

بررسی توالی و ساختار مرکزی در ساختمان الکل دهیدروژنازها

شمیلا درویش علیپور آستانه^۱، زرین مینوچهر^{۲*}، آرمین مددکار سبحانی^۳، مهران میراویلیایی^۴

تاریخ دریافت: ۹۳/۰۷/۲۸

تاریخ تصویب: ۹۵/۱۲/۱۱

چکیده

با هدف بررسی و شناسایی اسیدهای آمینه مهم از نظر ساختاری و عملکردی در خانواده الکل دهیدروژنازها که در طول تکامل محافظت شده‌اند، تجزیه و تحلیل کامپیوتری پیشرفته‌ای بر روی توالی و ساختار سه بعدی این آنزیم‌ها انجام شد. آنالیز توالی با توجه به برخی خصوصیات اسیدهای آمینه از جمله همسانی، تشابه، انتروپی، هیدروفوبیسیته برای تعیین سطح محافظت شدگی در هر جایگاه انجام گرفت. بر اساس محاسبات انجام شده، مناطق بسیار محافظت شده‌ای در ساختار و توالی الکل دهیدروژنازها شناسایی شدند. با آنالیز بیشتر توالی‌ها و جایگاه‌های محافظت شده، معلوم شد که در این مناطق، خصوصیات مشترکی مشاهده می‌شود. نتایج محاسبات در این مطالعه نشان داد که این امکان وجود دارد که در جایگاه‌های ویژه‌ای از ساختار این آنزیم، الگوهای محافظت شده خاصی تعریف شود و بر آن اساس، پایداری ساختاری آنزیم که عامل غلبگی آن در طول تکامل است را توجیه کند. در این مطالعه دو منطقه کاملاً محافظت شده که الگوی مشخصی را در توالی خود در طول تکامل حفظ کرده‌اند، مشخص شدند. به نظر می‌رسد این موتیف‌ها در تغییر آرایش

۱- گروه پزشکی مولکولی، پژوهشگاه ملی مهندسی ژنتیک و زیست فناوری، تهران، ایران

*۲- دانشیار، گروه زیست فناوری سامانه‌ای، پژوهشگاه ملی مهندسی ژنتیک و زیست فناوری، تهران، ایران
(نویسنده مسئول: minucheher@nigeb.ac.ir)

۳- دانشیار، گروه بیوانفورماتیک، مرکز تحقیقات بیوشیمی بیوفیزیک، دانشگاه تهران، تهران، ایران

۴- دانشیار، گروه زیست شناسی، دانشکده علوم، دانشگاه اصفهان، اصفهان، ایران

فضایی آنزیم بعد از اتصال به سوبسترا یک نقش مکانیکی به عنوان لولا داشته باشند.

واژه‌های کلیدی: ساختار مرکزی، ADHs، تجزیه و تحلیل چند متغیره

مقدمه

دارد. این آنزیم‌ها در ژنوم خود دارای تشابه بالاتر از ۲۰ درصد هستند و تنوع زیادی را در ساختمان و نیز مکانیسم‌های عملکردی خود نشان می‌دهند (Julian-Sanchez et al., 2003).

ساختار بسیاری از الکل دهیدروژنازها با روش‌های کریستالوگرافی X-ray و یا NMR مشخص شده است. شناسایی و بررسی این ساختارها با توجه به عملکرد آنها می‌تواند اطلاعات زیادی در ارتباط با واحدهای عملکردی و جایگاه‌های کلیدی در مکانیسم و ساختار الکل دهیدروژنازها در اختیار ما بگذارد.

در حوزه پژوهش‌های بیوانفورماتیکی، بررسی‌های مقایسه‌ای و آنالیزهای کامپیوتری ابزار قدرتمندی در زمینه شناسایی ژنوم و ساختار پروتئین‌ها می‌باشند. برای شکل‌گیری یک پروتئین با یک ساختار سه بعدی پایدار، همه اسیدهای آمینه نقش اختصاصی دارند، اما اهمیت آنها ممکن است متفاوت باشد. الگویی که اسیدهای آمینه بر طبق آن در یک توالی حضور می‌یابند رویدادی است که با گذشت زمان حاصل شده و محصول فشار تکاملی، جهش، نو ترکیبی،

الکل دهیدروژناز، گلیکوپروتئینی حاوی نقش‌های بسیار گسترده در طبیعت است. الکل دهیدروژنازها به طور وسیعی در طبیعت وجود دارند و در بسیاری از واکنش‌های ضروری برای حیات، ایفای نقش می‌کنند. در حضور NAD، الکل دهیدروژنازها سوبسترای الکی را به آلدئید تبدیل می‌کنند (Eklund et al., 1976). این آنزیم‌ها بر اساس اندازه مولکول و نیاز به کوفاکتور در سه طبقه اصلی تقسیم‌بندی می‌شوند. سومین گروه الکل دهیدروژنازها که در این مطالعه مورد توجه قرار گرفته‌اند گروهی از آنزیم‌های وابسته به روی (Zn) هستند که دهیدروژنازهای با زنجیر متوسط (MDRs) نام دارند. این گروه از الکل دهیدروژنازها به طور متوسط از ۳۷۵ اسید آمینه تشکیل شده و به طور متوسط دارای وزن مولکولی ۴۰۰۰ کیلو دالتون هستند. آنزیم‌ها در این گروه غالباً دارای دو اتم روی در هر زیر واحد هستند که یکی از آنها در کاتالیز نقش داشته و دیگری نقش ساختاری

الگوهای محافظت شده در ساختمان پروتئین کمک کرده و جایگاه‌های کلیدی در پایداری، عملکرد و فولدینگ پروتئین را مشخص کند (Grant, et al., 2007). مطالعات گسترده در سطح ساختار پروتئین‌ها نشان داده است که اعضای یک خانواده پروتئینی که عملکرد مشابه دارند، حتی در صورت پایین بودن میزان تشابه در سطح توالی، می‌توانند دارای ساختارهای سه بعدی محافظت شده باشند. (Chothia and Lesk, 1986, Gerstein and Altman, 1995).

بر اساس مطالعات و بررسی‌های دقیق ساختمانی می‌توان مجموعه‌ای از اتم‌ها را در ساختار پروتئین‌ها شناسایی نمود که از نظر فضایی، ثابت‌ترین جایگاه‌های ساختاری در تمامی اعضای یک خانواده را تشکیل داده و از نظر ژئومتریکی در طول دوره تکامل محافظت شده باقی مانده‌اند. این مجموعه را می‌توان با عنوان ساختار مرکزی^۱ در پروتئین تعریف نمود (Gerstein and Altman, 1995).

بررسی ساختار مرکزی در پروتئین‌ها، اولین بار برای مطالعه محافظت شدگی و تغییرات ساختاری آلوستریکی در ایمینوگلوبین‌ها مورد مطالعه قرار گرفت (Gelfand et al., 1998). پس از آن در مطالعه‌ای، تاثیر تغییرات اسیدآمینیه روی ساختار سه بعدی پروتئین در ۸ خانواده

و دریفتهای^۱ ژنتیکی است (Valdar, 2002, Hayward, 2004). شناسایی اسیدهای آمینه کلیدی و واحدهای عملکردی مهم که تحت فشار انتخاب طبیعی در طول تکامل محافظت شده باقی مانده‌اند؛ برای شناسایی واحدهای عملکردی پروتئین ابزار بسیار قدرتمندی محسوب می‌شوند (Grant et al., 2007).

اهمیت اسیدهای آمینه در حفظ پایداری و عملکرد پروتئین، معمولاً از میزان محافظت شدگی آنها برآورد می‌شود. یافتن مقیاس معتبر برای کمی نمودن سطح محافظت شدگی تاکنون بسیار مورد بررسی و توجه قرار گرفته و روشهای زیادی جهت تحلیل محافظت شدگی استفاده شده‌اند. برای نمونه می‌توان روشهای مبتنی بر تعداد تکرارها^۲ (Chothia et al., 1998)، روشهای مبتنی بر خصوصیات فیزیکی-شیمی، روشهای مبتنی بر بررسی انتروپی^۳ (Schneider, 1997) و روشهای مبتنی بر بررسی جهش (Valdar, 2002) را ذکر کرد. با به کار بردن مجموع این روش‌ها، انتظار می‌رود که به روش منسجم‌تری برای تحلیل محافظت شدگی دست یابیم (Grant et al., 2007).

در کنار بررسی توالی‌ها، مقایسه ساختارهای سه بعدی پروتئین‌های همولوگ نیز می‌تواند به شناسایی

- 1- Drift
- 2- Frequency
- 3- Entropy

و نیز میزان تماس با حلال نیز بدست می‌دهد که به تنهایی با مطالعه و آنالیز توالی‌ها قابل استحصال نیست (Gerstein and Altman, 1995., Cuff et al., 2013). بنابراین مقایسه و آنالیز همزمان توالی و ساختار سه بعدی در پروتئین‌ها کمک می‌کند که درک بهتری از این واقعیت داشته باشیم که چرا اسیدهای آمینه خاصی در مکان‌های خاصی دیده می‌شوند و این اسید آمینه‌ها چه نقشی در عملکرد یا پایداری این پروتئین‌ها دارا می‌باشند (Yao et al., 2016).

هدف ما در این مطالعه مقایسه توالی‌ها و ساختارهای MD-الکل دهیدروژنازهایی با تاریخچه تکاملی متفاوت با هدف شناسایی جایگاه‌هایی با کمترین میزان تغییرپذیری^۲ و در ادامه بررسی بیشتر این جایگاه‌ها برای شناسایی خصوصیات فیزیکی و شیمیایی آنهاست. نتایج این مطالعه، استنباط روشنی از روابط ساختار-توالی در مناطق محافظت شده ارائه می‌کند و جایگاه‌های مهم و موثر در پایداری ساختاری و خصوصیت دینامیک آنزیم را مشخص می‌کند.

مواد و روشها

۱- هم‌ردیفی توالی‌ها

برای گردآوری توالی الکل دهیدروژنازهایی که ساختار سه بعدی آنها تعریف شده باشد، پایگاه‌های اطلاعاتی مربوط به

پروتئینی بررسی شد و نشان داده شد که در پروتئین‌های هم خانواده که عملکرد مشابه دارند حتی در مناطقی از پروتئین که میزان تشابه در سطح توالی زیر ۱/۵% است، از نظر ساختار فضایی کاملاً محافظت شده باقی مانده‌اند (Chothia and Lesk, 1986). بررسی ساختار مرکزی در مطالعات زیادی برای شناسایی ساختارهای سه بعدی محافظت شده (Hubbard and Thornton, 1993; Berman et al., 2000, Mirny and Shakhnovich, 2001)، یافتن زمین‌های عملکردی (Gerstein and Krebs 1998, Chew et al., 1999)، مطالعه تغییرات آلوستریک (Yao et al., 2016) و نیز طبقه بندی پروتئین‌ها (Cuff et al., 2013) استفاده شده است. برای بررسی و جستجوی ساختار مرکزی تاکنون چندین ابزار بیوانفورماتیکی معرفی شده‌اند (Grant et al., 2006, Bottini et al. (2013; Vakulenko et al., 2016 که غالباً براساس منطبق نمودن ساختارهای سه بعدی در دفعات متوالی و بررسی تغییرات RMSD در کربن‌های آلفا عمل می‌کنند (Gerstein and Altman, 1995).

مطالعه ساختار سه بعدی پروتئین‌ها غیر از شناسایی ساختار مرکزی، اطلاعات مهم دیگری از جمله وضعیت تماس^۱ اسید آمینه‌ها در ساختار سه بعدی، وضعیت ژئومتریک اسیدهای آمینه

بر اساس همسانی، در هر موقعیت تنوع اسیدهای آمینه بررسی شد و فرکانس حضور برای تمام اسیدهای آمینه محاسبه گردید. اسید آمینه‌ای که دارای بالاترین فرکانس حضور در هر موقعیت بود به عنوان اسید آمینه شایع در آن موقعیت انتخاب شد. مکان‌هایی که هیچکدام از اسیدهای آمینه فرکانس بالای ۶۰ درصد نداشتند، به عنوان Gap در نظر گرفته شد. در بررسی درجه تشابه، با توجه به صف‌آرایی انجام شده، بر اساس ماتریکس BLOSUM 62 (Henikoff and Henikoff, 1992) امتیازدهی انجام شد

$$S = - \sum_i^N P_i \log_2 P_i$$

برای محاسبه انتروپی، اسیدهای آمینه بر اساس خصوصیات فیزیکی - شیمی به ۶ گروه تقسیم شدند (جدول ۱) و انتروپی بر اساس انتروپی شانون محاسبه شد (Shenkin et al., 1991). بر اساس این طبقه‌بندی، جهش‌های داخل یک گروه برای مثال ($V \rightarrow L$) در نظر گرفته نمی‌شدند و تنها جهش‌هایی که منجر به تغییر گروه اسید آمینه مربوط به آن جایگاه می‌شد، در محاسبات در نظر گرفته شدند (Mirny and Shakhnovich, 2007, Grant et al., 2001).

Pi = Amino acid frequency, i = Amino

توالی‌ها (UniProt) version 2013-08 released on 2013-08-01 (Bairoch and Apweiler 2000) و ساختار سه بعدی پروتئینها (PDB) (Berman et al., 2000) جستجو شدند. در نهایت ۲۲ ساختار متعلق به چهار ابرخانواده^۲ انتخاب شدند. توالی مربوط به این ساختارها استخراج شد و توالی‌های با بیش از ۹۰% تشابه^۳ حذف شدند. در نهایت ۱۵ ساختار و توالی مربوط به آنها برای آنالیز انتخاب شدند. توالی‌ها ابتدا با ClustalW (Thompson et al., 1994).

در نرم افزار (Mega) Tamura et al. 2007 هم‌ردیف^۴ شدند. برای اطمینان از میزان صحت این هم‌ردیفی، یک مدل مخفی مارکوف^۵ (Eddy, 1998)، از هم‌ردیفی توالی‌ها ساخته شد. توالی‌های انتخابی با استفاده از این مدل، مجدداً هم‌ردیف شدند.

۲- آنالیز محافظت شدگی

برای بررسی سطح محافظت شدگی، در هر جایگاه در یک هم‌ردیفی، همسانی^۶، تشابه^۷، انتروپی و هیدروفوبیسیته در هر موقعیت محاسبه شد.

برای محاسبه میزان محافظت شدگی

-
- 1- Protein data bank
 - 2- Super family
 - 3- Identity
 - 4- align
 - 5- hidden Markov model (HMM)
 - 6- Identity
 - 7- similarity

acid position

gap 1 + Groups 6) 7 = N

پروتئین قرار می گیرند. گروه دوم شامل اسیدهای آمینه هیدروفیل است (R, K, E, D, Q, N) که این گروه به احتمال بالا در سطح پروتئین یافت می شوند و گروه سوم اسیدهای آمینه خنثی هستند (P, H, Y, G, A, S, T) که دارای شانسی برابر، برای حضور در سطح یا داخل پروتئین می باشند.

در مورد هیدروفوبیسیته، اسیدهای آمینه در سه گروه بر اساس هیدروفوبیسیته نسبی خود قرار گرفتند (جدول ۲). گروه اول شامل اسیدهای آمینه هیدروفوب (C, V, L, I, M, F, W) است که به خاطر هیدروفوبیسیته بالا به احتمال بسیار زیاد در داخل (مرکز)

جدول ۱: اسیدهای آمینه بر اساس خصوصیات فیزیک- شیمی به شش گروه تقسیم شدند.

۱	A, V, L, I, M, C	Aliphatic
۲	F, W, Y, H	Aromatic
۳	S, T, N, Q	Polar
۴	K, R	Basic
۵	D, E	Acidic
۶	G, P	Special

جدول ۲: براساس هیدروفوبیسیته نسبی، اسیدهای آمینه به سه گروه تقسیم شدند.

Class ۱	C, V, L, I, M, F, W (Hydrophobe)
Class ۲	R, K, E, D, Q, N (Hydrophile)
Class ۳	P, H, Y, G, A, S, T (Neutral)

مورد تحلیل قرار گرفت. ابتدا میزان RMSD بین اتمهای کربن آلفای متناظر در تمام این ساختمانها محاسبه شد. سپس مجموعه ساختار مرکزی^۳ به عنوان مجموعه ای از اتمها که دارای کمترین حد تغییرات ساختاری در بین مجموع اتمها بودند، با روش آلتمن^۴ و گریستن (Gerstein and Altman, 1995) شناسایی شدند. در این مجموعه، میزان تماس با حلال^۵ با نرم افزار Hubbard and NACCESS (Thornton, 1993) و نیز تعداد تماسها^۶ با

چنانچه در یک جایگاه امتیاز مربوط به هر کدام از موارد بررسی شده از حد آستانه ۰/۶ بیشتر بود آن جایگاه به عنوان جایگاه محافظت شده در نظر گرفته می شود.

۳- بررسی ساختار

برای مقایسه ساختارهای الکل دهیدروژناز، ساختارهایی که با روش X-ray بدست آمده بودند و دارای قدرت تفکیک^۱ بهتر از ۲/۵ Å بودند، انتخاب شدند. ساختارها ابتدا با نرم افزار موستانگ^۲ (Konagurthu et al., 2006) به روی هم منطبق شدند و سپس با نرم افزار Grant et al., 2006 (Bio3d) انطباق حاصل

3- Core Structure

4-Altman

5-Solvent exposure

6-Contact number

1- resolution

2- Mustang

نتایج Map (Chung et al., 2007) Con-Struct و Bio3d

محاسبه شد. در بررسی میزان تماس با حلال به طور متوسط ۴۰٪ دسترسی به عنوان آستانه در نظر گرفته شد و در محاسبه تعداد تماس‌ها، چنانچه فاصله بین هر کدام از اتم‌های سنگین در دو اسید آمینه کمتر از ۵ انگستروم بود در تماس با هم در نظر گرفته شدند.

بر اساس محاسبات انجام شده، مجموعاً ۸۷ جایگاه محافظت شده از بین ۳۷۴ جایگاه مورد بررسی شناسایی شدند که ۲۷ درصد کل منطقه را شامل می‌شدند. در این مجموعه ۳۴٫۹۶٪ اسیدهای آمینه هیدروفوب، ۱۹٫۳٪ آنها هیدروفیل و ۴۴٪ جزو اسیدهای آمینه خنثی بودند. ۷۳٪ این جایگاه‌ها در ساختار ثانویه واقع شده بودند (جدول ۳).

جدول ۳: جایگاه‌های محافظت شده در سطح آنالیز توالی

<i>Secondary Structure</i>	<i>Number of Positions</i>	
<i>Sheet</i>	۳۵	۷۳٪
<i>Helix</i>	۲۲	
<i>Turn</i>	۱۳	
<i>Loop</i>	۲۶	۲۷٪
<i>Total</i>	۸۷	۱۰۰٪

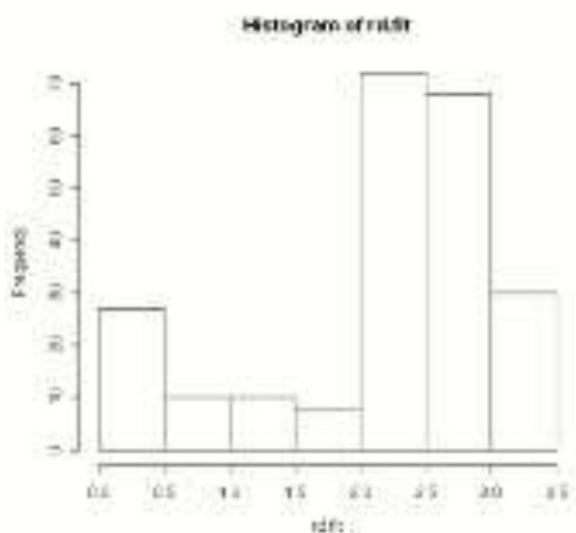
آنالیز ساختاری
منطبق شدند (شکل ۱) و انطباق حاصل برای آنالیز ساختاری، ساختارهای سه بعدی الکل دهیدروژنازها به روی هم



شکل ۱: انطباق ساختمانی ۱۵ ساختار کریستالوگرافی الکل دهیدروژناز. این انطباق با نرم افزار موستانگ انجام شد و توسط Bio3d مورد آنالیز قرار گرفت.

بر اساس انطباق انجام شده، میزان RMSD بین اتم‌های کربن آلفای متناظر در تمام این ساختمانها محاسبه شد که در تمام موارد بین $2/5^{\circ}A$ - ۲ انگستروم بود. (۱)

همچنین بیشترین میزان Pairwise RMSD در ساختارها، بین ۲ الی ۳ انگستروم با متوسط حدود ۲/۱۲ انگستروم بود (نمودار ۱).



نمودار ۱: توزیع متوسط pairwise RMSD در ۱۵ ساختار ADH مورد مطالعه

با بررسی انطباق انجام شده و روش آلتمن و گریستین، ساختار مرکزی در این ساختمانها شناسایی شدند. این مجموعه شامل ۲۷ جایگاه با واریانس ساختمانی کمتر از ۰/۵ انگستروم بودند که در چهار گروه طبقه بندی شدند (Core A-D) (جدول ۴).

تمامی اسیدهای آمینه شرکت کننده در ساختار مرکزی در همین کاتالیتیک آنزیم قرار گرفته بودند. بررسی بیشتر اسیدهای آمینه شرکت کننده در این مجموع نشان داد که ۵۰٪ از اسیدهای آمینه در منطقه ساختار مرکزی از اسیدهای آمینه خنثی و ۳۳٪ هیدروفوب بودند. اسیدهای آمینه در این مجموعه، بیشترین تعداد تماس با سایر جایگاهها و کمترین میزان سطح تماس با محلول (به طور متوسط کمتر از ۴۰٪) را دارا بودند. به عبارتی تمام اسیدهای آمینه شرکت کننده در ساختار مرکزی در بطن پروتئین مدفون گشته اند. بررسی مجموعه ساختارهای مرکزی و مناطق محافظت شده در سطح توالی در ارتباط بایکدیگر

مقایسه نتایج بدست آمده از آنالیز توالی و آنالیز ساختاری نشان داد که جایگاههای محافظت شده بدست آمده از آنالیز در این دو سطح، دارای ۶۰٪ همپوشانی هستند. به عبارتی ۶۰٪ از جایگاههای شرکت کننده در ساختار مرکزی، در سطح توالی نیز محافظت شده بودند.

بررسی مجموعه اسیدهای آمینه شرکت کننده در ساختار مرکزی با توجه به وضعیت آنها در محاسبات آنالیز توالی، نتایج زیر را نشان داد: ساختار مرکزی یا هسته A (Core A) شامل اسیدهای آمینه ۳۴ الی ۴۴ است. همگی اسیدهای آمینه این منطقه در سطح توالی نیز بسیار محافظت شده بودند. از نظر فضایی این مجموعه در اطراف Cys-46 واقع شده است. در سطح ساختار فضایی، اسیدهای آمینه این منطقه در تماس با اسیدهای آمینه مربوط به هسته B (Core B) و هسته D (Core D) مشاهده شدند.

جدول ۴: جایگاه‌های ساختار مرکزی و خصوصیات آنها

	start	end	length	Average Contact Number	Average Solution. Exposure	Hydrophobicity
Core A	۳۴	۴۰	۷	۱۰,۶	۷,۴۰٪	۱,۴
Core B	۷۰	۷۷	۸*	۹,۴	۲۴,۰۶٪	۲
Core C	۸۵	۹۰	۶	۹,۳	۱۲,۵۰٪	۱,۴
Core D	۱۴۸	۱۵۳	۶	۱۱	۱۷٪	۳

*Include Zn-binding signature(66-80) with pattern

(GHEXXGXXXXXGXXV)(Atteia, van Lis et al.)

هسته B (Core B) شامل اسیدهای آمینه ۷۰ الی ۷۷ است. این منطقه مطابق با الگوی پروسایت (De Castro et al., 2006) به قرار گرفته است.

هسته D شامل اسیدهای آمینه ۱۴۸ تا ۱۵۳ می‌باشد. از شش اسید آمینه شرکت کننده در تشکیل این هسته، تنها دو اسید آمینه در سطح توالی محافظت شده بودند، حضور اسیدهای آمینه در این منطقه از الگوی زیر پیروی می‌کرد:

[Vlma]100[amv]53[K]60[la]100[36]46

برای بیان نحوه حضور اسیدهای آمینه از روش پروسایت استفاده شد. جایگاه‌هایی که در ۲۵٪ توالی‌ها هیچ اسید آمینه هسته A و C هستند. نام Zn-binding signature در الکل دهیدروژنازها می‌باشد. اسیدهای آمینه در ساختار مرکزی His-67، B، را احاطه می‌کنند و در تماس با اسیدهای آمینه در هسته A و C هستند. هسته C (Core C)، شامل اسیدهای آمینه ۸۴ الی ۹۰ است که همگی آنها در سطح توالی نیز بسیار محافظت شده‌اند. هسته C در اطراف اتم روی ساختاری، که توسط چهار اسید آمینه سیستمین در موقعیت‌های

۱۶۵ را شامل می‌شود که در سطح توالی بسیار محافظت شده بوده و دارای الگوی زیر می‌باشد:

[Lfp]86E60[X][va]80[CA]100[33]73[37]10
0[g]46[C]66[ga]73

جایگاه محافظت شده بعدی، در موقعیت ۱۶۰-۱۵۶ مشاهده شد. این موتیف که β نامیده شد، از الگوی زیر پیروی می‌کند:

[Vlma]100[amv]73[K]60[lal]100[36]86

سطح دسترسی به حلال در این دو موتیف مورد محاسبه قرار گرفت که در هر دو به طور متوسط زیر حد آستانه (۴۰ درصد) بود.

واحدی در آن جاحضور نداشت با X نشان داده شد. اعداد زیرنویس درصد تعداد تکرار را برای اسیدهای آمینه در هر جایگاه نشان می‌دهد.

براساس نتایج بدست آمده از بررسی محافظت شدگی در سطح توالی نیز دو منطقه بسیار محافظت شده در ساختار الکل دهیدروژناز شناسایی شدند. این مناطق در ساختار هسته مرکزی حضور نداشتند اما حضور اسیدهای آمینه در این مناطق از یک الگوی مشخص پیروی می‌کرد. این مناطق با نام های موتیف های α و β نامگذاری شدند (جدول ۵). موتیف α اسیدهای آمینه در جایگاه ۱۷۴-

جدول ۵: دو منطقه بسیار محافظت شده به نام موتیف های α و β در ساختار الکل دهیدروژناز.

Positions	Conserved Pattern	Motif
۱۵۶-۱۶۰	[Vlma] _{۱۰۰} [amv] _{۷۳} [K] _{۶۰} [lal] _{۱۰۰} [۳۶] _{۸۶}	Motif α
۱۶۵-۱۷۴	[[Lfp] _{۸۶} E _{۶۰} [X][va] _{۸۰} [CA] _{۱۰۰} [۳۳] _{۷۳} [۳۷] _{۱۰۰} [g] _{۴۶} [C] _{۶۶} [ga] _{۷۳}	Motif β

توسط Eklund، روی آنزیم الکل دهیدروژناز اسب انجام شد (Eklund et al, 1976). پس از آن مطالعات زیادی با هدف شناسایی نقش و اهمیت اسیدهای آمینه در عملکرد و نیز مطالعه خصوصیات ژئومتریکی آنها در تغییرات ساختاری آلوستریکی آنزیم‌های ADH در گونه های متفاوت انجام شد. (Eklund et al., 1984; Jornvall et al., 1984; Chen et al., 1991; Ryde,

بحث در این مطالعه تغییرات موجود در توالی و ساختار برای جستجوی روابط بین ساختار و عملکرد و نیز یافتن اسیدهای آمینه و تعاملاتی که نقش اساسی در محافظت شدگی ساختاری و عملکردی در این دسته از الکل دهیدروژنازها دارند، مورد بررسی قرار گرفتند. کشف ساختار سه بعدی آنزیم الکل دهیدروژناز در سال ۱۹۷۶ و

مشخص شد که تمامی اسیدهای آمینه شرکت کننده در این مجموعه مربوط به منطقه ای هستند که در مطالعات گذشته به عنوان دمین کاتالیتیک آنزیم شناسایی شده است (Eklund et al., 1981; Esposito et al., 2002).

با توجه به این نتایج به نظر می رسد که اسیدهای آمینه در ساختار مرکزی، مسئول نگهداری تمامیت ساختاری پروتئین بوده و در ساختمان جایگاه فعال آنزیم شرکت می کنند. حضور قسمت های ثابت از نظر ساختاری (ساختار مرکزی) در دمین کاتالیتیک آنزیم موجب کاهش تاثیر نوسانات دمایی و یا سایر شرایط نامطلوب که روی اتصال به سوبسترا و عملکرد آنزیم تاثیر دارند، می شود و کمک می کند که در شرایط مختلف، ساختار آنزیم در آرایش مناسبی برای عملکرد قرار داشته باشد.

نتایج مربوط به مناطق محافظت شده در این مطالعه تا حدود زیادی توسط نتایج گزارش شده از مطالعات قبلی تایید می شود.

منطقه ای که به عنوان ساختار مرکزی A در این مطالعه شناسایی شد، از نظر فضایی در اطراف سیتئین ۶۶ قرار گرفته است. بر اساس گزارشات قبلی Cys-۶۶ همراه با Cys-۱۷۴ و His-۶۷ مسئول نگهداری اتم Zn کاتالیتیک می باشند (Eklund et al., 1981). با توجه به اهمیت این گروه از

1995; Ramaswamy et al., 1999; Bogin et al., 2002; Esposito et al., 2002; Kim et al., 2004). اکثر این مطالعات با تکیه بر مطالعات جهش زایی و روی یک الکل دهیدروژناز خاص انجام شده است (LeBrun et al., 2004; Goihberg et al., 2007).

در این مطالعه، برای اولین بار توالی و ساختار به طور همزمان و در ارتباط با هم به کمک ابزارهای بیوانفورماتیک مورد آنالیز دقیق و همه جانبه قرار گرفته و جایگاه های محافظت شده و نیز مجموعه ای از اتم ها با کمترین میزان تغییر پذیری به عنوان ساختار مرکزی در الکل دهیدروژنازهای گروه MD شناسایی شدند.

در این مطالعه آنزیم های الکل دهیدروژناز متعلق به موجودات زنده از سطوح متفاوت تکاملی برای بررسی انتخاب شدند. ابتدا میزان محافظت شدگی در سطح توالی بررسی شد و برای جستجوی ساختار مرکزی، مطالعه ساختمانی در آنها انجام گرفت. سپس نتایج حاصله از آنالیز ساختاری با توجه به نتایج بدست آمده از مطالعه توالی ها، مورد بررسی مجدد قرار گرفت. نتایج بدست آمده نشان داد که بین جایگاه هایی که از نظر توالی محافظت شده باقی مانده بودند و جایگاه های شرکت کننده در ساختار مرکزی، ۶۰ درصد همپوشانی وجود دارد.

با بررسی بیشتر مجموعه ساختار مرکزی،

در این منطقه در حفظ ساختار آنزیم نقش عمده ایفا می‌کند. همچنین این منطقه در نگهداری و حفظ ساختار چهارم آنزیم نیز مهم است (Goihberg et al., 2007).

موتیف‌های محافظت شده آلفا و بتا مناطق دیگری هستند که با الگوی مشخصی در تمام آنزیم‌های بررسی شده در این مطالعه، حضور داشتند. این مناطق در مطالعات قبلی گزارش نشده‌اند. موتیف آلفا در C- ترمینال دمین کاتالیتیک آنزیم قرار دارد و شامل هلیکس شماره ۲ است که توسط یک پیچ^۱ در اسید آمینه شماره ۱۷۴ به هلیکس شماره ۵ متصل می‌شود. بر

اساس مطالعات گذشته گزارش شده است که دو منطقه انتهایی در دو هلیکس مجاور هم، می‌توانند تشکیل یک لولا بدهند. این نوع ساختار لولایی double-hinged α -helix گفته می‌شود (Hayward, 1999); بنابراین این موتیف محافظت شده در الکل دهیدروناز که دو منطقه انتهایی از هلیکس‌های ۲ و ۵ را در جایگاه ۱۷۴ شامل می‌شود، ممکن است در تغییرات آرایش فضایی آنزیم، بعد از اتصال به سوبسترا نقش داشته باشد و در این روند به عنوان لولا عمل کند. اسید آمینه ۱۷۳ نیز در این موتیف قرار گرفته است. این اسید آمینه بسیار مهم، که در نگهداری اتم روی کاتالیتیک شرکت می‌کند (Eklund and Nordstrm, 1976);

اسید های آمینه، محافظت شده بودن این منطقه از نظر فضایی، قابل توجیه است. منطقه مربوط به توالی پروسایت مربوط به الکل دهیدرونازهای گروه MD در اسیدهای آمینه ۶۶ الی ۸۰ که تحت عنوان Zn-binding signature شناخته می‌شود و دارای الگوی محافظت شدگی G H E X X G X X X X G X X V است (De Castro et al., 2006); در این مطالعه نیز به عنوان یک منطقه کاملاً محافظت شده هم از نظر توالی و هم از نظر ساختار سه بعدی شناخته شده و به عنوان هسته B معرفی شد.

اسیدهای آمینه گلوتامین ۶۸ و هیستیدین ۶۷ که در مطالعات قبلی به عنوان دو اسید آمینه کلیدی در عملکرد الکل دهیدرونازها شناسایی و گزارش شدند (Ryde, 1995) نیز در این منطقه محافظت شده قرار دارند. هسته C نیز به عنوان منطقه دیگری از ساختار مرکزی در این مطالعه شناسایی شده است. در مطالعات قبلی این منطقه به عنوان یک واحد عملکردی مهم و محافظت شده از نظر ساختاری در الکل دهیدرونازها گزارش شده بود.

(Korkhin et al., 1998; Esposito et al., 2003; Guy et al., 2002). هسته C به صورت یک لوپ جلوآمده در ساختار آنزیم دیده می‌شود. بر اساس گزارشات قبلی، در مواقعی که دسترسی به اتم روی محدود است، تشکیل یک باندهی سولفید

اعضای یک خانواده، به شناسایی الگوها در خصوصیات توالی و ساختاری و نیز تعاملاتی که برای پایداری و عملکرد پروتئین مهم هستند کمک می‌کند، بنابراین بر اساس نتایج این گونه مطالعات می‌توان الگوهای برای سنتز آنزیم‌های جدید و یا بهبود ساختاری آنزیم‌ها طراحی کرد که دارای پایداری بیشتر در شرایط خاص و نیز عملکرد بهتر باشند. روند صورت گرفته در این مطالعه و این قبیل مطالعات می‌تواند برای سایر گروه‌های مهم پروتئینی نیز صورت گیرد.

سیاسگزاری

این مطالعه در پژوهشگاه ملی مهندسی ژنتیک و در راستای طرح پژوهشی ۳۴۶ صورت گرفته است، در اینجا از سرکار خانم دکتر شکیبا درویش‌علیپور که مقاله را بدقت مطالعه نموده و نظرات ارزشمندی در بهبود مقاله ارائه نمودند کمال تشکر و قدر دانی را داریم

بر اساس (Eklund and Samms, 1981) نتایج بدست آمده در این مطالعه در تماس^۱ با تعداد زیادی از اسیدهای آمینه در ساختار آنزیم مشاهده شد. با توجه به پایین بودن سطح دسترسی به حلال (زیر ۴۰٪) در این موتیف، استنباط می‌شود که موتیف α در مرکز مولکول پنهان بوده و با محیط اطراف در تماس نیست. موتیف بتا توالی محافظت شده دیگری است که از نظر موقعیت فضایی و ساختار ثانویه، خصوصیت شبیه موتیف آلفا را نشان داد. این موتیف دارای دو انتها از هلیکس ۱ و sheet شماره ۹ می‌باشد. بر اساس گزارشات قبلی مناطق انتهایی یک آلفا هلیکس و یک sheet نیز در پروتئین به صورت یک ساختار لولایی عمل می‌کند (Hayward, 1999; Hayward, 2004).

این دو موتیف که در این مطالعه توسط محاسبات بیوانفورماتیکی، در تمام ساختارهای مورد بررسی الکل دهیدروژنازها از سطوح مختلف تکاملی به صورت کاملاً محافظت شده شناسایی شده‌اند، به نظر می‌رسد که در تغییرات ساختاری آنزیم و در کنترل دقیق تحرک زمین‌های آنزیمی به دنبال اتصال سوبسترا و یا کوفاکتورها نقش اساسی بازی کنند.

مطالعه روی طبیعت اسیدهای آمینه در

منابع

- Atteia, A., van Lis, R., Henze, K., Martin, W., Riveros-Rosas, H., Gonzalez-Halphen, D. (2003) Bifunctional aldehyde/alcohol dehydrogenase (ADHE) in chlorophyte algal mitochondria. *Plant Molecular Biology* 53: 175-188
- Bairoch, A. and Apweiler, R. (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research* 28: 45-48
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, TN., Weissig, H., Shindyalov, I.N., Bourne, PE. (2000) The Protein Data Bank. *Nucleic Acids Research* 28: 235-242
- Bogin, O., Levin, I., Hacham, Y., Teler, S., Peretz, M., Frolov, F., Burstein, Y. (2002) Structural basis for the enhanced thermal stability of alcohol dehydrogenase mutants from the mesophilic bacterium *Clostridium beijerinckii*: contribution of salt bridging. *Protein science* 11: 2561-2574
- Bottini, S., Bernini, A., De Chiara, M., Garlaschelli, D., Spiga, O., Dioguardi, M., Vannucini, E., Tramontano, A., Niccolai, N. (2013) ProCoCoA: A quantitative approach for analyzing protein core composition. *Computational biology and chemistry* 43: 29-34
- Briffeuil, P., Baudoux, G., Lambert, C., De Bolle, X., Vinals, C., Feytmans, E., Depiereux, E. (1998) Comparative analysis of seven multiple protein sequence alignment servers: clues to enhance reliability of predictions. *Bioinformatics* 14: 357-366
- Chen, Z., Lee, W. R., Chang, S. H. (1991) Role of aspartic acid 38 in the cofactor specificity of *Drosophila* alcohol dehydrogenase. *European Journal Biochemistry* 202: 263-267
- Chen, Z., Lu, L., Shirley, M., Lee, W.R., Chang, S.H. (1990) Site-directed mutagenesis of glycine-14 and two critical cysteinyl residues in *Drosophila* alcohol dehydrogenase. *Biochemistry* 29: 1112-1118
- Chew, L. P. and Huttenlocher, D. (1999) Fast detection of common geometric substructure in proteins. *Journal of Computational Biology* 6: 313-325
- Chothia, C., Gelfand, I., Kister, A. (1998) Structural determinants in the sequences of immunoglobulin variable domain. *J Mol Biol* 278: 457-479
- Chothia, C., Gelfand, I., Kister, A. (1986) The relation between the divergence of sequence and structure in proteins. *The EMBO journal* 4:823-5
- Chung, J.L., Beaver, J.E., Scheeff, E.D., Bourne, P.E. (2007) Con-Struct Map: a comparative contact map analysis tool. *Bioinformatics* 23: 2491-2492
- Cuff, A., Murzin, A., Orengo, C. (2013) *Classifying Proteins into Domain Structure*

Families. Protein Families Relating Protein Sequence, Structure and Function: 37-68

- De Castro, E., Sigrist, C.J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gastegger, E., Bairoch, A., Hulo, N. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Research* 34: W362-W36
- Eddy, S. R. (1998) Profile hidden Markov models. *Bioinformatics* 14: 755-763
- Eddy, S. R. (2004) Where did the BLOSUM62 alignment score matrix come from. *Nature biotechnology* 22: 1035-1036
- Eklund, H. and Nordström, B. (1976) Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution. *Journal of molecular biology* 102: 27-59
- Eklund, H. and Nordstrom, B. (1976) Three-dimensional structure of horse liver alcohol dehydrogenase at 2-4 A resolution. *Journal of Molecular Biology* 102: 27-59
- Eklund, H., Samama, J.P., Jones, T.A. (1984) Crystallographic investigations of nicotinamide adenine dinucleotide binding to horse liver alcohol dehydrogenase. *Biochemistry* 23: 5982-5996
- Eklund, H., Samma, J.P., Wallen, L., Branden, C.I., Akeson, A., Jones, T.A. (1981) Structure of a triclinic ternary complex of horse liver alcohol dehydrogenase at 2.9 A resolution. *Journal Molecular Biolog* 146: 561-587
- Esposito, L., Sica, F., Raia, C.A., Giordano, A., Rossi, M., Mazzarella, L., Zagari, A. (2002) Crystal structure of the alcohol dehydrogenase from the hyperthermophilic archaeon *Sulfolobus solfataricus* at 1.85 Å resolution. *Journal of molecular biology* 318: 463-477
- Gelfand, I., Kister, A., Kulikowski, C., Stoyanov, O. (1998) Geometric invariant core for the V (L) and V (H) domains of immunoglobulin molecules. *Protein Engineering* 11: 1015-1025
- Gerstein, M. and Altman, R. B.. (1995) Average core structures and variability measures for protein families: application to the immunoglobulins. *Journal of Molecular Biology* 251: 161-175
- Gerstein, M. and Altman, R. B. (1995) Using a measure of structural variation to define a core for the globins. *Computational Applied Bioscience* 11: 633-644
- Gerstein, M. and Krebs, W. (1998) A database of macromolecular motions. *Nucleic Acids Research* 26: 4280-4290
- Goihberg, E., Dym, O., Tel-Or, S., Levin, I., Peretz, M., Burstein, Y. (2007) A single proline

- substitution is critical for the thermostabilization of *Clostridium beijerinckii* alcohol dehydrogenase. *Proteins* 66: 196-204
- Goihberg, E. and Dym, O. (2007) A single proline substitution is critical for the thermostabilization of *Clostridium beijerinckii* alcohol dehydrogenase. *Proteins: Structure, Function, and Bioinformatics* 66: 196-204
- Grant, B. J. and McCammon, J. A. (2007) Multivariate analysis of conserved sequence-structure relationships in kinesins: coupling of the active site and a tubulin-binding sub-domain. *Journal Molecular Biology* 368: 1231-1248
- Grant, B.J., Rodrigues, A.P., ElSawy, K.M., McCammon, J.A., Caves L.S. (2006) Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* 22: 2695-2696
- Guy, J.E., Isupov, M.N., Littlechild, J.A. (2003) The structure of an alcohol dehydrogenase from the hyperthermophilic archaeon *Aeropyrum pernix*. *Journal of Molecular Biology* 331: 1041-1051
- Hatley, M.E., Lockless, S.W., Gibson, S.K., Gilman, A.G., Ranganathan, R. (2003) Allosteric determinants in guanine nucleotide-binding proteins. *Proceedings of the National Academy of Sciences* 100: 14445-50
- Hayward, S. (1999) Structural principles governing domain motions in proteins. *Proteins: Structure, Function, and Bioinformatics* 36: 425-435
- Hayward, S. (2004) Identification of specific interactions that drive ligand-induced closure in five enzymes with classic domain movements. *Journal Molecular Biology* 339: 1001-1021
- Henikoff, S. and Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences* 89: 10915-10919
- Hubbard, S.J. (1993) NACCESS. Department of Biochemistry and Molecular Biology, University College London
- Jornvall, H., von Bahr-Lindstrom, H., Jeffery J. (1984) Extensive variations and basic features in the alcohol dehydrogenase-sorbitol dehydrogenase family. *European Journal Biochemistry* 140: 17-23
- Kim, S.J., Kim, M.R., Bedgar, D.L., Moinuddin, S.G., Cardenas, C.L., Davin, L.B., Kang, C., Lewis, N.G. (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. *Proceedings of the National Academy of Sciences* 101: 1455-1460
- Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J., Lesk, A.M. (2006) MUSTANG: a multiple structural alignment algorithm. *Proteins* 64: 559-574
- Korkhin, Y., Kalb, A.J., Peretz, M., Bogin, O., Burstein, Y., Frolow, F. (1998) NADP-de-

- pendent bacterial alcohol dehydrogenases: crystal structure, cofactor-binding and cofactor specificity of the ADHs of *Clostridium beijerinckii* and *Thermoanaerobacter brockii*. *Journal of molecular biology* 278: 967-981
- LeBru, L.A., Park, D.H., Ramaswamy, S., Plapp, B.V. (2004) Participation of histidine-51 in catalysis by horse liver alcohol dehydrogenase. *Biochemistry* 43:3014-26.
- Mirny, L. and Shakhnovich, E. (2001) Evolutionary conservation of the folding nucleus. *Journal Molecular Biology* 308: 123-129
- Ramaswamy, S., Park, D.H., Plapp, B.V. (1999) Substitutions in a flexible loop of horse liver alcohol dehydrogenase hinder the conformational change and unmask hydrogen transfer. *Biochemistry* 38: 13951-13959
- Riveros-Rosas, H., Julian-Sanchez, A., Villalobos-Molina, R., Pardo, J.P., Pina, E. (2003) Diversity, taxonomy and evolution of medium-chain dehydrogenase/reductase superfamily. *European Journal Biochemistry* 270: 3309-3334
- Ryde, U. (1995) On the role of Glu-68 in alcohol dehydrogenase. *Protein science* 4: 1124-1132
- Schneider, D. (1997) Information content of individual genetic sequences. *Journal of Theoretical Biology* 189: 427-441
- Shenkin, P.S. Erman, B. Mastrandrea, L.D. (1991) Information theoretical entropy as a measure of sequence variability. *Proteins: Structure, Function, and Bioinformatic* 11: 297-312
- Tamura, K., Dudley, J., Nei, M., Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology Evolution* 24: 1596-1599
- Thompson, J.D., Higgins, D.G., Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673-4680
- Vakulenko, Y.A., Nagaev, B., Alexeevski, A., Karyagina, A., Spirin, S. (2016) A new program for detecting the geometrical core of a set of structures of macromolecular complexes. *Biochemistry (Moscow)* 81: 428-431
- Valdar, W. S. (2002) Scoring residue conservation. *Proteins* 48: 227-241
- Yao, X.Q., Skjærven, L., Grant, B.J. (2016) Rapid Characterization of Allosteric Networks with Ensemble Normal Mode Analysis.