

کاربرد الگوریتم ژنتیک و پرسپترون چند لایه در پیشگویی ساختارهای مؤثر در اتصال اتوآنتیبادی‌ها

محمد طاهرزاده ثانی

دانشجوی کارشناسی ارشد

گروه برق و کامپیوتر، دانشکده فنی، دانشگاه تهران

مجتبی میرصالحی

دانشیار

گروه برق، دانشکده مهندسی، دانشگاه فردوسی مشهد

مرتضی خادمی

استادیار

گروه برق، دانشکده مهندسی، دانشگاه فردوسی مشهد

محمود محمودی

استادیار،

دانشکده پزشکی، دانشگاه علوم پزشکی مشهد

محمد باقر نقیبه سیستانی

دانشجوی دکتری

گروه برق، دانشکده مهندسی، دانشگاه فردوسی مشهد

اسد عازمی

دانشیار

دانشکده مهندسی برق، دانشگاه ایالتی پنسیلوانیا

چکیده

لوپوس اریتماتوس سیستمیک (SLE) از مهمترین بیماری‌های خودایمنی است و یکی از علائم آن در انسانها حضور آنتیبادی‌ها ضد DNA است. در این مقاله با استفاده از شبکه عصبی پرسپترون چند لایه و الگوریتم ژنتیک، نواحی مؤثر و اسیدهای آمینه مهم در آنتیبادی‌های DNA که باعث اتصال آنها به DNA می‌شوند، تعیین می‌گردد، نشان داده می‌شود که بیشترین تأثیر در اتصال آنتیبادی به DNA مربوط به نواحی CDR2 و CDR3 است. همچنین نتایج نشان می‌دهد که اسیدهای آمینه دارای بار الکتریکی، اهمیت زیادی در اتصال آنتیبادی به DNA دارند و تأثیر این اسیدهای آمینه در نواحی CDR2 و CDR3 بیشتر است. شمار زیادی از نتایج فوق با تحقیقات آزمایشگاهی پزشکی تایید شده است. این نتایج که به وسیله روشهای مبتنی بر شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک به دست می‌آیند، می‌توانند تحقیقات پزشکی را به سمت بررسی آزمایشگاهی تأثیر ساختارهای مشخص شده هدایت کنند. بدین ترتیب در هزینه و زمان تحقیق در زمینه طراحی داروهای جدید برای بیماری‌های خودایمنی، صرفه جویی قابل توجهی حاصل خواهد شد.

کلمات کلیدی

شبکه عصبی، پرسپترون چند لایه، الگوریتم ژنتیک، بیماری‌های خودایمنی، آنتی بادی

Application of Genetic Algorithm and Multilayer Perceptron in Prediction of Effective Structures of Autoantibodies Binding

M. Taherzadeh
Graduate Student
ECE Department, Faculty of Engineering,
University of Tehran

M. Mirsalehi
Associate Professor
Electrical Engineering Department,
Ferdowsi University of Mashhad

M. Khademi
Assistant Professor
Electrical Engineering Department,
Ferdowsi University of Mashhad

M. Mahmoudi
Assistant Professor
School of Medicine,
Mashhad University of Medical Sciences

M. B. Naghibi
Ph.D. Student
Electrical Engineering Department,
Ferdowsi University of Mashhad

A. Azemi
Associate Professor
Electrical Engineering Department,
Penn. State University

Abstract

Systemic lupus erythematosus (SLE or 'lupus') is a major autoimmune rheumatic disease. One of the indicators of the lupus disease in human is the presence of antibodies against double-stranded DNA (dsDNA). In this paper, using a multilayer perceptron neural network and genetic algorithm, the effective regions and important amino acids of antibodies that cause the binding to DNA are predicted. Our results indicate that the most influential regions in the antibodies binding are CDR3 and CDR2. Moreover, the amino acids with electrical charges are important in binding antibodies to DNA. The effects of these amino acids are more evident at CDR3 and CDR2 regions. Clinical experimental studies confirm the correctness of the proposed results.

Results obtained in this paper can be used to guide clinical studies in this subject, which reduces the cost of designing new drugs for autoimmune diseases.

Keywords

Neural network, Multilayer perceptron, Genetic algorithm, Autoimmune diseases, Antibody

مقدمه

آنتیبادی‌ها^۱ (پادتن) ملکول‌های گلیکو پروتئینی هستند که دارای اعمال متفاوتی در مواجهه با سیستم دفاعی بدن دارند. این پروتئین‌ها به آنتیژن‌ها^۲ (پادگن، جسم خارجی) متصل می‌شوند و باعث فعال شدن روندی می‌شوند که به حذف آنتیژن می‌انجامد. هر آنتیژن، دارای آنتیبادی‌های اختصاصی در بدن می‌باشد. گونه‌های از آنتیبادی‌ها که به آنتیژن‌های خودی متصل می‌شوند، اتوآنتیبادی نامیده می‌شوند. این آنتیبادی‌ها باعث حمله به آنتیژن‌های خودی، اغلب باعث بیماری می‌شوند. یکی از این بیماری‌ها لوپوس اریتماتوی سیستمیک^۳ نام دارد که در اثر آنتیبادی علیه DNA^۴ ایجاد می‌شود [۱]. در این تحقیق، این آنتیبادی مورد بررسی قرار خواهد گرفت. واضح است که آنتیبادی‌هایی که به یک نوع آنتیژن متصل می‌شوند، دارای ساختار یا ساختارهای خاص در شکل، بار الکتریکی و یا سایر خواص فیزیکوشیمیایی هستند. شناخت این ساختارهای خاص در

آنتیبادی‌های علیه DNA، ما را در راه از بین بردن آنها و کمک به درمان لوپوس اریتماتوی سیستمیک یاری می‌کند. با وجود زمان زیادی که از شناسایی آنتیبادی‌های علیه DNA و نقش آنها در بیماری لوپوس اریتماتوی سیستمیک می‌گذرد و با وجود تحقیقات دامنه داری که در این زمینه شده است، هنوز اتفاق نظر قابل ملاحظه‌ای درباره ساختار آنتیبادی‌های علیه DNA و نواحی متصل شونده این آنتیبادی به DNA، (خصوصاً در آنتیبادی‌های انسانی) حاصل نشده است [۲]. در این تحقیق برای پی بردن به این ساختارهای خاص، آنتیبادی‌های انسانی از نظر اتصال به DNA، به کمک شبکه عصبی پرسپترون چند لایه^۵، دسته‌بندی شدند. سپس با استفاده از الگوریتم ژنتیک و شبکه آموزش یافته، آنتیبادی‌های با قدرت اتصال بیشتر به DNA، به دست آمد. نسل جدید آنتیبادی‌های علیه DNA دارای تشابهات ساختاری مشخص تری در رشته پروتئینی آنها بودند. چون نسل جدید آنتیبادی‌های علیه DNA دارای ساختارهای مؤثر بیشتری در اتصال به DNA می‌باشند، تشابهات نسل جدید به عنوان ساختارهای مؤثر در اتصال آنتیبادی به DNA مورد بررسی بیشتر قرار گرفت.

۱- ساختار آنتیبادی

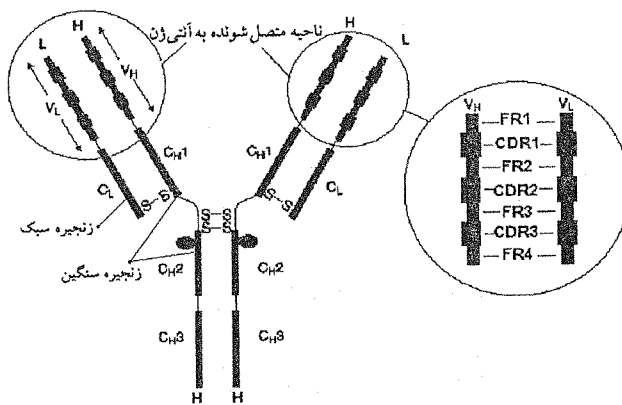
از دیدگاه ساختاری، آنتیبادی یک ملکول گلیکو پروتئینی قرینه است که از دو زنجیره^۶ سنگین و سبک از اسیدهای آمینه تشکیل شده است (شکل ۱). هر اسید آمینه در این زنجیره‌ها می‌تواند یکی از بیست نوع ممکن باشد که فهرست آنها در ستون اول جدول ۱ آمده است. قطعه‌های V_L و V_H که به ترتیب قسمت‌های متغیر زنجیره‌های سبک و سنگین هستند، ناحیه متصل شونده به آنتیژن را تشکیل می‌دهند [۳]. این دو قطعه، روبروی هم قرار گرفته‌اند. ناحیه متصل شونده به آنتیژن، در هر دو نیمه آنتیبادی به طور مشابه، قرار دارد. رشته‌های V_L و V_H هر کدام از هفت ناحیه FR1، CDR1، FR2، CDR2، FR3، CDR3 و FR4 تشکیل شده‌اند (شکل ۱).

تحقیقات آزمایشگاهی، نشان دهنده نقش بیشتر V_H در اتصال به DNA است [۴-۹]. در این تحقیق نیز تنها از V_H جهت آموزش شبکه عصبی استفاده گردیده است. قطعه پروتئینی V_H حداکثر دارای ۱۳۳ اسید آمینه است. قطعه V_H آنتیبادی‌های مورد نیاز، از بانک اطلاعاتی Kabat [۱۰] استخراج شده است.

۲- کد گذاری اطلاعات

هر اسید آمینه دارای مجموع‌های از خواص فیزیکوشیمیایی مانند pH ایزو الکتریک، جرم ملکولی، آب دوستی^۸ و... است که به ازای هر اسید آمینه، یک تعداد عدد، متناظر با پارامترهای مختلف، به آن اختصاص می‌یابد. به عنوان مثال، معیار آب دوستی که توسط افراد مختلف محاسبه شده است (جدول ۱)، نشان دهنده این می‌باشد که یک اسید آمینه در یک رشته پروتئینی و در محیط آبی، در صورت به هم پیچیدن پروتئین، تا چه حد میل به قرار گرفتن در سطح خارجی پروتئین دارد. به ازای هر توالی از اسیدهای آمینه، برای کد گذاری اطلاعات و ورود به شبکه عصبی، می‌توان یک رشته عددی شامل خواص فیزیکوشیمیایی اسیدهای آمینه داشت. در این طرح از دو پارامتر pH ایزوالکتریک و آب دوستی در اسیدهای آمینه جهت ورودی به شبکه عصبی استفاده شده است.

در مورد pH ایزوالکتریک، میانگین این پارامتر در بیست اسید آمینه، از pH ایزوالکتریک هر اسید آمینه کم شد و سپس بیست عدد بین -۱ و ۱ نرمالیزه شدند. پارامتر به دست آمده را pH ایزوالکتریک نرمالیزه شده (جدول ۱) می‌نامیم. در مورد آب دوستی محاسبات مختلفی توسط محققین به انجام رسیده است که به علت تفاوت روش‌ها، نتایج نیز متفاوتند. در این تحقیق از سه محاسبه معتبر هاپ و وودس [۱۱]، پارکر [۱۲] و کاپت و دولیتل [۱۳] استفاده شده است. سه محاسبه مختلف فوق را در مورد پارامتر آب دوستی نیز به طور جداگانه و به همان صورتی که در مورد pH ایزوالکتریک عمل شد، نرمالیزه کرده‌ایم. با این تفاوت که به جای تقسیم اعداد بر عددی مشخص برای نرمال شدن بین -۱ و ۱ برای برابر شدن انرژی سه پارامتر، اعداد را بر واریانس آنها تقسیم کرده‌ایم. سپس سه مقدار آب دوستی به دست آمده را باهم جمع و بین -۱ و ۱ نرمالیزه نموده‌ایم. پارامتر به دست آمده را آب دوستی نرمالیزه شده (جدول ۱) نام گذاری می‌کنیم. واضح است که در صورت استفاده دو پارامتر pH ایزوالکتریک نرمالیزه شده و آب دوستی نرمالیزه شده، به ازای هر آنتیبادی، یک ماتریس 2×133 و یا یک بردار با ۲۶۶ عنصر بدست می‌آید.



شکل (۱) ساختار آنتیبادی.

جدول (۱) میزان pH ایزوالکتریک و سه نوع محاسبه مختلف آب دوستی و مقادیر نرمالیزه شده آنها در اسیدهای آمینه.

نام اسید آمینه		pH ایزوالکتریک	آبدوستی Kyte و	آبدوستی Parker	آبدوستی Hopp و	آبدوستی نرمالیزه	pH ایزوالکتریک
نام سه حرفی	نام تک						
Ala	A	۶/۰۰	-۱/۸	۲/۱	-۰/۵	-۰/۱۸۷۲	-۰/۰۴۲
Cys	C	۵/۰۷	-۲/۵	۱/۴	-۱/۰	-۰/۳۲۵۳	-۰/۲۰۴
Asp	D	۲/۷۷	۳/۵	۱۰/۰	۳/۰	۱/۰۰۰۰	-۰/۶۸۵۷
Glu	E	۳/۲۲	۳/۵	۷/۸	۳/۰	۰/۹۱۵۳	-۰/۵۹۰۷
Phe	F	۵/۴۸	-۲/۸	-۹/۲	-۲/۵	-۰/۹۴۷۶	-۰/۱۱۳۹
Gly	G	۵/۹۷	۰/۴	۵/۰۷	۰	۰/۱۷۸۶	-۰/۰۱۰۵
His	H	۷/۵۹	۳/۲	۲/۱	-۰/۵	۰/۲۲۹۱	۰/۳۲۱۲
Ile	I	۶/۰۲	-۴/۵	-۸/۰	-۱/۸	-۰/۹۵۱۴	۰
Lys	K	۹/۷۴	۳/۹	۵/۷	۳/۰	۰/۸۶۶۹	۰/۷۸۴۸
Leu	L	۵/۹۸	-۳/۸	-۹/۲	-۱/۸	-۰/۹۴۰۶	-۰/۰۸۴
Met	M	۵/۷۴	-۱/۹	-۴/۲	-۱/۳	-۰/۵۳۰۱	-۰/۰۵۹۱
Asn	N	۵/۴۱	۳/۵	۷/۰	۰/۲	۰/۵۳۰۷	-۰/۱۲۸۷
Pro	P	۶/۳۰	۱/۶	۲/۱	۰	۰/۱۶۱۹	۰/۰۵۹۱
Gln	Q	۵/۶۵	۳/۵	۶/۰	۰/۲	۰/۴۹۲۲	-۰/۰۷۸۱
Arg	R	۱۰/۷۶	۴/۵	۴/۲	۳/۰	۰/۸۵۸۰	۱/۰
Ser	S	۵/۶۸	۰/۸	۶/۵	۰/۳	۰/۳۰۴۲	-۰/۰۷۱۷
Thr	T	۵/۶۰	۰/۷	۵/۲	۰/۴	۰/۱۵۷۵	-۰/۰۸۸۶
Val	V	۵/۹۶	-۴/۲	-۳/۷	-۱/۵	-۰/۷۲۳۴	-۰/۰۱۲۷
Trp	W	۵/۸۹	۰/۹	-۱۰/۰	-۳/۴	-۰/۷۹۰۷	-۰/۰۲۷۴
Tyr	Y	۵/۶۶	۱/۳	-۱/۹	-۲/۳	-۰/۳۰۷۲	-۰/۰۷۵۹

۳- دسته‌بندی آنتیبادی‌ها بر اساس اتصال به DNA

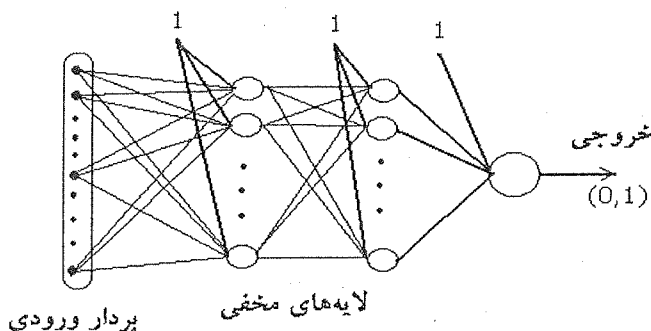
شبکه عصبی پرسپترون چند لایه، یکی از ابزارهای قدرتمند در دسته‌بندی داده‌ها است [۱۴] و به همین دلیل در این تحقیق مورد استفاده قرار گرفته است. رشته‌های عددی محاسبه شده، جهت آموزش یک شبکه عصبی پرسپترون چند لایه (شکل ۲) استفاده شد. آموزش این شبکه به روش پس انتشار خطا صورت گرفت. تعداد ورودی‌های شبکه عصبی ۲۶۶، تعداد نرون‌ها در لایه مخفی اول و دوم ۹ عدد، تابع فعالیت نرون‌های لایه مخفی تانژانت هیپربولیک و تابع فعالیت نرون خروجی سیگموئید انتخاب شد. تعداد لایه‌ها و نوع تابع فعالیت آنها به صورت سعی و خطا و جهت رسیدن به بهترین دقت انتخاب شده‌اند. تعداد آنتیبادی‌های علیه DNA موجود جهت آموزش شبکه عصبی ۴۸ مورد است، ولی در بانک‌های اطلاعاتی هیچ صحبتی از آنتیبادی‌هایی که به DNA متصل نمی‌شوند، نشده است. چون به دلایل ایمنولوژیک، احتمال اتصال آنتیبادی‌های غیر اتوآنتیبادی به DNA بسیار کم است، در این تحقیق از آنها به عنوان آنتیبادهایی که به DNA متصل نمی‌شوند استفاده

شد. تعداد این آنتیبادیها ۳۲۴ مورد است. به علت تعداد کم نمونه‌های موجود جهت آموزش شبکه عصبی، نمی‌توان برای به دست آوردن دقت شبکه، تعدادی از نمونه‌ها را برای آموزش و تعدادی را برای آزمایش سیستم استفاده کرد. به همین منظور برای به دست آوردن دقت شبکه عصبی، در هر دوره آموزش یکی از آنتیبادی‌ها از مجموعه آموزش، حذف و شبکه با بقیه آنتیبادی‌ها آموزش داده شد. پس از آموزش، آنتیبادی حذف شده به شبکه ارائه شد تا شبکه، میزان قدرت اتصال آن را به DNA تعیین کند. در صورت بیشتر بودن میزان قدرت اتصال به DNA از سطح آستان‌های مشخص، آنتیبادی مذکور، از نوع علیه DNA دسته‌بندی شد.

برای بررسی دقت عمل مدل، دو نوع دقت به شرح زیر تعریف گردید:

$$\times 100 = \frac{\text{تعداد آنتیبادی‌های علیه DNA که درست تشخیص داده شده‌اند}}{\text{تعداد کل آنتیبادی‌های علیه DNA}} \text{ دقت در تعیین آنتیبادی‌های علیه DNA (درصد)}$$

$$\times 100 = \frac{\text{تعداد آنتیبادی‌هایی که به DNA متصل نمی‌شوند و درست تشخیص داده شده‌اند}}{\text{تعداد آنتیبادی‌هایی که به DNA متصل نمی‌شوند}} \text{ دقت در تعیین آنتیبادی‌هایی که به DNA متصل نمی‌شوند (درصد)}$$



شکل (۲) شبکه عصبی پرسپترون چند لایه مورد استفاده.

با توجه به این که برای به دست آوردن ساختارهای خاص که باعث علیه DNA بودن آنتیبادی می‌شوند، نیازمند دقتی مناسب در هر دو مورد مذکور هستیم، بجای استفاده از سطح آستانه معمول ۰/۵ در خروجی شبکه عصبی، با تغییر سطح آستانه به دقتی مناسب در هر دو مورد مذکور رسیدیم. در سطح آستانه بهینه، دقت شبکه عصبی در هر دو مورد به 71 ± 1 درصد رسید.

۴- تولید آنتیبادی‌های علیه DNA قویتر و استخراج ساختارهای مؤثر آنها

الگوریتم ژنتیک یک روش جستجوی تصادفی برای بهینه‌سازی سیستم‌های پیچیده است که بر یک نسل داده اعمال می‌شود. در الگوریتم ژنتیک، به هر کدام از پارامترهای مورد بررسی، ژن و به مجموعه پارامترها کروموزوم گویند. هدف الگوریتم ژنتیک، یافتن مقادیری برای هر کدام از پارامترها است که به ازای آنها، تابع شایستگی^{۱۱} بیشینه شود. در الگوریتم ژنتیک، یک نسل کروموزوم اولیه با اعمال عملگرهایی (از جمله دورگه‌سازی^{۱۱} و جهش^{۱۲} ژنها) به یک نسل کروموزوم با تابع شایستگی بیشتر تبدیل می‌شود [۱۵].

در جستجوی ژنتیک برای رسیدن به یک نسل از آنتیبادی‌های علیه DNA قویتر، کروموزومی متشکل از ۱۳۳ ژن (به تعداد اسیدهای آمینه قطعه V_H در یک آنتیبادی) در نظر گرفته شده است. هر ژن می‌تواند یک عدد صحیح بین ۱ تا ۲۰ (به تعداد اسیدهای آمینه) باشد، اگرچه بعضی ژن‌ها به دلایل ساختاری فقط اسیدهای آمینه خاصی را می‌پذیرند. برای حل این مسأله ابتدا برای هر ژن، مجموعه‌های از اسیدهای آمینه مجاز تعیین شد و در جهش ژن‌ها، تنها یکی از اسیدهای آمینه این مجموعه به صورت تصادفی در ژن مربوطه جایگزین شد.

پس از آموزش شبکه عصبی پرسپترون با آنتیبادی‌های موجود، آنتیبادی‌های علیه DNA، به عنوان نسل اولیه انتخاب شدند. از خروجی شبکه عصبی به عنوان تابع شایستگی استفاده شد و با الگوریتم ژنتیک، فرایند به سمت یک نسل از آنتیبادی‌های علیه DNA قویتر هدایت شد. پس از چند نسل حرکت به سمت آنتیبادی‌های علیه DNA قویتر، مشاهده شد که آنتیبادی‌های جدیدی که با الگوریتم ژنتیک به دست آمده‌اند دارای ساختارهای ثابت مشخص تری در رشته پروتئینی خود هستند. با توجه به خاصیت روش‌های جستجوی تکاملی در یافتن ساختارهای موثر در افزایش تابع شایستگی و تکرار آنها در نسل‌های بعد، ظاهراً این ساختارهای مشترک همان ساختارهای مؤثر در آنتیبادی‌های علیه DNA است. برای حصول اطمینان از این مساله، از دو روش زیر استفاده شد:

اول این که با همان الگوریتم ژنتیک قبل و با انتخاب [(خروجی شبکه عصبی)-1] به عنوان تابع شایستگی، آنتیبادی‌هایی که دارای قدرت کم در اتصال به DNA هستند، را ایجاد کنیم. بدیهی است که این نسل‌ها دارای ساختارهای مذکور نیستند. از مقایسه دو نسل به دست آمده، می‌توان ساختارهایی که فقط در آنتیبادهای علیه DNA جهش یافته وجود دارد را به عنوان ساختارهای مؤثر در آنتیبادهای علیه DNA معرفی کرد.

روش دیگر برای حصول اطمینان کامل از صحت ساختارهای مؤثر به دست آمده، حذف آنها از آنتیبادی‌های علیه DNA و جای‌گزینی با تمام حالات ممکن و محاسبه میانگین تغییر خروجی شبکه عصبی آموزش یافته برای آنتیبادی فاقد ساختار خاص است. در صورتی که این تغییر، منفی (که نشان‌دهنده کاهش قدرت اتصال به DNA در آنتیبادی فاقد ساختار خاص است) و قابل ملاحظه باشد آن ساختار مشترک همان ساختار خاصی است که در اتصال به DNA تاثیر دارد.

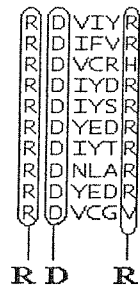
۵- نتایج شبیه‌سازی

در ابتدا دو خروجی الگوریتم ژنتیک مورد استفاده، مربوط به ناحیه CDR3 بررسی شد (شکل ۳). CDR3 ناحیه‌ای از V_H است که در بانک‌های اطلاعاتی، در اسید آمینه ۱۰۴ تا حداکثر اسید آمینه ۱۲۲ قطعه V_H قرار دارد. CDR3 بر اساس مطالعات ایمونولوژی دارای نقش مهمی در اتصال به DNA است [۱۶ و ۱۷] و به همین دلیل در این تحقیق نیز مورد توجه ویژه قرار گرفته است.

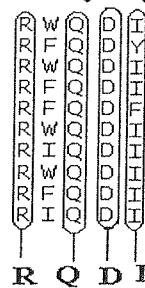
در به شکل ۳ مشاهده می‌شود که اسید آمینه R(Arg) در دو جایگاه ۱۰۴ و ۱۰۹ و چهار اسید آمینه مانده به آخرین جایگاه اسید آمینه در CDR3 در آنتیبادی‌های علیه DNA جهش یافته با فرکانس بالا وجود دارد. همچنین مشاهده می‌شود که این اسید آمینه در آنتیبادی‌هایی که قدرت بسیار ضعیفی در اتصال به DNA دارند در دو جایگاه فوق اصلاً وجود ندارد. این مساله نشان‌دهنده اثر مثبت حضور این اسید آمینه، در جایگاه‌های فوق قطعه V_H یک آنتیبادی، در افزایش قدرت اتصال به DNA است. اثر حضور R(Arg) در CDR3 در افزایش قدرت اتصال به DNA مورد تایید مطالعات ایمونولوژی است [۱۸ و ۱۹]. همچنین حضور اسید آمینه D(Asp) در جایگاه دومین اسید آمینه و یکی مانده به آخرین جایگاه اسید آمینه در CDR3، نقش مثبتی در افزایش قدرت اتصال به DNA بازی می‌کند. اسیدهای آمینه I(Ile) و Q(Gln) در جایگاه‌های انتهایی CDR3 نیز دارای نقش مثبتی در افزایش قدرت اتصال به DNA است.



شش اسید آمینه ابتدایی
در ناحیه CDR3



پنج اسید آمینه انتهایی
در ناحیه CDR3



(a)

ERHRVD
EREVID
ERDRVD
ERQVAD
ERDRGD
ERDSAD
ERDGYD
ELQYYD
ERNRVD

DFFFD
DRFFD
DEFFD
DKMFD
DEMFD
DEMFD
DRFFD
DRFFD
DRFFD

(b)

شکل (۳) خروجی الگوریتم ژنتیک برای قسمتی از قطعه V_H بنام CDR3 در مورد (a) آنتیبادی‌های

با قدرت اتصال زیاد به DNA، (b) آنتیبادی‌های با قدرت اتصال کم به DNA.

حال به همان ترتیبی که قبلاً ذکر شد میتوان میزان تأثیر هر کدام از اسیدهای آمینه مشترک را در اتصال به DNA بدست آورد. حذف R(Arg) در دو جایگاه ۱۰۴ و ۱۰۹ به ترتیب ۱۶ درصد و ۱۷ درصد از قدرت اتصال به DNA را در آنتیبادی‌های علیه DNA کاست. همچنین حذف اسیدآمینه مشترک R(Arg) در آنتیبادی‌های علیه DNA، ۳۵ درصد از قدرت اتصال آنها به DNA را کاست. این نتایج نشان دهنده اثر بسیار زیاد حضور این اسیدآمینه در افزایش قدرت اتصال به DNA است. در اغلب آنتیبادی‌های علیه DNA، دو جایگاه ۱۰۹ و چهار اسیدآمینه مانده به آخرین جایگاه اسیدآمینه در CDR3، در میانه CDR3 واقعند و نقش مثبت حضور R(Arg) در میانه CDR3 مورد تأیید مطالعات ایمونولوژیک است [۱۹].

همان طور که قبلاً ذکر شد، اسیدآمینه D(Asp) در دومین و یکی مانده به آخرین جایگاه در CDR3، در آنتیبادی‌های علیه DNA جهش یافته، به صورت مشترک وجود دارد. با حذف این اسید آمینه در دو جایگاه فوق، در آنتیبادی‌های علیه DNA، به ترتیب ۱۴ درصد و ۳۰ درصد از قدرت اتصال آنها به DNA کاسته شد که نشان دهنده اثر مثبت حضور این اسیدآمینه در دو جایگاه فوق است. در صورت ادامه جستجوی ژنتیک در رسیدن به آنتیبادی‌های علیه DNA قویتر، الگوریتم ژنتیک به توالی خاصی همگرا شد که این توالی قویترین آنتیبادی علیه DNA بدست آمده توسط الگوریتم ژنتیک است. (شکل ۴). حال در تمام اسیدهای آمینه مشترک، در آنتیبادی‌های علیه DNA جهش یافته و در ۱۳۳ جایگاه قطعه V_H ، کاهش فوق را محاسبه می‌نماییم. نتایج برجسته دیگری که در اثر حذف در آنتیبادی‌های علیه DNA، همگی آنان کاهش بیش از ۱۰ درصد در قدرت اتصال به DNA ایجاد کردند، به صورت زیر است:

حضور اسیدآمینه W(Trp) در دو اسیدآمینه مانده به انتهای CDR1.

حضور اسیدآمینه R(Arg) در یازدهمین اسیدآمینه CDR2 و یا هشت اسیدآمینه مانده به انتهای CDR2.

حضور اسیدآمینه R(Arg) در دوازدهمین اسیدآمینه CDR2 و یا هفت اسیدآمینه مانده به انتهای CDR2.

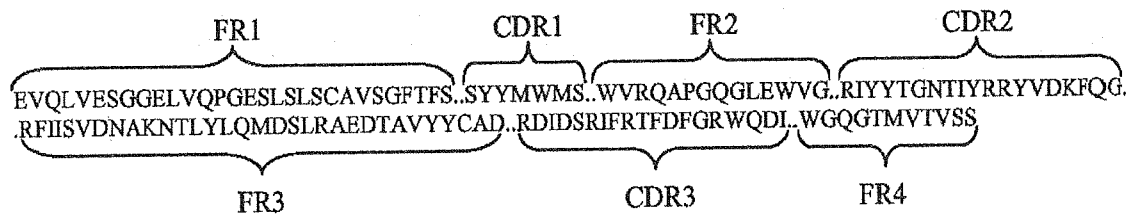
حضور اسیدآمینه D(Asp) در چهار اسیدآمینه مانده به انتهای CDR2.

حضور اسیدآمینه I(Ile) در هفتمین اسیدآمینه CDR3.

حضور اسیدآمینه I(Ile) در آخرین اسیدآمینه CDR3.

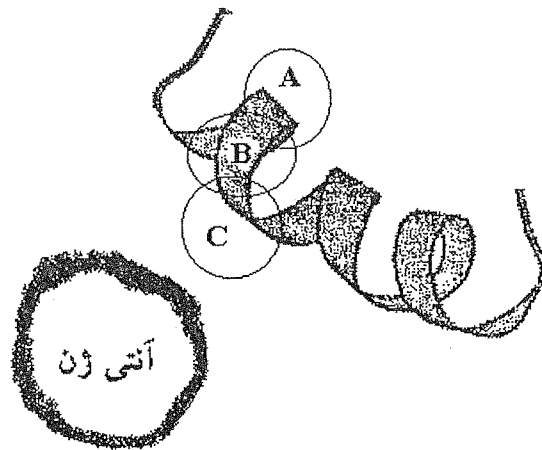
مسئله قابل توجه حضور مؤثر دو اسیدآمینه D و R در جایگاه‌های مختلف در اتصال به DNA است. اسیدهای آمینه D و R

به ترتیب دارای بار منفی و مثبت هستند و DNA دارای بار منفی است. بار منفی DNA، می‌تواند اثر اسیدآمینۀ R را که دارای بار مثبت است توجیه کند ولی اثر اسیدآمینۀ D باید به دلایل ساختاری دیگری باشد که تا کنون مورد توجه قرار نگرفته است. حضور موثر دو اسیدآمینۀ D و R ثابت‌کننده نقش مؤثر بار الکتریکی اسیدهای آمینۀ در اتصال به DNA است. از طرفی تجمع نتایج فوق در سه ناحیۀ CDR، نشان‌دهندۀ اثر بیشتر این نواحی در اتصال به DNA است که این مسأله نیز با روش‌های آزمایشگاهی و کریستالوگرافی به اثبات رسیده است [۲۰ و ۲۱]. همچنین تجمع بیشتر این نتایج در CDR3 نشان از اهمیت بسیار زیاد این ناحیه در اتصال به DNA دارد که تأیید‌کننده نظرات متخصصین این امر است [۱۶ و ۱۷].



شکل (۴) توالی قویترین آنتیبادی علیه DNA که توسط الگوریتم ژنتیک به دست آمده است.

اسیدهای آمینۀ W و I که دو اسید آمینۀ آب‌گریز هستند نیز تاثیر بسزایی در اتصال به DNA بر عهده دارند. مسأله جالب حضور مؤثر اسیدآمینۀ I در هر دو مورد مذکور در نزدیکی اسیدآمینۀ دارای بار الکتریکی است. در ۵۰ درصد موارد، این اسید آمینۀ در کنار اسید آمینۀ دارای بار الکتریکی قرار دارد و در ۴۴ درصد موارد این اسیدآمینۀ با اسیدآمینۀ دارای بار الکتریکی دو اسیدآمینۀ فاصله دارد. این مسأله ممکن است به این دلیل باشد که حضور یک اسیدآمینۀ آب‌گریز باعث افزایش اثر اسیدآمینۀ آب دوست مجاور می‌شود (R و D هر دو آب دوستند). با توجه به شکل CDRها که به صورت فنری^{۱۳} است (شکل ۵)، اسیدآمینۀ آب‌گریز (اسیدآمینۀ A در شکل ۵) باعث قرار گرفتن آن ناحیه در قسمت داخلی (دور از آب) در آنتیبادی می‌شود و این باعث قرار گرفتن اسیدهای آمینۀ مجاور (اسیدهای آمینۀ B و C در شکل ۵) در سطح خارجی آنتیبادی که با آنتیژن برخورد می‌کند، خواهد شد. همان‌طور که در شکل ۵ مشاهده می‌شود اسیدآمینۀ C کاملاً در سطح خارجی CDR3 قرار می‌گیرد و دارای اثر بیشتری در اتصال به آنتیژن خواهد بود. درباره اسیدآمینۀ I، که در هفتمین اسید آمینۀ CDR3 دارای نقش مثبتی در افزایش قدرت اتصال به DNA است، در ۸۵ درصد موارد این اسید آمینۀ با اسیدآمینۀ دارای بار الکتریکی دو اسیدآمینۀ فاصله دارد و اثر اسیدآمینۀ دارای بار الکتریکی به دلیل اسید آمینۀ آب‌گریز مجاور شدیداً افزایش خواهد یافت. با توجه به اینکه این اسیدآمینۀ دارای بار الکتریکی معمولاً در میانه CDR3 است، می‌توان گفت که اسیدهای آمینۀ دارای بار الکتریکی که در میانه CDR3 واقعند، تاثیر به سزایی در اتصال به DNA بر عهده دارند.



شکل (۵) ساختار فنری در پروتئینها، CDRها دارای این ساختار میباشند.

۶- نتیجه گیری

مهمترین نتایج بدست آمده در این تحقیق عبارتند از:

- اهمیت نواحی CDR در اتصال به DNA.
 - تاثیر مثبت اسیدهای آمینه دارای بار منفی در کناره‌های CDR2 و CDR3 در افزایش قدرت اتصال به DNA.
 - تاثیر مثبت اسیدهای آمینه دارای بار مثبت در جایگاه‌های خاص خصوصاً در میانه‌های CDR2 و CDR3 در افزایش قدرت اتصال به DNA.
 - تاثیر مثبت اسیدهای آمینه آب‌گریز در اطراف اسیدهای آمینه دارای بار الکتریکی در افزایش قدرت اتصال به DNA.
- البته به دست آوردن یک توالی با قدرت زیاد در اتصال به DNA، که در مقاله مورد بررسی قرار گرفت، دارای کاربردهای تحقیقاتی دیگری است که در [۲۲] بررسی شده‌اند.
- به طوری که ملاحظه می‌شود قسمتی از این نتایج، مؤید نظرات قبلی و قسمتی نیز بیان کننده ایده‌های نو است که در صورت تمرکز تحقیقات ایمونولوژیک بر این نتایج، می‌توان راه رسیدن به جواب واقعی را کوتاه‌تر کرد.

زیر نویس‌ها

- | | |
|---------------------------------------|---------------------------|
| 1. antibody | 8. hydrophilicity |
| 2. antigen | 9. error back propagation |
| 3. systemic lupus erythematosus (SLE) | 10. fitness function |
| 4. anti-DNA antibody | 11. crossover |
| 5. multilayer perceptron (MLP) | 12. mutation |
| 6. chain | 13. helix |
| 7. region | |

مراجع

- [1] M. H. Foster, B. Cizman and M. P. Madaio. "Nephritogenic autoantibodies in systemic lupus erythematosus: immunochemical properties, mechanisms of immune deposition, and genetic origins," *Lab. Invest.* Vol. 69, pp. 494-507 (1993).
- [2] D. Eilat and Y. Naparstek. "Anti-DNA autoantibodies: a puzzle of autoimmune phenomena." *Immunology today*, Vol. 20, No. 8, pp. 339-342, 1999.
- [3] Abbas, A. K., Lichtman, A. H. and Pober, J. S., *Cellular and molecular immunology*, 4th ed., WB Saunders co., Philadelphia, (2000).
- [4] M. Z. Radic, M. A. Mascelli, J. Erikson, H. Shan and M. Weigert. "Ig H and L chain contributions to autoimmune specificities." *J. Immunol.*, Vol. 146, pp. 176-182 (1991).
- [5] C. I. Mockridge, C. J. Chapman, M. B. Spellerberg, D. A. Isenberg and F. K. Stevenson. "Use of phage surface expression to analyze regions of human V4-34 (VH4-21)- encoded IgG autoantibody required for recognition of DNA: no involvement of the 9G4 idiotope." *J. Immunol.*, Vol. 157, pp. 2449-2454 (1996).
- [6] M. Polymenis and B. D. Stollar. "Critical binding site amino acids of anti-single chain Fv molecules. Role of heavy and light chain CDR3 and relationship Z-DNA to autoantibody activity." *J. Immunol.*, Vol. 152, pp. 5318-5329 (1994).
- [7] S. M. Barbas, H. J. Ditzel, E. M. Salonen, W. P. Yang, G. J. Silverman and D. R. Burton. "Human autoantibody recognition of DNA." *Proc. Natl. Acad. Sci. USA*, Vol. 92, pp. 2529-2533 (1995).
- [8] P. Roben, S. M. Barbas, L. Sandoval, J. M. Lecerf, B. D. Stollar, A. Solomon and G. J. Silverman. "Repertoire cloning of lupus anti-DNA autoantibodies." *J. Clin. Invest.* Vol. 98, pp. 2827-2837 (1996).
- [9] M. A. Rahman, C. A. Kettleborough, D. S. Latchman and D. A. Isenberg. "Properties of whole human IgG molecules produced by the expression of cloned anti-DNA antibody cDNA in mammalian cells." *J. Autoimmun.* Vol. 11, pp. 661-669 (1998).
- [10] <http://www.biochem.ucl.ac.uk/~martin/abs/simkab.html>
- [11] T. P. Hopp and K. R. Woods. "Prediction of protein antigenic determinations from amino acid sequences."

- Proceeding of the National Academy of Sciences of the United States of America*, Vol. 78, pp. 3824-3828 (1981).
- [12] J. Kyte and R. F. Doolittle. "A simple method for displaying the hydrophathic character of a protein." *Journal of Molecular Biology*, Vol. 157, pp. 105-132 (1982).
- [13] J. M. R. Parker, D. Guo and R. S. Hodges. "New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites." *Biochemistry*, Vol. 25, pp. 5425-5432 (1986).
- [14] S. Haykin, *Neural Networks*, 2nd ed., Prentice Hall, New Jersey, (1999).
- D. E. Goldenberg, *Genetic algorithms in search, optimization and machine learning*, Addison Wesley, Reading MA, (1989).
- [15] M. Z. Radic, J. Mackle, J. Erikson, C. Mol, W. F. Anderson and M. Weigert. "Residues that mediate DNA binding of autoimmune antibodies." *J. Immunol.*, Vol. 150, pp. 4966-4977 (1993).
- [16] F. K. Stevenson, C. Longhurst, C. J. Chapman, M. Ehrenstein, M. B. Spellerberg, T. J. Hamblin, C. T. Ravirajan, D. Latchman and D. Isenberg. "Utilization of the VH4-21 gene segment by anti-DNA antibodies from patients with systemic lupus erythematosus." *J. Autoimmun.*, Vol. 6, pp. 809-825 (1993).
- [17] T. H. Winkler, H. Fehr and J. R. Kalden. "Analysis of immunoglobulin variable region genes from human IgG anti-DNA hybridomas." *Eur. J. Immunol.*, Vol. 22, pp. 1719-1728 (1992).
- [18] Z. Li, E. W. Schettino, E. A. Padlan, H. Ikematsu and P. Casali. "Structure-function analysis of a lupus anti-DNA autoantibody: central role of the heavy chain complementarity-determining region 3 Arg in binding of double- and single-stranded DNA." *Eur. J. Immunol.*, Vol. 30, pp. 2015-2026 (2000).
- [19] Y. Pewzner-Jung, T. Simon and D. Eilat. "Structural elements controlling anti-DNA antibody affinity and their relationship to anti-phosphorylcholine activity." *J. Immunol.*, Vol. 156, pp. 3065-3073 (1996).
- [20] E. Eilat, M. Dayan, H. Zinger and E. Mozes. "The mechanism by which a peptide based on complementarity-determining region-1 of a pathogenic anti-DNA auto-Ab ameliorates experimental systemic lupus erythematosus." *Proc. Natl. Acad. Sci. USA* 98: 1148-1153 (2001).
- [21] M. Laura Cerutti, Juan M. Centeno, Fernando A. Goldbaum, and Gonzalo de Prat-Gay. "Generation of Sequence-specific, High Affinity Anti-DNA Antibodies." *J. of Biological Chemistry*, Vol. 276, No. 16, Issue of April 20, pp. 12769-12773 (2001).

