بررسی و آنالیز بیوانفورماتیک ژن و پروتئین اسکوالن سنتاز در سویه‌ی بومی *آئورانتیوکیتریوم*

مجتبی مرتضوی، شهریار شاکری*، محمود ملکی و فرشاد خوش بصیرت

کرمان، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، گروه بیوتکنولوژی

تاریخ دریافت: ۹۶/۳/۲۲ تاریخ پذیرش: ۹۶/۱۱/۱۶

چکیده

اسکوالن یک تری ترپن غیر اشباع می باشد که کاربردهای گسترده ای در داروسازی دارد. در این تحقیق تولید اسکوالن و بررسی بیوانفورماتیکی ژن آن در سویه بومی *آئورانتیوکیتریوم* مورد بررسی قرار گرفت. این سویه به ترتیب به میزان ۳/۷ و ۱/۶ گرم بر لیتر بیومس و روغن غنی از اسکوالن تولید می کند. مقدار اسکوالن تولیدی توسط این سویه ۴۸/۹ میلی گرم بر لیتر می باشد. همچنین ژن اسکوالن سنتاز از این سویه توسط آنالیز مولکولی شناسایی و تعیین توالی شد. بررسی ویژگی کدونی این توالی از نظر تعداد و درصد آنها، مجموع پورین و پیریمیدین‌ها و رسم نمودار گرافیکی آن به کمک نرم‌افزارهای بیوانفورماتیک انجام شد. به کمک نرم‌افزارها وپایگاه‌های مدل‌سازی، فرایند مدل‌سازی این آنزیم صورت گرفت. مدل به دست آمده، تأیید کننده نقش اسیدهای آمینه آلانین، گلوتامیک اسید، لوسین در ایجاد ساختار آلفا هلیکس می‌باشد. این نوع از مطالعات به شناسایی مکانیسم واکنش آنزیم اسکوالن سنتاز کمک می‌کند. ارزیابی این اطلاعات پنهان می تواند دانش فرایند فولدینگ اسکوالن سنتاز و چالش-های بیان پروتئین را بهبود داده و در طراحی جدید از این آنزیم، موثر واقع شود.

واژه های کلیدی: اسکوالن، اسکوالن سنتاز، *آئورانتیوکیتریوم*

* نویسنده مسئول، تلفن: ۰۹۱۳۳۴۷۰۴۵۳ ، پست الکترونیکی: sh.shakeri@kgut.ac.ir

مقدمه

اَسکوالن یک هیدروکربن تری ترپن می باشد که در روغن زیتون، پالم، تاج‌خروس و سبوس برنج یافت می شود. این ترکیب غیراشباع و شامل شش پیوند دوگانه می باشد که غنی‌ترین منبع تولید آن، روغن کبد کوسه ماهیان می‌باشد که بصورت سنتی استفاده‌های متعددی دارد. در انسان اسکوالن در سبوم و غدد ترشحی پوست یافت می شود. این ترکیب یک حدواسط بسیار مهم جهت سنتز استرول ها و هوپانوئیدها در سلول است. همچنین اسکوالن پیش ماده مهم برای سنتز تری‌ترپن‌ها می باشد. از کاربردهای اسکوالن می توان به استفاده از آن به عنوان ادجوانت واکسن ها، مهارکننده سلول های سرطانی و خاصیت آنتی اکسیدانتی آن اشاره کرد. میکروارگانیسم‌های متعددی توانایی تولید اسکوالن را دارند که شامل باکتری‌ها، مخمرها، قارچ‌ها، ریزجلبک ها و ترائوستوکیتریدها می باشند (۱، ۱۲ و ۴). ترائوستوکیتریدها آغازیان تک‌سلولی هتروترفی هستند که توانایی تولید اسیدهای چرب غیر اشباع را به میزان زیاد با استفاده از منابع کربن آلی دارند، باتوجه به اینکه شرایط کشت برای این سویه ها باتوجه به مصرف منابع کربن آلی و همچنین عدم نیاز به نور، نسبتا ساده است، بیشتر از منابع دیگر تولیدکننده‌ی اسکوالن مورد توجه قرار گرفته اند (۱۷). آئورانتیوکیتریوم یک جنس از خانواده‌ی ترائوستوکیتریدهاست که توانایی تولید بالای اسکوالن را دارد. مسیر تولید اسکوالن در سویه های ترائوستوکیترید در قسمت هایی با هم اشتراک دارند.

استفاده از منبع کربن توسط این سویه ها سبب افزایش تولید استیل کوآ در سلول می شود. این ترکیب در مسیر مالونیل کوآ به دوکوزاهگزانوئیک اسید تبدیل می شود (۲۰). همچنین استیل کوآ با ورود به مسیر موالونات به فارنسیل دی فسفات تبدیل می شود. فارنسیل دی فسفات سپس به پراسکوالن دی فسفات و در نهایت تحت تاثیر ژن اسکوالن سنتاز به اسکوالن تبدیل می شود. فارنسیل دی فسفات همچنین می تواند وارد مسیر بیوسنتز کاروتنوئیدها شده و به بتاکاروتن و در نهایت به آستازانتین تبدیل شود. ژن اسکوالن سنتاز مهم‌ترین ژن در مسیر بیوسنتز اسکوالن در ترائوستوکیتریدها می‌باشد. این ژن دارای دو اگزون و یک اینترون است که قالب باز خواندنی آن ۱۱۶۴ جفت باز طول دارد، ژن اسکوالن سنتاز در ترائوستوکیتریدها تولید آنزیمی با ۳۸۷ اسیدآمینه می‌کند که طی یک واکنش دومرحله‌ای دو مولکول فارنسیل دی فسفات را به اسکوالن تبدیل می‌کند (۲۲). هدف کلی این مطالعه انجام آنالیزهای بیوانفورماتیکی بر روی توالی نوکلئوتیدی و ساختار آنزیم اسکوالن سنتاز می‌باشد. در این زمینه مطالعات ویژگی کدونی از قبیل فراونی، ارجحیت کدونی، درصد در کل توالی و در جایگاه اول انجام شد (۸). همچنین، ساختار سه بعدی آنزیم اسکوالن سنتاز به کمک پایگاه محاسباتی I-TASSER، پیش‌بینی شده و ویژگی‌های ساختاری آن تعیین گردید (۲۱).

مواد و روشها

سویه *آئورانتیوکیتریوم*: در این تحقیق از سویه‌ی بومی *Aurantiochytrium* sp. F4 که توسط فکرت و همکاران در سال ۱۳۹۴ جداسازی و غربالگری شده بود، استفاده گردید (۱). برای تعیین وزن خشک سلولی و روغن در این سویه از محیط کشت شامل ۲۰gr/L گلوکز، ۵ gr/L پپتون، ۵ gr/L عصاره مخمر، ۲۰ درصد آب دریا و ۰/۳ gr/L از آنتی بیوتیک های پنی سیلین و استرپتومایسین استفاده شد.

استخراج روغن: استخراج روغن از بیومس خشک سلولی برطبق پروتکل استخراجی بلایت و دایر با کمی تغییرات انجام شد (۵). در این روش مقدار مشخصی از بیومس سلولی وزن و سپس به آن مقدار ۲ ml آب مقطر افزوده شد .پس از تخریب بیومس سلولی با دستگاه التراسونیک، ml ۲/۵ از محلول کلروفرم و ۵ ml از محلول متانول به آن اضافه شد و به شدت همزده شد. سپس مقدار ۲/۵ ml آب مقطر و ۲/۵ ml کلروفرم اضافه و ورتکس انجام شد سپس نمونه سانتریفیوژ شده و به دنبال آن دو فاز تشکیل شد . فاز آلی را برداشته و عمل تبخیر حلال کلروفرم انجام شد. در نهایت روغن حاصل جهت نگه‌داری به فریز ۲۰- درجه سانتی‌گراد منتقل شد.

آنالیز کروماتوگرافی: برای بررسی تولید اسکوالن توسط سویه‌ی *آئورانتیوکیتریوم* از کروماتوگرافی لایه نازک (TLC) استفاده شد. کاغذ کروماتوگرافی با مشخصات Spherical silica gel 60; Wako Pure Chemicals در ابعاد ۱,۵ در ۸ cm برش داده شد و به میزان ۱۰ میکرولیتر از روغن به صورت یک لکه‌ی گرد روی آن قرار داده شد. کاغذهای TLC درون بشر حاوی حلال‌های کلروفرم و متانول (۱:۹) به مدت ۳۰ دقیقه قرار داده شد. سپس کاغذها بوسیله‌ی پنس از درون بشر به آرامی برداشته شدند و برروی آن‌ها اسید سولفوریک ۲۰ درصد اسپری شد، در نهایت به مدت ۳۰ دقیقه درون آون ۶۰ درجه قرار داده شدند تا رنگ بگیرند. لکه‌ی اسکوالن بوسیله کاغذ شاهد حاوی اسکوالن استاندارد شناسایی شد و Rf آن طبق معادله‌ی ۱ اندازه گیری شد (۱۳).

معادله ۱- فرمول به دست آوردن فاکتور نگه‌داری

$$= \frac{Z_s}{Z_f - Z_0} \text{ (Rf)}$$

که در این فرمول، Z_s :فاصله‌ی بین باند ظاهر شده با نقطه شروع نمونه (mm)، Z_f :فاصله‌ای که فاز مایع از سطح

حلال پیموده است (mm) و Z_D فاصله‌ی بین سطح حلال و نقطه شروع (mm) می باشد.

کروماتوگرافی مایع با کارایی بالا (HPLC) : میزان mg ۵۰ از روغن استخراجی را با ml ۴ از محلول KOH اتانولی مخلوط شد (۱۳). نمونه به مدت ۳۰ ثانیه ورتکس شده و به مدت یک ساعت در حمام آب ۹۰ درجه قرار داده می‌شود. سپس ml ۲ هگزان به همراه ml ۰/۵ آب مقطر و ml ۰/۵ اتانول به آن افزوده شد. چند ثانیه نمونه را با دست تکان داده و پس از آن فاز رویی که مربوط به هگزان و مواد صابونی نشده را برداشته و بعد از تبخیر حلال ml ۱ از محلول استونیتریل-تتراهیدروفوران (۱:۹) اضافه شد، نمونه ۳۰ ثانیه ورتکس شد و پس از فیلتر در فریزر ۲۰- نگه داری شد. ستون دستگاه HPLC ZOR Bax,Sb-C18 (4.6×250mm, 5-micron) بود. اسکوالن در طول موج ۲۱۰nm ردیابی شد و نرخ جریان دستگاه ml/min ۱ و با حجم تزریق ۱۰ میکرولیتر و دمای ستون ۳۰ درجه سانتی گراد قرار داده شد. فاز متحرک ستون از محلول استونیتریل/تتراهیدروفوران (۲۰:۸۰) تشکیل شده بود.

پرایمرها و تکثیر ژن به روش PCR : در این تحقیق از پرایمر هایی که توسط Hong won-kyung در سال ۲۰۱۳ (۱۰) برای بررسی ژن اسکوالن سنتاز در سویه‌ی *Aurantiochytrium* sp. KRS101 طراحی شده بود استفاده شد (جدول۱). برای تکثیر ژن به روش PCR میکروتیوب های حاوی ۱۲ میکرولیتر آب دوبار تقطیر، ۱۰ میکرولیتر مستر میکس (Takapozist) ، ۱ میکرولیتر DNA استخراج شده توسط کیت Bioneer با غلظت ng/µl ۵۰، ۱ میکرولیتر پرایمر رفت و ۱ میکرولیتر پرایمر برگشت از استوک با غلظت pmol ۱۰ آماده شدند، و به دستگاه PCR منتقل شدند. برنامه PCR در دماهای مختف انجام شد و در نهایت طبق جدول ۲ در ۳۵ سیکل بهینه شد.

جدول۱- پرایمرهای استفاده شده در این پژوهش

نام پرایمر	توالی	طول محصول	دمای اتصال (°C)
SQF	GCTAGCATGCCTAACAAGCCT	۱۱۶۴	۵۶
SQR	GGATCCGTCAGAGTGGGTTTGGC	۱۱۶۴	۵۶

آنالیز توالی نوکلئوتیدی: به منظور بررسی ویژگی کدونی آنزیم اسکوالن سنتاز، توالی نوکلئوتیدی این آنزیم در پایگاه محاسباتی http://ppuigbo.me/programs/CAIcal آنالیز گردید. در این پایگاه، میزان هر نوکلئوتید، درصد آنها، مجموع پورین و پیریمیدین‌ها و همچنین فراوانی آنها در موقعیت کدون شماره یک را گزارش می‌دهد. همچنین به کمک پایگاه داده‌ای http://gcua.schoedl.de/seqoverall_v2.html (۸) نمودار گرافیکی ویژگی کدونی این آنزیم بررسی گردید.

مدل‌سازی، ارزیابی مدل و آنالیز ساختاری: برای بررسی ویژگی‌های ساختاری و کدون‌های نادر در آنزیم اسکوالن سنتاز، به کمک سرور I-TASSER، ساختار سه بعدی این آنزیم از طریق تکنیک همولوژی مدلینگ مدل سازی شد (۲۱).

جدول۲- برنامه ی بهینه شده‌ی PCR

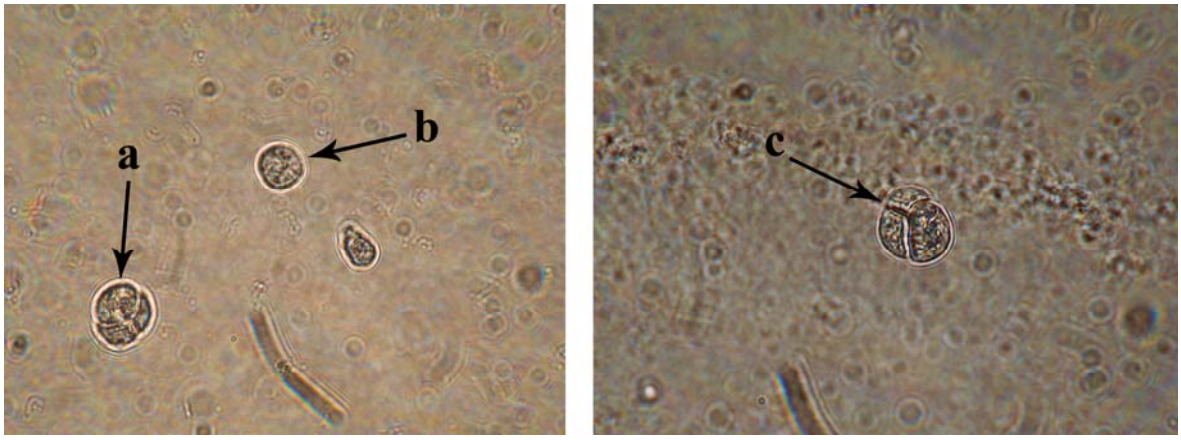
دما (°c)	۹۴	۹۴	۵۶	۷۲	۷۲
زمان	۱ دقیقه	۳۰ ثانیه	۳۰ ثانیه	۹۰ ثانیه	۷ دقیقه

در این سرور برای ایجاد مدل هر پروتئین در مجموع پنج مدل تولید می‌شود. در این وب سرور، مدل های سه بعدی بر اساس بازآرایی چندگانه LOMETS ساخته می‌شوند (۱۹). مدلهایی با بهترین امتیاز Z-score توسط این سرور انتخاب می‌شوند. در ادامه به کمک سرور خصوصات فیزیکی و شیمیایی این آنزیم مطالعه و بررسی شد (۱۳). همچنین، به کمک نرم‌افزارهای SPDBV و سیستم

گرافیکی مولکولی PyMOL این ساختار مدل سازی شده مطالعه و ساختار آن از نظر شباهت با آنزیم های مشابه بررسی شد (۷, ۱۱). ویژگی‌های پیوندهای هیدروژنی به کمک سرورهای What IF و PIC نیز محاسبه شدند (۱۸, ۱۶).

سویه‌ی *آئورانتیوکیتریوم* F4 در محیط کشت مایع GYP کشت داده شد. . نتایج مطالعات میکروسکوپی نشان دادند که این سویه دارای دیواره‌ی پلی‌ساکاریدی کاملا مشخص است و همچنین اندازه سلول‌ها ۱۵ میکرومتر می باشد (شکل۱).

نتایج



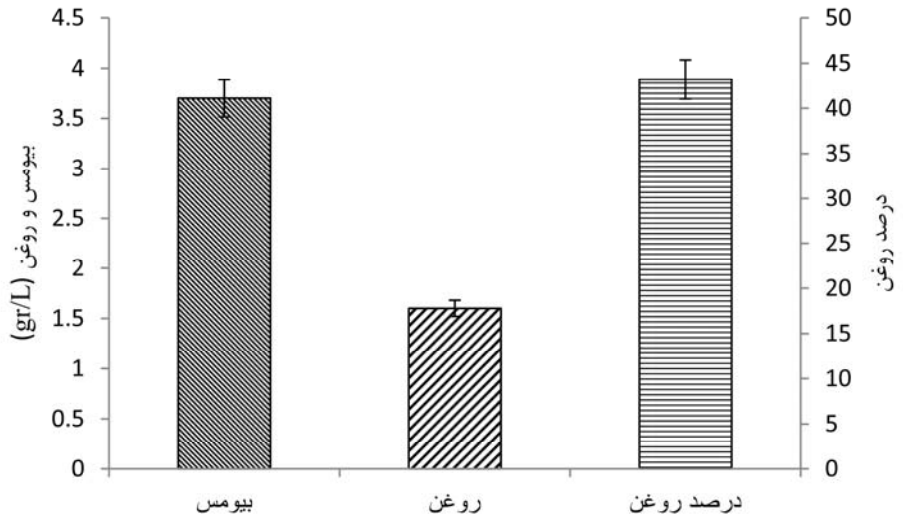
شکل۱- سویه F4 در روز چهارم از کشت: a: سلول درحال تقسیم سلولی، b: سلول منفرد دارای دیواره پلی‌ساکاریدی کاملا مشخص، c: سلول در حال تقسیم و تشکیل کلاستر سلولی

اندازه گیری بیومس، استخراج روغن و انجام آنالیز TLC : نتایج اندازه گیری میزان بیومس و روغن تولیدی توسط سویه‌ی F4 نشان داد که این سویه در روز چهارم از کشت میزان ۳/۷ gr/L بیومس و ۱/۶ gr/L روغن تولید می‌کند، که این میزان تولید روغن معادل ۴۳/۲ درصد از بیومس می‌باشد (شکل۲).

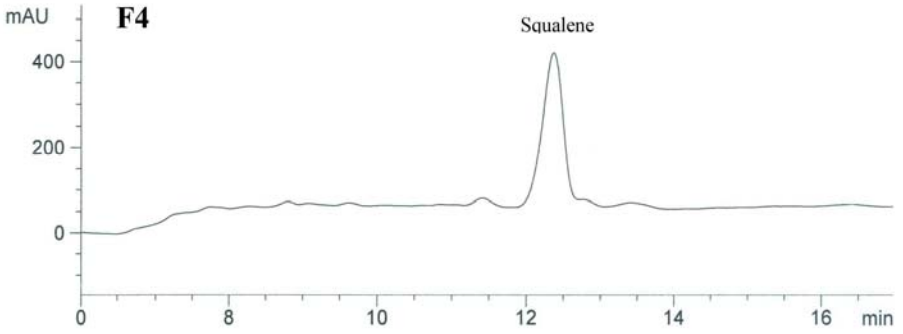
همچنین نتایج کروماتوگرافی لایه نازک نشان داد که سویه-ی F4 اسکوالن تولید میکند. لکه‌ی اسکوالن در این سویه بوسیله کاغذ شاهد حاوی اسکوالن استاندارد شناسایی و Rf آن طبق فرمول اندازه گیری شد (شکل۳). Rf اسکوالن در این تحقیق ۰/۹۷ اندازه‌گیری شد.

آنالیز HPLC : صابونی کردن روغن سلولی و آماده سازی نمونه برای HPLC انجام شد، از اسکوالن استاندارد به عنوان شاهد برای یافتن زمان خارج شدن اسکوالن از دستگاه استفاده شد. نتایج نشان دادند که پیک مربوط به اسکوالن از سویه F4 در دقیقه‌ی ۱۲/۳ ظاهر می‌شود (شکل۴)، با استفاده از منحنی استاندارد اسکوالن و سطح زیر منحنی بدست آمده از پیک اسکوالن در این سویه، مقدار کمی اسکوالن تولیدی اندازه گیری شد. نتایج نشان دادند که سویه F4 در روز چهارم از کشت میزان mg/L ۴۸/۹ اسکوالن تولید می‌کند که این میزان از اسکوالن ۱/۲ درصد از بیومس و ۳ درصد از روغن تولیدی توسط این سویه می‌باشد (شکل۵).

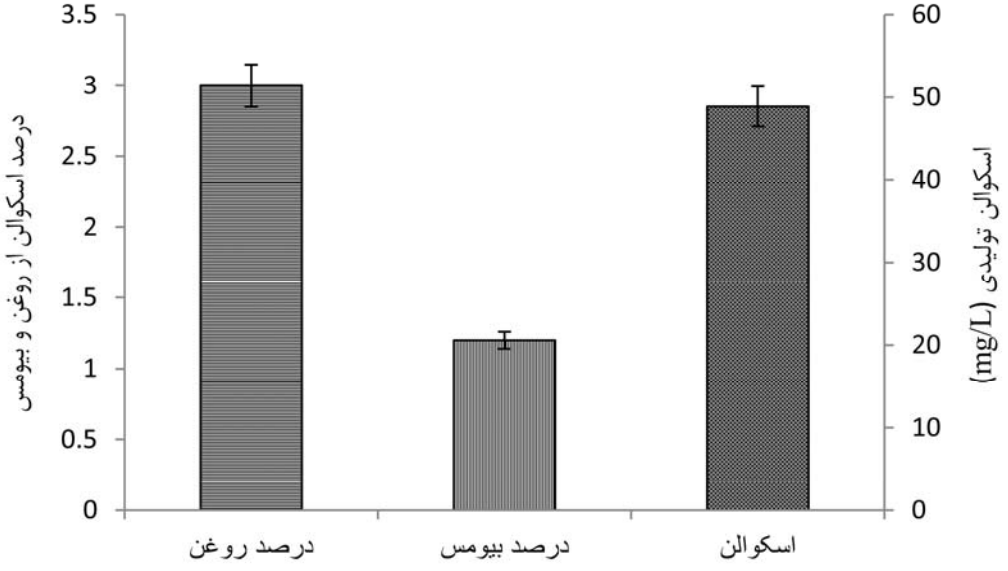
PCR و تولی‌یابی : طول توالی ژن اسکوالن سنتاز ۱۱۶۴ جفت باز می‌باشد که با انجام PCR این توالی به خوبی تکثیر شد (شکل۶)، ژن اسکوالن سنتاز بوسیله شرکت یکتا تجهیز آزما از ژل الکتروفورز ۲ درصد استخراج شد و توالی این ژن با استفاده از همان پرایمرهای رفت و برگشت مورد استفاده در PCR بدست آمد.



شکل۲- میزان بیومس و روغن و درصد روغن تولیدی توسط سویه F4 در روز چهارم از کشت

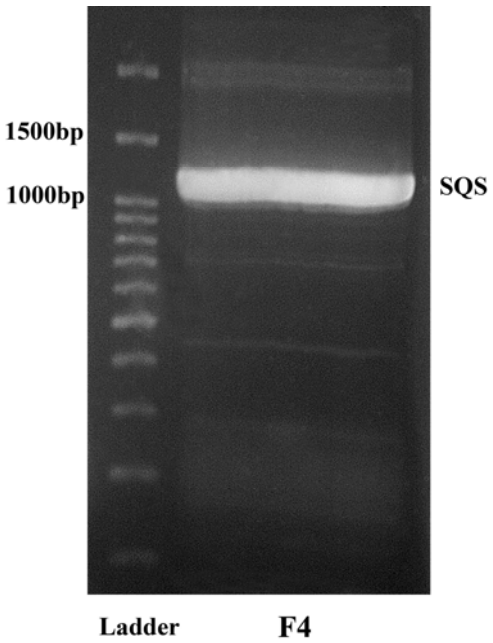


شکل۴- پیک مربوط به اسکوالن از سویه F4



شکل۵- میزان اسکوالن تولیدی توسط سویه F4 و درصد اسکوالن از روغن و بیومس از این سویه

شکل۳- لکه اسکوالن از سویه F4 در مقایسه با لکه اسکوالن استاندارد به خوبی قابل مشاهده است.



شکل۶- باند مربوط به ژن اسکوالن سنتاز در سویه F4

مطالعه توالی نوکلئوتیدی: به منظور بررسی ویژگی کدونی آنزیم اسکوالن سنتاز، توالی نوکلئوتیدی این آنزیم در پایگاه محاسباتی http://ppuigbo.me/programs/CAIcal بررسی گردید (شکل ۷). همانطور که در شکل ۷ نشان داده شده، طول این ژن حدود ۱۱۶۱ نوکلئوتید است و تعداد و درصد نوکلئوتیدهای آدنین، سیتوزین، تیمین و گوانین، درصد نوکلئوتیدهای گوانین با سیتوزین، گوانین با آدنین، ، گوانین با تیمین، آدنین با تیمین، آدنین با سیتوزین و سیتوزین با تیمین نشان داده شده است. همچنین، تعداد و درصد نوکلئوتیدهای آدنین، سیتوزین، تیمین و گوانین در موقعیت اول، دوم و سوم کدون سه تایی و درصد نوکلئوتیدهای گوانین با سیتوزین، گوانین با آدنین، ، گوانین با تیمین، آدنین با تیمین، آدنین با سیتوزین و سیتوزین با تیمین در موقعیت اول، دوم و سوم کدون سه تایی نشان داده شده است (شکل ۷).

در ادامه نمودار گرافیکی ویژگی کدونی این آنزیم نیز به کمک پایگاه http://gcua.schoedl.de/seqoverall_v2.html رسم گردید. در این آنالیز به کمک داده‌های مربوط به توالی کدونی گونه *Schizosaccharomyces pombe* میزان شباهت این توالی با داده های کدونی این گونه آنالیز شده و نمودار گرافیکی آن رسم گردید (شکل ۸). در شکل ۸، درصد شصت و چهار نوکلئوتیدهای سه تایی مربوط به کدون های بیست اسید آمینه و کدون توقف توالی نوکلئوتیدی آنزیم اسکوالن سنتاز در مقایسه با گونه نزدیک به آن با نام *Schizosaccharomyces pombe* که دادهای ژنی آن در این پایگاه موجود است نشان داده شده است. این مقایسه نشان می‌دهد که در تعدادی از کدونهای شصت وچهارتایی فراوانی آنها در توالی نوکلئوتیدی آنزیم اسکوالن سنتاز از گونه *Schizosaccharomyces pombe* بیشتر بوده و در بعضی از کدونها همانند کدون ACT از ترئونین و کدون TGG متعلق به اسید آمینه تریپتوفان این فراوانی **کمتر شده است**.

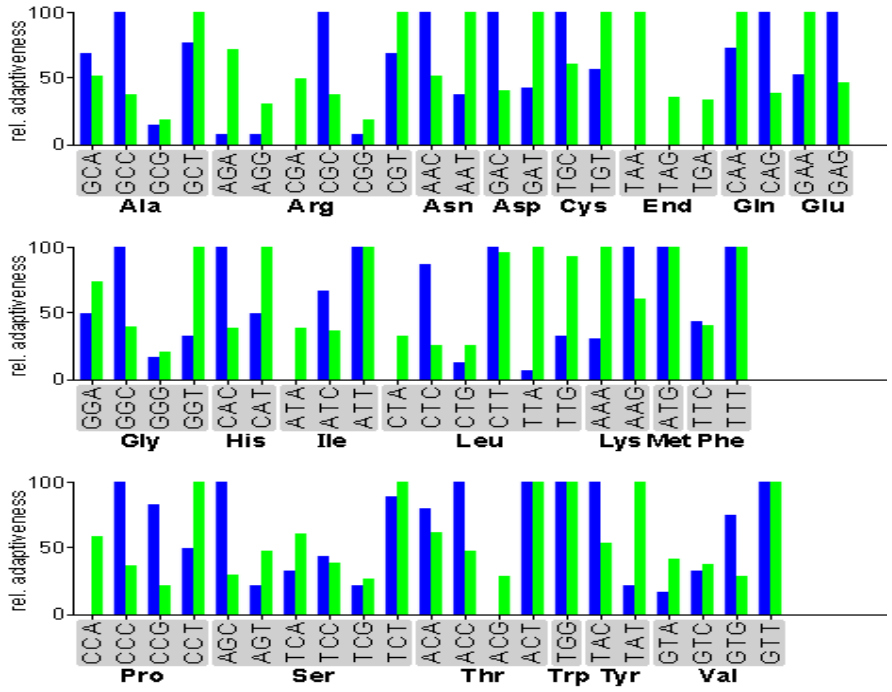
مطالعه ساختار آنزیم: برای مطالعه دقیق ساختار این آنزیم، مدل سه بعدی این آنزیم مورد نیاز است. در این زمینه، ساختار سه بعدی این آنزیم با ثبت توالی پروتئینی آن در سرور I-TSSAR به دست آمد. نتایجی که از وب

سرورهای I-TSSAR و Swiss-Model بدست می آیند محدود بوده و می توان جهت کارهای دقیق تر از روشهای Molecular Dynamics Simulation استفاده نمود. این سرور از ساختار کریستالوگرافی ده آنزیم مشابه با اسکوالن سنتاز جهت مدل‌سازی استفاده می‌کند. کد ثبت ساختار این آنزیم‌ها در سایت PDB (http://www.rcsb.org/pdb/home/home.do) و ویژگی‌های آنها و همچنین هم‌ترازی ناحیه N–ترمینال این آنزیم‌ها در شکل زیر نشان داده شده است (شکل ۹). در این شکل، شماره کد پروتئینی این ده آنزیم الگو در سمت چپ و میزان همسانی این توالی‌ها با توالی آنزیم اسکوالن سنتاز مورد نظر، میزان پوشش این هم ترازی و همچنین z-score نرمالیز شده را نشان می‌دهد. بخش سمت راست تصویر هم هم ترازی توالی آنزیم اسکوالن سنتاز مورد نظر با توالی پروتئینی این ده آنزیم الگو را نشان می‌دهد.

SEQUENCES \ PARAMETERS	length	A	C	T	G	%A	%C	%T	%G	% G+C	% G+A	%G+T	%A+T	% A+C	%C+T	A1	C1	T1	G1	%A1	%C1	%T1	%G1	% G1+C1	% G1+A1	% G1+T1	% A1+T1	% A1+C1	% C1+T1
	1161	257	318	283	303	22.14	27.39	24.38	26.10	53.49	48.23	50.47	46.51	49.53	51.77	88	95	63	141	22.74	24.55	16.28	36.43	60.98	59.17	52.71	39.02	47.29	40.83

SEQUENCES \ PARAMETERS	A2	C2	T2	G2	%A2	%C2	%T2	%G2	% G2+C2	% G2+A2	% G2+T2	% A2+T2	% A2+C2	% C2+T2	A3	C3	T3	G3	%A3	%C3	%T3	%G3	% G3+C3	% G3+A3	% G3+T3	% A3+T3	% A3+C3	% C3+T3	% G3s+C3s
	123	79	109	76	31.78	20.41	28.17	19.64	40.05	51.42	47.80	59.95	52.20	48.58	46	144	111	86	11.89	37.21	28.68	22.22	59.43	34.11	50.90	40.57	49.10	65.89	57.45

شکل ۷– مطالعه ویژگی کدونی آنزیم اسکوالن سنتاز



شکل ۸– نمودار گرافیکی ویژگی کدونی این آنزیم اسکوالن سنتاز (رنگ سبز) و سویه *Schizosaccharomyces pombe* (رنگ آبی)

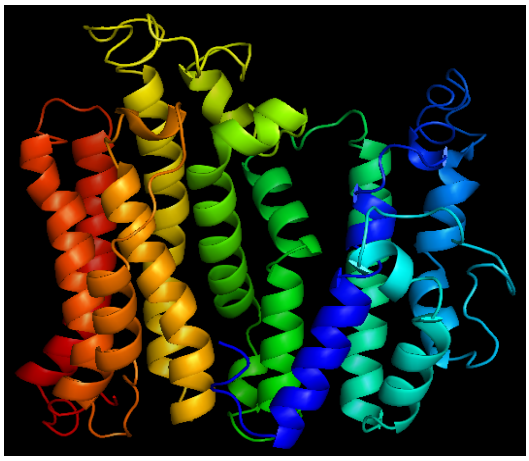
Rank	PDB Hit	Iden1	Iden2	Cov	Norm. Z-score
1	3wcaA	0.38	0.34	0.82	2.02
2	3vj8A	0.39	0.34	0.80	4.18
3	3wcaA	0.38	0.34	0.82	2.18
4	3wcaA	0.39	0.34	0.82	6.17
5	3wcaA	0.39	0.34	0.82	4.67
6	3wcaA	0.39	0.34	0.81	2.90
7	3wcaA	0.37	0.34	0.82	7.31
8	3vj9A	0.39	0.34	0.81	4.28
9	3wcaA	0.38	0.34	0.82	1.84
10	1ezfA	0.00	0.33	0.78	1.64

```
                    20                  40                  60
                     |                   |                   |
Sec.Str CCCCCCHHHHHHHHHHHHHHHHCCCCCCCCCCCCHHHHHHHHHHHHHHHHCCCHHHHHHHHHHHHHHHHHCCCCC
Seq     MPNKPDAPIRLAVGIFYIVLRALDTVEDDMNLSSFDSYVLEEDKKDVEDARSAAKQRLLCQFAQRLSDSVEGNA
1       I-MELDEEMRDAVCIFYLVLRALDTVEDDMSIP------------------VEFKLRELPKFHEHLHD------
2       --QALDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVP------------------LLHNFHSFLYQPDWRFM
3       VIMELDEEMRDAVCIFYLVLRALDTVEDDMSIP------------------VEFKLRELPKFHEHLHD------
4       VIMELDEEMRDAVCIFYLVLRALDTVEDDMSIPV--E----------------FKLRELPKFHEHLHDTTWC--
5       IM-ELDEEMRDAVCIFYLVLRALDTVEDDMSIPVE------------------FKLRELPKFHEHLHDTTWC-M
6       ----LDEEMRDAVCIFYLVLRALDTVEDDMSIP------------------VEFKLRELPKFHEHLHDT-----
7       IMELDE-EMRDAVCIFYLVLRALDTVEDDMSIPVEFKLR------------------ELPKFHEHLHDT-TWCM
8       VIQALDGEMRNAVCIFYLVLRALDTLEDDMTIS------------------VEKKVPLLHNFHSFLYQPDWRFM
9       I-MELDEEMRDAVCIFYLVLRALDTVEDDMSIP------------------VEFKLRELPKFHEHLHDTTWCMS
10      ------------------------------------------------------------------------
```

شکل ۹- کد ساختاری، ویژگیها و همچنین هم‌ترازی ناحیه N-ترمینال آنزیم‌های مورد استفاده در مدل‌سازی

این سرور پنج مدل ایجاد کرد که بهترین آنها را انتخاب کرده و به کمک نرم‌افزارهای PDB و PyMOL مطالعه شدند. میزان C-score این مدل برابر با ۰/۲۱- و میزان TM-Score آن برابر ۰/۶۹±۰/۱۲ می باشد. همچنین اندازه RMSD آزمایشگاهی برابر با ۴/ ۲ Å ۷/۲± Å بدست آمد. این آنزیم ۳۸۷ آمینواسید دارد. C-score یک نمره نشان دهنده میزان اعتماد به کیفیت مدل پیش بینی شده توسط I-TASSERاست. این محاسبه بر اساس هم ترازی الگوهای استفاده شده است. C-score معمولا در محدوده ۲- تا ۵+ بوده و بالاتر این عدد نشان دهنده اطمینان بیشتر به مدل می‌باشد. TM-score یک مقیاس برای اندازه گیری شباهت ساختاری بین دو ساختار است. هدف از TM-score برای حل مسئله RMSD می‌باشد. از آنجا که RMSD میانگین فاصله تمام جفت های باقیمانده در دو ساختار است، یک خطای کوچک، مقدارRMSD را افزایش می دهد. نمره حساسیت به خطای مدل سازی محلی. TM-score بیشتر از ۰/۵ نشان دهنده توپولوژی درست مدل و کمتر از ۰/۱۷ به معنی شباهت تصادفی مدل با ساختار الگو است. شکل ۱۰ تصویر ساختار مدل سازی شده آنزیم اسکوالن سنتاز مورد نظر را نشان می‌دهد (شکل ۱۰). همانطور که شکل نشان می‌دهد ساختار آنزیم اسکوالن سنتاز مدل‌سازی شده فقط از زنجیره‌های آلفا هلیکس تشکیل شده و صفحات بتا در آن وجود ندارد (شکل ۱۰).

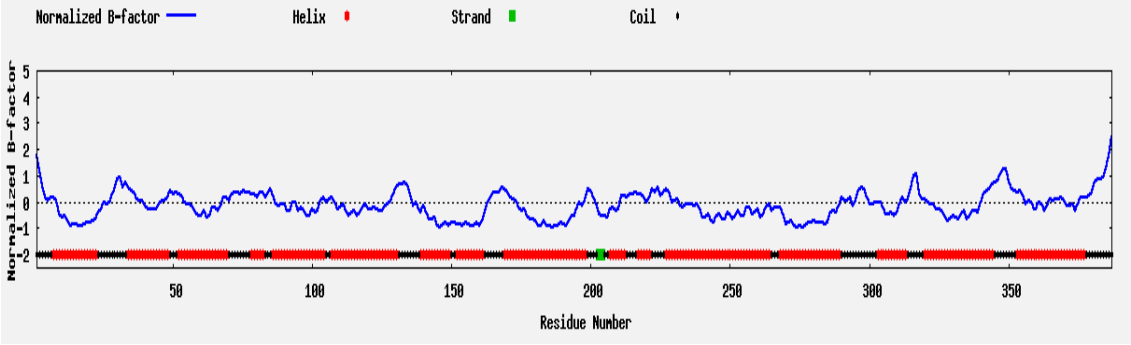
همانطور که در شکل ۱۱ نشان داده شده است در این آنزیم میزان بتا فاکتور (β-factor) پروتئین به شکل زیر می‌باشد. بتافاکتور (β-factor) عاملی است که میزان تحرک پذیری حرارتی ذاتی اسیدهای آمینه در ساختار پروتئین‌ها را نشان می‌دهد. در I-TASSER، این مقدار از هم‌ترازی پروتئین‌های PDB در ترکیب با پروفایل توالی‌های به دست آمده از پایگاه‌های ارائه کننده داده‌های توالی به دست می‌آید. نمودار آبی رنگ در این شکل نشان دهنده بتافاکتور نرمالیز شده و بخش‌های قرمز رنگ، نواحی از توالی که در توالی آمینواسیدی ساختار آلفا هلیکس تشکیل می دهند را نشان می‌دهد. همچنین، نواحی سبز نشان دهنده قسمت‌های تشکیل دهنده صفحات بتا و نواحی مشکی بخش‌هایی است که تشکیل لوپ می‌دهند را نشان می‌دهد.



شکل ۱۰- تصویر ساختار سه بعدی آنزیم مدل‌سازی شده اسکوالن سنتاز

البته این قسمت بسیار محدود بوده و صرفا احتمال تشکیل را نشان می دهد که ممکن است به صورت بتا دیده نشوند و بصورت لوپ در ساختر نمایش داده شوند.

در ادامه خصوصیات فیزیکوشیمیایی این آنزیم به کمک سرور ExPASy - ProtParam tool با بارگذاری اطلاعات توالی در این سرور محاسبه شد که در جدول زیر نشان داده شده است (۱۳) (جدول ۳). همچنین میزان اسیدهای آمینه این آنزیم محاسبه شد و نتایج نشان می دهد که اسیدهای آمینه لوسین، آلانین و آسپارتیک اسید بیشترین فراوانی را در این آنزیم دارند (جدول ۴).



شکل ۱۱- میزان بتا فاکتور در ساختار آنزیم اسکوالن سنتاز

جدول۳- خصوصیات فیزیکوشیمیایی آنزیم اسکوالن سنتاز

No.	Parameters	CDase
1	Theoretical pI	5.16
2	Molecular weight	43285.19
3	Sequence length	387
4	Extinction coefficients (M-1 cm-1at 260 nm)*	44515- 43890
5	Asp + Glu	56
6	Arg + Lys	42

جدول ۴- میزان اسیدهای آمینه آنزیم اسکوالن سنتاز

Amino acid composition:

Ala	(A)	34	8.8%
Arg	(R)	25	6.5%
Asn	(N)	11	2.8%
Asp	(D)	33	8.5%
Cys	(C)	11	2.8%
Gln	(Q)	19	4.9%
Glu	(E)	23	5.9%
Gly	(G)	24	6.2%
His	(H)	9	2.3%
Ile	(I)	20	5.2%
Leu	(L)	36	9.3%
Lys	(K)	17	4.4%
Met	(M)	13	3.4%
Phe	(F)	13	3.4%
Pro	(P)	14	3.6%
Ser	(S)	28	7.2%
Thr	(T)	14	3.6%
Trp	(W)	5	1.3%
Tyr	(Y)	11	2.8%
Val	(V)	27	7.0%
Pyl	(O)	0	0.0%
Sec	(U)	0	0.0%
	(B)	0	0.0%
	(Z)	0	0.0%
	(X)	0	0.0%

بحث

آنزیم اسکوالن سنتاز طی یک واکنش دو مرحله‌ای، با اتصال دو مولکول فارنسیل دی فسفات، سنتز یک مولکول اسکوالن را کاتالیز می‌کند. اخیراً نشان داده شده است که آنزیم‌های دیگری نیز درکنترل شار کربن به سمت ایزوپرنوئیدهای مختلف در این مسیر دخیل هستند (۲۲). یکی از این آنزیم‌ها، آنزیم اسکوالن سنتاز می‌باشد که اولین مرحله در بیوسنتز کلسترول در پستانداران و ارگوسترول در قارچ‌ها را کاتالیز می‌کند. اسکوالین سنتاز و یا فارنسیل-دی فسفات: فارنسیل-دی فسفات ترانسفراز فارنسیل یک آنزیم متصل به غشاء شبکه آندوپلاسمی بوده و در مسیر

بیوسنتز ایزوپرنوئید شرکت می‌کند. در این مسیر تسریع یک واکنش دو مرحله است که دو مولکول یکسان از فارنسیل پیرو فسفات (FPP) با مصرف NADPH به اسکوالن تبدیل می‌شوند. به تازگی ژن جدید آنزیم اسکوالن سنتاز از سویه های دریازی ترائوستوکیترید کلون شده و توالی ژنی آن در بانک ژن ثبت گردیده است (۲۲). در این مطالعه انجام آنالیزهای بیوانفورماتیکی ژنی و پروتئینی انجام شد. قبلا نیز چند آنالیز بیوانفوراتیکی بر روی پروتئین‌های ویروس هپاتیت C و B و همچنین آنزیم لوسیفراز انجام شده است (۱۴). به منظور ارزیابی بیوانفورماتیکی، توالی نوکلئوتیدی این آنزیم در برنامه CAIcal و .seqoverall_v2 مورد ارزیابی قرار گرفت. نتایج نشان می‌دهد که میزان همپوشانی این توالی آنزیمی با گونه *Schizosaccharomyces pombe* حدود ۳۵٪ می‌باشد که همپوشانی پایینی است. همچنین مدل‌سازی ساختاری نیز در سایت انجام شد که آنالیز این مدل نشان می‌دهد که تمام این ساختار از زنجیره‌های آلفا تشکیل شده و صفحات بتا در آن مشاهده نمی‌شود. مدل‌سازی ساختار پروتئین به معنای ساخت مدل سه‌بعدی پروتئین از روی دنباله آمینواسیدهای آن و تعیین ساختار دوم و سوم از روی ساختار اولیه پروتئین است. تعیین ساختار پروتئین از مبنا با مسأله طراحی یک پروتئین متفاوت است. مدل‌سازی پروتئین یکی از مسائل مهم در حوزه بیوانفورماتیک و شیمی تئوری است و اهمیت زیادی در پزشکی (دارو) و بیوتکنولوژی دارد. بعد از اتمام فرایند مدل‌سازی، ویژگی های ساختاری این مدل نیز بررسی گردید. مطالعات نشان می‌دهد که زنجیره های جانبی مختلف تمایل متفاوتی برای حضور در مارپیچ آلفا دارند. بنابر این آلانین، گلوتامیک اسید، لوسین و متیونین تشکیل دهنده های خوب مارپیچ آلفا هستند در حالیکه پرولین، گلایسین، تایروزین و سرین از این نظر مناسب نیستند (۱۵). این تمایلات در پیشگویی ساختمان دوم از توالی اسیدهای آمینه نقطه کانونی به حساب آمده اما به اندازه کافی قوی نیستند که باعث پیشگویی دقیقی شوند. نتایج محاسبه میزان اسیدهای آمینه این آنزیم نشان می دهد که اسیدهای آمینه لوسین، آلانین و آسپارتیک اسید بیشترین فراوانی را در این آنزیم دارند و با توجه به رابطه فراوانی اسیدهای آمینه در تشکیل ساختارهای ثانویه (۱۵, ۶) بنابراین احتمال تشکیل مارپیچ‌های آلفا در این ساختار زیاد بوده و به نظر می‌رسد که احتمالا کیفیت مدل ساختاری آنزیم اسکوالن سنتاز مناسب باشد. همچنین لازم است که پروموتور این ژن توسط روش های بیوانفورماتیکی شناسایی شده (۲) و سپس توسط کلونینگ ژن مورد نظر و بیان آنزیم اسکوالن سنتاز، پایداری آن در شرایط فیزیکی و شیمیایی مورد بررسی قرار گیرد (۳). این مطالعه شناخت جامعی از توالی ژنی، توالی پروتئینی و ساختار آنزیم را فراهم کرده و در استفاده و کاربردهای بیوتکنولوژی این آنزیم بسیار حائز اهمیت است.

منابع

۱. فکرت ف. و شاکری ش. ۱۳۹۴. بررسی و تولید روغن امگا۳ غنی از دوکوزاهگزانوئیک اسید توسط سویه بومی آئورانتیوکیتریوم TA4 . زیست شناسی میکروارگانیسم ها، جلد ۴، شماره ۱۴، ص ۹-۲۴.

۲. زینتی ز. و شاملو دشت پاگردی ر. ۱۳۹۶. بررسی مقایسهای پروموترهای ژنهای *Catalase 2* و*Thioredoxin H5* در آرابیدوپسیس به منظور تعیین نحوه کارکرد آنها در پاسخ به تنشهای زیستی و غیر زیستی. مجله پژوهش های سلولی و مولکولی (مجله زیست شناسی ایران)، جلد ۳۰. شماره ۴. ص ۴۸۸-۴۹۸.

۳. خالقی نژاد س.ح.، کارخانه ع.ا.، مطلب غ.، امین زاده س. و یخچالی ب. ۱۳۹۴. کلونینگ، بیان و تعیین خصوصیات لیپاز کایمریک باسیلوس ترموکاتنولاتوس در باکتری *Escherichia coli*. مجله پژوهش های سلولی و مولکولی (مجله زیست شناسی ایران)، جلد ۲۸. شماره ۲. ص ۲۰۲-۲۱۰.

4. Ačimovič, J. and Rozman, D., 2013 "Steroidal triterpenes of cholesterol synthesis," *Molecules,* vol. 18, pp. 4002-4017.

5. Bligh, E. G. and Dyer, W. J., 1959 "A rapid method of total lipid extraction and purification," *Canadian journal of biochemistry and physiology,* vol. 37, pp. 911-917.

6. Crick, F. H., 1953 "The packing of α-helices: simple coiled-coils," *Acta crystallographica,* vol. 6, pp. 689-697.

7. DeLano, W. L., 2002 "The PyMOL molecular graphics system."

8. Fuhrmann, M., Hausherr, A., Ferbitz, L., Schödl, T., Heitzer, M., and Hegemann, P., 2004 "Monitoring dynamic expression of nuclear genes in Chlamydomonas reinhardtii by using a synthetic luciferase reporter gene," *Plant molecular biology,* vol. 55, pp. 869-881.

9. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., *et al.*, *Protein identification and analysis tools on the ExPASy server*: Springer, 2005.

10. Hong, W.-K., Heo, S.-Y., Park, H.-M., Kim, C. H., Sohn, J.-H., Kondo, A., *et al.*, 2013 "Characterization of a squalene synthase from the Thraustochytrid Microalga Aurantiochytrium sp. KRS101," *J Microbiol Biotechnol,* vol. 23, pp. 759-765.

11. Kaplan, W. and Littlejohn, T. G., 2001 "Swiss-PDB viewer (deep view)," *Briefings in bioinformatics,* vol. 2, pp. 195-197.

12. Langdon, R. G. and Bloch, K., 1953 "The biosynthesis of squalene," *Journal of Biological Chemistry,* vol. 200, pp. 129-134.

13. Matsuura, H., Watanabe, M. M., and Kaya, K., 2012 "Squalene quantification using octadecylbenzene as the internal standard," *Procedia Environmental Sciences,* vol. 15, pp. 43-46.

14. Mortazavi, M., Zarenezhad, M., Alavian, S. M., Gholamzadeh, S., Malekpour, A., Ghorbani, M., *et al.*, 2016 "Bioinformatic Analysis of Codon Usage and Phylogenetic Relationships in Different Genotypes of the Hepatitis C Virus," *Hepatitis Monthly,* vol. 16, pp. 1-8.

15. Prothero, J., 1966 "Correlation between the distribution of amino acids and alpha helices," *Biophysical journal,* vol. 6, pp. 367-370.

16. Tina, K., Bhadra, R., and Srinivasan, N., 2007 "PIC: protein interactions calculator," *Nucleic acids research,* vol. 35, pp. 473-476.

17. Toborek, M., Lee, Y. W., Garrido, R., Kaiser, S., and Hennig, B., 2002 "Unsaturated fatty acids selectively induce an inflammatory environment in human endothelial cells," *The American Journal of Clinical Nutrition,* vol. 75, pp. 119-125.

18. Vriend, G., 1990 "WHAT IF: a molecular modeling and drug design program," *Journal of molecular graphics,* vol. 8, pp. 52-56.

19. Wu, S. and Zhang, Y., 2007 "LOMETS: a local meta-threading-server for protein structure prediction," *Nucleic acids research,* vol. 35, pp. 3375-3382.

20. Zhang, X., Agrawal, A., and San, K. Y., 2012 "Improving fatty acid production in escherichia coli through the overexpression of malonyl coA-Acyl carrier protein transacylase," *Biotechnology progress,* vol. 28, pp. 60-65.

21. Zhang, Y., 2008 "I-TASSER server for protein 3D structure prediction," *BMC bioinformatics,* vol. 9, pp. 1-8.

22. Zhao, M., Liang, W., Zhang, D., Wang, N., Wang, C., and Pan, Y., 2007 "Cloning and characterization of squalene synthase (SQS) gene from Ganoderma lucidum," *Journal of microbiology and biotechnology,* vol. 17, pp. 1106-1112.

Investigation and bioinformatics analysis of squalene synthase gene and protein in native strain of *Aurantiochytrium*

Mortazavi M., Shakeri Sh., Maleki M. and Khoshbasirat F.

Biotechnology Dept., Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, I.R. of Iran

Abstract

Squalen is a polyunsaturated triterpene with wide range of applications in pharmacy. In this study, production of squalen and bioinformatic analysis of corresponding gene were investigated in native strain of *Aurantiochytrium*. This strain produced 3.7 and 1.6 g/l biomass and oil rich in squalene, respectivelly. The amount of produced squalene by this strain was assessed as 48.9 mg/l. Also, *Aurantiochytrium* squalen synthase gene was identified and sequenced by molecular analysis. Using the bioinformatics software's, some of codon parameters were evaluated and graphical chart was drawn. By use of modeling databases and their software's, the modeling process of this enzyme was conducted. The results of modeling confirmed the role of some amino acids as alanine, glutamic acid, leucine in construction of alpha helix structure. These types of studies help in identification of squalene synthase reaction mechanism. Evaluation of these hidden information can improve the knowledge of the squeal-like scalene folding process and the challenges of protein expression and will be effective in the new design of the enzyme.

Keywords: Squalene, Squalene synthase, *Aurantiochytrium*