

طراحی بیوانفورماتیکی واکسن علیه بیماری ویبریوکلرا با استفاده از رویکرد معکوس

Bioinformatics design of a vaccine against *Vibrio cholerae* using the reverse approach

مینا عقیبی طلب، محمدجواد دهقان عصمت آبادی^{*}، علیرضا سعیدنی نیا، محمدعلی یعقوبی مقدم

Mina Oghbatalab, Mohammad Javad Dehgha Esmatabadi^{*}, Alireza Saeedinia,
Mohammad Ali Yaghobi Moghaddam

مجتمع دانشگاهی شیمی و مهندسی شیمی، دانشگاه صنعتی مالک اشتر، تهران، ایران

Faculty of Chemistry and Chemical Engineering, Malek Ashtar University of
Technology, Tehran, Iran

^{*} نویسنده مسئول مکاتبات، پست الکترونیکی: mohammad_dehghan@mut.ac.ir

(تاریخ دریافت: ۹۹/۶/۷ - تاریخ پذیرش: ۹۹/۸/۲۶)

چکیده

واژه‌های کلیدی

طراحی واکسن در گذشته تا به امروز براساس روش‌های کلاسیک صورت می‌گیرد که زمان‌بر، هزینه‌بر و دارای ملاحظات هستند، اما با کشف ژنوم و تعیین توالی ژن، روش‌های مدرنی برای طراحی واکسن ابداع شده که از جمله آن می‌توان به روش طراحی واکسن به روش معکوس اشاره کرد. در این پژوهش از سرور NCBI (ORF Finder) به منظور یافتن توالی‌های Coding و برای مکان‌یابی پروتئین از سرور PSORTb استفاده شده و بررسی آلرژنیک بودن پروتئین از سرور Algpred و به منظور پیش‌بینی اپی توپ خطی و فضایی از سرورهای LBtope - ABCpred - IEDB CBTOPE - Disco Tope 2.0 - IEDB (Ellipro) انتخاب شده و سپس تجزیه و تحلیل سازه پلی توپی را توسط سرور ExPasy (ProtParam) بررسی کردیم و در نهایت با سرور JCat توالی نوکلئوتیدی بهینه شده را بدست آوردیم. با استفاده از این روش ۶۰۷۴ چارچوب خوانش و همچنین ۴۳ پروتئین سطحی و ترشحی شناسایی گردیده که در نهایت، ۱۰ اپی توپ مناسب به عنوان کاندیدای واکسن قادر به ایجاد ایمنی در برابر طیف وسیعی از سویه‌ها می‌باشند، امید است که راه پیشرفت تولید این واکسن را بسیار آسان کنند.

بیوانفورماتیک،

پروتئین،

سازه پلی توپی،

عامل وبا،

واکسن‌سازی به روش معکوس

مقدمه

بر اساس شواهد تاریخ، هیچ بیماری هولناک‌تر از وبا نبوده است. در مورد مرگ ناشی از وبا که قدمت آن به دوران بقراط و نوشته‌های سانسکریت باز می‌گردد، اشاره شده است. گمان می‌رود که منشأ باستانی این بیماری از هند و دلتاهای رود گنگ است. ضعف شرایط بهداشتی در آنجا و استفاده از استخرهای آلوده‌ای که آب در آنها ساکن بود، باعث شد که از عهد باستان، موارد این بیماری در هند رایج شود. بیماری نخست از طریق راه‌های تجاری خشکی و دریایی در سال ۱۸۱۷ به روسیه رسید. سپس از روسیه به موازات پیشرفت‌های فناوری در این کشور و مبادلات تجاری، به اروپا و سپس آمریکای شمالی رسید.

در فاصله زمانی ۱۵۰ ساله هفت پاندمی اصلی در جهان رخ داده: - همه‌گیری اول: در قسمت بنگال هند بین سال‌های ۱۸۱۷ تا ۱۸۲۴ رخ داد. بیماری از هند به جنوب شرقی آسیا و سپس به چین، ژاپن، خاورمیانه و جنوب روسیه رسید.

- همه‌گیری دوم: بین سال‌های ۱۸۲۷ تا ۱۸۳۵ رخ داد و عمدتاً آمریکا و اروپا را متأثر کرد. دلیل آن هم پیشرفت‌های فناوری و میزان زیاد مبادلات تجاری و مهاجرت‌های گسترده‌ای بود که رخ می‌داد.

- همه‌گیری سوم: بین سال‌های ۱۸۳۹ تا ۱۸۵۶ بود و به آفریقا و جنوب آمریکا رسید به خصوص در برزیل غوغا کرد.

- همه‌گیری چهارم: بین سال‌های ۱۸۳۶ تا ۱۸۷۵ رخ داد و منطقه زیر صحرای بزرگ آفریقا را گرفتار کرد.

- همه‌گیری‌های پنجم و ششم: به ترتیب بین سال‌های ۱۸۸۱ تا ۱۸۹۶ و ۱۸۹۹ تا ۱۹۲۳ رخ دادند. این دو آمار مرگ و میر کمتری داشتند، چون در این زمان باکتری وبا کشف شده بود و راه پیشگیری از آن مشخص شده بود. اما متأسفانه در مصر، ایران، هند و فیلیپین وبا در همین زمان به طرز گسترده‌ای شایع شد و افراد زیادی به خاطر ابتلا به آن درگذشتند.

- آخرین همه‌گیری: در سال ۱۹۶۱ در اندونزی گزارش شد و به خاطر ظهور گونه جدید El Tor بود که این گونه هنوز هم در

کشورهای در حال توسعه شیوع می‌یابد. از قرن نوزدهم به بعد، بیماری وبا باعث مرگ ده‌ها میلیون نفر شده است. در روسیه به تنهایی بین سال‌های ۱۸۴۷ تا ۱۸۵۱، بیش از یک میلیون نفر به خاطر ابتلا به آن درگذشتند و در طی همه‌گیری دوم ۱۵۰ هزار آمریکایی و ۸ میلیون هندی جان باختند (Dehghannejad & Kasiri, 2011). در سطح جهانی، تعداد واقعی وبا بسیار بیشتر از موارد گزارش شده به WHO، تخمین زده می‌شود. در سطح جهان، ۱/۳ میلیارد نفر در کشورهای بومی مبتلا به وبا در معرض خطر هستند. در سال ۲۰۱۰، بزرگترین شیوع وبا در جهان در نیم قرن گذشته رخ داد که بیش از ۶۰۰۰۰۰ فرد را آلوده کرد و حدود ۸۰۰۰ نفر را به قتل رساند (Nair & Takeda, 2014). شیوع بیماری وبا در یمن از اکتبر سال ۲۰۱۶ آغاز شد، از ژانویه سال ۲۰۱۷، ۲۷۵/۹۸۷ مورد، مشکوک به وبا و ۱۶۳۴ مرگ ناشی از وبا ثبت شده است. سازمان بهداشت جهانی و صندوق کودکان سازمان ملل متحد در بیانیه مشترکی اعلام کردند که نوع بیماری وبا رایج در یمن از بدخیم‌ترین نوع آن است و در نتیجه این بیماری تاکنون بیش از ۱۳۰۰ نفر جان خود را از دست داده‌اند و آمار تلفات رو به افزایش است. با آنکه بیماری وبا کاملاً شناخته شده است و در صورت استفاده به موقع از روشهای مناسب، درمانی ساده و موثر دارد ولی همچنان بی‌وقفه به صورت همه‌گیر گسترش می‌یابد و این در حالیست که حتی رعایت حداقل اصول بهداشتی نیز از بروز آن پیشگیری می‌نماید (Razmjoo, 2018).

اگرچه تا سال‌ها استفاده از روش‌های سنتی در تولید واکسن مرسوم بود، اما با پیدایش علوم مانند مهندسی ژنتیک و بیوانفورماتیک، فرصتی برای توسعه واکسن‌های جدید و بهبود یافته فراهم شده است. در این روش، ژنوم کامل یک پاتوژن به کمک تجزیه و تحلیل‌های کامپیوتری بررسی میشود تا آنتی‌ژنهایی که بیش‌ترین احتمال ایمنی‌زایی را دارند، شناسایی شوند (Shaikh et al, 2020). بیوانفورماتیک دانش استفاده از علوم کامپیوتر، آمار و احتمال در شاخه زیست‌شناسی مولکولی است. در چند دهه گذشته، پیشرفت در زیست‌شناسی مولکولی و تجهیزات مورد نیاز تحقیق در این زمینه باعث افزایش سریع تعیین توالی ژنوم بسیاری از گونه‌های موجودات شد، تا جایی که پروژه‌های تعیین توالی ژنوم‌ها از پروژه‌های بسیار رایج به حساب می‌آیند. گسترش

کرده و در نهایت به چند آنتیژن (پلی توپ) دست یافتیم که قادر است بیشترین ایمنی‌زایی را برای بدن ایجاد کند.

روش‌ها و نتایج

شناسایی توالی‌های چارچوب خوانش باز (ORF): در ابتدای توالی کامل باکتری ویبریولرا از سایت NCBI استخراج شده (NC_002505.1) و توسط پایگاه داده ORFfinder توالی‌های چارچوب خوانش باز یک میلیون انتهایی ژنوم باکتری ویبریولرا مورد خوانش قرار گرفته، نمونه‌ای تعداد چارچوب‌ها و شماره نوکلئوتید خوانده شده که در جدول ۱ نشان داده شده است.

مکان‌یابی توالی‌های چارچوب خوانش باز: یکی از لازمه‌های اپی توپ مناسب دسترسی بالا و سطحی بودن آن است. هدف ما در این بخش از طراحی واکنش، پیدا کردن پروتئین‌هایی است که سطحی و ترش‌خی می‌باشند. بدین منظور برای مکان‌یابی از نرم افزار PSORTb استفاده کردیم که در این پروژه ۴۳ کاندید پروتئینی در غشای بیرونی و خارج سلولی قرار داشتند که به عنوان پروتئین‌های کاندید مناسب اپی توپ انتخاب کردیم.

حذف سیگنال پپتید: با توجه به اینکه سیگنال پپتید قسمتی از توالی پروتئین می‌باشد که بعد از ترشح شناسایی و حذف می‌شود، اپی توپ‌هایی که در این ناحیه از پروتئین هستند را به عنوان کاندید اپی توپ در نظر نمی‌گیریم. به همین جهت باید قبل از بررسی‌های بعدی طراحی واکنش نسبت به حذف توالی ذکر شده اقدام کنیم. باید توجه داشت که ممکن است بعضی از پروتئین‌ها فاقد سیگنال پپتید باشند. بدین منظور سیگنال پپتید پروتئین‌های کاندید توسط سرور SignalP شناسایی و سپس حذف شده‌اند.

بررسی حساسیت زایی برای انسان: در ادامه با استفاده از سرور ALGpred پروتئین‌ها از لحاظ آلرژن بودن یا نبودن بررسی شدند که در نهایت پروتئین non allergen را انتخاب کردیم.

روزافزون حجم عظیمی از داده‌های ژنومی و همچنین نیاز به ذخیره، بازیابی و تحلیل مناسب این داده‌ها، موجب پیدایش علم بیوانفورماتیک شد (Del Tordello, et al, 2017). با وجود پیشرفت‌هایی که در درمان بیماری‌های عفونی صورت گرفته، میکروارگانیسم‌های پاتوژن هنوز مهم‌ترین عامل تهدیدکننده‌ی سلامت عمومی‌اند و هرچند واکنش‌های مرسوم، در درمان یا ریشه‌کنی برخی عوامل بیماری‌زا نقش اساسی داشته‌اند اما با پیدایش بیماری‌های نوظهور نیازمند روش‌های جدیدی در زمینه‌ی طراحی واکنش می‌باشد. روش‌های مرسوم در تولید واکنش که در گذشته به کار گرفته می‌شد نیازمند کشت میکروارگانیسم پاتوژن و شناسایی اجزای ایمنی‌زایی آن بود که روشی زمان‌بر بوده و تنها قادر به شناسایی آن دسته از آنتی‌ژن‌هایی است که به میزان زیادی تولید شده و قابلیت تخلیص بالایی دارند، در حالی که فراوانی پروتئین همواره به معنای ایمنی‌زایی نبوده و یا گاهی آنتی‌ژن‌هایی که در شرایط سلول زنده در طی بیماری‌زایی تولید می‌شوند، در شرایط آزمایشگاهی تولید نخواهند شد. از طرف دیگر چنین روش‌هایی در مورد میکروارگانیسم‌های غیرقابل کشت کارایی نخواهد داشت. امروزه پیشرفت در زمینه‌ی تعیین توالی ژنوم و پیدایش فناوری‌های زیستی وابسته به کامپیوتر روش‌های جدیدی هستند که به منظور بررسی آنتی‌ژن‌های محافظ، از جمله طراحی واکنش به روش معکوس را پیش رو قراردادده است (Nandy & Basak, 2019).

در این روش پروتئین‌های سطحی و ترش‌خی احتمالی با یک روش معکوس که به جای میکروارگانیسم از ژنوم شروع می‌شود و با استفاده از روش‌های محاسباتی و تشخیص الگو، شناسایی خواهند شد. بنابراین، از دیگر مزیت‌های این روش، علاوه بر شناسایی تمامی آنتی‌ژن‌هایی که با روش‌های مرسوم قابل بررسی‌اند، قادر به شناسایی آنتی‌ژن‌های جدید نیز می‌باشد که نقش مهمی در ایمنی‌زایی واکنش‌های نسل جدید دارند (Del Tordello, et al, 2017).

ما در این پروژه مناسب‌ترین اپیتوپ‌های باکتری ویبریولرا از نظر ایمنی‌زایی توسط نرم افزارهای مختلف بیوانفورماتیک شناسایی

جدول ۱- شماره خوانش ها و تعداد خوانش یافت شده در یک میلیون انتهایی کروموزوم ویبریولا

Table1. Number of readings and number of readings found in one million ends of *Vibrio cholerae* chromosome

تعداد خوانش	شماره نوکلئوتید خوانده شده	شماره خوانش
۲۹۹	۲۰۴۹۰۰۰-۱۹۹۹۰۰۰	۱
۲۹۹	۲۰۹۹۰۰۱-۲۰۴۹۰۰۱	۲
۳۳۱	۲۱۴۹۰۰۲-۲۰۹۹۰۰۲	۳
۳۶۱	۲۱۴۹۰۰۳-۲۱۹۹۰۰۳	۴
۳۵۳	۲۱۹۹۰۰۴-۲۲۴۹۰۰۴	۵
۳۵۳	۲۲۴۹۰۰۵-۲۲۹۹۰۰۵	۶
۳۳۰	۲۲۹۹۰۰۶-۲۳۴۹۰۰۶	۷
۳۲۶	۲۳۴۹۰۰۷-۲۳۹۹۰۰۷	۸
۲۹۸	۲۳۹۹۰۰۸-۲۴۴۹۰۰۸	۹
۳۴۲	۲۴۴۹۰۰۹-۲۴۹۹۰۰۹	۱۰
۲۹۸	۲۴۹۹۰۱۰-۲۵۴۹۰۱۰	۱۱
۳۰۴	۲۵۴۹۰۱۱-۲۵۹۹۰۱۱	۱۲
۲۹۹	۲۵۹۹۰۱۲-۲۶۴۹۰۱۲	۱۳
۳۱۸	۲۶۴۹۰۱۳-۲۶۹۹۰۱۳	۱۴
۳۴۰	۲۶۹۹۰۱۴-۲۷۴۹۰۱۴	۱۵
۳۱۸	۲۷۴۹۰۱۵-۲۷۹۹۰۱۵	۱۶
۳۴۴	۲۷۹۹۰۱۶-۲۸۴۹۰۱۶	۱۷
۳۲۵	۲۸۴۹۰۱۷-۲۸۹۹۰۱۷	۱۸
۳۰۰	۲۸۹۹۰۱۸-۲۹۴۹۰۱۸	۱۹
۷۴	۲۹۴۹۰۱۹-۲۹۹۹۰۰۰	۲۰

پروتئین های کانید اپی توپ، هیچ گونه همولوژی با انسان و موش یافت نشد.

بررسی دقیق تر جایگاه پروتئین های کانید: به منظور مکان یابی دقیق توالی های آمینو اسیدی در سلول از پایگاه داده TMHMM استفاده شد که در واقع توسط این سرور قسمت های ترانس ممبرین را شناسایی و حذف کردیم زیرا قسمت های ترانس ممبرین به علت در دسترس نبودن اپی توپ های مناسبی نمی باشند. TMHMM در آدرس الکترونیکی <http://www.cbs.dtu.dk/services/TMHMM-2.0> یک روش پیشرفته است که از مدل های ریاضی پیچیده مدل های مخفی

در حقیقت این سرور به ما این امکان را می دهد که پروتئین هایی که برای موجود زنده (انسان یا جانور آزمایشگاهی) تولید حساسیت می کند را حذف کنیم. به عبارت دیگر پروتئین هایی که برای طراحی واکسن انتخاب می شوند نباید آلرژن یا توانایی بالقوه تولید حساسیت را داشته باشند. در طی این بررسی ۳۲ پروتئین از پروتئین ها کانید آلرژن و ۱۱ پروتئین غیر آلرژن بوده است.

بررسی همولوژی پروتئین های کانید: لازم است همولوژی پروتئین های کانید، با انسان و موش بررسی شود، ضرورت انجام این مرحله به علت این است که اگر با انسان و موش همولوژی داشته باشند سیستم ایمنی آن ها را به عنوان پروتئین های خودی در نظر گرفته و سیستم ایمنی فعال نمی شود که پس از بررسی تمام

مارکوف (HMM) برای پیش بینی نواحی هلیکس درون غشایی (ترانس ممبرین) استفاده می‌شود.

بررسی پروتئین های کاندید در نمودار رامچاندرا: نمودار رامچاندرا وضعیت مجاز هر زاویه را برای ساختارهای پروتئین نشان می دهد. Procheck برای ساختار ورودی نمودار رامچاندرا استفاده می‌شود. این نرم افزار در بررسی ایده آل بودن طول پیوندها، زاویه ی پیوندها و مقادیر ϕ و ψ کاربردی منطقی دارد. در نتیجه با پایگاه داده PROCHECK در ناحیه مجاز بررسی کرده و به عنوان گزینه مناسبی برای طراحی واکسن انتخاب کردیم.

جداسازی و انتخاب اپی توپ های نهایی برای طراحی سازه پلی توپی: برای بدست آوردن اپی توپ های کاندید نهایی به

صورت خطی از پایگاه داده های ABCpred، LBtope، BCPRED و IEDB و اپی توپ های کاندید نهایی به صورت ساختار فضایی از پایگاه داده های CBTOPE، Disco Tope و Elli Pro استفاده شده است. سپس اپی توپ های بدست آمده از پایگاه داده های مختلف را به صورت تک آمینواسیدی با هم مورد مقایسه قرار داده و توالی هایی به عنوان اپی توپ مناسب انتخاب می‌شوند که دارای دو شاخص اصلی باشند: ۱- توالی هدف بین تمامی سرورهای پیشگویی اپی توپ B مشترک باشند ۲- اپی توپ های انتخابی Score بالاتری داشته باشند. در نهایت اپی توپ های مشترک را انتخاب و درصد score آنها را محاسبه کردیم و توالی ها با score بالا به عنوان اپی توپ نهایی مناسب انتخاب شدند که در این پروژه به ۱۰ اپی توپ های مناسب دست یافتیم. اپیتوپ های بدست آمده در جدول ۲ نشان داده شده است.

جدول ۲- توالی اپی توپ های خطی و موقعیت اپی توپ های کاندید

Table 2. The sequence of linear epitopes and the position of candidate epitopes

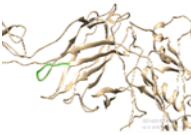
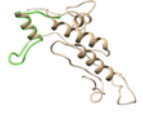
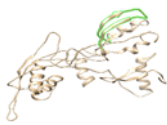
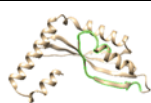
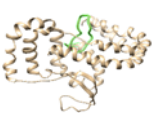
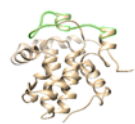
بخشی از توالی اپی توپ	Score	شماره چارچوب	شماره خوانش
TLTQDDLQDEQ	۰/۵۹	ORF۲۲۹	۲۰۹۹۰۰۲-۲۱۴۹۰۰۲
LYGTTAPSEQGKVG DY	۰/۵۲۹	ORF۲۳۲	
IQVDDYDPEEEYLDE	۰/۴۷	ORF۲۱۵	۲۲۴۹۰۰۴-۲۱۹۹۰۰۴
SSAKMPSNTRS CSSTELKRLG	۰/۶۲۸	ORF۱۸۰	۲۲۹۹۰۰۵-۲۲۴۹۰۰۵
VQQQSAGQQRYTGQE	۰/۶۲۳	ORF۲۶۱	
MINIQTAKRMNDNE	۰/۵۸	ORF۱۷۴	۲۳۴۹۰۰۶-۲۲۹۹۰۰۶
PKMGFTIEERNQSQGT	۰/۵۵	ORF۳۲۲	
PIYWLLKNSFQRRRPQ	۰/۵۱	ORF۱۹۲	۲۶۹۹۰۱۳-۲۶۴۹۰۱۳
LKADEKPKPKSVL	۰/۵۴	ORF۲۸۹	۲۸۴۹۰۱۶-۲۷۹۹۰۱۶
EQGQLNKNGEQV	۰/۵۸	ORF۲۰۴	۲۸۹۹۰۱۷-۲۸۴۹۰۱۷

موقعیت اپی توپ‌های کاندید بر روی پروتئین: موقعیت اپی توپ‌های کاندید توسط نرم افزار Chimera شناسایی و در جدول ۳ با رنگ سبز نمایش داده شده‌اند. البته این نکته حائز اهمیت است که اپی توپ‌های کاندید، تمام ویژگی‌های اپی توپ مناسب از جمله: انعطاف‌پذیری (flexibility)، در دسترس بودن سطحی (Surface accessibility)، آب‌دوستی (hydrophilicity)، پیچ‌های بتا (Beta-turns)، آنتی‌ژنیسیته (Antigenicity) را شامل می‌شوند.

ساخت سازه پلی توپی: بعد از تعیین اپی توپ‌های کاندید باتوجه به خواص مختلف، اپی توپ‌ها را در کنار یکدیگر قرار

داده و سازه پلی توپی طراحی میشود. برای طراحی سازه پلی توپی مناسب، اپی توپ‌های مناسب را از نظر امتیاز از درصد بالا به پایین کنار هم قرارداده و بین آنها لینکر مناسب قرارداده می‌شود. در این مطالعه از لینکر معروف GGSGG به دلیل خاصیت هیدروفیلی و انعطاف‌پذیری بالا لوپ تشکیل نمی‌دهند و طول مناسبی را در سازه‌ی پلی توپی ایجاد می‌کنند مورد استفاده قرار گرفته است.

جدول ۳- موقعیت اپی توپ‌های کاندید بر روی پروتئین
Table3. Position of candidate epitopes on protein.

شماره خوانش	کانفورماسیون فضایی	شماره خوانش	کانفورماسیون فضایی
ORF 232 (114-129)		ORF 229 (154-165)	
ORF 174 (617-621)		ORF 180 (130-156)	
ORF 322(254-275)		ORF 261(622-636)	
ORF 289 (128-146)		ORF 204 (21-225)	
ORF 215 (231-252)		ORF 192 (81-101)	

جدول ۴- توالی سازه پلی توپی طراحی شده در این مطالعه

Table 4. Sequence of polytopic structures designed in this study

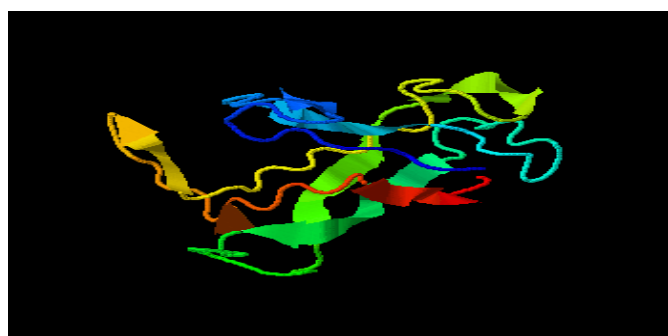
توالی سازه پلی توپی
SSAKMPNSTRSCSSTELKRLGGGSGSSAKMPNSTRSCSSTELKRLGGGSGSSAKMPNSTRSCSSTELKRLGGGSGGVQQQSAGQQRYTGQEGGSGGTLTQDDLQDEQGGSGGEQQLNKNGEQVGGSGGMINIQTAKRMNDNEGGSGGPKMGFTIEERNQSQGTGGSGGLKADEKPKPKSVLGGSGGLYGTTPAQGKVGDDYGGSGGPIYWLLKNSFQRRRPQGGSGG IQVDDYDPEEEYLD

جدول ۵- آنالیز سازه پلی توپی

Table 5. Analysis of polytopic structures

The part of Construct seq	Number Of amino acid	Molecular weight	PI	Asp+Glu	Arg+lys	Half-life in vitro	Half-life in vivo (yeast)	Half-life in vivo (E.coli)	Aliphatic index	GRAVY	Instability index
SSAKMPNSTRSCSSTELKRLGLVLVLVLVLVGGSGGSSAKMPNSTRSCSSTELKRLGLVLVLVGGSGGSSAKMPNSTRSCSSTELKRLGLVLVLVLVGGSGG VQQQSAGQQRYTGQELVLVLVLVGGSGG TLTQDDLQDEQGGSG	312	31852/09	6/39	27	27	1.9h	20h	10h	59.48	-1.038	42/55

هدف از این مرحله، آگاهی و شناخت پروتئین ساخته شده می- باشد که می توان باتوجه به ویژگی های آن با کارآمدی بیشتری از آن استفاده کنیم. در نهایت مدل سازه ی پلی توپی را با نرم افزار I tasser بدست آوردیم که در شکل ۱ قابل مشاهده است.



شکل ۱- نمایی شماتیک پلی توپ طراحی شده با سرور I TASSER

Fig1. Schematic view of a poly ball designed with I TASSER server

تجزیه و تحلیل سازه پلی توپی

تجزیه و تحلیل توسط پایگاه داده ProtParam صورت گرفته و پارامترهای محاسبه توسط پایگاه داده شامل: وزن مولکولی، ترکیب آمینواسیدی، ترکیب اتمی، ضریب خاموشی و ... است. نتایج این تجزیه و تحلیل منجر به رسیدن چندین پلی توپ نهایی می شود. توالی مورد نظر شامل ۳۱۲ اسید آمینه است که وزن مولکولی آن ۳۱۸۵۲/۰۹ دالتون و pH ایزوالکتریک آن ۶/۳۹ می باشد. فرمول مولکولی این پروتئین C1048H1694N334O395S9 و مجموع اتم های آن ۳۴۸۰ است که به این ترتیب شاخص ناپایداری (II) آن برابر با ۴۲/۵۵ است بنابراین پروتئین مورد نظر جزء پروتئین های پایدار دسته بندی می شود.

برای ساخت سازه پلی توپی نیاز به توالی نوکلئوتیدی می باشد که توسط پایگاه داده Backtranseq صورت می گیرد.

بهینه سازی توالی نوکلئوتیدی

به منظور عملکرد بهتر سازه در تست های آزمایشگاهی توالی نوکلئوتیدی سازه، توسط پایگاه داده JCat بهینه سازی شده است

مدل سازی سازه پلی توپی

خلاصه ای از نرم افزار و پایگاه داده های مورد استفاده

جدول ۶- نرم افزار، سرور و پایگاه داده های مورد استفاده

Table 6. Software, server and database used in this study

نرم افزار/ سرور/ پایگاه داده	آدرس	کاربرد
NCBI	https://www.ncbi.nlm.nih.gov/	استخراج ژنوم
NCBI (ORF Finder)	https://www.ncbi.nlm.nih.gov/orffinder	یافتن توالی هایی از ژنوم که پروتئین را کد میکند
PSORTb	https://www.psort.org/psortb	مکان یابی پروتئین
SignalP-5.0	http://www.cbs.dtu.dk/services/SignalP/	حذف سیگنال پپتید
Algpred	http://crdd.osdd.net/raghava/algpred/	بررسی آلرژیک بودن پروتئین
NCBI (BLAST)	https://blast.ncbi.nlm.nih.gov/Blast.cgi	بررسی همولوژی
TMHMM – ExPASy(TMpred)	http://www.cbs.dtu.dk/services/TMHMM/ - https://embnet.vital-it.ch/software/TMPRED_form.html	مکان یابی دقیق آمینواسیدها
SWISS-MODEL	https://swissmodel.expasy.org	مدل سازی
LBtope - ABCpred IEDB (Bepipred, Kolaskar and Tongaonkar, Emini, Parker, Chou and Fasman, Karplus)	- http://crdd.osdd.net/raghava/lbtope/ - http://crdd.osdd.net/raghava/abcpred/ - http://tools.iedb.org/bcell/	پیش بینی اپی توپ خطی
CBTOPE - Disco Tope 2.0 – IEDB (Ellipro)	http://crdd.osdd.net/raghava/cbtope/ - http://www.cbs.dtu.dk/services/DiscoTope/ - http://tools.iedb.org/ellipro/	پیش بینی اپی توپ فضایی
PROCHECK	https://servicesn.mbi.ucla.edu/PROCHECK/	بررسی نمودار رامچاندرا پروتئین
Chimera 1.8.1	Software	ایجاد تغییرات روی مدل پروتئین
ExPasy (ProtParam)	https://web.expasy.org/protparam	تجزیه و تحلیل سازه پلی توپی
Backtranseq	https://www.ebi.ac.uk/Tools/st/emboss_backtranseq/	ترجمه توالی آمینواسیدی به نوکلئوتیدی
JCat	http://www.jcat.de	بهینه سازی توالی نوکلئوتیدی

بحث

بررسی نمونه های کار شده روی ویبریکلرا

در سال ۲۰۱۴ وقارگانی نتتاری و همکارانش در دانشگاه SRM، که با هدف طراحی و توسعه یک کاندید واکسن علیه ویبریکلرا تحقیق کرده اند. در این پژوهش از سویه ویبریکلرا ۰۳۹۵ استفاده شده است که این سویه دارای ۱۲۸۶۵ توالی پروتئینی بیماری زا است که وی از ۵ توالی به صورت انتخابی استفاده کرده است. این محقق توالی پروتئین های ویبریکلرا به وسیله پایگاه داده TIGER و SDSC Biology Bench Work بررسی کرده و سپس

از پایگاه داده های ProPred، HLA Pred و PORTER به ترتیب برای غربالگری و ساخت ساختار دوم پروتئین های شناسایی شده، استفاده کرد. در نهایت آنتی ژن پنجم طراحی شده توسط وی با توالی LLAL به عنوان آنتی ژن برتر انتخاب شد (Qureshi, et al, 2018).

در سال ۲۰۱۶ پژوهش ها نیات نورین و همکارانش از OMP ها و پروتئین های فلاژله در سویه O۱ به عنوان پروتئین های کلیدی در طراحی واکسن استفاده شد. در این تحقیق ۳۸ سویه ویبریکلرا به دو زیرگروه PSC1 و PSC2 تقسیم شدند سپس این ۳۸ سویه از لحاظ ژنومیک، تکاملی و ساختاری به وسیله پایگاه داده های NCBI، CLC Sequence Viewer و EMBOSS بررسی شدند.

از مجموعه ۱۰۰ توالی ژنوم cholera Vibrio انتخاب شده اند که می تواند به عنوان عوامل درمانی موفق مورد استفاده قرار داد (Zed, et al, 2019).

مقایسه نتایج این مطالعه با مطالعات دیگر: در تحقیق وقار و همکارانش نقطه ضعف آن را می توان مربوط به طول کم اپی توپ نهایی دانست که برای فولد و در نهایت عملکرد بهینه مناسب نیست. در نتایج مربوط به این پروژه می توان طول های متفاوت و متنوعی مشاهده کرد که در نهایت به عملکرد و ساخت پلی توپ با کارایی بالا میتوان استفاده کرد.

در پژوهش مربوط به پرادیپتا و همکارانش پیش بینی یک اپی توپ با طول مناسب و همچنین توالی ایمنی را مشترک بین B cell و T cell می توان به عنوان نقطه قوت یاد کرد اما استفاده نکردن از همه ابزارهای موجود برای یافتن جامع ترین نتیجه از نقاط ضعف آن هاست. در این پروژه نگاه جامع تری نسبت به ابزارهای در دسترس دارد.

در یافته های نورین و همکارانش به صورت جامع تر نسبت به پرادیپا و همکارانش، پژوهش ها انجام داده اند اما تفکیک بین اپی توپ های خطی و فضایی از نقاط ضعف آن می باشد در صورتی که در این پروژه اپی توپی که کاندید شده است دارای تاییدیه پیش بینی خطی و فضایی است.

پلاوی و همکارانش شاید یکی از کامل ترین روش ها را امتحان کرده اند و استراتژی ترکیبی آن نقطه قوت و قابل اتکا آن ها می باشد.

نتیجه گیری کلی: به منظور شناسایی دقیق شاخص های آنتی ژنیک نیاز به بدست آوردن پروتئین مورد نظر از ارگانیسم و شناخت مناسب قطعات پپتیدی بدست آمده از آن و آزمایش بر روی آن ها جهت شناسایی فعالیت ایمنی زایی می باشد. این چنین مطالعاتی طاقت فرسا، زمان بر و پرهزینه بوده، بنابراین تلاش و توجه پژوهشگران به سمت توسعه روش های سریع تر نظیر پیشگویی، مبتنی بر توالی آمینو اسید پروتئین مورد بررسی، می باشد. پیچیدگی ذاتی شناسایی آنتی ژن، پیشگویی اپی توپ را مشکل می کند از این رو الگوریتم های متنوعی جهت پیشگویی اپی توپ

در مرحله بعد توالی ها به وسیله پایگاه داده های ABCPred، Ellipro، Bepipred و ProPred برای شناسایی اپی توپ های B cell و T cell بررسی شدند. در مورد B cell، ۴ اپی توپ فضایی برتر انتخاب شدند و به وسیله مدل سازی سه بعدی مورد بررسی قرار گرفته اند (Noureen, et al, 2016).

در تحقیق پلاوی و همکارانش در سال ۲۰۱۷ به وسیله واکسن شناسی معکوس، کاندیدهای واکسن مناسب علیه پاتوژن و ویروس انتخاب کردند. این پژوهشگران به وسیله استراتژی ریورس واکسن، OMP های ویبرو را که می توانستند کاندید مناسب واکسن باشند را بررسی کردند. در این بررسی ۲۳ آنتی ژن بررسی شد که از این تعداد دو عدد به عنوان لیپوپروتئین غشای خارجی شناسایی شدند. همچنین اپی توپ های B و T برای OMP های جدید به کار برده شده در این تحقیق پیش بینی شدند که می تواند در طراحی واکسن مبتنی بر اپی توپ مناسب به کار رود. لازم به ذکر است توالی نوکلئوتیدی سویه O1 ویبرو به وسیله پایگاه داده NCBI بازیابی شد. سپس با VaxiJen، PSORTb و EMBOSS مورد بررسی قرار دادند. سپس BLAST انجام شده است و در نهایت اپی توپ های خطی B cell بوسیله BCPred، AAP و FBCPred بررسی شدند. نتایج براساس طول و عدم هم پوشانی انتخاب و با یک B cell پایدار (consensus) ترکیب شدند. از این رو از اپی توپ های به دست آمده در جهت تهیه واکسن براساس پپتید استفاده کردند (Baliga, et al, 2018).

در پژوهش ها Rehman و همکارانش در سال ۲۰۱۹ با هدف یافتن آنتی ژن های جدید برای توسعه واکسن ویبرو کلرا، که در این مطالعه لیپوپروتئین NlpD، پروتئین غشای خارجی OmpU، فاکتور استعمار جانبی Porin، AcfA، پروتئین غشایی و غشایی بیرونی OmpW که این اپی توپ های پیش بینی شده را به عنوان کاندید واجد شرایط یک واکسن پپتیدی قابل اعتماد گزارش کردند (Rashid, et al, 2019).

Saadia Andleeb و همکارانش در سال ۲۰۱۹ از طراحی واکسن به روش معکوس برای انتخاب کاندیدهای جدید واکسن از پروتئوم cholera Vibrio استفاده کردند نتایج نشان داده شد که ۱۰ پروتئین به عنوان کاندیدای احتمالی واکسن از ۹۹۸ ژن هسته

مناطق با آنتی ژنسیتی بالا را در پروتئین پیش‌بینی می‌کنند تقریباً می‌توان با احتمال بالائی نواحی ایمونوژنیک را در پروتئین شناسایی کرد. براساس همولوژی به وسیله سرور SWISS MODEL، آنتی‌ژن‌ها غربال‌گری شدند. اپی توپ نهایی انتخاب شده در این پژوهش با معرفی ۱۰ پروتئین جدید به عنوان کاندیدای واکسن، راه پیشرفت واکسن را بسیار آسان می‌کند. داده‌های حاصل از مطالعه می‌توانند در تعیین راهکارهای درمانی دقیق به منظور کاهش آمار مبتلایان به وبا بسیار مفید باشند.

ها طراحی شده‌اند. در خصوص این که سبب تسهیل توسعه واکسن‌های نوین مبتنی بر داده‌های ژنومیک و پرتئومیک شده است. این چنین واکسن‌هایی، به خصوص آن‌هایی که بر علیه بیماری‌هایی عفونی هستند (بیماری‌های که سبب ابتلا و مرگ گروه زیادی در تمام جهان می‌شوند) حائز اهمیت استراتژیک از نظر سلامتی جهانی می‌باشند.

با در دست داشتن دانسته‌های کافی، دسترسی به تکنولوژی‌های جدید و الگوریتم‌های طبقه بندی شده ابزارهای محاسبه‌ای که

منابع

- Baliga, P., Shekar, M., & Venugopal, M. N. (2018).** Potential outer membrane protein candidates for vaccine development against the pathogen *Vibrio anguillarum*: a reverse vaccinology based identification. *Current microbiology*, 75(3), 368-377.
- Dehghannejad, M., Kasiri, M. (2011).** The problem of Quarantines in Iran during Naseroldin Shah's era (1847-1986), *Pajooohesh-haye Tarikhi*, 2(4), 1-14. (In Farsi with English abstract).
- Del Tordello, E., Rappuoli, R., & Delany, I. (2017).** Reverse vaccinology: exploiting genomes for vaccine design. In *Human vaccines* (pp. 65-86). Academic Press.
- Nair, G. B., & Takeda, Y. (Eds.). (2014).** Cholera outbreaks (Vol. 379). Springer.
- Nandy, A., & Basak, S. C. (2019).** Bioinformatics in Design of Antiviral Vaccines. *Encyclopedia of Biomedical Engineering*, 280.
- Noureen, N., Tariq, M., Farooq, A., Arif, A., & Bokhari, H. (2016).** In silico identification of receptor specific epitopes as potential vaccine candidates from *Vibrio cholerae* strains. *Gene Reports*, 4, 222-232.
- Qureshi, A., Tantray, V. G., Kirmani, A. R., & Ahangar, A. G. (2018).** A review on current status of antiviral siRNA. *Reviews in medical virology*, 28(4), e1976. (In Farsi with English abstract).
- Rashid, M. I., Rehman, S., Ali, A., & Andleeb, S. (2019).** Fishing for vaccines against *Vibrio cholerae* using in silico pan-proteomic reverse vaccinology approach. *PeerJ*, 7, e6223. (In Farsi with English abstract).
- Razmjoo, A., Razmjoo, M. (2018).** Legal review of the use of cholera bacteria in Yemen by Saudi, *Bi-Quarterly Journal of Politics and International Relations*, 2(4), 53-68. (In Farsi with English abstract).
- Shaikh, H., Lynch, J., Kim, J., & Excler, J. L. (2020).** Current and future cholera vaccines. *Vaccine*, 38, A118-A126.
- Zeb, S., Ali, A., Gulfam, S. M., & Bokhari, H. (2019).** Preliminary work towards finding proteins as potential vaccine candidates for *Vibrio cholerae* Pakistani isolates through reverse vaccinology. *Medicina*, 55(5), 195. (In Farsi with English abstract).

Genetic Engineering and Biosafety Journal
Volume 9, Number 2
2021

Bioinformatics design of a vaccine against *Vibrio cholerae* using the reverse approach

Mina Oghbatalab, Mohammad Javad Dehgha Esmatabadi*, Alireza Saeedinia, Mohammad Ali Yaghobi Moghaddam

Faculty of Chemistry and Chemical Engineering, Malek Ashtar University of Technology, Tehran, Iran

*Corresponding Author, Email: mohammad_dehghan@mut.ac.ir

Abstract

Vaccine design from the past to the present is based on classical methods, which are time consuming, costly and have some risks, but with the discovery of the genome and gene sequencing, the modern methods for vaccine design have been developed, including Referred to the reverse method of vaccine design (vaccine polytop). In this study we used NCBI (ORF Finder) to find the coding sequences, PSORTb server to locate the protein, Allpred server to check the allergenicity of the protein and LBTpe - ABCpred - IEDB CBTOPE servers - Disco Tope 2.0 - IEDB (Ellipro) to predict the linear and spatial epitopes. The final epitope was selected and then we analyzed the polypeptide structure by ExPasy server (ProtParam) and finally obtained the optimized nucleotide sequence with JCat server. Using this method, 6074 reading frameworks as well as 43 surface and secretory proteins were identified, from which, finally, 10 suitable epitopes that are able to provide immunity against a wide range of strains were selected as a vaccine candidate. It is hoped that this will facilitate the development of this vaccine.

Keywords: Bioinformatics, Cause of Cholera, Polytopic structure, Protein, Reverse vaccination