

مطالعه ایمونوبیوانفورماتیک پروتئین اتوترانزپورتر پادگن ۴۳ اشریشیا کلی انتروتوکسیژنیک جدا شده از گوساله

ریحانه قربانپور^۱، غلامرضا نیکبخت بروجنی^۱، امیر حسین جلالی^{۲،۳}

^۱گروه میکروبیولوژی و ایمنولوژی، دانشکده دامپزشکی دانشگاه تهران، تهران، ایران

^۲دانشکده منابع طبیعی دانشگاه صنعتی اصفهان، اصفهان، ایران

^۳پژوهشکده زیست فناوری و مهندسی زیستی، دانشگاه صنعتی اصفهان، اصفهان، ایران

(دریافت مقاله: ۵ شهریور ماه ۱۳۹۷، پذیرش نهایی: ۲۳ آبان ماه ۱۳۹۷)

چکیده

زمینه مطالعه: تلاش‌ها در زمینه تهیه واکسن‌های کارآمد برای مقابله با *E. coli* انتروتوکسیژنیک (ETEC) به سمت شناسایی پادگن‌های حراست شده ژنومی معطوف شده است. پادگن ۴۳ یکی از اعضای خانواده بزرگ پروتئین‌های ترشحی باکتری *E. coli* و سایر باکتری‌های گرم منفی با نام اتوترانزپورترها (ATs) است. پروتئین‌های اتوترانزپورتر ساختمان حراست شده مشابهی دارند و تعدادی از آن‌ها به دنبال عفونت‌های طبیعی و تجربی ETEC شناسایی شده‌اند. پادگن ۴۳ با داشتن خصوصیات حدت‌زای تجمع سلولی و تشکیل بیوفیلم و حضور در سرم انسان‌های مبتلا به اسهال ناشی از ETEC، هدف ارزشمندی در تهیه واکسن محسوب می‌شود.

هدف: در این مطالعه با بهره‌گیری از روش‌های تجربی و ایمونوبیوانفورماتیک به بررسی دقیق ساختار آنتی‌ژنی و ارزیابی ایمنی‌زایی پروتئین پادگن ۴۳ سویه ۵۱۰ *E. coli* جدا شده از گوساله‌ها پرداختیم.

روش کار: توالی اسیدآمینه‌ای، خصوصیات فیزیکی-شیمیایی، پایداری، ساختار دوم و سوم پروتئین و توانایی بالقوه در ایجاد پاسخ‌های ایمنی سلول‌های B و T با دارا بودن اپی‌توپ‌های مؤثر و خصوصیات آلرژی‌زا بودن آن از جمله شاخص‌های مورد توجه در این پژوهش هستند. **نتایج:** در این مطالعه ۱۵ قطعه پپتیدی محرک پاسخ یاخته‌های B و T شناسایی شده‌است که تعداد ۹ عدد از آن‌ها به‌عنوان پادگن معرفی گردیدند.

نتیجه‌گیری نهایی: نتایج حاصل از مطالعات مجازی (in-silico) انجام شده روی این پروتئین حاکی از مناسب بودن آن برای تهیه واکسن‌های کلی‌باسیلوز گاو است.

واژه‌های کلیدی: ایمونوبیوانفورماتیک، *E. coli* انتروتوکسیژنیک (ETEC)، اتوترانزپورتر، پادگن ۴۳، کلی‌باسیلوز

کپی‌رایت ©: حق چاپ، نشر و استفاده علمی از این مقاله برای مجله تحقیقات دامپزشکی محفوظ است.

*نویسنده مسئول: تلفن: ۰۲۱-۶۱۱۱۷۰۵۷، نامبر: ۰۲۱-۶۶۴۲۷۵۱۷، Email: nikbakht@ut.ac.ir

How to Cite This Article

Ghorbanpour, R., Nikbakht Brujeni, G., Jalali, S. (2019). Immuno-Bioinformatics Study of Autotransporter Protein, Antigen 43, in Enterotoxigenic *Escherichia coli* Isolated From Calves. *J Vet Res*, 74(1), 127-141. doi: 10.22059/jvr.2018.211721.2504



مقدمه

پروتئین‌های اتوترانزپورتر (Autotransporter)، بزرگترین گروه از پروتئین‌های لایه بیرونی و ترشحی را در باکتری‌های گرم منفی تشکیل می‌دهند. آن‌ها در همه موارد از یک ساختار مدولار مشابه تشکیل شده‌اند که شامل سه دامنه پپتید سیگنال (Signal peptide)، دامنه مسافر (Passenger domain)، دامنه ناقل (Transporter domain) می‌شود و منجر به ترشح خودبه‌خودی این پروتئین‌ها از لایه بیرونی باکتری به سطح می‌گردد (۴۶).

دامنه ناقل (β) در انتهای کربوکسی اتوترانزپورتر (C)، منفذی را در لایه بیرونی ایجاد می‌کند که از آن طریق دامنه مسافر (α) انتهای آمینی (N) به بیرون یاخته منتقل می‌شود. معمولاً به دنبال این انتقال، دامنه مسافر به وسیله یک پروتئاز لایه بیرونی شکافته شده و از طریق یک برهم کنش غیر کووالانسی با دامنه بتا در تماس با سطح باقی می‌ماند و یا به صورت یک پروتئین محلول آزاد می‌شود (۱۸).

پروتئین‌های اتوترانزپورتر بسیار متنوع هستند و عملکردهای مختلفی برعهده دارند که در مجموع حدت را تقویت می‌کنند. درحالی‌که دامنه ناقل و پپتید سیگنال در بین پروتئین‌های اتوترانزپورتر بسیار حراست (Conserve) شده‌اند، این دامنه مسافر است که به‌عنوان یک پروتئین بالغ ترشحی، عملکردهای مؤثر متنوعی را به اتوترانزپورترهای گوناگون اعطا می‌نماید. از آن جمله می‌توان به پروتئولیز کاتالیز کننده، ارائه به عنوان یک عامل اتصالاتی، میانجی‌گری حرکت با واسطه اکتین و ارائه به عنوان یک سایتوتوکسین اشاره نمود. علی‌رغم فراوانی آن‌ها و نقشی که در روند بیماری‌زایی باکتری دارند، بیشتر این پروتئین‌ها از لحاظ ساختاری مشخص نشده‌اند و اطلاعات اندکی در مورد نحوه عمل آن‌ها وجود دارد (۱۹، ۱۶).

در میان خانواده بزرگ اتوترانزپورترها، نوع AIDA-I (Involved in Diffuse Adherence) بزرگترین و متنوع‌ترین پروتئین‌های اتوترانزپورتر هستند. در اشرشیا کلی ۱۶ تحت گروه AIDA-I شناسایی شده‌است که در طول و توالی دامنه مسافر و خواص عملکردی با یکدیگر متفاوت هستند (۴۶). از این میان پادگن ۴۳ (Ag ۴۳) به خوبی مورد مطالعه قرار گرفته است. پادگن ۴۳ یک عامل اتصالاتی اتوترانزپورتر خودشناخت (Self-recognizing) است که به صورت یک پیش پروتئین ۱۰۳۹ اسیدآمینوای تولید می‌شود. پپتید سیگنال انتهای آمینی پروتئین (۵۲ اسیدآمینو) انتقال را از میان لایه سیتوپلاسمی به پری پلاسم هدایت می‌کند. ساختار رایج دامنه مسافر و ناقل پادگن ۴۳ به ترتیب از ۴۹۹ و ۴۸۸ اسیدآمینو تشکیل شده‌اند (۴۴، ۱۶).

پادگن ۴۳ در بیشتر پاتوتیپ‌های *E. coli* یافت می‌شود و گاهی به صورت آلل‌های چندگانه در یک سویه از *E. coli* وجود دارد. حضور آلل‌های ژن پادگن ۴۳ (agn⁴³) که با نام flu نیز شناخته می‌شود در نواحی مختلف کروموزوم و همچنین تنوع در توالی دامنه مسافر آن‌ها در میان

سویه‌های مختلف *E. coli* می‌تواند منجر به تفاوت در خواص عملکردی آن‌ها در هر سویه از *E. coli* گردد. از جمله عملکردهای به‌خوبی مشخص شده پادگن ۴۳، تجمع سلولی و تشکیل بیوفیلم هستند (۴۴، ۹).

حضور این پروتئین در سرم بیماران مبتلا به اسهال ناشی از *E. coli* انتروتوکسیژنیک (ETEC)، از بیان آن در عفونت‌های تجربی و طبیعی متعاقب ETEC و ایمنی‌زا بودن آن حکایت می‌کند. جالب توجه است که واکسیناسیون موش‌ها با دامنه مسافر نوترکیب پروتئین Ag ۴۳ حفاظت مطلوبی در مقابل استقرار روده‌ای ETEC انسانی فراهم می‌آورد (۱۸).

با وجود فراوانی اتوترانزپورترها و مقدار فزاینده داده‌هایی که پروتئین‌های AT را با حدت باکتریایی مرتبط می‌سازند، اطلاعات کمی در مورد ساختمان مولکولی این پروتئین‌ها در دسترس است. در حقیقت پروتئین‌های AT در بانک داده پروتئین (PDB)، به طور قابل توجهی فقط با یک ساختار تمام طول AT، ۷ دامنه مسافر AT، ۵ دامنه بتا AT، ۱۲ دامنه کوچک اتوترانزپورترهای سه تایی (Trimeric autotransporters) در میان حدود ۸۷۵۰۰ ساختار موجود در حال حاضر تعداد اندکی را تشکیل می‌دهند و هیچ یک از آن‌ها به گروه بزرگ پروتئین‌های AT نوع AIDA-I تعلق ندارند به استثنای ساختار سه بعدی دامنه مسافر پادگن ۴۳ سویه *Y. enterocolitica* انسانی (CFT۰۷۳) که به تازگی به این بانک اضافه شده است (۱۹).

با توجه به موارد ذکر شده و نیز عدم وجود اطلاعات کافی در مورد پروتئین‌های اتوترانزپورتر در سویه‌های *E. coli* بیماری‌زای دامی برآن شدید که در این مطالعه به بررسی بیوانفورماتیک پروتئین Ag ۴۳ سویه ۵۱۰ *E. coli* انتروتوکسیژنیک عامل کلی‌باسیلوز به‌عنوان یک کاندید بالقوه ایمنی‌زا در تهیه واکسن‌های مؤثر بر این بیماری بپردازیم.

مواد و روش کار

استخراج ترادف و انجام هم‌ردیفی: سویه ۵۱۰ که از سویه‌های *E. coli* ETEC گاوی با سروتیپ O1۰۱:H۴۴K۹۹O1۰۱- است در این پژوهش مورد استفاده قرار گرفت. ژن کدکننده دامنه مسافر پروتئین اتوترانزپورتر پادگن ۴۳ این سویه (ناحیه بین انتهای قسمت پپتید سیگنال و ابتدای دامنه ناقل) با استفاده از آغازگرهای ذکر شده توسط Harris و همکاران در سال ۲۰۱۱، (آغازگر رفت ۵'-gtggcgattgcgctgtct-۳ و آغازگر برگشت ۵'-tacaccggctgatggct-۳) افزوده گردید (۱۶). به طور خلاصه برای تکثیر این پروتئین، واکنش زنجیره‌ای پلیمرز در حجم نهایی ۲۵ μ l شامل ۲/۵ μ l بافر X1، ۱/۵ μ l کلرید منیزیم (۱ mM)، ۵۰ μ l dNTPs (۱/۲۵ mM)، ۰/۲ μ l آنزیم تک پلیمرز، ۰/۱۷۵ μ l از هر کدام از آغازگرهای اختصاصی (۲۵ nM)، DNA به میزان ۲ μ l و آب مقطر استریل ۱۵/۳ μ l صورت گرفت. مواد لازم جهت انجام PCR از



<http://www.imtech.res.in/raghava/betapred> و <http://www.imtech.res.in/raghava/betapred/index.html> (۲۱، ۲۲).

جهت پیشگویی ساختار سه بعدی پروتئین، از سرورهای متعدد ModBase، SWISS-MODEL، CPH models، Phyre2 (PS) و Rapturx، HHpred و استفاده گردید که مدل سازی همولوژی را بر اساس تشابه توالی ها انجام می دهند (۴۱، ۳۴، ۳۲، ۳۱، ۲۳، ۲۰، ۵). به عبارتی در این نوع مدل سازی، ساختمان سه بعدی پروتئینی که دارای ساختار نامشخص است بر اساس تشابه توالی با پروتئین دارای ساختار مشخص ترسیم می شود. پس از آن، ساختار کلی با مدل سازی *ab initio* یعنی فقط بر اساس اطلاعات مربوط به توالی پروتئین و بدون داشتن الگوی مشخص با استفاده از سرور I-Tasser ساخته شد (۵۰، ۱). در نهایت نیز نرم افزار Robetta که از ترکیب هر دو روش مدل سازی همولوژی و *ab initio* استفاده می کند، برای پیشگویی ساختار سوم پروتئین Ag۴۳ سویه ۵۱۰ *E. coli* (ETEC) گاوی به کار رفت (۲۴).

بهینه سازی انرژی و ارزیابی اعتبار و پایداری ساختار پیش بینی شده پروتئین: برای شناسایی و اصلاح خطاهای مدل انتخابی، بهینه سازی ساختار پروتئین با استفاده از سرور ModRefiner (<http://zhanglab.ccmb.med.umich.edu/ModRefiner>) انجام شد (۴۸). مختصات ساختار سه بعدی بهینه شده Ag۴۳ با فرمت PDB به وبسایت ProSA (<https://prosa.services.came.sbg.ac.at/prosa.php>) که اغلب برای تأیید اعتبار ساختار پروتئین به کار می رود، عرضه گردید (۴۷). کیفیت نتایج استریشیمی ساختار پروتئین بهینه شده نیز توسط نقشه رامانچاندرا در نرم افزارهای PROCHECK (<http://www.ebi.ac.uk/thornton-srv/software/PROCHECK>) و RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) بررسی شد (۲۷، ۲۸). ضمن اینکه از امتیاز C (*C-score*) حاصل از سرور I-TASSER به منظور تخمین کیفیت مدل پیشگویی شده آن استفاده گردید.

پیشگویی حالیت و در دسترس بودن سطحی پروتئین: حالیت پروتئین به وسیله Recombinant Protein Solubility Prediction (<http://www.biotech.ou.edu>) و وضعیت در دسترس بودن سطحی اسید آمینه ها با Emini Surface Accessibility (<http://tools.ieb.dcu.ie/~emini/>) ارزیابی شد (۱۰، ۱۴).

پیشگویی مکان درون یاختهای: جایگاه درون یاخته های پروتئین در یاخته پروکاریوتی (باکتری گرم منفی) با کمک سرورهای CELLO (<http://e093.life.nctu.edu.tw/index.html>) و PSLpred (<http://www.imtech.res.in/raghava/pslpred>) پیش بینی شد. PSLpred از ترکیب روش یکپارچه سازی PSI-BLAST (*Position-Specific Iterative Basic Local Alignment Search Tool*) و سه مدل (Support vector machine) SVM برای پیش بینی جایگاه

شرکت سیناژن تهیه شدند. چرخه های دمایی شامل یک مرحله واسرشتگی اولیه در دمای ۹۵ °C به مدت دو دقیقه، ۲۹ چرخه سه مرحله ای، شامل مرحله واسرشته سازی در دمای ۹۵ °C به مدت ۳۰ ثانیه، مرحله اتصال در دمای ۵۷/۳ °C به مدت ۳۰ ثانیه، مرحله بسط در دمای ۷۲ °C به مدت ۲۱۰ ثانیه و در انتها یک مرحله بسط انتهایی در دمای ۷۲ °C به مدت ۵ دقیقه بود. محصول افزوده سازی با استفاده از آغازگرهای رفت و برگشت توالی یابی شد. تعیین توالی توسط مؤسسه ماکروژن (سئول، کره) با دستگاه توالی یاب خودکار DNA (ABI ۳۷۳۰ XL) انجام شد.

در ادامه، این قطعه با استفاده از پلاسمید pJET (مطابق دستورالعمل کیت Thermo Scientific CloneJET PCR Cloning Kit) به منظور دستیابی به قطعه کاملتر (مکان آغازگرهای مورد استفاده) همسانه سازی (Clone) گردید. آزمون PCR با آغازگرهای پلاسمید pJET بر اساس دستور کیت انجام گرفت و همسانه های مقاوم به آمپی-سیلین حاوی قطعه با اندازه مشابه ژن مورد نظر مشخص گردیدند. تعیین توالی بر روی محصول افزوده سازی پلاسمید بعد از تخلیص با کیت Qiagen (QIAprep spin miniprep kit) از دو انتهای آغازگرهای رفت و برگشت پلاسمید صورت پذیرفت.

جستجو جهت یافتن توالی های نوکلئوتیدی و اسید آمینه ای مشابه در پایگاه داده بانک ژن NCBI و Uniprot به ترتیب به آدرس <http://www.ncbi.nlm.nih.gov/Blast.cgi> و <http://www.uniprot.org/blast> انجام شد.

تجزیه و تحلیل اولیه توالی پروتئین: شاخص های مختلف فیزیکی-شیمیایی پروتئین مانند نقطه ایزوالکتریک، وزن مولکولی، نیمه عمر، تعداد دنباله های منفی و مثبت، شاخص آلیفاتیک و مقدار (GRAVY) (*Grand average of hydropathicity*) با استفاده از ابزار ProtParam (<http://expasy.org/tools/protparam.html>) Expasy پایگاه محاسبه گردید (۱۵). از نرم افزارهای VaxiJen (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) و ANTIENpro (<http://scratch.proteomics.ics.uci.edu>) به منظور پیش بینی آنتی ژنیسیته پروتئین استفاده شد (۷، ۱۱). از سرور <http://emboss.bioinformatics.nl/cgi-bin/pepfind> (*emboss/pepfind*) برای پیش بینی توالی های PEST به عنوان جایگاه های بالقوه شکاف پروتئولیتیک استفاده گردید (۳۶).

پیشگویی ساختار دوم و سوم پروتئین: جهت پیشگویی ساختار دوم پروتئین از سرورهای مختلف Rapturx، GOR4، SPIDER2، Predictprotein، Psipred، jpred4 و CFSSP استفاده شد (۶، ۸، ۱۲، ۱۷، ۲۵، ۲۹، ۴۹). دوره های آلفا و همچنین انواع متفاوت دوره های بتا از قبیل I، II، IV، VIII نیز در پروتئین به ترتیب با استفاده از <http://www.imtech.res.in/raghava/ALPHAPRED>



ژن DRB۳ مورد بررسی قرار گرفت. بر طبق نتایج حاصل از پژوهش Nikbakht و همکاران در سال ۲۰۱۶، آل‌های دارای فراوانی بالای ۵ درصد جمعیت انتخاب شدند (۳۲) و از آنجا که تاکنون هیچ سروری آل‌های MHC-II گاوی را جهت پیشگویی اپی‌توپ‌های ایمنی‌زای یاخته T ارائه نکرده‌است، آل‌های MHC انسانی (HLA) و موشی (H-II) که مشابه با آل‌های MHC گاو (BoLA) بودند بررسی شدند. به این صورت که مجموعه آل‌های HLA و H-II به ترتیب در پایگاه داده IPD-IMGT/HLA (<http://www.ebi.ac.uk/ipd/imgt/hla>) و IMGT (<http://www.imgt.org/IMGTrepertoireMH/index.php?section=LocusGenes&repertoire=Chromosomal&localization&species=mouse>) مشخص شدند و پس از هم‌دیفی و کاوش توالی‌های انسانی و موشی مشابه با آل‌های گاوی موجود در جمعیت هلستاین ایران، آل‌های معادل آن‌ها برای تعیین اپی‌توپ مورد استفاده قرار گرفتند.

برای پیشگویی پپتیدهای متصل شونده به MHC کلاس II آل‌های انسانی و موشی از سرور IEDB (<http://tools.iecb.org/mhcii>) استفاده شد. همچنین آل‌های موشی در دو سرور دیگر Rankpep (<http://imed.med.ucm.es/Tools/rankpep.html>) و PREDmus (<http://dfci.harvard.edu/cvc/predmus/HTML/>) (۴) نیز بررسی شدند (۳۵، ۴۵، ۵۲).

پیشگویی جایگاه‌های آلرژی‌زا و بررسی شباهت احتمالی با ژن‌های گاو: برای مشخص کردن شباهت احتمالی پروتئین Ag۴۳ و اپی‌توپ‌های پیشگویی شده آن با ژن‌های گاو، توالی اسیدآمینه‌ای آن‌ها در پایگاه اطلاعاتی NCBI مورد مطالعه بلاست (BLASTp) قرار گرفت. پیشگویی آلرژن‌ها و مشخص کردن اپی‌توپ‌های IGE از وبسایت AlgPred (<http://www.imtech.res.in/raghava/algpred>) بدست آمد.

Algpred، پیش‌بینی آلرژن‌ها را براساس شباهت به اپی‌توپ‌های شناخته شده با هر ناحیه پروتئین با استفاده از روش ترکیبی ARPs+ MAST+BLAST (allergen-representative peptides) اپیتوپ IGE+ SVMc انجام می‌دهد (۳۸). همچنین آلرژن بودن با اجرای آزمون آلرژن‌زایی FAO/WHO در پایگاه داده SDAP (https://fermi.utmb.edu/cgi-bin/SDAP/sdap_01_edu/cgi-bin/SDAP/sdap) مورد تحلیل بیشتر قرار گرفت.

نتایج

استخراج مترادف و انجام هم‌دیفی: نتیجه تعیین توالی محصول افزوده‌سازی واکنش زنجیره‌ای پلیمرز قطعه‌ای به طول ۲۰۴۱ bp است. نتیجه حاصل از هم‌سانه‌سازی آن، قطعه‌ای با طول ۲۱۰۶ bp نوکلئوتید است که یک پپتید ۷۰۲ اسیدآمینه‌ای را می‌سازد. بلاست این قطعه بیشترین تطابق با توالی‌های مرتبط به ژن پادگن ۴۳ سویه‌های مختلف یدکولای

دورن یاخته‌ای استفاده می‌کند (۵۱، ۳).

پیشگویی جایگاه شکاف پروتئاز پروتئین: پیشگویی جایگاه‌های شکاف Ag۴۳ با استفاده از سرور NetChop۳/۱ (<http://www.cbs.dtu.dk/services/NetChop>) با روش Artificial neural networks (ANN) و وبسایت Pcleavage (<http://www.imtech.res.in/raghava/pcleavage>) با روش SVM بدست آمد. سرور NetChop۳/۱ از بهترین سرورهای موجود برای پیشگویی جایگاه‌های شکاف است و روش SVM سرور Pcleavage با آن قابل مقایسه است (۴، ۳۰).

پیشگویی اپی‌توپ‌های ایمنی‌زا (اپی‌توپ‌های پادگنی یاخته‌های B): یک پادگن برای اینکه کاندید خوب واکسن باشد باید هیدروفیل بوده و بتواند هر دو پاسخ ایمنی با واسطه یاخته‌های B و T را برانگیزد. بنابراین، ابتدا توالی کامل پروتئین برای پیش‌بینی اپی‌توپ یاخته B با استفاده از دو روش BCpreds و AAPpreds در پایگاه <http://ailab.ist.psu.edu/bcpred/predict.html> (۱۳). توالی‌های ۲۰ تایی اپی‌توپ بر سطح یاخته B مشخص شدند و با استفاده از سرور VaxiJen ([http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen](http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html)) با مشخصات (ACC output و حد آستانه ۰/۴) جهت بررسی پادگن‌زایی مورد تحلیل بیشتری قرار گرفتند. از طرف دیگر، اپی‌توپ‌های خطی یاخته B در سرورهای ABCpred (http://www.imtech.res.in/raghava/abcpred/ABC_submission.html) با استفاده از روش ANN، Bcepred (<http://www.imtech.res.in/raghava/bcepred>) با استفاده از BepiPred (<http://www.cbs.dtu.dk/services/BepiPred>) با استفاده از ترکیب روش‌های hidden Markov model و propensity scale نیز مشخص گردیدند (۲۶، ۳۷، ۳۹). ABCpred براساس ANN مطابق با نمره بدست آمده و بالاتر از حد آستانه (۰/۵)، اپی‌توپ‌ها را رتبه‌بندی می‌کند بدین ترتیب که هرچه نمره بیشتر باشد احتمال اپی‌توپ بودن آن توالی بیشتر است. سرور Bcepred از هفت معیار فیزیکی-شیمیایی شامل آبدوستی، اعطاف‌پذیری، تحرک، قابلیت دسترسی، قطبیت، سطح در معرض، دورها و پادگن‌زایی جهت پیشگویی اپی‌توپ‌های خطی موجود یاخته B استفاده می‌کند. در BepiPred، حد آستانه ۰/۳۵ انتخاب شد که حداکثر حساسیت و ویژگی را دارد. همچنین برای پیشگویی اپی‌توپ‌های فضایی یاخته B از توالی اولیه، سرور CBTOPE (<http://www.imtech.res.in/raghava/cbtope>) به کار رفت (۲).

اپی‌توپ‌های پادگنی یاخته‌های T: از آنجا که جمعیت هدف این مطالعه گاوهای هلستاین ایران است، جهت شناسایی اپی‌توپ‌های یاخته T، یعنی اپی‌توپ‌های متصل شونده به مولکول‌های MHC، در ابتدا فراوانی آل‌های MHC کلاس II گاوها با محوریت آگرون دوم



جدول ۱. پیشگویی ساختار دوم پروتئین.

نام پروتئین	تعداد اسید آمینه	سرور / PDBsum	هلیکس (H) درصد	صفحه‌های بتا (E) درصد	لوپ (L) درصد
۴۳ Ag	۷۰۲	Psipred	۲/۲۷	۶۳/۱۰	۳۴/۶۳
۴۳ Ag	۷۰۲	Rapturx	۲	۵۷	۳۹
۴۳ Ag	۷۰۲	SPIDER۲	۳/۴۱	۵۷/۵۶	۴۵/۰۳
Chain-A/۴kh۳ (۴kh۳A)	۴۹۲	PDBsum	۰	۵۶/۷	۴۳/۳

جدول ۲. مقایسه شباهت آل‌های غالب BoLA در جمعیت گاوهای هلستاین ایران با آل‌های MHC انسانی و موشی.

آل‌های گاو	فراوانی در جمعیت (درصد)	آل‌های انسانی مشابه	درصد شباهت	آل‌های موشی مشابه	درصد شباهت
۰۱۰*BoLA-DRB۳	۱۸/۱۶	۱۳:۱۱*HLA-DRB۱	۷۰	IE ۲-H	۶۴
۱۵۰*BoLA-DRB۳	۱۷/۵۸	۱۴:۸۲*HLA-DRB۱	۸۸	IE ۲-H	۶۴
۱۱۰*BoLA-DRB۳	۹/۲۱	۱۱:۸۹*HLA-DRB۱	۸۶	IE ۲-H	۷۱
۱۲۰*BoLA-DRB۳	۸/۴۲	۱۴:۱۴*HLA-DRB۱	۸۸	IE ۲-H	۷۰
۰۹۰۲*BoLA-DRB۳	۷/۱۱	۱۴:۵۱*HLA-DRB۱	۸۸	IE ۲-H	۷۵

شاخص برای پروتئین‌های پایدار نتایج کمتر از ۴۰ است. شاخص آلفاتیکی که نشان‌دهنده پایداری پروتئین در بازه دمایی وسیع است برای این پروتئین مقدار ۸۱/۹۸ گزارش گردید. مقدار GRAVY یک پپتید یا پروتئین از مجموع مقدار هیدروپاتی همه اسید آمینه‌ها تقسیم بر تعداد اسید آمینه‌های توالی پروتئین بدست می‌آید که در مورد ۴۳ Ag برابر ۰/۱۹۴- است. مقدار منفی این شاخص نشان می‌دهد که پروتئین غیر قطبی است. مجموع کل اسید آمینه‌ها با شارژ منفی (Asp+Glu) و مثبت (Arg+Lys) در این پروتئین به ترتیب ۶۰ و ۴۱ است. نیمه عمر این توالی در یاخته‌های پستانداران ۱۰۰ ساعت، در مخمر بیش از ۲۰ ساعت و در *D. discoideum* بیش از ۱۰ ساعت پیش‌بینی گردید. قطعه مورد نظر به‌عنوان پادگن پیش‌بینی شد و نتیجه آنتی‌ژنیسیته مرتبط با نرم افزارهای Vaxijen و ANTIGENpro به ترتیب ۱/۸۲۱۱ و ۰/۹۴۸۹۴۹ است. موتیف‌های PEST، توالی‌های غنی از اسید آمینه‌های پرولین، گلوتامیک اسید، سرین، ترئونین و به میزان کمتر اسید آسپارتیک هستند که بطور چشمگیری نیمه عمر درون یاخته‌ای پروتئین‌ها را کاهش می‌دهند. از این رو، آن‌ها پروتئین را برای خرد شدن پروتئولیتیکی هدف‌گذاری می‌کنند. سرور epestfind، موتیف‌های PEST را به دو صورت ضعیف (poor) و بالقوه (potential) در پروتئین مورد نظر به ترتیب بر اساس نمره پایین و بالای حد آستانه شناسایی می‌نماید. هرچه نمره بیشتر شود احتمال خرد شدن پروتئین با میانجیگری موتیف PEST یاخته یوکاریوتی قویتر است. در پروتئین ۴۳ Ag، پنج موتیف ضعیف PEST مشخص گردید که در تصویر ۱ قابل مشاهده است.

ساختار دوم پروتئین: چنانچه دامنه مسافر پروتئین ۴KH۳ (زنجیره A) را به‌عنوان، تنها پادگن ۴۳ ثبت شده در بانک داده پروتئین (PDB) مقیاس قرار دهیم، در میان سرورهای بکار رفته برای پیش بین ساختار دوم پروتئین، نتایج ۲ سرور psipred و rapturx بیشترین شباهت را به این پروتئین و به یکدیگر دارند. سومین سروری که نتایج نزدیکی دارد،

را نشان می‌دهد. به این ترتیب که نتیجه بلاست توالی نوکلئوتیدی این پروتئین در پایگاه داده NCBI، ژن کد کننده پروتئین پادگن ۴۳ سویه BVA باکتری *D. discoideum* (ژن کد لای اتروتوکسیژنیک انسانی به طول ۳۱۲۰ bp (شماره دسترسی: CP۰۰۵۹۹۸/۱) با هم‌پوشانی (coverage) ۱۰۰ درصدی و شباهت (similarity) ۹۹ درصدی است و همچنین حاصل بلاست توالی اسید آمینه‌ای آن در پایگاه داده UniprotKB، پروتئین پادگن ۴۳ باکتری *D. discoideum* (شماره دسترسی: A۰A۰B۱N۸۳۷) است که به ترتیب هم‌پوشانی و شباهت ۱۰۰ و ۹۹/۷ درصدی با قطعه مورد نظر دارد و اینگونه صحت آزمایش را آشکار می‌سازد. این قطعه در بانک ژن با شماره دسترسی KX۱۹۶۴۵ ثبت گردید.

در مقایسه هم‌ردیفی توالی اسید آمینه قطعه افزوده شده با ۵۲، ۴۹۹ و ۴۸۸ اسید آمینه تشکیل دهنده پپتید سیگنال، دامنه مسافر و دامنه ناقل پادگن ۴۳ سویه فرانس *D. discoideum* (۱۲-k) مشخص گردید که این قطعه به ترتیب از ابتدای توالی خود شامل ۱۷، ۴۹۹ و ۱۸۶ اسید آمینه تشکیل دهنده سه دامنه پروتئین پادگن ۴۳ می‌شود. این امر نشان دهنده تکثیر کامل ناحیه دامنه مسافر ۴۳ Ag سویه ۵۱۰ است که عملکرد پروتئین را برعهده دارد و بررسی مجازی ایمنی‌زایی آن مورد توجه مطالعه ایمونوبیوانفورماتیک پژوهش حاضر بوده است.

تجزیه و تحلیل اولیه توالی پروتئین: در این مطالعه شاخص‌های فیزیکی-شیمیایی مربوط به پروتئین پادگن ۴۳ مشخص گردید. متوسط وزن مولکولی محاسبه شده برای ۴۳ Ag، ۷۰۸۰۲/۵ Da است. نقطه ایزوالکتریک (PI) پروتئین ۴/۸۸ است. با توجه به اینکه اسیدیته پروتئین با $VPI >$ مشخص می‌شود این پروتئین اسیدی است. ضریب خاموشی (Extinction coefficient) در طول موج ۲۸۰ nm برابر ۱-cm 1.479×10^5 است. در مورد شاخص ناپایداری، ابزار ProtParam این پروتئین را جزء پایدارها (II ۸/۹۷) طبقه‌بندی کرده است. محدوده این



جدول ۳ (الف). اپی توپ‌های مشترک پیشگویی شده یاخته‌های B و T. - اعداد نشانگر شماره اسیدآمیننه قطعه پروتئینی Ag ۴۳ هستند.

ردیف	یاخته T						یاخته B (خطی)	یاخته B (فضایی)
	BoLA-DRB۳	BoLA-DRB۳	BoLA-DRB۳	BoLA-DRB۳	BoLA-DRB۳	BoLA-DRB۳		
	۱۵-۱*	۱۱-۱*	۱۲-۱*	۰۹-۲*	DRB۳	DRB۳		
	HLA-DRB۱	HLA-DRB۱	HLA-DRB۱	HLA-DRB۱	HLA-DRB۱	HLA-DRB۱	IEd ۲-H	
	۱۴:۸۲*	۱۱:۸۹*	۱۴:۱۴*	۱۴:۵۱*	۱۴:۵۱*	۱۴:۵۱*		
۱	-	۳-۱۹	۴-۱۸	-	-	-	۱۴	
۲	-	-	۳۸-۵۲	۳۸-۵۲	-	-	۵۲، ۳۸-۴۸	
۳	-	-	-	-	۶۸-۷۶	۶۴-۷۴	-	
۴	-	-	-	-	۱۳-۱۴۹	۱۴-۱۷۶	۱۴۰، ۱۴۹-۱۴۵	
۵	-	۱۷۷-۱۹۲	-	۱۷۷-۱۹۲	۱۸۳-۱۹۶	۱۷۴-۲۱۱	۱۹۱-۱۸۳، ۱۹۲-۱۸۶	
۶	-	-	-	-	۲۶۶-۲۷۹	۲۲۲-۲۷۷	۲۶۶-۲۷۴	
۷	۲۸۱-۳۰۱	۲۸۲-۳۰۱	۲۸۳-۳۰۲	۲۸۳-۳۰۲	-	-	۳۰۰-۲۸۴، ۳۰۱-۲۸۷	
۸	-	۳۴۸-۳۶۳	۳۴۸-۳۶۳	۳۴۸-۳۶۳	-	-	۳۵۷-۳۶۳	
۹	۳۷۷-۳۹۱	-	-	۳۷۷-۳۹۱	-	-	-	
۱۰	-	-	-	-	۴۰۶-۴۱۴	۳۹۷-۴۴۲	-	
۱۱	-	۴۹-۵۰۸	۴۹۱-۵۱۳	۴۹۱-۵۱۳	۴۹۱-۵۰۴	-	۴۹۶-۴۹۷	
۱۲	-	-	-	-	۵۲۸-۵۳۶	۵۱۴-۵۳۸	۵۲۸-۵۳۶	
۱۳	-	۵۴۶-۵۶۲	۵۴۵-۵۷۰	۵۴۶-۵۷۰	۵۴۶-۵۷۰	۵۴۲-۵۷۳	۵۵۴-۵۵۹	
۱۴	۵۹۹-۶۱۷	-	-	-	-	-	-	
۱۵	-	۶۶-۷۰۲	۶۶۱-۷۰۲	۶۵۹-۶۸۴	۶۴۷-۶۷۴	۶۴۹-۶۹۶	۶۷۴، ۶۶۷، ۶۶۹، ۶۷۲	

جدول ۳ (ب). اپی توپ‌های مشترک پیشگویی شده یاخته‌های B و T. - توالی‌های محدود به مکان ۵۱۶-۱۸ مربوط به دامنه مسافر (α) از قطعه Ag ۴۳ هستند. - نمره بالای حد آستانه (۰/۴) در سرور VaxiJen، پادگن (A) و نمره کمتر از آن غیرپادگن (N) است. ۱- اپی توپ NRAENAYADYVLDIGKRI در ایدیکولای HV:O۱۵۷ و میزبان گاو. ۲- اپی توپ SAPIAPNRAENAYADYVL پروتئین ECS-۰۸۴۷ در ایدیکولای HV:O۱۵۷ و میزبان گاو.

ردیف	مکان اپی توپ مشترک	طول اپی توپ	توالی اپی توپ	نمره VaxiJen	IEDB Epitope ID (Identity درصد)
۱	۴-۱۵	۱۲	ALSAAVTSVPA	(A) /۵۴۰۲	وجود ندارد
۲	۳۸-۵۲	۱۵	NQIVFGTANGMTIST	(A) /۵۷۹۲	وجود ندارد
۳	۶۸-۷۴	۷	WIQNGGT	(N) /۱۰۵۳	وجود ندارد
۴	۱۴-۱۴۹	۱۰	QAIKSGAVAT	(A) /۹۱۹۵	وجود ندارد
۵	۱۸۳-۱۹۲	۱۰	KNGRQIVAAE	(N) /۱۵۷۵	وجود ندارد
۶	۲۶۶-۲۷۷	۱۲	GTATNVTLKQGG	(A) /۱۹۳۵	وجود ندارد
۷	۲۸۳-۳۰۱	۱۹	TAATVTGINRLGAFSVVEG	(N) /۰۶۱۳	وجود ندارد
۸	۳۴۸-۳۶۳	۱۶	NGGVLLADSGAAVSGT	(N) /۲۴۳۸	وجود ندارد
۹	۳۷۷-۳۹۱	۱۵	ADALMLEKGSFTLN	(A) /۴۲۸۴	وجود ندارد
۱۰	۴۰۶-۴۱۴	۹	FTARGGTLA	(A) /۹۴۵۰	وجود ندارد
۱۱	۴۹۱-۵۰۴	۱۴	VTTDVIAQRGTALK	(A) /۵۵۹۳	وجود ندارد
۱۲	۵۲۸-۵۳۶	۹	WNIPDNATV	(N) /۰۳۱۸	وجود ندارد
۱۳	۵۴۶-۵۶۲	۱۷	AGQIHFTSTRTGKFPVA	(A) /۵۲۲۵	وجود ندارد
۱۴	۶۰۰-۶۱۶	۱۷	TGKTILNLVNAGNSASG	(N) /۰۳۳۲۱	وجود ندارد
۱۵	۶۶۱-۶۷۴	۱۴	ESWYLRSEAYRAE	(A) /۷۳۳۲۷	۵۵۴۶۸۴ (A) ۱ ۵۵۵۸۷۷ (A) ۲

کدام از آن‌ها تغییراتی جهت بهبود پیشگویی متحمل گشته است (جدول ۱). دورها محلی از ساختار دوم پروتئین هستند که در آن زنجیره پلی پپتیدی

SPIDER۲ است. نکته قابل توجه اساس کار هر سه این سرورهاست که بر پایه روش شبکه عصبی (Neural Network) عمل می‌نمایند که در هر



برابر با ۱/۱۵- را نشان می‌دهد. TM-score مورد انتظار این پروتئین برابر با 0.15 ± 0.05 است، با توجه به اینکه TM-score بالای ۰/۵ نشان‌دهنده دقت توپولوژی است. فایل pdb مدل سه بعدی مذکور برای به حداقل رساندن انرژی و بهینه کردن ساختار پروتئین به سرور ModRefiner ارائه گردید. نتیجه به‌دست آمده دارای TM-score و RMSD به ترتیب 0.9425 و $2/645$ نسبت به مدل ابتدایی است. جهت ارزیابی کیفیت کلی ساختار پروتئین فایل pdb آن در برنامه ProSA مورد بررسی قرار گرفت. بر این اساس Z-score ساختار ورودی مدل انتخابی با نمره $-1/15$ در بازه نمراتی است که به‌طور معمول برای پروتئین‌های طبیعی (Native) با اندازه مشابه یافت می‌شود (تصویر ۳).

پایداری ساختمان پروتئینی مدل انتخابی با نقشه رامچاندرا حاصل از سرورهای PROCHECK و Rampage بررسی شد (تصویر ۴). درصد اسیدآمینه‌ها در نواحی مطلوب‌ترین (Most favoured regions)، مجاز فرعی (Additional allowed regions)، مجاز سخاوتمندانه (Generously allowed regions) و غیرمجاز (Disallowed regions) سرور PROCHECK به ترتیب $69/1$ ، $22/4$ ، $4/3$ و $4/3$ در نواحی مطلوب (Favoured regions)، مجاز (Allowed regions) و خارج از محدوده (Outlier regions) سرور Rampage به ترتیب $81/7$ ، $13/9$ و $4/4$ گزارش گردید درحالی‌که این مقادیر قبل از بهینه شدن مدل به ترتیب در سرور PROCHECK برابر با $51/8$ ، $29/4$ ، $10/8$ و 8 و در سرور Rampage برابر با $67/7$ ، $17/6$ و $14/7$ بودند.

حلالیت و دسترسی به سطح پروتئین: بر اساس پیش‌بینی انجام شده Ag ۴۳ هنگامیکه در (دکولای بیان شود به میزان صفر درصد احتمال حلالیت دارد. میانگین دسترسی سطحی $1/00$ و میزان حداکثر و حداقل آن به ترتیب $5/812$ و $0/142$ است. در خصوص پپتیدهای پیش‌بینی شده با بیشترین دسترسی، از 702 اسیدآمینه کل، تعداد 251 اسیدآمینه میزان دسترسی به سطح بالای میانگین دارند و 15 توالی پپتیدی با حداقل طول 6 اسیدآمینه را می‌سازند. از میان این پپتیدها، پپتید 21 اسیدآمینه‌ای موجود در بازه مکانی $673-653$ (انتهای قطعه افزوده) بیشترین طول را دارد و بیشترین میزان دسترسی به سطح نیز متعلق به اسیدآمینه موقعیت 658 از همین پپتید است.

جایگاه درون یاخته‌ای پروتئین: هر دو پایگاه CELLO و PS1pred، مکان این پروتئین را در یاخته پروکاریوتی، خارج از یاخته (Extracellular) پیش‌بینی کردند.

جایگاه‌های شکاف پروتئین: نتایج پیشگویی جایگاه‌های شکاف دو سرور NetChop $3/1$ و Peleavage با در نظر گرفتن حد آستانه پیش فرض هر یک از آن‌ها، به ترتیب تعداد 176 و 285 جایگاه را نشان دادند که تعداد 92 جایگاه مشترک هستند.

اپی‌توپ‌های یاخته B: مهم‌ترین نتایج پیش‌بینی و تعیین اپی‌توپ

جهت کلی خود را معکوس می‌کند. آن‌ها ماریج‌های آلفا و صفحات بتا را به یکدیگر مرتبط می‌نمایند. در پیشگویی دوره‌های ساختار دوم پروتئین Ag ۴۳ سویه 510 ، 8 دور آلفا و 100 دور بتا تعیین گردید. هر دور آلفا از 5 اسیدآمینه و هر دور بتا از 4 اسیدآمینه متوالی تشکیل شده است.

ساختار سوم پروتئین: با استفاده از سرورهای مختلف یاد شده ساختار سوم پروتئین Ag ۴۳ سویه 510 پیشگویی گردید. نکته قابل توجه و مشترک در نتایج تمام سرورها، انتخاب خودکار زنجیره A پروتئین $4KH3$ به‌عنوان اولین الگو جهت مدل‌سازی محدوده 516 اسیدآمینه اول قطعه مورد نظر است. نتیجه ساختار پیشنهادی سرورهای Phyre 2 ، I-Tasser، $73-2$ (PS) و همچنین دو سرور Rapturx و Robetta شامل تمام 702 اسیدآمینه سازنده پروتئین است درحالی‌که، سرورهای Modeller، Swiss-model، CPH models و HHpred با توجه به الگوی انتخابی در مدل نهایی خود فقط اسیدآمینه‌های محدوده دامنه مسافر را لحاظ کردند. از این میان، سرورهای Rapturx و Robetta دو دامنه جدا از هم مطابق با دامنه مسافر و دامنه ناقل پروتئین (به ترتیب محدوده 500 اسیدآمینه اول و 200 اسیدآمینه آخر قطعه) و بقیه سرورها ساختار مونومری پیشنهاد دادند. از آنجا که این قطعه تحت مطالعه in-vivo قرار نگرفته است تا به صورت ساختار رایج خود یعنی تشکیل مستقل دامنه ناقل و دامنه مسافر بررسی گردد، تشکیل ساختار مونومر در مطالعه همسانه‌سازی آن مورد انتظار است. بنابراین در مرحله اول ساختار مونومر به‌دست آمده از سه سرور I-Tasser، Phyre 2 ، $73-2$ (PS) که از تمام 702 اسیدآمینه قطعه پروتئینی Ag ۴۳ تشکیل شدند، مورد بررسی قرار گرفت.

ارزیابی انتقادی پیشگویی ساختار پروتئین (CASP) یک تجربه جهانی است که با هدف ارتقاء روش‌های شناسایی ساختار سه بعدی پروتئین براساس توالی اسیدآمینه‌ای آن‌ها هر دو سال یکبار برگزار می‌گردد. طبق آخرین نتایج منتشر شده (CASP 11) در سال 2014 ، سرور I-Tasser به‌عنوان بهترین سرور در امر پیشگویی ساختار سوم پروتئین معرفی شد. در نهایت، بنا بر این موضوع و با توجه به اینکه در انتهای مدل پیشنهادی سرورهای Phyre 2 و $73-2$ (PS) که اساس کار آن‌ها مدل‌سازی برپایه تشابه توالی است ساختار رندم کوئل مشاهده می‌شود و عدم موفقیت این سرورها را در پیش‌بینی ساختار سه بعدی ناحیه مربوط نشان می‌دهد، اولین ساختار پیشنهادی سرور I-Tasser که اساس کار آن مدل‌سازی برپایه ab initio است با C-score برابر با $-1/15$ و TM-score و RMSD مورد انتظار به ترتیب 0.15 ± 0.057 و $4/6 \text{ \AA}$ و $10/8$ انتخاب گردید (تصویر ۲).

بهینه‌سازی انرژی و ارزیابی اعتبار و پایداری ساختار پیش‌بینی شده پروتئین: نمره اطمینان (C-score) برای تخمین کیفیت مدل پیش‌بینی شده توسط I-TASSER، است که به‌طور معمول در بازه $(2$ تا $-5)$ قرار دارد. هرچا که نمره C-score بیشترین است نشان‌دهنده اطمینان بالاتر آن ساختار است که بهترین مدل پیش‌بینی شده برای این پروتئین، C-score

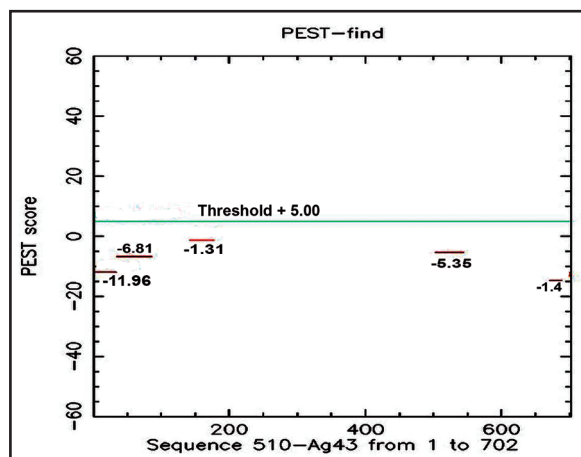


نمره ۰/۷) حاصل کار دو روش BCpreds و AAPpreds بود درحالیکه در هر دو روش توالی اپی توپ‌های پیشنهادی با یکدیگر همپوشانی ندارند. در مورد پیشگویی سرور ABCpred تعداد ۶۸۷ توالی ۱۶ تایی که همپوشانی دارند با رتبه‌بندی ۱ تا ۳۱ و نمره بالای حد آستانه ۰/۵ بدست آمد. از این تعداد، ۲۰ پپتید اول آن (رتبه ۹-۱ و نمره بالای ۰/۸۴) انتخاب شدند. سرور BepiPred هر اسیدآمیننه را به‌طور مستقل بررسی می‌کند و بین ۳ تا ۳ نمره می‌دهد. هرچه نمره بیشتر باشد احتمال اپی توپ بودن بیشتر می‌شود. در مورد توالی Ag۴۳، ۴۷ مکان شامل توالی‌های حداقل یک اسیدآمیننه تا حداکثر ۷۲ اسیدآمیننه معرفی شدند که بیشترین احتمال مربوط به مکان ۱۶۷-۱۵۸ با میانگین نمره ۲/۳۹۵۲ است. سرور BcePred نیز هفت معیار فیزیکی و شیمیایی را برای هر اسیدآمیننه با نمره جداگانه مشخص می‌کند و برای هر مورد، اسیدآمیننه با نمره بالای ۱/۸ را انتخاب می‌کند. سرور CBTOPE با مقیاس احتمال (۰-۹)، اسیدآمیننه‌هایی که نمره ۴ به بالا دارند را اپی توپ در نظر می‌گیرد که برای توالی Ag ۴۳، ۵۲ مکان شامل حداقل یک اسیدآمیننه تا حداکثر ۲۰ اسیدآمیننه‌ای با بیشترین نمره ۴ و ۵ به عنوان اپی توپ‌های فضایی یا ناپیوسته یاخته B، پیش‌بینی شد. همانطور که قابل ملاحظه است پروتئین Ag ۴۳ تعداد زیادی اپی توپ خطی و فضایی یاخته B دارد. از میان این نتایج، اپی توپ‌های خطی مشترک شناسایی و با اپی توپ‌های فضایی مقایسه شدند.

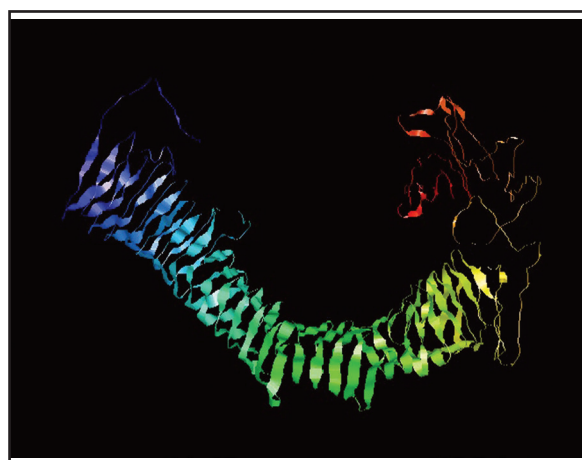
اپی توپ‌های یاخته T: براساس مطالعه Nikbakht و همکاران در سال ۲۰۱۶، پنج آلل MHC کلاس II گاوی با نام‌های BoLA-DRB۳، ۰۱۰۱×DRB۳، ۱۵۰۱×BoLA-DRB۳، ۱۱۰۱×BoLA-DRB۳ و ۱۲۰۱×BoLA-DRB۳ بیشترین فراوانی را در جمعیت مورد مطالعه گاوهای هلشتاین ایرانی به ترتیب با ۱۸/۱۶، ۱۱/۵۸، ۹/۲۱، ۸/۴۲ و ۷/۱۱ درصد دارند. فراوانی سایر آلل‌های جدا شده زیر ۵ درصد است. بنابراین بر طبق نتیجه بلاست آلل‌های انسانی و موشی با بیشترین شباهت به این پنج آلل مشخص شدند (جدول ۲).

در مرحله بعد با استفاده از سرور IEDB اپی توپ‌های پروتئین Ag ۴۳ برای هر کدام از آلل‌های انسانی و موشی مشابه پیشگویی شدند. در مورد آلل‌های موشی، ژن IE ۲-H به‌عنوان شبیه‌ترین برای هر ۵ مورد آلل گاوی انتخاب شد. جهت مشخص کردن اپی توپ‌های ایمینی‌زا به خاطر انتخاب نژاد Balb/c به‌عنوان مدل آزمایشگاهی این مطالعه جستجو بر اساس هاپلوتیپ این نژاد یعنی IEd انجام گرفت.

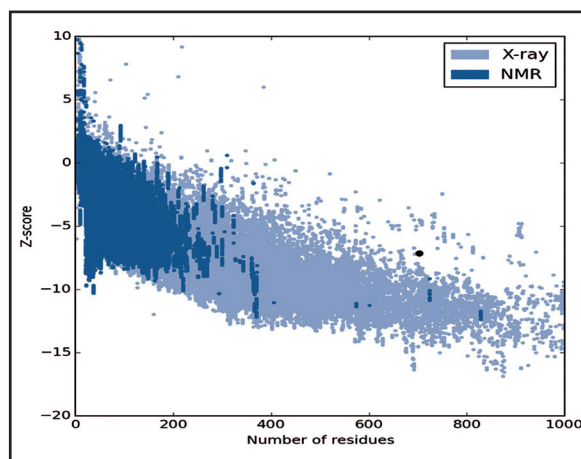
پیشگویی براساس روش پیشنهادی IEDB انجام شد. برای آلل ۱۳:۱۱×HLA-DRB۱ روش sturniolo و برای چهار آلل دیگر انسانی روش NetMHCIIpan را انتخاب کرد. نتایج بر حسب رتبه درصدی (percentile_rank) چیدمان شدند. در روش پیش‌بینی پیشنهادی IEDB، اپی توپ‌ها با percentile_rank کمتر، متصل‌شونده‌های بهتری هستند. بر همین اساس ۲۰ توالی ۱۵ تایی ابتدایی نتیجه هر آلل



تصویر ۱. موتیف‌های PEST روی توالی Ag ۴۳ سویه ۵۱۰.



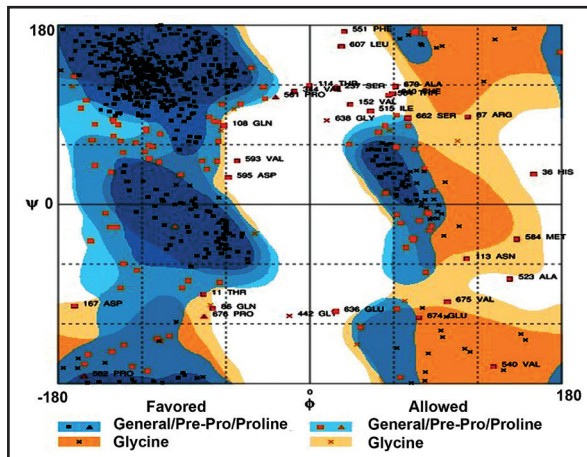
تصویر ۲. ساختار ۳ بعدی Ag ۴۳.



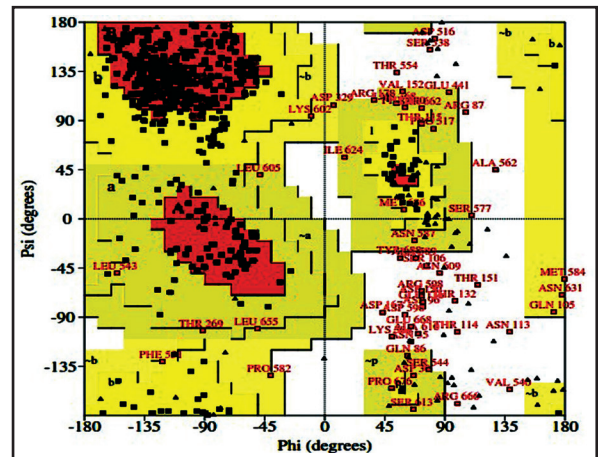
تصویر ۳. نقشه Z-score و بسایت ProSA. Z-score، کیفیت کلی مدل را نشان می‌دهد و بوسیله کریستالوگرافی اشعه X (آبی روشن) یا اسپکتروسکوپی NMR (آبی تیره) با توجه به طول زنجیره‌های پروتئین موجود در پایگاه PDB تعیین می‌گردد. نقشه تنها نتایج Z-score مشابه یا کمتر از ۱۰ را نشان می‌دهد. Z-score پروتئین Ag ۴۳ سویه ۵۱۰ به صورت یک نقطه بزرگ مشخص شده است. نمره آن در بازه پروتئین‌های طبیعی و برابر ۷-۱۵ است.

خطی و اپی توپ‌های فضایی با کمک سرورهای مختلف به شرح زیر است: ۱۹ توالی ۲۰ تایی (۱۸ توالی اول با نمره ۱ و توالی آخر با نمره ۰/۱۴) و ۱۶ توالی ۲۰ تایی (۱۴ توالی اول با نمره بین ۱ تا ۰/۸ و ۲ توالی آخر با





تصویر ۳. نقشه رامچاندرا - هر نقطه قرمز رنگ در هر دو نمودار نشان‌دهنده یک اسید آمینه با زاویه phi و psi غیرمجاز است. (الف) سرور PROCHECK: چهار ناحیه رنگی قرمز، زرد، کرم و سفید به ترتیب نواحی مطلوبترین، مجاز اضافی، مجاز سخاوتمندانه و غیرمجاز را نشان می‌دهند. (ب) سرور RAMPAGE: سه ناحیه رنگی تیره، روشن و سفید به ترتیب معرف نواحی مطلوب، مجاز و خارج از محدوده هستند.



آن‌ها غیر آلرژن معرفی شدند. در نهایت، آزمون آلرژن‌زایی براساس گزارش FAO/WHO نیز بر روی قطعه Ag ۴۳ صورت گرفت. براساس قانون آلرژن‌زایی FAO/WHO در گزارش سال ۲۰۰۱ این نهاد، در صورتی یک پروتئین واکنش متقاطع با آلرژن‌های شناخته شده دارد که:

۱. بیش از ۳۵ درصد identity در توالی پروتئینی با استفاده از جستجوی پنجره ۸۰ اسید آمینه‌ای متوالی یافت شود. نتیجه این مرحله در مورد توالی Ag ۴۳، ۶۲۳ توالی ۸۰ تایی متوالی بود که در بیشترین حالت شباهت ۳۲/۷۵ درصدی در پنجره ۸۰ تایی اسید آمینه ۲۳۸-۱۵۹ به دست آمد.

۲. شش اسید آمینه پشت سر هم شناسایی شود. حاصل این مرحله نیز در مورد توالی Ag ۴۳، ۵ توالی ۶ تایی است که به ترتیب در نواحی اسید آمینه ۱۹-۱۴، ۹۹-۹۴، ۳۲۱-۳۱۶، ۴۸۵-۴۸۰ و ۵۸۵-۵۸۰ قرار دارند.

بحث

در تلاش برای ایجاد مصنوعیت بر ضد سویه‌های ETEC مطالعات گسترده‌ای در حیطه‌های دامی و به خصوص انسانی صورت گرفته است. پیشرفت در مسیر دانش سازوکار بیماری‌زایی ETEC، انتخاب گزینه‌های کارآمدی را در توسعه تولید واکسن به دنبال داشته است. از جمله آن‌ها، شناسایی پادگن‌های حراست شده ژنومی (دکولای) است. بنابر مطالعات پیشین بر روی پاسخ ایمنی علیه عفونت ETEC، گروهی از مولکول‌های بیان شونده سطحی تحت عنوان اتوترانزپورترها شناسایی شدند. انواع متعددی از پروتئین‌های اتوترانزپورترها، در طی دهه گذشته در باکتری (دکولای) و سایر باکتری‌های گرم منفی معرفی شدند. از خصوصیات قابل توجه این پروتئین‌ها که اغلب با حدت باکتری مرتبط هستند، بیان سطحی دامنه مسافر آن‌هاست. این امر، آن‌ها را به اهداف مناسبی جهت ساخت واکسن تبدیل نموده است.

پادگن ۴۳ نمونه‌ای از پروتئین‌های این خانواده است که در بیشتر از ۹۰

انتخاب شدند. از آنجا که نتایج برخی آل‌ها فاقد اپی‌توپ‌هایی در محدوده مکانی ۵۱۶-۱۸ مربوط به دامنه α بود و چون در حالت طبیعی دامنه مسافر پروتئین به خارج از سلول باکتری ترشح شده و پاسخ ایمنی میزبان را تحریک می‌نماید، این ناحیه به صورت مجزا نیز با روش مشابه بررسی گردید.

سرور IEDB برای پیشگویی اپی‌توپ‌های ایمنی‌زا در آل IEd موشی، از روش smm استفاده کرد. نتایج آن برای کل قطعه Ag ۴۳ دامنه α به صورت مجزا از هم و در کنار ۲۰ اپی‌توپ ۹ تایی حاصل از دو سرور Rankpep و PREDmus که بالاترین امتیاز را داشتند، انتخاب شدند. قطعات اپی‌تویی مشترک یاخته T آل‌های انسانی و موشی به طور جدا از هم بررسی شدند. از میان آن‌ها، توالی‌های مشترک در حداقل دو آل انسانی و دو سرور موشی تعیین شد و مقایسه‌ای بین آن‌ها و اپی‌توپ‌های یاخته B صورت گرفت و اپی‌توپ‌هایی که قادرند هر دو پاسخ یاخته‌های B و T را القاء نمایند مشخص شدند (جدول ۳ الف). آنگاه در نهایت امر، جهت بررسی بیشتر و تعیین اعتبار، توسط سرور VaxiJen برای مشخص کردن میزان پادگن‌زایی و در پایگاه اطلاعاتی IEDB (http://www.php.iedb.org/home_v2) جهت یافتن اپی‌توپ‌های مشابه تجربی ثبت شده، واکاوی شدند (جدول ۳ ب).

جایگاه‌های آلرژن‌زای پروتئین و میزان شباهت با ژن‌های گاوی: در جستجوی توالی‌های پروتئینی مشابه گاوی و موشی با قطعه Ag ۴۳ در پایگاه NCBI، نتیجه‌ای به دست نیامد. در مورد اپی-توپ‌های حاصل در جدول ۳ نیز، شباهت معنی‌داری در کتابخانه آلرژن SDAP و همچنین ژن‌های گاوی در NCBI یافت نشد. براساس دستاوردهای متفاوت آلرژن‌زایی در ابزار AlgPred، این پروتئین تنها با روش SVM به عنوان آلرژن بالقوه شناخته و با سایر روش‌ها آلرژن محسوب نشد و دستاورد هیبرید (مجموع روش‌ها) نیز آن را غیر آلرژن معرفی کرد و همین‌طور تمامی اپی‌توپ‌های مشخص شده جدول ۳ با تک تک روش‌ها و دست‌آورد هیبرید



شد. زوایای چرخش phi و psi برای همه باقیمانده‌های ساختار پروتئین در محورهای X و Y قابل مشاهده است. زاویه phi چرخش را حول باند Calpha-N اسیدآمینه نشان می‌دهد درحالیکه زاویه psi چرخش را حول باند Calpha-C مشخص می‌نماید. از آنجا که پلی‌پپتیدها فرایند تاخوردگی (folding) را متحمل می‌شوند، زوایای phi و psi در محدوده چرخش ۱۸۰- تا ۱۸۰+ درجه بسته به نوع زنجیره جانبی اسیدآمینه قرار می‌گیرند. ترکیباتی که اتم‌ها در آن در فاصله‌ای نزدیکتر از مجموع شعاع واندروالسی آن‌ها قرار می‌گیرند نواحی غیرمجاز یا پرت نقشه را می‌سازند. قرارگیری در این نواحی برای همه اسیدآمینه‌ها به غیر از گلیسین از نظر برخورد فضایی ممکن نیست. گلیسین به علت نداشتن زنجیره جانبی محدودیت سایر اسیدآمینه‌ها را نداشته و در تمام نقاط نقشه به خصوص نواحی مخصوص دورها در ساختار دوم، که برای سایرین ممنوع است، می‌تواند قرار بگیرد. مکان این اسیدآمینه در نقشه حاصل از سرور PROCHECK با علامت مثلث قابل تشخیص است و در محاسبات آماری لحاظ نمی‌شود در حالیکه در نقشه بدست آمده از سرور Rampage با علامت ضربدر قابل مشاهده است و تعداد آن در آمار محاسبه می‌گردد. درصد باقیمانده‌ها در ناحیه مطلوب راهنمای خوبی جهت ارزیابی کیفیت استریوشیمیایی ساختار است. این مقدار در سرور PROCHECK بالای ۹۰ درصد و در سرور Rampage حدود ۹۸ درصد برای یک ساختار پایدار با بررسی کریستالوگرافی ساختارهای پروتئینی شناخته شده، مورد انتظار است. در مدل بهینه ساختار سوم Ag ۴۳ یعنی بعد از به حداقل رساندن انرژی و اصلاح باندهای هیدروژن، وضعیت مکانی ستون فقرات و موقعیت زنجیره‌های جانبی توسط سرور ModRefiner، در مجموع بیش از ۹۵ درصد از اسیدآمینه‌ها در نواحی مطلوب و مجاز نقشه رامچاندرا قرار گرفتند. با توجه به فقدان الگوی مناسب جهت دامنه ناقل این پروتئین در پایگاه PDB، بیشتر اسیدآمینه‌های تشکیل دهنده ناحیه مجاز نقشه رامچاندرا متعلق به قسمت انتهایی قطعه هستند و همین امر منجر به کاهش نسبی پایداری کل ساختار شده است.

توانایی تشخیص اپی‌توپ‌ها در پاسخ ایمنی، کاربردهای مهمی در تشخیص بیماری دارد و در نتیجه از گام‌های اساسی در مطالعات ایمونوبیوانفورماتیک شناسایی و نقشه‌برداری آن‌ها محسوب می‌شود. اپی‌توپ‌ها، گروه‌های تعریف شده‌ای از اسیدهای آمینه هستند که در ترکیب یک پروتئین پادگنی قرار دارند و پاسخ ایمنی را به وسیله برهمکنش با گیرنده‌های یاخته T و B فعال می‌کنند (۴۲، ۴۳). الگوریتم‌های پیشگویی کننده اپی‌توپ‌های یاخته B و T به دو دسته تقسیم می‌شوند. روش‌های بر پایه ساختار که اطلاعات را از ساختمان سه بعدی پروتئین‌ها کسب می‌کنند و روش‌های بر پایه توالی که توالی اسیدآمینه‌ای را مد نظر دارند (۳۳). از آنجا که ساختار سوم پیشگویی شده in-silico هرچند با کیفیت بالا با ساختار سوم واقعی در شرایط in-vivo می‌تواند

درصد سویه‌های ایدکولای بیماری‌زا در دستگاه ادراری و اسهال‌زا و نزدیک به ۵۶ درصد جدایه‌های همزیست حضور دارد. واکسیناسیون موشها با دامنه مسافر نوترکیب این پروتئین که از سویه انسانی جدا شده است، با ایجاد حفاظت مطلوب در مقابل استقرار روده‌ای ETEC‌ها، آن را به عنوان کاندید بالقوه ایمنی‌زا در طراحی واکسن‌های جدید اسهال ناشی از ETEC نوزادان جالب توجه کرده است (۱۸، ۴۴).

عدم وجود اطلاعات کافی در مورد سویه‌های بیماری‌زای ETEC دامی و توجه به رشد و پیشرفت روز افزون واکسن‌شناسی معکوس از طریق مطالعات ایمونوبیوانفورماتیک هدف ما را در این مطالعه بررسی ساختاری و ایمنی‌زایی پروتئین پادگن ۴۳ سویه ۵۱۰ باکتری ETEC قرار داده است. دامنه مسافر پروتئین نوترکیب Ag ۴۳ سویه ۵۱۰ از ۴۹۹ اسیدآمینه تشکیل شده است که با باقیمانده‌های ۱۷ تایی و ۱۸۶ تایی قطعات پپتید سیگنال و دامنه ناقل، توالی ۷۰۲ اسیدآمینه‌ای را تشکیل می‌دهد. در بررسی خصوصیات فیزیکی- شیمیایی، شاخص بی‌ثباتی (Instability index) کمتر از ۴۰ نشان دهنده مقاوم بودن پروتئین است و همچنین بالا بودن شاخص آلیفاتیک که با نسبت حضور اسیدآمینه‌های دارای زنجیره آلیفاتیک (آلانین، والین، ایزولوسین و لوسین) ارتباط دارد حاکی از احتمال پایداری پروتئین در بازه دمایی وسیع است (۱۷). با توجه به ساختار دوم پروتئین میزان بیشتر آن را صفحات بتا تشکیل می‌دهد که به خوبی خاصیت تجمع‌ی این پروتئین را توضیح می‌دهد (۴۰). به‌ویژه هنگامیکه وقتی که در نگاه جزئی‌تر مشخص می‌شود قطعه مرتبط با دامنه مسافر شامل هیچ ساختار آلفا-هلیکسی نیست.

در مورد پیشگویی ساختار سوم پروتئین، در حال حاضر یک ساختار PDB از پروتئین دامنه مسافر Ag ۴۳a سویه یوروپاتوزنیک انسانی (CFT۰۷۳) در پایگاه داده پروتئین وجود دارد و نتیجه هم‌ردیفی Ag ۴۳ با آن، ۷۹ درصد هم‌پوشانی و ۸۱ درصد شباهت است. از این پروتئین در اکثر سرورهای پیشگویی کننده ساختار سه بعدی، به عنوان اولین الگوی انتخابی استفاده گردید. در مورد اسیدآمینه‌های انتهایی قطعه Ag ۴۳ نوترکیب که هم‌ردیف دامنه ناقل پادگن ۴۳ سویه ۱۲ k قرار دارند، در اکثر سرورهای پیشگویی کننده برپایه تشابه توالی، ساختاری برای آن تعیین نشد. در بین سه سرور I-Tasser، Phyre۲ و PS۲-۷۳ که یک ساختار کلی با همه اسیدآمینه‌های شرکت کننده ارائه دادند نیز تنها مدل پیشنهادی سرور I-Tasser بود که یک ساختار مشخص واحد برای هر دو دامنه قطعه پیشگویی کرد و با توجه به تأیید جهانی نتایج این سرور، آن مدل انتخاب گردید. اساس عملکرد سرور I-Tasser مدل‌سازی ab initio است، ضمن آنکه جستجو توالی مشابه برای توالی اسیدآمینه‌ای مورد تحقیق و نیز جستجوی ساختار پروتئینی مشابه با ساختار پیشگویی شده از جمله روندهای اجرایی این سرور است.

به منظور ارزیابی کیفیت مدل مورد نظر از نقشه رامچاندرا استفاده



باسیلوز و اعتبار طرح پژوهشی طرح نوع ششم به شماره ۷۵۰۲۰۱۵/۶/۲۶ تأمین شده است.

تعارض در منافع

بین نویسندگان هیچ گونه تعارض در منافع گزارش نشده است.

References

1. Abbass, J., Nebel, J. C., Mansour, N. (2013). Ab initio protein structure prediction: methods and challenges. *Biological Knowledge Discovery Handbook*. Hoboken, John Wiley & Sons, Inc. New Jersey, USA. p.703-724.
2. Ansari, H. R., Raghava, G. P. (2010). Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome Res*, 6, 6. <https://doi.org/10.1186/1745-7580-6-6> PMID: 20961417
3. Bhasin, M., Garg, A., Raghava, G. (2005). PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics*, 21, 2522-2524. <https://doi.org/10.1093/bioinformatics/bti309> PMID: 15699023
4. Bhasin, M., Raghava, G. (2005). Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. *Nucleic acids Res*, 33, W202-W207. PMID: 15988831
5. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*, 42, W252-W258. <https://doi.org/10.1093/nar/gku340> PMID: 24782522
6. Buchan, D. W., Minnici, F., Nugent, T. C., Bryson, K., Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res*, 41, W349-W357. <https://doi.org/10.1093/nar/gkt381> PMID: 23748958
7. Cheng, J., Randall, A. Z., Sweredoski, M. J., Baldi, P. (2005). SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res*, 33, W72-W76. <https://doi.org/10.1093/nar/gki396> PMID: 15980571

تفاوت‌های قابل ملاحظه‌ای داشته باشد در این مطالعه تنها به سرورهای پیشگویی کننده اپی توپ بر پایه توالی پروتئین بسنده نمودیم. تعداد قابل ملاحظه‌ای اپی توپ یاخته B مشخص شدند و در مقایسه نتایج سرورهای مختلف اپی توپ‌های مشابه پیشنهاد گردیدند.

در خصوص اپی توپ‌های یاخته T ضروری است که پپتیدهای پادگنی به MHC متصل شوند تا یاخته‌های T بتوانند آن‌ها را بشناسند. بنابراین، شناسایی پپتیدهای متصل شونده به MHC بخش اساسی هر الگوریتمی است که اپی توپ‌های یاخته T را پیش‌بینی می‌کند (۳۳). در این تحقیق از آنجاییکه هدف ما شناخت کاندیدهای مناسب واکسن کلی‌باسیلوز بوده و با توجه به اینکه باکتری *E. coli* انتروتوکسیژنیک به طور غالب یک پاتوژن خارج یاخته‌ای است، اپی توپ‌های متصل شونده به MHC کلاس II مد نظر قرار گرفتند.

بدین منظور بر اساس نتایج حاصل از مطالعه پیشین Nikbakht و همکاران در سال ۲۰۱۶ که حاوی اطلاعات لازم برای تخمین فراوانی آل‌های BoLA-DRB۳ در جمعیت گاوهای هلشتاین ایران بود، ۵ آل غالب جمعیت (فراوانی بالای ۵ درصد) انتخاب شدند. این در حالی است که تا کنون هیچ سرور آل‌های MHC-II گاو را جهت تعیین اپی توپ‌های یاخته T ارائه نداده است. لذا اپی توپ‌های یاخته T برای آل‌های مشابه انسانی و مدل حیوان آزمایشگاهی موش تعیین شدند. سرانجام به ۱۵ توالی پپتیدی در قطعه کلون شده Ag۴۳ سویه ۵۱۰ دست یافتیم که ۱۰ توالی پپتیدی آن در دامنه مسافر قرار دارند و بالقوه قادرند باعث تحریک هر دو یاخته B و T شوند. همه این توالی‌ها مورد تحلیل سرور VaxiJen قرار گرفتند. از آنجا که حد آستانه این سرور ۰/۴ است، قطعات پپتیدی با نمره بالای ۰/۴ توسط این سرور پادگن محسوب می‌شوند. بنابراین در نهایت از مجموع ۱۵ قطعه پپتیدی حاصل، ۹ اپی توپ پادگنی شناخته شدند که تعداد ۶ اپی توپ از دامنه مسافر هستند. درحالیکه هیچکدام از آن‌ها در بردارنده نواحی آل‌ژن‌زای شناسایی شده بر طبق قانون WHO نیستند.

در مجموع شایان ذکر است که با توجه به قدرت تحریک ایمنی، پایداری، حلالیت مناسب و اپی توپ‌های یافت شده در این مطالعه می‌توان پروتئین پادگن ۴۳ سویه ۵۱۰ *E. coli* گاو به‌عنوان یک پادگن حراست شده در بین سویه‌های *E. coli* را جهت تهیه واکسن‌های نو ترکیب با استفاده از دامنه مسافر آن و همچنین در طراحی واکسن‌های بر پایه اپی توپ با معرفی اپی توپ‌های بالقوه ایمنی‌زا در مقابله با بیماری اسهال گوساله‌ها در نظر گرفت. نتایج حاصل از تحقیق حاضر برای طراحی واکسن بر مبنای اپی توپ کاربرد داشته و می‌توان از توالی‌های پیشنهادی برای تهیه واکسن‌های ترکیبی نیز استفاده نمود.

تشکر و قدردانی

هزینه‌های این طرح از بودجه طرح تولید دانش فنی واکسن کلی



8. Chou, P. Y., Fasman, G. D. (1974). Prediction of protein conformation. *Biochemistry*, 13, 222-245. <https://doi.org/10.1021/bi00699a002>
9. De Luna, M. d. G., Scott-Tucker, A., Desvaux, M., Ferguson, P., Morin, N. P., Dudley, E. G., Henderson, I. R. (2008). The *Escherichia coli* biofilm-promoting protein Antigen 43 does not contribute to intestinal colonization. *FEMS Microbiol Lett*, 284, 237-246. <https://doi.org/10.1111/j.1574-6968.2008.01207.x> PMID: 18507683
10. Diaz, A. A., Tomba, E., Lennarson, R., Richard, R., Bagajewicz, M. J., Harrison, R. G. (2010). Prediction of protein solubility in *Escherichia coli* using logistic regression. *Biotechnol Bioeng*, 105, 374-383. <https://doi.org/10.1002/bit.22537> PMID: 19739095
11. Doytchinova, I. A., Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8, 4. <https://doi.org/10.1186/1471-2105-8-4>
12. Drozdetskiy, A., Cole, C., Procter, J., Barton, G. J. (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Res*, 43, W389-W394. <https://doi.org/10.1093/nar/gkv332> PMID: 25883141
13. EL-Manzalawy, Y., Dobbs, D., Honavar, V. (2008). Predicting linear B-cell epitopes using string kernels. *J Mol Recogn*, 21, 243-255. <https://doi.org/10.1002/jmr.893> PMID: 18496882
14. Emini, E. A., Hughes, J. V., Perlow, D., Boger, J. (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol*, 55, 836-839. PMID: 2991600
15. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D., Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook*. Springer. p. 571-607.
16. Harris, J. A., Roy, K., Woo-Rasberry, V., Hamilton, D. J., Kansal, R., Qadri, F., Fleckenstein, J. M. (2011). Directed evaluation of enterotoxigenic *Escherichia coli* autotransporter proteins as putative vaccine candidates. *PLoS Negl Trop Dis*, 5, e1428. <https://doi.org/10.1371/journal.pntd.0001428> PMID: 22163060
17. Heffernan, R., Paliwal, K., Lyons, J., Dehzangi, A., Sharma, A., Wang, J., Zhou, Y. (2015). Improving prediction of secondary structure, local backbone angles, and solvent accessible surface area of proteins by iterative deep learning. *Sci Rep*, 5, 11476. <https://doi.org/10.1038/srep11476> PMID: 26098304
18. Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C., Ala'Aldeen, D. (2004). Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev*, 68, 692-744. <https://doi.org/10.1128/MMBR.68.4.692-744.2004> PMID: 15590781
19. Heras, B., Totsika, M., Peters, K. M., Paxman, J. J., Gee, C. L., Jarrott, R. J., Schembri, M. A. (2014). The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping. *Proc Nat Acad Sci*, 111, 457-462. <https://doi.org/10.1073/pnas.1311592111> PMID: 24335802
20. Huang, T.-T., Hwang, J.-K., Chen, C.-H., Chu, C.-S., Lee, C.-W., Chen, C.-C. (2015). (PS)2: protein structure prediction server version 3.0. *Nucleic Acids Res*, 43, W338-W342. <https://doi.org/10.1093/nar/gkv454> PMID: 25943546
21. Kaur, H., Raghava, G. (2004). A neural network method for prediction of β -turn types in proteins using evolutionary information. *Bioinformatics*, 20, 2751-2758. <https://doi.org/10.1093/bioinformatics/bth322>
22. Kaur, H., Raghava, G. (2004). Prediction of α -turns in proteins using PSI-BLAST profiles and secondary structure information. *Proteins: Structure, Function, and Bioinformatics*, 55, 83-90. <https://doi.org/10.1002/prot.10569> PMID: 14997542
23. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, 10, 845-858. <https://doi.org/10.1038/nprot.2015.053> PMID: 25950237
24. Kim, D. E., Chivian, D., Baker, D. (2004). Protein structure prediction and analysis using the



- Robetta server. *Nucleic Acids Res*, 32, W526-W531. <https://doi.org/10.1093/nar/gkh468> PMID: 15215442
25. Kloczkowski, A., Ting, K. L., Jernigan, R., Garnier, J. (2002). Combining the GOR V algorithm with evolutionary information for protein secondary structure prediction from amino acid sequence. *Proteins: Structure, Function, and Bioinformatics*, 49, 154-166. <https://doi.org/10.1002/prot.10181> PMID: 12210997
 26. Larsen, J. E. P., Lund, O., Nielsen, M. (2006). Improved method for predicting linear B-cell epitopes. *Immunome Res*, 2, 2. <https://doi.org/10.1186/1745-7580-2-2> PMID: 16635264
 27. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*, 26, 283-291. <https://doi.org/10.1107/S0021889892009944>
 28. Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, D. C. (2003). Structure validation by α geometry: ϕ , ψ and $C\beta$ deviation. *Proteins: Structure, Function, and Bioinformatics*, 50, 437-450. <https://doi.org/10.1002/prot.10286>
 29. Ma, J., Wang, S., Zhao, F., Xu, J. (2013). Protein threading using context-specific alignment potential. *Bioinformatics*, 29, i257-i265. <https://doi.org/10.1093/bioinformatics/btt210> PMID: 23812991
 30. Nielsen, M., Lundegaard, C., Lund, O., Keşmir, C. (2005). The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics*, 57, 33-41. <https://doi.org/10.1007/s00251-005-0781-7> PMID: 15744535
 31. Nielsen, M., Lundegaard, C., Lund, O., Petersen, T.N. (2010). CPHmodels-3.0—remote homology modeling using structure-guided sequence profiles. *Nucleic Acids Res*, 38, W576-W581. <https://doi.org/10.1093/nar/gkq535> PMID: 20542909
 32. Nikbakht Brujeni, G., Ghorbanpour, R., Esmailnejad, A. (2016). Association of BoLA-DRB3. 2 Alleles with BLV Infection Profiles (Persistent Lymphocytosis/Lymphosarcoma) and Lymphocyte Subsets in Iranian Holstein Cattle. *Biochem Genet*, 54, 194-207. <https://doi.org/10.1007/s10528-016-9712-6> PMID: 26782666
 33. Patronov, A., Doytchinova, I. (2013). T-cell epitope vaccine design by immunoinformatics. *Open Biol*, 3, 120139. <https://doi.org/10.1098/rsob.120139> PMID: 23303307
 34. Pieper, U., Webb, B. M., Dong, G. Q., Schneidman-Duhovny, D., Fan, H., Kim, S. J., Hammel, M. (2014). ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res*, 42, D336-D346. <https://doi.org/10.1093/nar/gkt1144> PMID: 24271400
 35. Reche, P. A., Reinherz, E. L. (2007). Prediction of peptide-MHC binding using profiles. *Immunoinformatics*, Springer. p. 185-200. https://doi.org/10.1007/978-1-60327-118-9_13 PMID: 18450001
 36. Rechsteiner, M., Rogers, S. W. (1996). PEST sequences and regulation by proteolysis. *Trends Biochem Sci*, 21, 267-271. PMID: 8755249
 37. Saha, S., Raghava, G. (2004). BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. In, *International Conference on Artificial Immune Systems*, Springer. p. 197-204. https://doi.org/10.1007/978-3-540-30220-9_16
 38. Saha, S., Raghava, G. (2006). AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res*, 34, W202-W209. <https://doi.org/10.1093/nar/gkl343> PMID: 16844994
 39. Saha, S., Raghava, G. (2006). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins: Structure, Function, and Bioinformatics*, 65, 40-48. <https://doi.org/10.1002/prot.21078> PMID: 16894596
 40. Shivu, B., Seshadri, S., Li, J., Oberg, K. A., Uversky, V. N., Fink, A. L. (2013). Distinct β -sheet structure in protein aggregates determined by ATR-FTIR spectroscopy. *Biochemistry*, 52, 5176-5183. <https://doi.org/10.1021/bi400625v>



- PMID: 23837615
41. Söding, J., Biegert, A., Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res*, 33, W244-W248. <https://doi.org/10.1093/nar/gki408> PMID: 15980461
42. Soni, B., Seniya, C., Misra, R., Vyas, V. (2013). Immuno-informatic approach of in silico T cell epitope prediction. *Int J Adv Technol Eng Res*, 3, 9-18.
43. Tomar, N., De, R. K. (2010). Immunoinformatics: an integrated scenario. *Immunology*, 131, 153-168. <https://doi.org/10.1111/j.1365-2567.2010.03330.x> PMID: 20722763
44. Vander Woude, M. W., Henderson, I. R. (2008). Regulation and function of Ag43 (flu). *Ann Rev Microbiol*, 62, 153-169. <https://doi.org/10.1146/annurev.micro.62.081307.162938> PMID: 18785838
45. Wang, P., Sidney, J., Kim, Y., Sette, A., Lund, O., Nielsen, M., Peters, B. (2010). Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinform*, 11, 568. <https://doi.org/10.1186/1471-2105-11-568> PMID: 21092157
46. Wells, T. J., Tree, J. J., Ulett, G. C., Schembri, M. A. (2007). Autotransporter proteins: novel targets at the bacterial cell surface. *FEMS Microbiol Lett*, 274, 163-172. <https://doi.org/10.1111/j.1574-6968.2007.00833.x>
47. Wiederstein, M., Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*, 35, W407-W410. <https://doi.org/10.1093/nar/gkm290> PMID: 17517781
48. Xu, D., Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J*, 101, 2525-2534. <https://doi.org/10.1016/j.bpj.2011.10.024> PMID: 22098752
49. Yachdav, G., Klopman, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., Hönigschmid, P., Schafferhans, A., Roos, M., Bernhofer, M. (2014). PredictProtein—an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res*, 42, W337-W343. <https://doi.org/10.1093/nar/gku366> PMID: 24799431
50. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat Methods*, 12, 7-8. <https://doi.org/10.1038/nmeth.3213> PMID: 25549265
51. Yu, C. S., Lin, C. J., Hwang, J. K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Science*, 13, 1402-1406. <https://doi.org/10.1110/ps.03479604> PMID: 15096640
52. Zhang, G. L., Srinivasan, K. N., Veeramani, A., August, J. T., Brusica, V. (2005). PRED^{BALB/c}: a system for the prediction of peptide binding to H2d molecules, a haplotype of the BALB/c mouse. *Nucleic Acids Res*, 33, W180-W183. <https://doi.org/10.1093/nar/gki479> PMID: 15980450



Immuno-Bioinformatics Study of Autotransporter Protein, Antigen 43, in Enterotoxigenic *Escherichia coli* Isolated From Calves

Reyhaneh Ghorbanpour¹, Gholamreza Nikbakht¹, Amir Hossein Jalali²

¹Department of Microbiology and Immunology, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran

²Department of Natural Resources Engineering, Isfahan University of Technology, Isfahan, Iran

³Institute of Biotechnology & Bioengineering, Isfahan University of Technology, Isfahan, Iran

(Received 27 August 2018, Accepted 14 November 2018)

Abstract:

BACKGROUND: Extensive effort is focused on identifying genomic conserved antigens in development of effective vaccine against Enterotoxigenic *Escherichia coli*. Antigen 43 is one of the members of a large secreted protein family named autotransporters in the *E. coli* and other gram negative bacteria. Autotransporter proteins have a similar conserved structure. Some of them are recognized during both experimental and naturally occurring ETEC infections. Antigen 43 is represented as a potential target in vaccine development because of its virulence functions such as cell aggregation, biofilm formation and its presence in convalescent sera from human patients with ETEC diarrhea.

OBJECTIVES: In this study, we carefully investigate antigenic structure and immunogenicity of the Antigen 43 protein of strain 510 *E. coli* isolated from calves by experimental methods and immunoinformatics tools.

METHODS: Amino acid sequence, physico-chemical parameters, stability, secondary and tertiary protein structure, the ability of induction the B and T cell immune responses by having the effective epitopes and also the allergenicity assessment were analyzed.

RESULTS: In this study, we identified 15 peptide sequences that can potentially induce B and T cell immune responses and finally, 9 of them were introduced as antigens.

CONCLUSIONS: The results of in-silico analysis on this protein suggested that it can be used in bovine colibacillosis vaccine development.

Keywords:

Immuno-Bioinformatics, Enterotoxigenic *Escherichia coli* (ETEC), Autotransporter, Antigen 43, Colibacillosis

Figure Legends and Table Captions

Table 1. Secondary structure prediction.

Table 2. The similarity comparison of the most frequent of BoLA alleles in Iranian Holstein cattle with human and mouse MHC alleles.

Table 3 (A, B). The common B and T cell predicted epitopes.

Figure 1. PEST motifs of Ag43 sequence in strain 510.

Figure 2. Tertiary structure of Ag43.

Figure 3. Z-score plot of ProSA website. The z-score indicates overall model quality and is determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) of protein chains in current PDB with respect to their length. The plot shows results with a z-score ≤ 10 . The z-score of Ag43 protein is indicated by a large dot. The score is in the range of native proteins with the value of -7.15.

Figure 4. The Ramachandran plot- every red dot in both plots indicates an amino acid with illegal phi and psi angle. (A) PROCHECK server: 4 colored regions red, yellow, cream and white display most favourable, additional allowed, generously allowed and disallowed regions, respectively. (B) RAMPAGE server: 3 colored regions dark, light and white shows respectively the favoured, allowed and outlier regions.

