روشن به سبز تیره، ۹۹/۷٪، ۹۵٪ و ۸۰٪ اسید آمینه ها را در محدوده مجاز شامل می شوند.

Figure 4- Ramachandran plot of a structured model for the CsUGT protein. The white areas are in the unauthorized range, and each red dot represents an amino acid with an unauthorized Phi and Sai angle. Colored areas from light green to dark green, 99.7%, 95%, and 80% of amino acids are within the allowable range (Bertoni et al., 2017).

1 - Phi ( $\Psi$ )  
 2 - Psi ( $\Phi$ )

Alignment: |

```

.....|.....|.....|.....|.....|.....|
          10      20      30      40      50
5tmb.1   GMGSMSTPAA SANGGQVLLL PFPAAQGHTN PMLQFGRRLA YHGLRPTLV
CsGIs    -----ML NGNKCHILL PCPAQG-HIN PILQFGKRLA SHNLLTTLVN
CsUGT74AD1 ----MEQKDV NGNKCHILL PCPAQG-HIN PILQFGKRLA SHNLLTTLVN
UGTCs4   ----MEQKDE NGNNCHILL PYPAQG-HIN PMLQFGKRLA SRNLRTTLVN
GLT2_CROSA -----ML NGNKCHILL PCPAQG-HIN PILQFGKRLA SHNLLTTLVN
Clustal Co ..*  :*** * ** * * *:*****:***  :.* .***.

.....|.....|.....|.....|.....|.....|
          60      70      80      90     100
5tmb.1   TRYVLSTPP PGDPFRVAAI SDGFDDASGM AALPDPGEYL RTLEAHGART
CsGIs    TRFLSNSTKS EPGPVNIECI SDGFDSG-GM NAAPSRRAYF DRLESVSGSET
CsUGT74AD1 TRFLSNSTKS EPGPVNIECI SDGFDSG-GM NAAPSRRAYF DRLESVSGSET
UGTCs4   TRFISNSTKP DPGPVNTECI SDGFDSG-GL NAAPSREAYF DRLESVSGSET
GLT2_CROSA TRFLSNSTKS EPGPVNIQCI SDGFDPG-GM NAAPSRRAYF DRPQSRSGQK
Clustal Co ***: :.* .  .*. . * ***** . *: * *. *:  :: ....

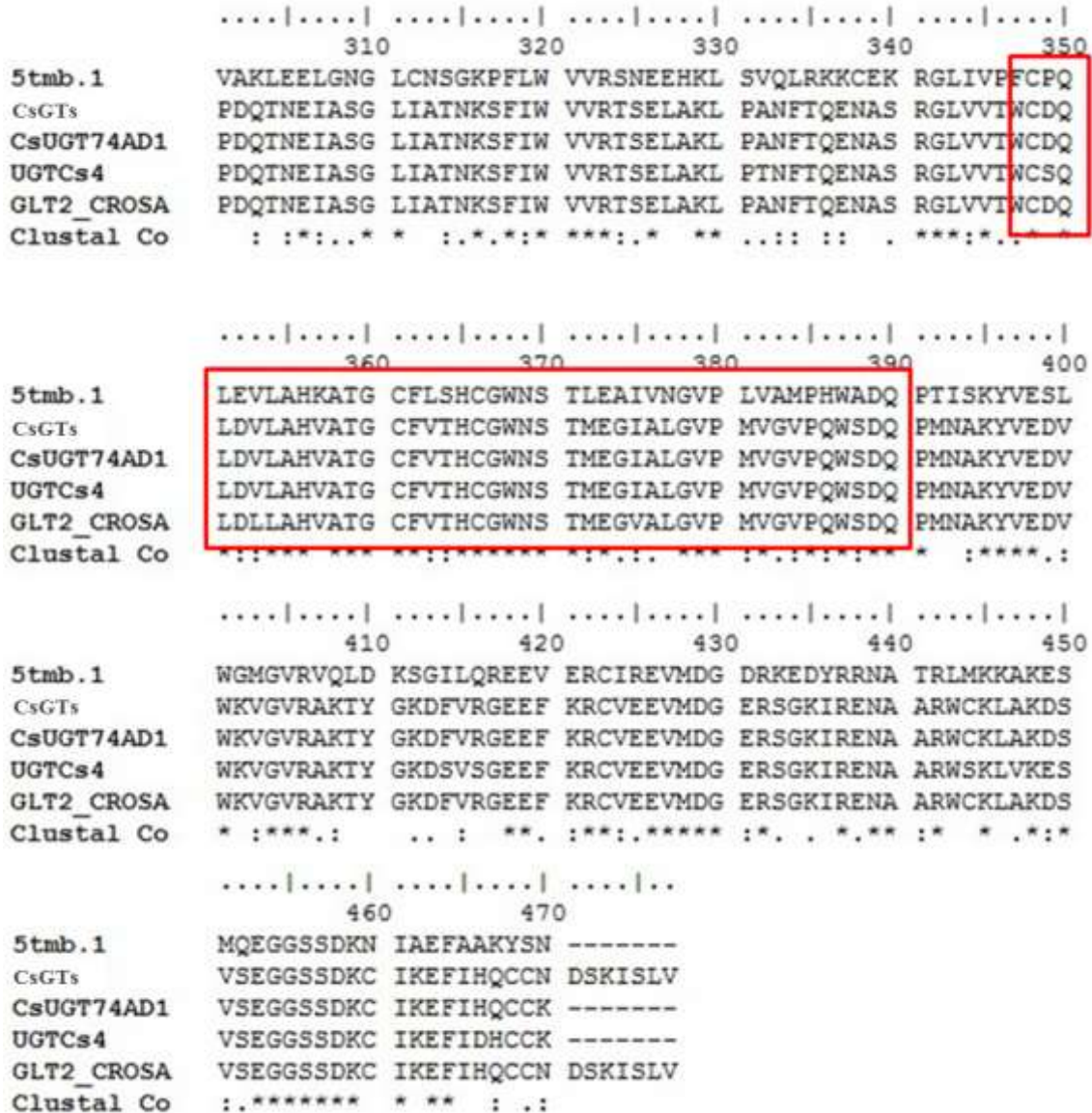
.....|.....|.....|.....|.....|.....|
          110     120     130     140     150
5tmb.1   LAELLSEAR AGRPARVLVY DPHLPWARRV ARAAGVATAA FLSQPCAVDL
CsGIs    LFGLIESLRS RGRPAHVLVY DPFLPWAMNV AERSGLRSVA FFTQPCAVDT
CsUGT74AD1 LFGLIESLRS RGRPAHVLVY DPFLPWAMNV AERSGLRSVA FFTQPCAVDT
UGTCs4   LFGLIESLRS RGRPAHVLVY DPFLPWATNV AERSGLRSVA FFTQPCAVDT
GLT2_CROSA HVGLIESLRS RGRPGACFGL RPVPLWAMNV AERSGLRSVA FFTQPCAVDT
Clustal Co *: *   ***. :  * ** . * . **: :.* **:*****

.....|.....|.....|.....|.....|.....|
          160     170     180     190     200
5tmb.1   IYGEVCARRL ALPVIPTDAR GLYARGVLGV ELGPDDVPPF VAAPELTPAF
CsGIs    IYRHVWEGRI KVPVTEPVR- ---LPGLPPL EPSDLPSFVT DSDPVINPDL
CsUGT74AD1 IYRHVWEGRI KVPVTEPVR- ---LPGLPPL EPSDLPSFVT DSDPVVNPDL
UGTCs4   IYRLVWEGQI KVPVTEPVR- ---LPGFPPL EPSDLPSFVT DSHPVINPEL
GLT2_CROSA IYRHVWEGRI KVPVAEPVR- ---LPGLPPL EPSDLPCVRN GFGRVVNPDL
Clustal Co ** *   :: :***. .  * . : * .  :.* :

.....|.....|.....|.....|.....|.....|
          210     220     230     240     250
5tmb.1   CEQSIEQFAG LEDDDVVLVN SFSDLPEKEA AYMES--TWR AKTIGPSLPS
CsGIs    LPLLVNQHKN LDKADMVLIN SIYELEHEEF DWMESRLPLP VKSIGPTVPS
CsUGT74AD1 LPLLVNQHKN LDKADMMLIN SIYELEHEEF DWMESRLPLP VKSIGPTVPS
UGTCs4   LPLLVNQHKN LEKADMMLIN SVYELEHEEF DWMKSRLPLP VKSIGPTVPS
GLT2_CROSA LPLRVNQHKN LDKADMGRN SIYELEADLL DGSR--LPLP VKSIGPTVPS
Clustal Co **: .  *:. * : * *. :** .  .  .  .*:*****

.....|.....|.....|.....|.....|.....|
          260     270     280     290     300
5tmb.1   FYLDDGRLRS NTAYGFNLFR -STVPCMEWL DKQPPRSVVL VSYGTVSTFD
CsGIs    TYLDN-RIPS DSHYGFNLYT PDITPYLDWL DSKAPNSVIY VSFGLSSSL
CsUGT74AD1 TYLDN-RIPS DSHYGFNLYT PDITPYLDWL DSKAPNSVIY VSFGLSSSL
UGTCs4   TYLDN-RIPS DSHYGFNLYT PDITPYLDWL DSKDPNSVIY VSFGLSSSL
GLT2_CROSA TYLDN-RIPS DSHYGFNLYT PDITPYLDWL DSKAPNSVIY VSFGLSSSL
Clustal Co ***: * : * :: *****:  .*. * : ** *.: *.*: **:*:.....

```



شکل ۵- همردیفی چندگانه پروتئین‌های خانواده CsGTs و 5tmb.1.A کادر رنگی مربوط به جعبه PSPG حفاظت شده در خانواده گلیکوزیل ترانسفرازها می‌باشد. اسید آمینه‌های تریپتوفان (W)، گلوتامیک اسید (E) و گلوتامین (Q) موجود در موتیف PSPG بخوبی حفظ شده‌اند.

Figure 5- Multiple alignments of proteins of the CsGTs family and 5tmb.1.A the color box corresponds to the PSPG box protected in the glycosyltransferase family. The amino acids tryptophan (W), glutamic acid (E), and glutamine (Q) in the PSPG motif are well conserved.

سازای ساختار اول رشته‌های پپتیدی این پروتئین و توالی اسید آمینه‌ای پروتئین گلیکوزیل ترانس سفراز برج انجام پذیرفت و براساس همولوژی بخش‌هایی از این پروتئین، شناخت بهتری از عملکرد آن حاصل شد. دمین انتهای C پروتئین CsUGT دارای

کاربرد و عملکرد ویژه هر بخش از پروتئین، توسط شکل فضایی آن بخش از پروتئین تعیین می‌گردد و هر نوع تغییر در این ساختار می‌تواند باعث اختلال در عملکرد آن شود. در این تحقیق ساختار سه‌بعدی پروتئین CsGUT شبیه سازی شد (شکل ۳). بررسی ساختار سه بعدی این پروتئین با همردیف

ندارد که این امر در نتایج حاصل از این تحقیق نیز صدق می‌کند. به دلیل وجود تنوع در ساختار پروتئین‌های CsUGT و همچنین سوبستراهای مختلف آنها، تعیین اساس شیمیایی آنها غیرممکن شده است (Mozos, 2013). وجود انواع مختلفی از متابولیت‌های ثانویه در کلاله زعفران می‌تواند به دلیل وجود گستردگی در توالی‌های مرتبط با UGT در مقایسه با سایر گیاهان باشد (Hosseinpour Azad et al., 2020). زعفران حاوی گروه‌های پیچیده از متابولیت‌های می‌باشد که بیشتر آنها به فرم گلیکوزیده هستند. وجود ساختارهای گلیکوزیده برای هر یک از متابولیت‌های ثانویه، وجود خانواده‌های مختلف از UGTها را که منجر به الحاق و حذف مولکول‌های قندی به آنها می‌شوند را اثبات می‌نماید. همچنین وضعیت تریپلوئیدی گیاه زعفران نیز می‌تواند از دیگر علل وجود تعداد بیشتر خانواده‌های UGT در گیاه زعفران نسبت به دیگر گیاهان باشد (Hosseinpour Azad et al., 2020).

### نتیجه‌گیری

گلیکوزیل ترانسفرازها که از متابولیت‌های ثانویه به عنوان پیش ماده استفاده می‌کنند، از اجزای اصلی سلول‌های گیاهی هستند (Vogt & Jones, 2000). اگرچه بسیاری از این آنزیم‌ها از برخی گونه‌های گیاهی جداسازی شده و مورد آزمایش قرار گرفته‌اند، اما در بسیاری از موارد نقش و ساختار آنها در گیاهان هنوز ناشناخته است.

پروتئین CsUGT مانند بسیاری از UGTها که در متابولیسم ثانویه گیاه نقش دارند، متعلق به خانواده گلیکوزیل ترانسفراز می‌باشد. ژن کد کننده این آنزیم ۱۳۸۹ جفت باز دارد و تجزیه و تحلیل تعداد و جایگاه اینترون، در ژن *UGT*، حاکی از آن است که فاقد اینترون می‌باشد. از رونویسی و ترجمه ژن *CsUGT*

موتیفی با ۴۴ اسیدآمینه است که با اسیدآمینه‌های توالی 'PSPG' یافت شده در بسیاری از گلیکوزیل ترانسفرازها (*CsUGT74AD1*، *UGTCs4* و *GLT2*) مشابهت دارد (کادر قرمز رنگ شکل ۵).

تمامی خانواده UGT در انتهای C زنجیره اسیدآمینه‌ای دارای سیگنال موتیف PSPG می‌باشند که این موتیف با اهداکننده قند (UDP-Glc) میان‌کنش دارد و محل فعالیت آنزیم می‌باشند (Trapero et al., 2012). جایگاه جعبه PSPG در انتهای C بوده و شامل توالی‌های حفاظت شده می‌باشند که منجر به فعال سازی گروه هیدروکسی مولکول پذیرنده گلیکوزیل، جهت اتصال و شکل‌گیری ارتباطات گلیکوزیدی می‌گردند (Hosseinpour Azad et al., 2017). این موتیف حاوی اسیدآمینه‌هایی است که نقش کلیدی در اتصال قند نوکلئوتیدی فسفات دارند (Trapero et al., 2012; Mozos, 2013). نشان داده شده است که اسید آمینه‌های تریپتوفان (W)، گلوتامیک اسید (E) و گلوتامین (Q) موجود در توالی UGTهای چندین گونه گیاهی (به ترتیب در جایگاه‌های ۲۲، ۴۳ و ۴۴ موتیف PSPG) که بخوبی حفظ شده‌اند، UDP-Glu را به عنوان پیش ماده اهدا کننده قند ترجیح می‌دهند (Shao et al., 2005). در حالی که در این مطالعه تنها دو اسید آمینه تریپتوفان و گلوتامین به ترتیب در جایگاه‌های ۲۲ و ۴۴ موتیف PSPG حفظ شده‌اند و اسید آسپارتیک (D) در جایگاه ۴۳ این موتیف جایگزین شده است که با نتایج سایر مطالعات مطابقت دارد (Hosseinpour Azad et al., 2020; Ahrazem et al., 2015). درحالی‌که اهرازم و همکاران (Ahrazem et al., 2015) دریافتند اسیدآمینه اسید آسپارتیک (D) در جایگاه ۹ (Asp-9) در تمام توالی‌های UGT2 (*UGT74AD*) حفظ شده است و در جعبه PSPG دیگر گروه‌های مختلف UGT وجود

همچنین در دسترش بودن توالی cDNA آنزیم‌های گلیکوزیل ترانسفراز می‌تواند منجر به یک منبع جایگزین و رقابتی برای تولید طبیعی کرو سین در یک میزبان هترولوگ همانند *E. coli* یا مخمر شود.

### سپاسگزاری

از ستاد توسعه زیست فناوری به دلیل حمایت مالی از این پروژه تحقیقاتی، گروه زیست فناوری مواد غذایی مؤسسه پژوهشی علوم و صنایع غذایی به دلیل در اختیار قراردادن امکانات آزمایشگاهی و تحقیقاتی و کلیه بزرگوارانی که در انجام این پژوهش همکاری نموده‌اند، صمیمانه تشکر و قدردانی می‌گردد

پروتئینی ۴۶۲ اسید آمینه‌ای با وزن مولکولی حدود ۵۱ کیلوالتون حاصل می‌شود. این پروتئین با شاخص ناپایداری ۴۸/۹۶ ناپایدار بوده و مانند گلیکوزیل ترانسفرازهایی که تاکنون شنا سایی شده‌اند، فاقد توالی سیگنال و فاقد سیگنال غشایی برای هدف قرار دادن غشا و فرآیندهای غشایی است و این نتایج پیشنهاد می‌کند که این آنزیم‌ها در سیتوزول عمل می‌کنند.

پروتئین CsUGT غیرقطبی و دارای ثبات در دماهای بالا بوده و فاقد دمین آب‌گریز و ناحیه درون غشایی است و تنها دارای یک ماریج پوششی می‌باشد لذا پیش بینی می‌شود که این آنزیم در سیتوزول عمل نموده هرچند ممکن است بواسطه قسمت‌هایی به عنوان اجزای جانبی با سیستم درون غشایی در ارتباط باشند.

### منابع

- Ahmad baba, Sh., Jain, D., Abbas, N., & Ashraf, N. (2015). Overexpression of *Crocus* Carotenoid cleavage dioxygenase, *CsCCD4b*, in *Arabidopsis* imparts tolerance dehydration, salt and oxidative stresses by modulating ROS machinery. *Journal of Plant Physiology*, 189, 114-125. <https://doi.org/10.1016/j.jplph.2015.11.001>.
- Ahrazem, O., Rubio-Moraga, A., Jimeno, M.L., & Gomez-Gomez, L. (2015). Structural characterization of highly glucosylation crocins and regulation of the biosynthesis during flower development in *crocus*. *Frontiers in Plant Science*, 6, 1-14. <https://doi.org/10.3389/fpls.2015.00971>.
- Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, 37(4), 420-423. <https://doi.org/10.1038/s41587-019-0036-z>.
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannindis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I., & Stockinger, H. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research*, 40(W1), W597-W603.
- Bathaie, S.Z., Ashrafi, M., Bolhasani, A., Etemadikia, B., & Moosavi-movahedi, A.A. (2006). Purification of carotenoid and monoterpene aldehydes from Iranian saffron and investigation of their effect on the structure of DNA. Histone H1 and H1-DNA Complex. *Iranian Journal of Medicinal and Aromatic Plants*, 22(2), 85-97. (In Persian with English Summary).
- Beiki, A., Keify, F., & Mozafari, J. (2011). Rapid genomic DNA isolation from corm of *Crocus* species for genetic diversity analysis. *Journal of Medicinal Plants Research*, 5, 4596-4600.
- Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., &

- Schwede, T. (2017). Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Scientific Reports*, 7(1), 10480. <https://doi.org/10.1038/s41598-017-09654-8>.
- Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., & Schwede, T. (2017). The SWISS-MODEL Repository - new features and functionality. *Nucleic Acids Research*, 45, D313-D319. <https://doi.org/10.1093/nar/gkw1132>.
- Bouvier, F., Suire, C., Mutterer, J., & Camara, B. (2003). Oxidative remodeling of chromoplast carotenoids: identification of the carotenoid dioxygenase *CsCCD* and *CsZCD* genes involved in *Crocus* secondary metabolite biogenesis. *Plant Cell*, 15(1), 47-62. <https://doi.org/10.1105/tpc.006536>.
- Brown, T.A. (2016). Gene Cloning and DNA Analysis: an Introduction. 7<sup>th</sup> ed, Wiley Blackwell.
- Combet, C., Blanchet, C., Geourjon, C., & Deleage, G. (2000). NPS@: Network Protein Sequence Analysis. *Trends in Biochemical Science*, 25(3), 147-150. [https://doi.org/10.1016/S0968-0004\(99\)01540-6](https://doi.org/10.1016/S0968-0004(99)01540-6).
- Cote, F., Cormier, F., Dufresen, C., & Willemot, C. (2000). Properties of a glucosyltransferase involved in Crocin synthesis. *Plant Science*, 153, 55-63. [https://doi.org/10.1016/S0168-9452\(99\)00248-4](https://doi.org/10.1016/S0168-9452(99)00248-4).
- Cote, F., Cormier, F., Dufresen, C., & Willemot, C. (2001). A highly specific glucosyltransferase is involved in the synthesis of crocetin glucosylesters in *Crocus sativus* cells. *Journal of plant physiology*, 158, 553- 560. [https://doi.org/10.1016/S0168-9452\(99\)00248-4](https://doi.org/10.1016/S0168-9452(99)00248-4).
- Cunningham, FX., & Gantt, E. (1998). Genes and enzymes of carotenoid biosynthesis in plants. *Annual Review of Plant Physiology and Molecular Biology*, 49, 557-583. <https://doi.org/10.1146/annurev.arplant.49.1.557>.
- Demurtas, O.S., Frusciante, S., Ferrante, P., Diretto, G., Hosseinpour Azad, N., Pietrella, M., e Aprea, G., Taddei, A.R., Romano, E., Mi, J., Al-Babili, S., Frigerio, L., & Giuliano, G. (2018). Candidate enzymes for saffron crocin biosynthesis are localized in multiple cellular compartments. *Plant Physiology*, 177(3), 990-1006. <https://doi.org/10.1104/pp.17.01815>.
- Dong, T., & Hwang, I. (2014). Contribution of ABA UDP-glucosyltransferases in coordination of ABA biosynthesis and catabolism for ABA homeostasis. *Plant Signal and Behavior*, 9, e28888. <https://doi.org/10.4161/psb.28888>.
- Dufresene, C., Cormier, F., & Dorion, S. (1997). In vitro formation of crocetin glucoyl esters by *Crocus sativus* callus extract. *Planta Medica*, 63(2), 150-153. <https://doi.org/10.1055/s-2006-957633>.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, J.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., & Finn, R.D. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research*, 48(D1), D427-D432. <https://doi.org/10.1093/nar/gky995>.
- Frusciante, S., Diretto, G., Bruno, M., Ferrante, P., Pietrella, M., Prado-cabrero, A., Rubio-Morega, A., Beyer, P., Gomez-Gomez, L., Al-Babili, S., & Giuliano, G. (2014). Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *Proceeding of the National Academy Sciences of the USA*, 111(33), 12246- 12251. <https://doi.org/10.1073/pnas.1404629111>.
- Gasteiger, E., Hoogland, Ch., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook. Humana Press, New York,



- 571-607. <https://doi.org/10.1385/1-59259-584-7:531>.
- Gomez-Gomez, L., Parra-Vega, V., Rivas-sendra, A., Segui-Simmaro, J. M., Molina, R.V., Pallotti, C., Rubio-Moraga, A., Diretto, G., Prieto, A., & Ahrazem, O. (2017). Unraveling massive crocins transport and accumulation through proteome and microscopy tools during the development of saffron stigma. *International Journal of Molecular Sciences*, 18(1), E76. <https://doi.org/10.3390/ijms18010076>.
- Hosseinpour Azad, N., Nematzadeh, G.A., Giuliano, G., Ranjbar, G.A., & Yamchi, A. (2017). Identification of Apo- Carotenoids' crocin and crocetin isomers in saffron crude extracts by HPLC coupled to atmospheric pressure chemical ionization and high resolution orbitrap mass spectrometry. *Saffron Agronomy and Technology* 4(4): 291-300. <https://doi.org/10.22048/jsat.2016.38670>.
- Hosseinpour Azad, N. (2020). Identification and recombinant expression of crocin synthesis isoform gene in saffron stigma. *Saffron Agronomy and Technology*, 7(4): 481-490. (In Persian with English Summary). <https://doi.org/10.22048/jsat.2019.133922.1302>.
- Hughes, J., & Hughes, M.A. (1994). Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA Sequencing and Mapping*, 5(1), 41-49. <https://doi.org/10.3109/10425179409039703>.
- Jones, PR., Moller, BL., & Hoj, PB. (1999). The UDP-glucose:p-hydroxymandelonitrile-O-glucosyltransferase that catalyses the last step in synthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor*. *Journal of Biological Chemistry*, 274(50), 35483-35491. <https://doi.org/10.1074/jbc.274.50.35483>.
- Kite, J., & Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1), 105-32. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0).
- Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, EL. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, 305(3): 567-580. <https://doi.org/10.1006/jmbi.2000.4315>.
- Kumar, R., Sangwan, R.S., Mishra, S., Sabir, F., & Sangwan, N.S. (2012). In silico motif diversity analysis of the glycon preferentiality of plant secondary metabolic glycosyltransferases. *Plant Omics*, 5(3), 200-210.
- Lee, H-I., & Raskin, I. (1999). Purification, cloning, and expression of a pathogen inducible UDP-glucose: salicylic acid glucosyltransferase from tobacco. *Journal of Biological Chemistry*, 274(51), 36637-36642. <https://doi.org/10.1074/jbc.274.51.36637>.
- Lim, E-K., Baldauf, S., Li, Y., Elias, L., Worrall, D., Spencer, SP., Jackson, RG., Taguchi, G., Ross, J., & Bowles, DJ. 2003. Evolution of substrate recognition across a multigene family of glycosyltransferases in Arabidopsis. *Glycobiology*, 13(3), 139-145. <https://doi.org/10.1093/glycob/cwg017>.
- Lomize, A.L., Pogozheva, I.D., Lomize, M.A., & Mosberg, H.I. (2007). The role of hydrophobic interactions in positioning of peripheral proteins in membranes. *BMC Structural Biology*, 7, 44. <https://doi.org/10.1186/1472-6807-7-44>.
- Mai, N., Terasaka, K., Owaki, M., Sota, M., Inukai, T., Nagatsu, A., & Mizukami, H. (2012). UGT75L6 and UGT94E5 mediate sequential glucosylation of crocetin to crocin in *Gardenia jasminoides*. *FEBS Letters*, 586(7), 1055. <https://doi.org/10.1016/j.febslet.2012.03.003>.
- Maniatis, T., Fritsch, E., & Sambrook, F. (1995). Molecular cloning. A laboratory Manual. Cold Spring Harbor Laboratory, New York.
- Moraga, A.R., Nohales, P.F, Perez, J.A., & Gomez-

- Gomez, L. (2004). Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from *Crocus sativus* stigmas. *Planta*, 219, 955–66. <https://doi.org/10.1007/s00425-004-1299-1>.
- Mozos, A.T. (2013). Isolation and characterization of enzymes involved in the biosynthesis of secondary metabolites with phytotherapeutic interest. PhD Thesis. Departamento de Ciencia y Tecnología Agroforestal y Genética, University of Castilla-La Mancha Facultad de Farmacia, Spain.
- Mirhoseini, S.Z., Pezeshkian, Z., & Ghovvati, Sh. (2016). Phylogenetic and *in silico* analysis of interferon Beta-1b Protein. *Journal of Mazandaran University of Medical Science*, 26(145), 70-82. (In Persian with English Summary).
- Nagatoshi, M., Terasaka, K., Owaki, M., Sota, M., Inukai, T., Nagatsu, A., & Mizukami, H. (2012). UGT75L6 and UGT94E5 mediate sequential glucosylation of crocetin to crocin in *Gardenia jasminoides*. *FEBS Lett*, 586(7), 1055–1061. <https://doi.org/10.1016/j.febslet.2012.03.003>.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., & Ferrin, T.E. (2004). UCSF Chimera-avisualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605-1612. <https://doi.org/10.1002/jcc.20084>.
- Rubio, A., Rambla, J.L., Santaella, M., Gomez, D., Orzaez, D., Granell, A., & Gomez-Gomez, L. (2008). Cytosolic and plastoglobule-targeted carotenoid dioxygenases from *Crocus sativus* are both involved in  $\beta$ -Ionone release. *The Journal of Biological Chemistry*, 283(36), 24816–24825. <https://doi.org/10.1074/jbc.M804000200>.
- Rubio Moraga, A., Nohales, P.F., Peres, J.A., & Gomez-Gomez, L. (2004). Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from *Crocus sativus* stigmas. *Planta*, 219, 955- 966. <https://doi.org/10.1007/s00425-004-1299-1>.
- Rubio Moraga, A., Rambla, J.L., Ahrazem, O., Granell, A., & Gomea-Gomez, L. (2009). Metabolite and target transcript analyses during *Crocus sativus* stigma development. *Phytochemistry*, 70, 1009- 1016. <https://doi.org/10.1016/j.phytochem.2009.04.022>.
- Rubio-Moraga, A., Rambla, J.L., Fernández-de-Carmen, A., Trapero-Mozos, A., Ahrazem, O., Orzáez, D., Granell, A., & Gómez-Gómez, L. (2014). New target carotenoids for CCD4 enzymes are revealed with the characterization of a novel stress-induced carotenoid cleavage dioxygenase gene from *Crocus sativus*. *Plant Molecular Biology*, 86(4-5), 555-569. <https://doi.org/10.1016/j.phytochem.2009.04.022>.
- Shao, H., He, X., Achnine, L., Blount, J.W., Dixon, R.A., & Wang, X. (2005). Crystal structures of a multifunctional triterpene/flavonoid glucosyltransferase from *Medicago truncatula*. *Plant Cell*, 17, 3141–3154. <https://doi.org/10.1105/tpc.105.035055>.
- Trapero, A., Ahrazem, O., Rubio-Moraga, A., Jimeno, M.L., Gomez, M.D., & Gomez-Gomez, L. (2012). Characterization of glucosyltransferase enzyme involved in the formation of kaempferol and quercetin sophorosides in *Crocus sativus*. *Plant Physiology*, 159(4), 1335- 1354. <https://doi.org/10.1104/pp.112.198069>.
- Vogt, T., & Jones, P. (2000). Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends in Plant Science*, 5(9), 380-386. [https://doi.org/10.1016/S1360-1385\(00\)01720-9](https://doi.org/10.1016/S1360-1385(00)01720-9).
- Wang, W., He, P., Zhao, D., Ye, L., & Dai, L. (2019). Construction of Escherichia coli cell

- factories for crocin biosynthesis. *Microbial Cell Factories*, 18:120. <https://doi.org/10.1186/s12934-019-1166-1>.
- Wetterhorn, K.M., Newmister, S.A., Caniza, R.K., Busman, M., McCormick, S., Berthiller, F., Adam, G., & Rayment, I. (2016). Crystal structure of OS79 (Oso4go2oo66oo) from *Oryza sativa*: a UDP-glucosyltransferase involved in the detoxification of deoxynivalenol. *Biochemistry*, 55(44), 6175- 6186. <https://doi.org/10.2210/pdb6BK3/pdb>.
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: homology modeling of protein structures and complexes. *Nucleic Acids Research*, 46(W1): W296-W303. <https://doi.org/10.1093/nar/gky427>.
- Xi, L., & Qian, Z. (2006). Pharmacological properties of crocetin and crocin (digentiobiosyl ester of crocetin) from saffron. *Natural Product Communications*, 1(1), 65-75.
- Yousefi Javan, I., & Gharari, F. (2017). The structure of the protein and gene expression of PIC2 affecting blooming flowers (*Crocus sativus* L.). *Saffron Agronomy and Technology*, 5(1), 73-90. (In Persian with English Summary). <https://doi.org/10.22048/jsat.2017.63141.1200>.