



## ارزیابی مدل‌های شبکه ارتجاعی در پیش‌بینی تغییرات ساختاری پروتئین‌ها

سیروس احمدی طوسی<sup>1</sup>، رضا سهیلی فرد<sup>2\*</sup>

1- دانشجوی کارشناسی ارشد، مهندسی پزشکی گرایش بیومکانیک، دانشگاه حکیم سبزواری، سبزوار

2- استادیار، گروه مهندسی مکانیک، دانشگاه حکیم سبزواری، سبزوار

\* سبزواری، صندوق پستی 397، r.soheilifard@hsu.ac.ir

### اطلاعات مقاله

مقاله پژوهشی کامل

دریافت: 25 شهریور 1394

پذیرش: 23 آبان 1394

ارائه در سایت: 28 آذر 1394

کلید واژگان:

تغییرات ساختاری

مدل شبکه ارتجاعی

آنالیز مود نرمال

شعاع حدی

### چکیده

تغییرات ساختاری پروتئین‌ها که در هنگام اتصال به لیگاند یا پروتئین‌های دیگر صورت می‌گیرد، نقش حیاتی در پدیده‌های بیولوژیکی ایفا می‌نماید. این گونه حرکات به صورت تجمعی و با فرکانس پایین بوده و آنالیز مود نرمال روش متداول جهت یافتن فرکانس‌ها و شکل مودها می‌باشد. مطالعات گسترده‌ای جهت پیش‌بینی این حرکات با استفاده از مدل‌های مختلف پروتئین صورت گرفته است. از این میان، مدل‌های شبکه ارتجاعی با توجه به عدم نیاز به کمیته‌سازی انرژی پتانسیل و ساده بودن از مطلوبیت بالایی برخوردار است. با این وجود، تاکنون مطالعه جامعی در خصوص بررسی تأثیر پارامترهای مختلف این مدل‌ها در پیش‌بینی تغییرات ساختاری صورت نگرفته است. در این مطالعه تعداد 20 پروتئین با تغییرات ساختاری مشخص انتخاب شده و با ایجاد مدل‌های مختلف شبکه ارتجاعی، میزان موفقیت و اعتبار هر یک در پیش‌بینی ساختار نهایی مورد بررسی قرار گرفته است. براساس نتایج این مطالعه سه مود اول هر مدل اغلب بیشترین نقش را در پیش‌بینی تغییرات ساختاری پروتئین‌ها ایفا می‌کنند. علاوه بر این، انتخاب شعاع حدی مناسب تأثیر بیشتری نسبت به تابع پتانسیل در مدل شبکه ارتجاعی دارد. همچنین نتایج حاصله نشان می‌دهد از میان مدل‌های مختلف در نظر گرفته شده، مدل‌های غیر نمایی با شعاع حدی 10 آنگستروم با وجود محاسبات مقرون‌به‌صرفه تر از دقت بیشتری در پیش‌بینی تغییرات ساختاری برخوردار می‌باشند.

## Evaluating elastic network models in prediction of conformational changes of proteins

Cyrus Ahmadi Toussi, Reza Soheilifard\*

Department of Mechanical Engineering, Hakim Sabzevari University, Sabzevar, Iran

\* P.O.B. 397, Sabzevar, Iran, r.soheilifard@hsu.ac.ir

### ARTICLE INFORMATION

Original Research Paper  
Received 16 September 2015  
Accepted 14 November 2015  
Available Online 19 December 2015

#### Keywords:

conformational changes  
normal mode analysis  
elastic network model  
cutoff radius

### ABSTRACT

Conformational changes during protein-protein or ligand-protein docking play an important role in the biological processes. These changes involve low frequency collective motions, and normal mode analysis is generally used for finding the frequencies and mode shapes of the proteins. Many studies have focused on the prediction of these transitions using different protein models. Among them, elastic network models are popular, as they are simple and do not require energy minimization. However, so far no systematic study has been undertaken regarding the effect of different parameters in prediction of these conformational changes. In this study, 20 proteins with pre-determined conformational changes were selected and the success and validation of each elastic network model in predicting the bound state were tested. According to the results, the first three modes play the major role in predicting the conformational changes. Moreover, choosing the proper cutoff radius is more effective than the potential function. Results also show that non-exponential models with 10 angstrom cutoff are more accurate in predicting conformational changes, in spite of their simplicity and being less time consuming.

### 1- مقدمه

حرکت پروتئین‌ها و رفتار دینامیکی آن‌ها منجر به عملکرد خاص پروتئین‌ها در بدن شده و نقش حیاتی در بسیاری از پدیده‌های بیولوژیکی ایفا می‌کند [1]. از جمله این حرکات، تغییرات ساختاری یک پروتئین در هنگام اتصال به پروتئینی دیگر و یا دارو می‌باشد که در نتیجه آن پیکربندی<sup>1</sup> پروتئین از یک حالت (ساختار اولیه) به حالتی دیگر (ساختار نهایی) تغییر می‌کند (شکل 1). نتیجه این عمل ایجاد یک ترکیب<sup>2</sup> مشتمل بر پروتئین بزرگتر به‌عنوان

گیرنده<sup>3</sup> و دارو یا پروتئین کوچکتر به‌عنوان لیگاند است.

پیش‌بینی تغییرات ساختاری و در نظر گرفتن انعطاف‌پذیری گیرنده و لیگاند در هنگام اتصال، اهمیت بالایی در طراحی دارو مبتنی بر ساختار<sup>4</sup> (SBDD) دارد [2-5]، چرا که توصیف دقیق گذار از ساختار اولیه به نهایی می‌تواند منجر به شناخت مکانیزم‌های پنهان و طراحی داروهای الوستریک<sup>5</sup> شود. داروی الوستریک با اتصال به قسمتی از پروتئین باعث

3- receptor

4- Structural Based Drug Design

5- allosteric

1- configuration

2- complex

Please cite this article using:

C. Ahmadi Toussi, R. Soheilifard, Evaluating elastic network models in prediction of conformational changes of proteins, *Modares Mechanical Engineering*, Vol. 16, No. 1, pp. 81-88, 2016 (in Persian)

برای ارجاع به این مقاله از عبارت ذیل استفاده نمایید:

شده را محاسبه نمودند [14,13]. همچنین کیم و همکاران مدل هیبریدی را ارائه کردند که در آن قسمت‌هایی که منعطف‌تر می‌باشند با دقت تصویر بیشتری مدل می‌شود [16,15]. هینسن نیز استفاده از مدل شبکه ارتجاعی با سختی نمایی [18,17] را مناسب دانسته و شعاع حدی را حذف نمود.

جهت اعتبارسنجی و یافتن پارامترهای بهینه در این مدل‌ها در بسیاری از مطالعات از مقایسه معیاری از تحرک اتم‌ها به نام بی‌فاکتور<sup>3</sup> بین مدل محاسبه شده و نتایج تجربی استفاده شده است [1,19]. گرچه اخیراً تحقیقات محققین نشان داده است که حرکت صلب پروتئین در محیط کریستالی بیشتر از حرکت داخلی پروتئین در بی‌فاکتور تاثیرگذار است، که این امر استفاده از بی‌فاکتورها را بدین منظور خدشه‌پذیر می‌نماید [21,20]. روش دیگری که جهت ارزیابی مدل‌ها می‌توان از آن‌ها بهره برد، استفاده از تغییرات ساختاری تجربی پروتئین‌ها است، زیرا که مطالعات نشان داده است تغییرات ساختاری بزرگ دامنه پروتئین‌ها با تعداد کمی موده‌های نرمال کم فرکانس، به‌خوبی قابل پیش‌بینی است [7,8,22].

در این مطالعه جهت ارزیابی و اعتبارسنجی مدل‌های مختلف شبکه ارتجاعی به‌جای استفاده از بی‌فاکتورها از معیار پیش‌بینی تغییرات ساختاری پروتئین‌ها استفاده شده است. بدین منظور مجموعه‌ای از ترکیب پروتئین‌ها در دو حالت اولیه و نهایی انتخاب شده و مدل‌های مختلف شبکه ارتجاعی آن‌ها ایجاد می‌گردد. سپس با انجام آنالیز مود نرمال روی شبکه‌های مختلف و مقایسه موده‌های با فرکانس پایین با تغییرات ساختاری مدل‌ها، پارامترهای مدل بررسی شده و بهترین مدل‌ها انتخاب می‌گردند.

## 2- مواد و روش‌ها

### 2-1- پروتئین‌های مورد مطالعه

در این تحقیق 14 ترکیب پروتئین- پروتئین موجود در بنچمارک داکینگ<sup>4</sup> شماره 4 انتخاب شده است [23]، که هر کدام از پروتئین‌ها شامل ساختار اولیه و ساختار نهایی (در حالت ترکیب) می‌باشند.

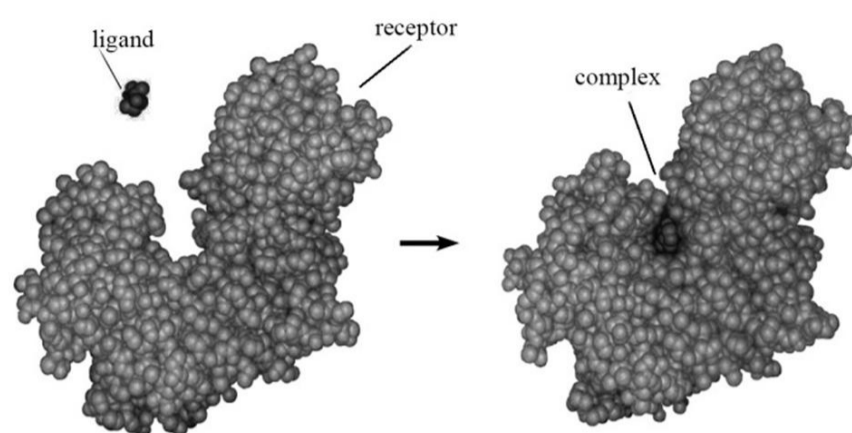
از تعداد 28 پروتئین موجود در این مجموعه 8 پروتئین دارای ساختار اولیه و نهایی یکسان بوده‌اند که در مطالعه حاضر مورد اهمیت نمی‌باشند، بنابراین نهایتاً 20 پروتئین از این مجموعه به‌عنوان پروتئین‌های مورد مطالعه در نظر گرفته شده‌اند. جدول 1 نام و مشخصات این ترکیبات را نمایش می‌دهد.

### 2-2- مدل شبکه ارتجاعی

در مدل شبکه ارتجاعی متداول، پروتئین به‌صورت مجموعه‌ای از نقاط یا گره‌هایی که توسط فنرهایی با پتانسیل هوکی به یکدیگر متصل می‌شوند در نظر گرفته می‌شود (شکل 2). البته معمولاً به جهت سادگی بیشتر محاسبات به‌جای در نظر گرفتن تمام اتم‌ها، اتم‌های آلفا کربن هر آمینواسید به‌عنوان گره‌های این شبکه انتخاب می‌شوند. در این مدل دو اتم آلفا کربن در صورتی به یکدیگر متصل می‌گردند که فاصله آن‌ها کمتر از یک شعاع حدی انتخابی  $R_c$  باشد. همچنین فرض بر آن است که ساختار شبکه در این حالت در وضعیت انرژی مینیمم باشد، یعنی فنرها تحت کشش نباشند. انرژی پتانسیل کل مدل شبکه ارتجاعی را می‌توان به‌صورت زیر نوشت.

$$E_{\text{network}} = \frac{1}{2} \sum f(R_{ij}^0) (R_{ij} - R_{ij}^0)^2 H(R_{ij}^0 - R_c) \quad (1)$$

که در آن  $R_{ij}$  فاصله دو اتم  $i$  و  $j$  از یکدیگر و  $R_{ij}^0$  فاصله آن‌ها در حالت تعادلی



**Fig. 1** conformational change of a receptor from initial to final structure due to ligand binding

**شکل 1** تغییرات ساختاری گیرنده از ساختار اولیه به نهایی در اثر اتصال به لیگاند

تغییر شکل در قسمتی دیگر شده و منجر به تنظیم عملکرد آن پروتئین می‌شود. تحقیقات آزمایشگاهی تاکنون محدود به مشاهدات ساختارهای اولیه، نهایی و تعدادی از ساختارهای میانی بوده و از فهم اتفاقات حالات گذار ناتوان می‌باشند، از این‌رو استفاده از روش‌های محاسباتی جهت مطالعه تغییرات ساختاری مورد توجه ویژه‌ای قرار گرفته است.

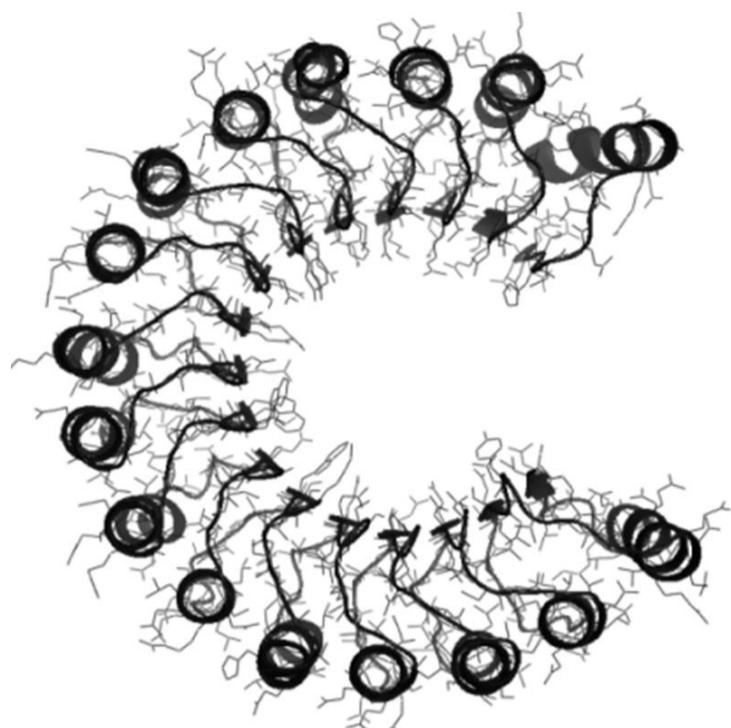
نکته قابل توجه در تغییرات ساختاری، تجمعی بودن حرکت اجزاء پروتئین است که فرکانس این حرکات پایین و دامنه آن‌ها بزرگ می‌باشد [7,6].

از جمله روش‌های محاسباتی جهت پیش‌بینی این‌گونه حرکات آنالیز مود نرمال<sup>1</sup> است. این آنالیز مبتنی بر کمینه‌سازی اولیه انرژی پتانسیل، محاسبه ماتریس هشن<sup>2</sup> (مشتق دوم انرژی پتانسیل نسبت به مختصات اتم‌ها) و نهایتاً یافتن موده‌ها و فرکانس‌های ارتعاش جسم است. با وجود این‌که این روش قادر است به‌خوبی انعطاف یا تغییرات ساختاری پروتئین‌ها را در هنگام اتصال به یکدیگر محاسبه کند، دارای نقاط ضعفی نیز هست که از جمله آن می‌توان به زمان‌بر بودن آن به دلیل پیچیدگی اندرکنش خاص بین اتم‌ها و فرایند کمینه‌سازی انرژی اشاره نمود. این امر بخصوص در پروتئین‌های بزرگ هزینه محاسبات را به‌شدت افزایش داده و یا غیرممکن می‌سازد. علاوه بر این، کمینه‌سازی انرژی به دلیل ایجاد اعوجاج چند آنگسترومی در ساختار اولیه مشاهده شده توسط کریستالوگرافرها، منجر به کاهش دقت در پیش‌بینی جهت حرکت می‌شود [8].

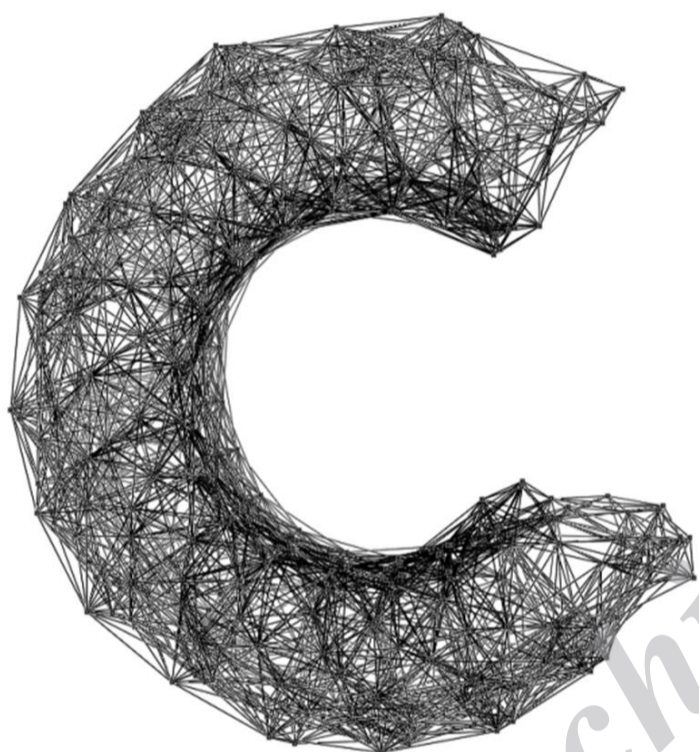
روش شبکه ارتجاعی با جایگزینی یک پتانسیل هوکی ساده بجای این اندرکنش‌ها، توانسته است مشکلات آنالیز مود نرمال را تا حد مطلوبی مرتفع سازد. اولین مدل شبکه ارتجاعی توسط تریون [9] ارائه شد که در آن اندرکنش ساده هوکی بین هر جفت اتم که در فاصله یک شعاع حدی مشخص قرار دارند، در نظر گرفته شده است. آتیلگان و همکاران در سال 2001 همان مدل را با در نظر گرفتن اتم‌های آلفا کربن به‌تنهایی درشت‌دانه‌سازی نموده‌اند [10]. محققین مدل‌های مختلفی برای شبکه ارتجاعی ارائه نمودند که در آن‌ها تلاش‌هایی جهت افزایش دقت مدل و یا کاهش هزینه محاسبات صورت گرفته است [12,11] از جمله، شوپلر و همکاران مدل ارتجاعی کلاستر را پیشنهاد نمودند. آن‌ها پروتئین را به‌صورت شبکه‌ای از اجسام صلب که توسط پتانسیل هارمونیک به یکدیگر متصل شده‌اند، در نظر گرفتند و سپس آن را با مدل تمام اتمی مقایسه کرده و زمان صرفه‌جویی

3- B-factors  
4- docking benchmark

1- normal mode analysis  
2- Hessian matrix



**Fig. 2-a** cartoon representation of RNase A in initial structure  
 شکل 2- الف نمای کارتونی از ساختار اولیه ی پروتئین ریبونوکلاز آ



**Fig. 2-b** RNase A elastic network model with 10 Å cutoff radius

شکل 2- ب شبکه ی ارتجاعی پروتئین ریبونوکلاز آ به ازای شعاع حدی 10 انگستروم

ضریب مقیاس 3، 4 و 5 انگستروم مورد بررسی قرار گرفته و سپس میزان موفقیت هر مدل در تخمین جهت تغییرات ساختاری و پیش‌بینی ساختار نهایی ارزیابی شده است. جهت مدل‌سازی انواع شبکه‌های ارتجاعی و سایر محاسباتی که در ادامه ذکر می‌شود از نرم‌افزار متلب استفاده شده است. شایان ذکر است مقادیر ذکر شده برای شعاع حدی و ضریب مقیاس، مقادیر متداول استفاده شده در مطالعات پیشین است. با این وجود، مطالعه حاضر برای مقادیر بیشتر و کمتر از این حد نیز محاسبه شده است، گرچه این مقادیر منجر به بهبود مدل در پیش‌بینی تغییرات ساختار نشده است. به‌عنوان نمونه شعاع 7 انگستروم باعث به وجود آمدن صفرهای بیشتر در مقادیر ویژه به‌دست آمده از قطری‌سازی ماتریس هشن و عدم ثبات شبکه ارتجاعی شده و شعاع حدی بیشتر از 18 انگستروم نیز در عمل پروتئین را به یک جسم صلب تبدیل می‌کند.

### 2-3- آنالیز مود نرمال

انجام آنالیز مود نرمال مستلزم انجام سه مرحله است:

جدول 1 نام ترکیبات و مشخصات پروتئین‌های آن‌ها در حالت اولیه و نهایی

**Table 1** complexes' names and their pdb ID in initial and final structure.

ترکیب شناسه	تعداد رزیجو	شناسه گیرنده	تعداد رزیجو	شناسه لیگاند	توضیح (نام لیگاند / نام گیرنده)
1CGI(E:I)	245	1CHG	56	1HPT	Alpha-chymotrypsinogen / human pancreatic secretory trypsin inhibitor
1BVK(DE:F)	116	1BVL(LH)	125	3LZT	Antibody Hulys11 Fv / Lysozyme
2JEL*(LH:P)	226	2JEL(LH)	85	1POH	Jel42 Fab Fragment / A06 Phosphotransferase
1DFJ(I:E)	124	2BNH	456	7RSA	Ribonuclease inhibitor / Ribonuclease A
1MLC(AB:E)	218	1MLB(AB)	129	1LZA	IgG1 D44.1 Fab fragment / Lysozyme
1PPE*(E:I)	223	2PTN	29	1PPE(I)	Trypsin / CMT-1 Virus Uracil-DNA glycosylase / inhibitor
1UDI*(E:I)	244	1UDH	82	1UDI(I)	camel VHH / Pancreatic alpha-amylase
1KXQ (H:A)	496	1KXQ_H	120	1PPI	Barnase/Barstar
1AY7_A:B	96	1RGH_B	89	1A19_B	Xylanase / Xylanase inhibitor
2B42_B:A	185	2DCY_A	385	1T6E_X	SRP GTPase / Cell division protein FtsY
2J7P_A:D	108	1NG1_A	17	2IYL_D	Fab / HIV-1 capsid protein p24
1E6J_HL:P	96	1E6O_HL	89	1A43	PIGF receptor binding domain / Flt1 protein domain 2
1RV6_VW:X	93	1FZV_AB	91	1QSZ_A	Streptogrisim / Ovomuroid inhibitor third domain
3SGQ_E:I	185	2QA9_E	56	2OVO_A	

است.  $H$  تابع پله‌ای هوی ساید به صورت

$$H(R_{ij}^0 - R_c) = \begin{cases} 1: & R_{ij}^0 < R_c \\ 0: & R_{ij}^0 \geq R_c \end{cases} \quad (2)$$

تعریف شده که در آن  $R_c$  شعاع حدی می‌باشد. در این مطالعه چهار تابع مختلف برای سختی فنرها با روابط (a-3) الی (d-3) در نظر گرفته شده است.

$$f(r) = 1 \quad (a-3)$$

$$f(r) = \frac{1}{r} \quad (b-3)$$

$$f(r) = \frac{1}{r^2} \quad (c-3)$$

$$f(r) = \exp\left(-\frac{r^2}{l_0^2}\right) \quad (d-3)$$

در این روابط  $f$  بیانگر نحوه تغییر سختی فنرها بافاصله دو اتم است. در رابطه (d-3)،  $l_0$  ضریب مقیاس در توابع نمایشی می‌باشد که تا حدودی نقش همان شعاع حدی را در این مدل ایفا کرده و  $R_c$  در این مدل بی‌نهایت در نظر گرفته می‌شود. در این مطالعه برای مدل‌های شبکه ارتجاعی غیرنمایی (a-3 تا c-3) سه شعاع‌های حدی 10، 15 و 18 انگستروم و در مدل نمایشی (d-3) نیز سه

با پروتکل‌های بیان شده در بخش 2-2 برای آن‌ها ساخته شد. سپس اورلپ بین جهت تغییرات ساختاری به‌دست آمده از داده‌های تجربی و جهت پیش‌بینی شده برای هر پروتئین به ازای 10 مود اول هر مدل محاسبه شده و مودی که بیشترین اورلپ را با تغییرات ساختاری داشته است، شناسایی گردید.

جدول 2 نتایج آماری مربوط به مود غالب در مجموعه پروتئین‌ها را نشان می‌دهد. مطابق جدول 2، مود حرکتی اول در 53% موارد بهتر از بقیه مودها جهت حرکت را پیش‌بینی می‌کند. مود حرکتی دوم و سوم نیز به ترتیب با 27% و 15% از اهمیت خاصی برخوردارند. تنها در تعداد بسیار ناچیزی از پروتئین‌ها (5%)، مودهای 4 الی 10 پیش‌بینی بهتری از حرکت را داشته‌اند. نتایج جدول 2 به ویژه در محاسبات داکینگ انعطاف‌پذیر اهمیت می‌یابد، چراکه در آن طبق محاسبات فوق، کافی است تنها سه مود اول را در نظر گرفت و این امر هزینه محاسبات را شدیداً کاهش می‌دهد.

جدول 3 میزان اورلپ هر یک از مدل‌ها را به همراه شماره مود غالب به ازای 20 پروتئین مورد بررسی نشان می‌دهد. با توجه به میانگین مقادیر اورلپ در 20 پروتئین به ازای هر مدل، مشاهده می‌گردد که به‌طور کلی مدل نمایی با ضریب مقیاس 5 و با میانگین اورلپ 0.47 دارای بیشترین تطابق حرکتی با تغییرات ساختاری است. البته این مدل بدون شعاع حدی بوده و در آن تمامی اندرکنش‌ها در نظر گرفته می‌شود که خود هزینه محاسبات را به‌طور چشمگیری افزایش می‌دهد.

در میان مدل‌های غیر نمایی، مدل  $1/r$  با شعاع حدی 10 آنگستروم بیشترین میزان تطابق را داشته و پس از آن مدل سختی ثابت و سپس مدل  $1/r^2$  با شعاع‌های حدی 10 قرار دارند. کمترین میزان اورلپ مربوط به مدل سختی ثابت با شعاع حدی 18 آنگستروم است. به‌طور کلی، مدل‌های غیرنمایی با شعاع حدی 10 آنگستروم عملکرد بهتری را در پیش‌بینی جهت تغییرات ساختاری دارا می‌باشند که این مطلب با مطالعات سانجواند [24] نیز تطابق دارد [25].

همچنین براساس نتایج میانگین اورلپ در جدول 2 می‌توان نتیجه گرفت در پیش‌بینی جهت تغییرات ساختاری پروتئین‌ها، تاثیر شعاع حدی در مدل‌های ارتجاعی غیرنمایی بیشتر از نوع تابع پتانسیل سختی فنر می‌باشد. در اغلب پروتئین‌های جدول 1 تفاوت قابل ملاحظه‌ای بین مدل‌های مختلف وجود دارد. شکل 3 جهت بردار مود غالب شبکه ارتجاعی و همچنین تغییرات ساختاری زیلاناز را نمایش می‌دهد. این پروتئین در مجاورت مهار کننده زیلاناز<sup>1</sup> به‌صورت زاویه‌دار از حالت بسته به حالت باز تغییر شکل داده و دو پروتئین به یکدیگر متصل می‌گردند. شکل 3- الف حالت اولیه و نهایی این پروتئین و شکل 3- ب جهت بردار مود غالب از مدل شبکه ارتجاعی این پروتئین را نشان می‌دهد.

جدول 2 تعداد و درصد موفقیت هر مود در پیش‌بینی صحیح جهت تغییرات ساختاری در مجموعه پروتئین‌ها

**Table 2** population and percentage of success for each mode in prediction of the conformational change in the protein dataset.

شماره مود	تعداد موفقیت در مجموعه	در صد موفقیت
1	129	53%
2	64	27%
3	35	15%
بقیه مودها	12	5%

1- xylanase

1 - کمینه سازی انرژی پتانسیل برحسب مختصات اتم‌ها

2 - محاسبه ماتریس هشن  $H$

3- قطری‌سازی ماتریس هشن و استخراج فرکانس‌ها و مودهای نرمال در مدل شبکه ارتجاعی، انرژی پتانسیل انتخاب شده به‌خودی‌خود در حالت کمینه بوده و بنابراین تنها مراحل 2 و 3 لازم است که صورت گیرد. عناصر ماتریس هشن در رابطه (1) به‌صورت معادله (4) به دست می‌آید:

$$H_{ij} = \left( \frac{\partial^2 E_{\text{network}}}{\partial r_i \partial r_j} \right) \quad (4)$$

که در آن  $r_i$  و  $r_j$  به ترتیب مختصات اتم‌های  $i$  و  $j$  می‌باشند. پس از آن قطری‌سازی ماتریس هشن منجر به مسئله مقدار ویژه به فرم

$$H U_k = \lambda_k \omega_k^2 U_k \quad (5)$$

می‌گردد که در آن  $\omega_k$  و  $U_k$  به ترتیب فرکانس و بردار مود نرمال  $k$  ام پروتئین می‌باشند. شایان ذکر است با توجه به این‌که پروتئین‌های مورد بررسی مقید نشده‌اند، مدل دارای (3 مود انتقالی و 3 مود چرخشی) با فرکانس صفر است. مودهای غیر صفر با فرکانس پایین مربوط به حرکت جمعی پروتئین می‌باشند که تاثیر بیشتری در پیش‌بینی تغییرات ساختاری پروتئین دارند.

## 2-4- ارزیابی مدل‌های شبکه ارتجاعی

### 2-4-1- میزان تطابق مود غالب و تغییر ساختاری

به‌منظور ارزیابی تطابق جهت مود غالب به دست آمده از محاسبات با جهت تغییرات ساختاری که از اختلاف داده‌های کریستالوگرافی پروتئین در حالت اولیه و نهایی به‌دست آمده، معمولاً از شاخص اورلپ<sup>1</sup> استفاده می‌شود:

$$\frac{|\sum_i^{3N} a_{ij} \Delta r_i|}{[\sum_i^{3N} a_{ij}^2 \Delta r_i^2]^{1/2}} = O_j \quad (6)$$

که در آن  $O_j$  ضریب اورلپ،  $a_{ij}$  مؤلفه  $i$  ام از بردار مود  $j$  ام و  $\Delta r_i$  مؤلفه  $i$  ام بردار تغییرات ساختاری است. اورلپ واحد به‌معنای تطابق کامل و اورلپ صفر به معنی پیش‌بینی کاملاً اشتباه جهت حرکت می‌باشد.

### 2-4-2- پیش‌بینی ساختار نهایی با استفاده از مودهای ساختار اولیه

به‌عنوان دومین معیار جهت ارزیابی مدل‌های مختلف، این مدل‌ها برای ساختار اولیه پروتئین‌ها ایجاد شده و آنالیز مود نرمال روی آن‌ها صورت می‌گیرد، سپس ترکیب خطی 10 مود اول با فرکانس پایین از این مدل‌ها در نظر گرفته شده و ساختار نهایی با دگرگون‌سازی ساختار اولیه در راستای این 10 مود و کمینه‌سازی RMSD<sup>2</sup> (جذر مجموع مربعات اختلاف مختصه‌های مکانی دو ساختار) بین ساختار نهایی و ساختار پیش‌بینی شده محاسبه می‌گردد.

$$\text{RMSD}(\text{Ref}_i, \text{Tar}) = \sum_{i=1}^N ((x_{\text{Ref}})_i - (x_{\text{Tar}})_i)^2 \quad (7)$$

که در آن،  $(x_{\text{Ref}})_i$  و  $(x_{\text{Tar}})_i$  به ترتیب مؤلفه‌های  $i$  ام بردار مکانی ساختار مرجع و ساختار هدف و  $N$  تعداد کل نقاط می‌باشند. مطالعات قبلی حاکی از کاهش RMSD برحسب تعداد مودهای در نظر گرفته شده بوده و هرچه تعداد مود بیشتر شود، محاسبات دقیق‌تر اما با زمان بیشتری همراه خواهد بود.

## 3- نتایج

### 3-1- میزان تطابق مود غالب و تغییر ساختاری در مجموعه

#### پروتئین‌ها

جهت ارزیابی پیش‌بینی صورت گرفته از تغییرات ساختاری، حالت‌های اولیه و نهایی 20 پروتئین از جدول 1 که به‌صورت گیرنده یا لیگاند می‌باشند، مورد مطالعه قرار گرفته است. بدین منظور مدل‌های مختلف شبکه ارتجاعی مطابق

جدول 3 ضریب اورلپ مدل‌های شبکه‌های ارتجاعی مختلف برای مجموعه پروتئین‌ها. عدد داخل پرانتز مربوط به شماره مودی است که بیشترین اورلپ را داشته است.

**Table 3** Overlap coefficient for different elastic network models in the dataset. The number in parenthesis shows the most involved mode.

$F(r_{ij}^0) = 1$			$F(r_{ij}^0) = \frac{1}{r}$			$F(r_{ij}^0) = \frac{1}{r^2}$			$F(r_{ij}^0) = \exp(-\frac{r^2}{r_c^2})$			نام مدل
10	15	18	10	15	18	10	15	18	3	4	5	شعاع حدی {Å}
0.5855	0.3848	0.3709	0.6074	0.3822	0.3742	0.6212	0.3838	0.3781	0.6217	0.666	0.6655	نام پروتئین
(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	2J7P_receptor
0.7888	0.6062	0.5444	0.7917	0.6181	0.5635	0.7939	0.6273	0.5807	0.7834	0.8052	0.8087	2B42_receptor
(1)	(2)	(2)	(1)	(2)	(2)	(1)	(2)	(2)	(1)	(1)	(1)	1AY7_receptor
0.6464	0.5044	0.4555	0.6432	0.4876	0.5128	0.2244	0.4769	0.4768	0.5453	0.6545	0.5178	1AY7_ligand
(3)	(2)	(2)	(3)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(2)	1RV6_ligand
0.2377	0.3146	0.2597	0.2332	0.3278	0.2883	0.2300	0.3337	0.3111	0.2087	0.2925	0.3246	3SGQ_ligand
(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(1)	(2)	(2)	2E6J_receptor
0.4652	0.3259	0.2915	0.4641	0.3354	0.3304	0.4606	0.3501	0.3332	0.4142	0.4502	0.4548	1BVK_receptor
(3)	(2)	(2)	(3)	(2)	(2)	(3)	(3)	(2)	(1)	(4)	(4)	1BVK_ligand
0.5531	0.2996	0.3013	0.5516	0.3527	0.3005	0.5432	0.3489	0.3085	0.4843	0.4176	0.4514	1CGI_receptor
(1)	(2)	(2)	(1)	(2)	(2)	(1)	(2)	(2)	(1)	(1)	(8)	1CGI_ligand
0.506	0.5003	0.4697	0.4745	0.4651	0.4776	0.5286	0.4768	0.4781	0.7034	0.6855	0.6753	1DFJ_receptor
(1)	(1)	(2)	(2)	(1)	(2)	(2)	(2)	(2)	(1)	(1)	(1)	1DFJ_ligand
0.4681	0.461	0.4661	0.4635	0.4643	0.4699	0.4575	0.4673	0.4728	0.3814	0.4338	0.4660	1MLC_receptor
(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	1MLC_ligand
0.2707	0.3134	0.2749	0.294	0.3207	0.3118	0.3101	0.3223	0.3267	0.3235	0.3531	0.3481	1PPE_receptor
(3)	(3)	(3)	(3)	(3)	(3)	(3)	(3)	(3)	(6)	(3)	(3)	1PPE_ligand
0.3349	0.4598	0.4463	0.2574	0.4703	0.4512	0.2649	0.4824	0.4581	0.3621	0.3803	0.4323	1KXQ_receptor
(2)	(1)	(1)	(3)	(1)	(1)	(3)	(1)	(1)	(1)	(2)	(1)	2JEL_ligand
0.1454	0.1418	0.1525	0.149	0.1463	0.1538	0.1476	0.1482	0.1545	0.1460	0.1396	0.1098	1UDI_ligand
(2)	(1)	(1)	(2)	(1)	(1)	(2)	(1)	(1)	(3)	(3)	(1)	میانگین
0.3366	0.301	0.3319	0.3526	0.3449	0.3267	0.364	0.3224	0.3242	0.3877	0.5116	0.4831	
(1)	(2)	(2)	(1)	(2)	(2)	(1)	(2)	(2)	(9)	(10)	(10)	
0.7967	0.8143	0.8199	0.795	0.8129	0.8163	0.7962	0.8139	0.8183	0.8015	0.7942	0.8012	
(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	
0.4382	0.4394	0.4503	0.4369	0.4393	0.4489	0.4361	0.4388	0.4469	0.4650	0.4442	0.4339	
(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	
0.2986	0.1709	0.2629	0.3036	0.2062	0.2799	0.3067	0.269	0.2861	0.3614	0.3132	0.2859	
(1)	(1)	(3)	(3)	(3)	(3)	(3)	(3)	(3)	(9)	(3)	(3)	
0.5866	0.5879	0.5998	0.5874	0.5901	0.6038	0.5918	0.5937	0.6067	0.4100	0.5852	0.6034	
(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(3)	(1)	(1)	
0.4857	0.4132	0.4249	0.5003	0.5062	0.4774	0.4634	0.5332	0.4904	0.4954	0.4494	0.4295	
(4)	(1)	(1)	(4)	(4)	(1)	(4)	(4)	(4)	(5)	(1)	(1)	
0.4075	0.4324	0.4116	0.4054	0.4258	0.4335	0.4021	0.424	0.4268	0.3635	0.3828	0.4049	
(1)	(1)	(2)	(2)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	
0.163	0.1788	0.1627	0.3554	0.2927	0.1891	0.3736	0.2596	0.3519	0.395	0.273	0.448	
(1)	(2)	(2)	(6)	(8)	(2)	(6)	(8)	(10)	(1)	(7)	(9)	
0.3599	0.3309	0.2406	0.3361	0.366	0.3081	0.3082	0.3566	0.353	0.3259	0.3481	0.3275	
(1)	(3)	(3)	(1)	(2)	(2)	(1)	(2)	(2)	(1)	(2)	(2)	
0.4437	0.3990	0.3868	0.4501	0.4177	0.4058	0.4312	0.4214	0.4191	0.4489	0.4690	0.4735	

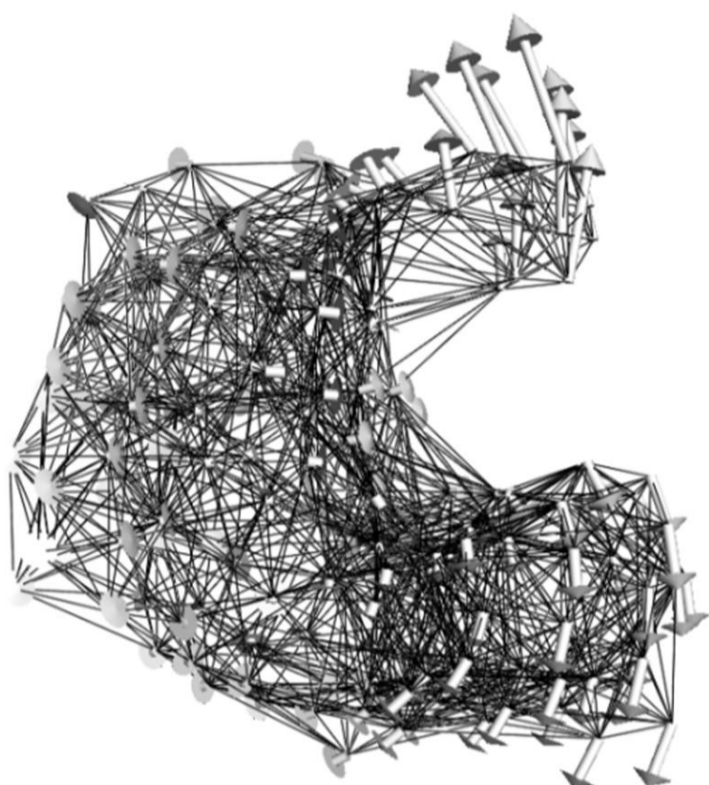
ب جهت بردار مود غالب از مدل شبکه ارتجاعی این پروتئین را نشان می‌دهد. همان‌طور که در شکل دیده می‌شود مود اول شبکه ارتجاعی در این پروتئین به‌طور مطلوب جهت حرکت تغییرات ساختاری را پیش‌بینی می‌کند.

در شکل 6 بردار یکه جابجایی مدل سختی ثابت با شعاع حدی 18 آنگستروم و مدل نمایی با ضریب مقیاس 5 (به ترتیب، کمترین و بیشترین میانگین اورلپ در مجموعه) همراه با بردار یکه تغییرات ساختاری به ازای هر آمینواسید رسم شده است. البته مدل  $1/r$  با شعاع حدی 10 (که بالاترین میانگین اورلپ در حالت غیر نمایی را دارا می‌باشد) نیز در این شکل در نظر گرفته شده که به علت یکی بودن جابجایی آن با مدل نمایی از رسم آن صرف‌نظر شده است. مشاهده می‌شود که در این پروتئین هر دو مدل تقریب نسبتاً مناسبی از این تغییرات را ارائه می‌دهند.

تفاوت در میزان تطابق بین مدل‌ها در این پروتئین به 0.25 می‌رسد. بردار یکه جابجایی دو مدل سختی ثابت با شعاع حدی 10 آنگستروم و مدل نمایی با ضریب مقیاس 5 به همراه بردار یکه تغییرات ساختاری در شکل 4 رسم شده است.

در این شکل به‌خوبی دیده می‌شود که مدل سختی ثابت با شعاع حدی 18 آنگستروم تطابق مناسبی با تغییرات نداشته، گرچه مدل نمایی این تغییرات را به‌خوبی پیش‌بینی می‌کند. در تعداد محدودی از پروتئین‌ها نیز میزان تطابق در تمامی مدل‌ها در یک حدود است. به‌عنوان نمونه، در پروتئین ریبونوکلاز میزان اورلپ در تمام مدل‌ها در حدود 0.8 می‌باشد.

این پروتئین در مجاورت مهار کننده ریبونوکلاز از حالت بسته به حالت باز تبدیل می‌شود. شکل 5- الف حالت اولیه و نهایی این پروتئین و شکل 5-



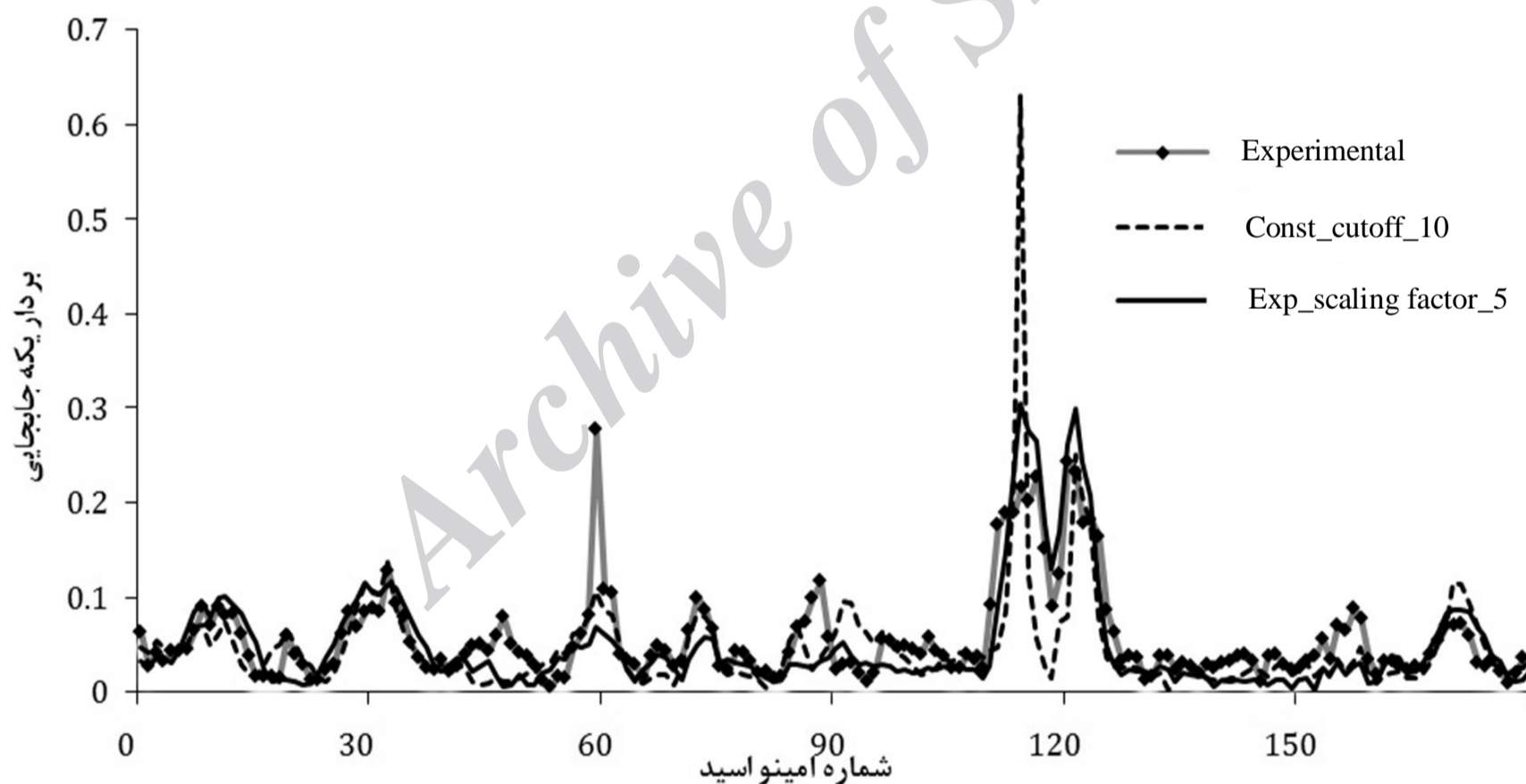
**Fig. 3-b** conformational changes predicted by NMA of ENM with 10Å cutoff radius in its first mode

شکل 3-ب جهت پیش‌بینی شده به وسیله آنالیز مود نرمال شبکه الاستیک در مود شماره 1 برای شبکه الاستیک با شعاع حدی 10 آنگستروم.



**Fig. 3-a** conformational change's direction observed from experiments in Xylanase. Xylanase in initial state is shown in black and in bound state in gray. This protein changes from initial to final state with a small angle.

شکل 3-الف جهت تغییرات ساختاری حاصل از داده‌های تجربی در زیلاناز. زیلاناز در حالت اولیه به صورت تیره و در حالت نهایی به صورت روشن ترسیم شده و در حضور مهارکننده زیلاناز از حالت بسته به حالت باز و با کمی زاویه تغییر ساختاری داده است.



**Fig. 4** Unit displacement vector of Xylanase. Thin line with markers shows crystallographic data. Dash line represents the ENM with constant stiffness and 10 Å cutoff radius and the thick line shows the exponential ENM with scaling factor 5

شکل 4 بردار یکه جابجایی در زیلاناز. خطوط کمرنگ با شناسه نقطه از داده‌های تجربی حاصل از کریستالوگرافی می‌باشد. خط چین متعلق به مدل شبکه ارتجاعی با سختی ثابت و شعاع حدی 10 آنگسترومی و خط پر مربوط به مدل نمایی با ضریب مقیاس 5 است.

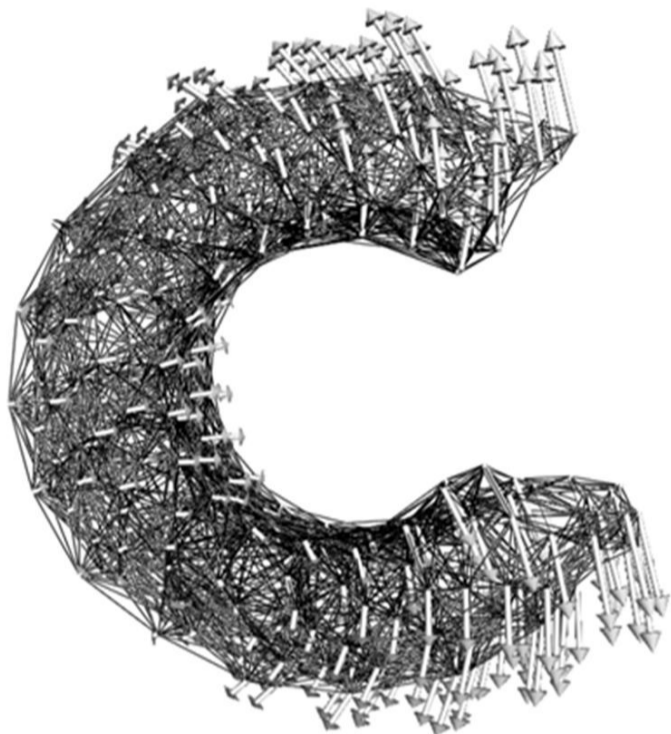
تعداد مود در نظر گرفته شده نشان می‌دهد.

مطابق شکل مشاهده می‌شود که با افزودن مدهای نرمال به ساختار اولیه، ساختار پیش‌بینی شده به ساختار نهایی نزدیک شده و مقدار RMSD این دو ساختار کاهش می‌یابد.

نتایج به دست آمده از نمودار شکل 7 حاکی از بهتر بودن مقدار RMSD در مدل‌های شبکه ارتجاعی نمایی و غیر نمایی با شعاع حدی 10 است که بر تاثیر بسزای شعاع حدی تأکید می‌کند. البته با توجه به هزینه بالای محاسبات

### 2-3- پیش‌بینی ساختار نهایی با استفاده از مدهای ساختار اولیه در مجموعه پروتئین‌ها

به عنوان دومین روش برای ارزیابی مدل‌های مختلف، از ترکیب بهینه‌ای از مدهای فرکانس پایین استفاده شده است. بدین منظور میزان RMSD ساختار اولیه دگرگون شده توسط ترکیب خطی بهینه از چند مود اول هر پروتئین با تغییرات ساختاری محاسبه شده و میانگین این مقدار برای مجموعه پروتئین مورد بررسی محاسبه شده است. شکل 7 میزان RMSD میانگین را به ازای



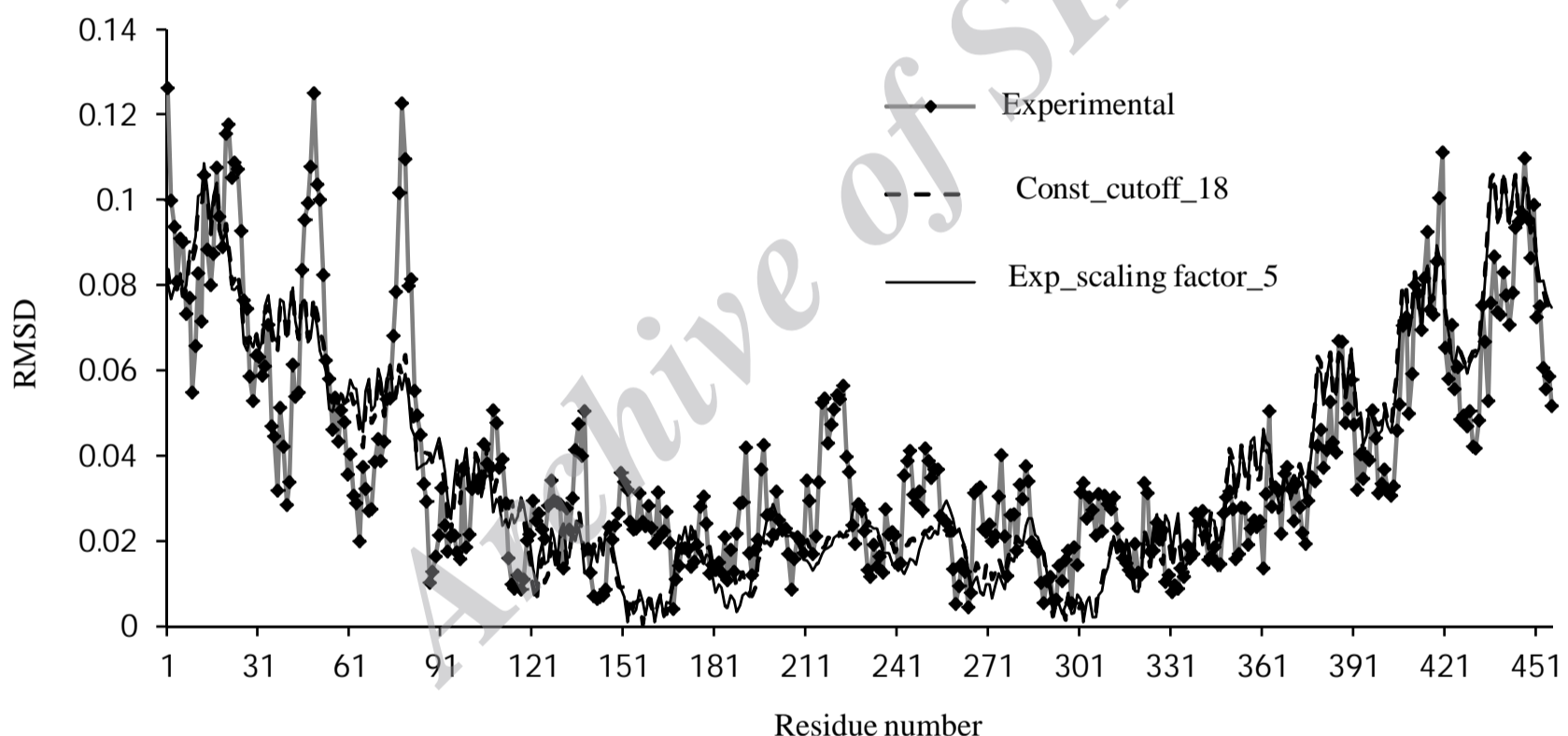
**Fig. 5-b** represents conformational changes predicted by NMA of ENM with 10 Å cutoff radius for its first mode

شکل 5-ب جهت پیش‌بینی شده به وسیله آنالیز نرمال مود شبکه ارتجاعی را در مود شماره 1 برای شبکه ارتجاعی با شعاع حدی 10 انگستروم نشان می‌دهد.



**Fig. 5-a** conformational change's direction observed from experiments in RNase A. RNase A is shown in black in initial state and in gray for the final state. This protein changes from initial to final state in the presence of RI.

شکل 5-الف جهت تغییرات ساختاری حاصل از داده‌های تجربی در ریبونوکلاز آ. ریبونوکلاز آ در حالت اولیه به صورت تیره و در حالت نهایی به صورت روشن‌تر ترسیم گردیده و مشخصاً در حضور مهارکننده ریبونوکلاز از حالت بسته به حالت باز تغییر ساختاری می‌دهد.



**Fig. 6** unit displacement vector for Ribonuclease A. thin line with markers represents experimental crystallographic data. Dash line represents the constant stiffness model with 10 Å cutoff and thick line shows the exponential model with scaling factor 5.

شکل 6 بردار یکه جابجایی را در ریبونوکلاز آ. خطوط کم‌رنگ با شناسه نقطه از داده‌های تجربی حاصل از کریستالوگرافی می‌باشد. خط چین متعلق به مدل شبکه ارتجاعی با سختی ثابت و شعاع حدی 18 انگسترومی و خط پر مربوط به مدل نمایی با ضریب مقیاس 5 است.

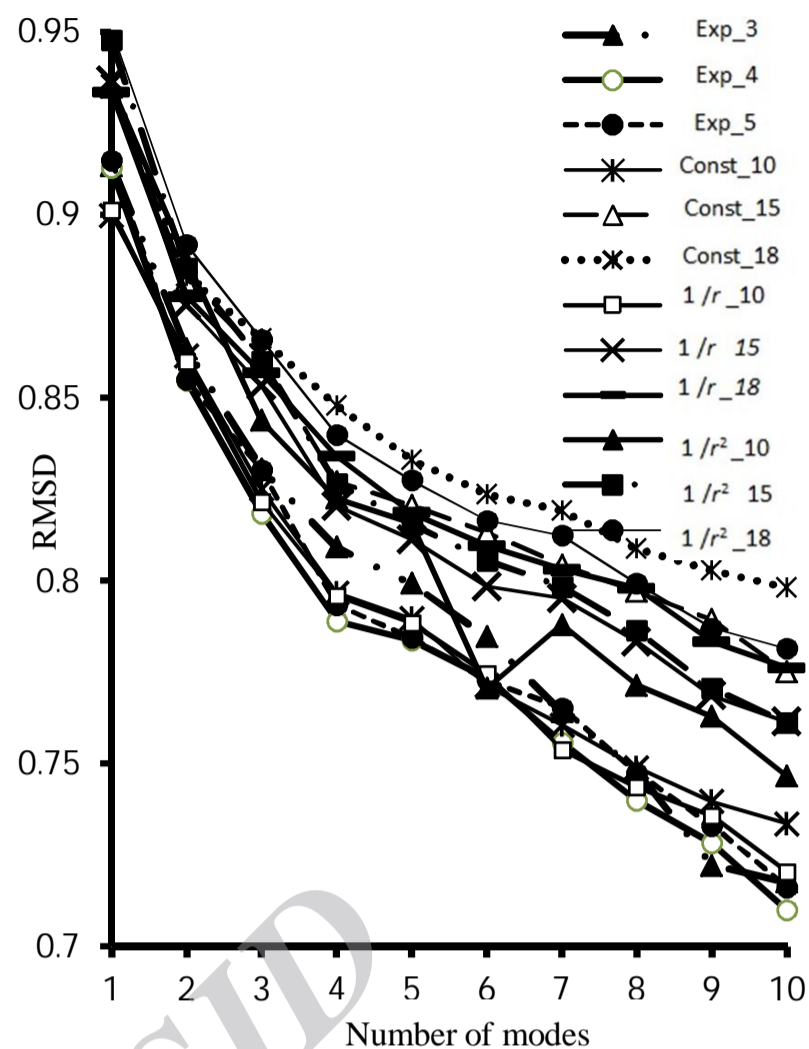
محاسبه شده و مود غالب و میزان اورلپ مربوط به هر مود محاسبه گردید. همچنین میزان RMSD ساختار دگرگون شده اولیه با استفاده از ترکیب خطی مدهای با فرکانس پایین در مقایسه با ساختار نهایی محاسبه گردید. نتایج حاصله نشان داد غالباً سه مود اول بیشترین نقش را در پیش‌بینی تغییرات ساختاری دارند. همچنین مدل‌های نمایی و مدل‌های غیرنمایی با شعاع حدی 10 بهترین پیش‌بینی را از تغییرات ساختاری ارائه می‌نمایند. البته در مدل نمایی شعاع حدی وجود ندارد که این امر هزینه محاسبات را بسیار بالا برده و از کارآمدی مدل می‌کاهد. با توجه به این امر، مدل‌های غیرنمایی (بخصوص مدل 1/2) با شعاع حدی 10 انگستروم به‌عنوان مدل بهینه در پیش‌بینی تغییرات ساختاری توصیه می‌گردد.

در مدل‌های نمایی به علت در نظر گرفتن تمامی اندرکنش‌ها بدون شعاع حدی، مدل غیر نمایی با شعاع حدی 10 انگستروم از مطلوبیت بیشتری برخوردار هستند، علاوه بر این که نتایج آن‌ها بسیار نزدیک به مدل نمایی می‌باشد.

#### 4- نتیجه‌گیری

در این مطالعه، مدل‌های مختلف شبکه ارتجاعی برای تعداد 20 پروتئین از بنچمارک داکینگ 4 با پارامترهای مختلف به‌منظور شبیه‌سازی تغییرات ساختاری در نظر گرفته شد. با استفاده از آنالیز مود نرمال و محاسبه مدهای با فرکانس پایین، میزان اورلپ 10 مود اول هر مدل با تغییرات ساختاری

- [6] S. Mahajan, Y.H. Sanejouand, On the relationship between low-frequency normal modes and the large-scale conformational changes of proteins, *Archives of Biochemistry and Biophysics*, Vol. 567, pp. 59–65, 2015.
- [7] M. Li, J. Z. H. Zhang, F. Xia, Heterogeneous elastic network model improves description of slow motions of proteins in solution, *Chemical Physics Letters*, Vol. 618, pp. 102–107, 2015.
- [8] F. Tama, Conformational change of proteins arising from normal mode calculations, *Protein engineering*, Vol. 14, No. 1, pp. 1–6, 2001.
- [9] M. Tirion, Large Amplitude Elastic Motions in Proteins from a Single-Parameter, *Atomic Analysis*, Vol. 14, No. 9, pp. 1905–1908, 1996.
- [10] A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, *Biophysical Journal*, Vol. 80, No. 1, pp. 505–15, 2001.
- [11] H. Na, G. Song, Conventional NMA as a better standard for evaluating elastic network models, *Proteins Structure Function and Bioinformatics*, Vol. 83, No. 2, pp. 259–267, 2015.
- [12] P. Doruker, R. L. Jernigan, I. Bahar, Dynamics of large proteins through hierarchical levels of coarse grained structures, *Journal of Computational Chemistry*, Vol. 23, No. 1, pp. 119–127, 2002.
- [13] A. D. Schuyler, G. S. Chirikjian, Normal mode analysis of proteins: a comparison of rigid cluster modes with C  $\alpha$  coarse graining, *Journal of Molecular Graphics and Modelling*, Vol. 22, No. 3, pp. 183–193, 2004.
- [14] A. D. Schuyler, G. S. Chirikjian, Efficient determination of low-frequency normal modes of large protein structures by cluster-NMA, *Journal of Molecular Graphics and Modelling*, Vol. 24, No. 1, pp. 46–58, 2005.
- [15] M. K. Kim, R. L. Jernigan, G. S. Chirikjian, Rigid-cluster models of conformational transitions in macromolecular machines and assemblies, *Biophysical Journal*, Vol. 89, No. 1, pp. 43–55, 2005.
- [16] S. Seo, Y. Jang, P. Qian, W. Kam, J. Choi, B. Soo, M. Ki, Journal of Molecular Graphics and Modelling Efficient prediction of protein conformational pathways based on the hybrid elastic network model, *Journal of Molecular Graphics and Modelling*, Vol. 47, pp. 25–36, 2014.
- [17] K. Hinsen, G. R. Kneller, A simplified force field for describing vibrational protein dynamics over the whole frequency range, *Journal of Chemical Physics*, Vol. 111, No. 24, pp. 10766–10769, 1999.
- [18] K. Hinsen, A. J. Petrescu, S. Dellerue, M. C. Bellissent-Funel, G. R. Kneller, Harmonicity in slow protein dynamics, *Chemical Physics*, Vol. 261, No. 1, pp. 25–37, 2000.
- [19] E. Eyal, L.-W. Yang, I. Bahar, Anisotropic network model: systematic evaluation and a new web interface, *Bioinformatics*, Vol. 22, No. 21, pp. 2619–2627, 2006.
- [20] M. F. Thorpe, Comment on elastic network models and proteins, *Physical Biology*, Vol. 4, No. 1, p. 60, 2007.
- [21] R. Soheilifard, D. E. Makarov, G. J. Rodin, Critical evaluation of simple network models of protein dynamics and their comparison with crystallographic B-factors, *Physical Biology*, Vol. 5, No. 2, 026008, 2008.
- [22] S. E. Dobbins, V. I. Lesk, M. J. E. Sternberg, Insights into protein flexibility: The relationship between normal modes and conformational change upon protein-protein docking, Vol. 105, No. 30, pp. 10390–10395, 2008.
- [23] Z. Hwang, H. Vreven, T. Janin, J. Weng, Protein-Protein Docking Benchmark 4, *Proteins: Structure, Function, and Bioinformatics*, Vol. 78, No. 15, pp. 3111–3114, 2010.
- [24] M. Delarue, Y.-H. Sanejouand, Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model, *Journal of Molecular Biology*, Vol. 320, No. 5, pp. 1011–1024, 2002.
- [25] Y.-H. Sanejouand, *Elastic network models: Theoretical and empirical foundations*, in *Biomolecular Simulations*, Biomolecular Simulations, Springer, New York, pp. 601–616, 2013.



**Fig. 7** average RMSD for the best predicted structure with bound structure with respect to number of modes considered

شکل 7 مقدار میانگین RMSD بهترین ساختار پیش‌بینی شده با ساختار نهایی بر حسب تعداد مود در نظر گرفته شده

نکته حائز اهمیت دیگر در این مطالعه این است که تاثیر شعاع حدی بسیار بیشتر از نوع تابع پتانسیل می باشد. به عنوان مثال با وجود این که مدل  $1/r$  با شعاع حدی 10 بهترین مدل در میان مدل‌های فوق می باشد، در حالت شعاع حدی 18 از جمله ناکارآمدترین مدل‌ها به حساب می آید. لازم به ذکر است مدل‌های شبکه ارتجاعی زیادی در مطالعات پیشین ارائه شده است که در این مطالعه تنها متداول‌ترین این مدل‌ها مورد بررسی قرار گرفته‌اند. گرچه این نتیجه که شعاع حدی مهم‌تر از نوع اندر کنش در مدل شبکه ارتجاعی می باشد، می تواند قابل تعمیم به مدل‌های دیگر شبکه ارتجاعی دارای شعاع حدی نیز باشد.

## 5- مراجع

- [1] M. Hyeok, B. Ho, M. Ki, Robust elastic network model: A general modeling for precise understanding of protein dynamics, *Journal of Structural Biology*, Vol. 190, No. 3, pp. 338–347, 2015.
- [2] S. J. Teague, Implications of protein flexibility for drug discovery, *Nature Reviews Drug Discovery*, Vol. 2, No. 7, pp. 527–541, 2003.
- [3] K. L. Meagher, H. A. Carlson, Incorporating protein flexibility in structure-based drug discovery: using HIV-1 protease as a test case, *Journal of the American Chemical Society*, Vol. 126, No. 41, pp. 13276–13281, 2004.
- [4] H. A. Carlson, Protein flexibility is an important component of structure-based drug discovery, *Current Pharmaceutical Design*, Vol. 8, No. 17, pp. 1571–1578, 2002.
- [5] V. Khandan, B. Firozabadi, M. Saidi, Study of inhibition of A $\beta$ 42 aggregation and oligomer forming with Docking and MD simulation, *Modares Mechanical Engineering*, Vol. 15, No. 1, pp. 39–48, 2015. (in Persian)