

Research Article

Aslı Yenenler*, Umut Gerlevik and Ugur Sezerman

Understanding the impacts of self-shuffling approach on structure and function of shuffled endoglucanase enzyme via MD simulations**Kendi üzerine gerçekleştirilen DNA karma yöntemi ile elde edilen endoglukanaz enzimlerinin yapı ve fonksiyon değişikliklerinin MD simülasyonları ile anlaşılması**<https://doi.org/10.1515/tjb-2018-0180>

Received May 8, 2018; accepted May 24, 2019

Abstract

Objective: We identify the impacts of structural differences on functionality of EG3_S2 endoglucanase enzyme with MD studies. The results of previous experimental studies have been explained in details with computational approach. The objective of this study is to explain the functional differences between shuffled enzyme (EG3_S2) and its native counterpart (EG3_nat) from *Trichoderma reesei*, via Molecular Dynamics approach.

Materials and methods: For this purpose, we performed MD simulations along 30 ns at three different reaction temperatures collected as NpT ensemble, and then monitored the backbone motion, flexibilities of residues, and intramolecular interactions of EG3_S2 and EG3_nat enzymes.

Results: According to MD results, we conclude that EG3_S2 and EG3_nat enzymes have unique RMSD patterns, e.g. RMSD pattern of EG3_S2 is more dynamic than that of EG3_nat at all temperatures. In addition to this dynamicity, EG3_S2 establishes more salt bridge interactions than EG3_nat.

***Corresponding author: Aslı Yenenler**, Department of Biostatistics and Medical Informatics, School of Medicine, Acibadem Mehmet Ali Aydınlar University, Istanbul, Turkey, e-mail: asliyenenler@gmail.com. <https://orcid.org/0000-0002-9169-388X>

Umut Gerlevik: Department of Biostatistics and Medical Informatics, School of Medicine, Acibadem Mehmet Ali Aydınlar University, Istanbul, Turkey; and Department of Molecular Biology and Genetics, Istanbul Technical University, Istanbul, Turkey, e-mail: umut.gerlevik@gmail.com

Ugur Sezerman: Department of Biostatistics and Medical Informatics, School of Medicine, Acibadem Mehmet Ali Aydınlar University, Istanbul, Turkey, e-mail: ugur.sezerman@acibadem.edu.tr

Conclusion: By taking these results into an account with the preservation of catalytic Glu residues in a proper manner, we explain the structural basis of differences between shuffled and native enzyme via molecular dynamic studies.

Keywords: Molecular Dynamics (MD) simulations; Endoglucanase; *Trichoderma reesei*; DNA shuffling; EG3 (Cel12A).

Öz

Amaç: Bir önceki deneysel çalışmamızda elde ettiğimiz EG3_S2 endoglukanaz enziminin yapısal farklılıklarının fonksiyonel etkilerini MD simülasyonları ile inceledik. Bu çalışmanın amacı DNA karma yöntemi ile elde edilen EG3_S2 endoglukanaz enzimi ile tabii EG3_nat enzimi arasındaki yapısal ve fonksiyonel farklılıkların moleküler dinamik simülasyonları ile ortaya konulmasıdır. Elde edilen fonksiyonel değişiklikleri açıklayabilecek teorik bir altyapı elde etmek bu çalışmanın amaçlarından biridir.

Gereç ve Yöntem: Bu sebeple ilk olarak; 3 farklı reaksiyon sıcaklığında koşturulan MD simülasyonları sonucunda rezidülerin esneklik analizleri, temel protein zinciri hareketi ve rezidüler arası etkileşimler incelenmiştir.

Bulgular: MD sonuçlarına göre, EG3_S2 ve EG3_nat enzimleri kendilerine özgü bir RMSD örüntüsü göstermektedir. Örnek olarak, EG3_S2 enzimi EG3_nat'e göre daha dinamik bir protein zinciri hareketine sahiptir. Bu dinamik yapıya ek olarak; EG3_S2 tabii enzime göre daha fazla tuz köprüsü oluşturmaktadır. Yani DNA karma yöntemi ile elde ettiğimiz EG3_S2 enziminin yapısal kararlılığı ve moleküler arası etkileşimleri tabii endoglukanaz enziminin farklıdır.

Sonuç: Tüm bu MD sonuçları ile birlikte katalitik Glu rezidüsünün EG3_S2'de korunduğu bilgisini göz önünde



buldurduğumuz zaman, EG3_S2'nin sahip olduğu dinamik yapı ile EG3_nat'den nasıl farklı davrandığını açıklayan bir teorik alt yapıyı elde etmiş oluyoruz. EG3_S2 ve EG3_nat'in göstermiş olduğu farklı iç dinamikler, elde ettiğimiz fonksiyonel farkların temel gerekçelerini ortaya koymaktadır.

Anahtar Kelimeler: Moleküler Dinamik simülasyonları; Endoglukanaz; *Trichoderma reesei*; DNA karma metodu; EG3 (Cel12A).

Introduction

Among various biological molecules, enzymes are considered as the most remarkable products of natural evolution due to being involved in almost every cycle of life domain. Naturally occurred enzymes are mostly limited to adapt the harsh conditions of biotechnological applications since the needs and conditions of natural reactions are different than that of biotechnologically related ones. Hence, it creates a need for the invention of new methods to generate better catalysts than natural ones [1, 2]. Among many existed methods, DNA shuffling is considered as one of promising approach, enabling us to perform directed evolution of enzymes in laboratory environment [3, 4], in terms of generation of novel enzymes [2, 5–9].

Cellulase enzymes are able to convert available cellulose and lingo-cellulose to biomass, and this particular ability makes them more attractive and promising to solve the problem about available green energy source problem in all over the world [10]. They have great potential to be used in several fields such as food, brewery and wine, animal feed, textile and laundry, biofuel, and pulp and paper industry [11, 12]. They are divided into two broad categories as endoglucanase (EG) and cellobiohydrolase (CB) in terms of mode of actions [13].

In our previous study, we performed self-shuffling approach with Cel12A (EG3) gene belonging to a fungal GH family 12 enzymes [14], being considered as one of five endoglucanase produced by *Trichoderma reesei* [13, 15]. During the self-shuffling approach, 84, 240 and 419 bp, and 80, 220 and 420 bp DNA fragments of EG3_nat reassembled to end up with EG3_S1 and EG3_S2 enzymes [2]. Due to the newly established structure and the full conservation of catalytic residues, we reported EG3_S2 as a promising catalyst than its native counterpart with enhanced enzymatic and thermal stability, and altered kinetic properties [2]. The basis of this study is to elucidate how

EG3_S2 is more promising catalyst than its native counterpart from the structural point of view.

Materials and methods

Through I-TASSER, the secondary structure of EG3_S2 enzyme was modeled [16, 17]. In crystal structure, several residues in N-terminus of EG3_nat were also missing in PDB ID: 1H8V [14], and this missing part was also modeled through I-TASSER [16, 17]. Among generated several models, the best one was selected based on C-scores [16]. Followed by the solvation and ionization of enzymes in Visual Molecular dynamics (VMD) [18], these systems were subjected to model refinement through 2 ns equilibration at 298 K, collected as NPT, after performing 20,000-step energy-minimization with Greedy Algorithm. At the end of refinements, refined models were subjected to 10,000-step energy-minimization with Greedy Algorithm, and then it was followed by 1 ns equilibration at 298 K, collected as NPT ensemble. Thirty nanoseconds production simulations were performed at three different temperatures, 313 K, 323 K and 33 K, with 2-fs time step. All simulations were repeated twice using different velocity seeds and the average results were reported for all cases, e.g. RMSD (backbone), RMSF, the evolution of salt bridges, and the radius of gyration.

All molecular dynamics simulations were performed using the NAMD [19] program with CHARMM 36 all-atom force field [20, 21] including the correction maps [22]. Water was described as the TIP3P model [23]. During the NPT simulations, the pressure was kept constant at 1 atm using Langevin pressure coupling and periodic boundary conditions were applied at all dimensions (x, y, z). To calculate the electrostatic interactions, the particle-mesh Ewald (PME) method was used [24, 25]. Through VMD, the visualization and analysis of MD trajectories, e.g. RMSD (backbone), RMSF, salt bridge interactions and radius of gyration, were performed [18].

The validation of completed EG3_nat and EG3_S2 structures were realized after modelling in I-TASSER (crude models), after 10,000-step energy minimization and after equilibration for 1 ns, separately, via SAVES v5.0 (<http://servicesn.mbi.ucla.edu/SAVES/>) by employing the results of VERIFY 3D [26, 27], ERRAT [28], PROVE [29], PROCHECK [30] and WHATCHECK [31]. VERIFY 3D compares a protein model with 3D profile of its amino acid sequences that was generated by using good quality structures. ERRAT does a statistical comparison for the non-bonded interactions between different atom types that were claimed as

distributed non-randomly. PROVE tests the models' statistical Z-score deviations from good quality structures in Protein Data Bank (PDB) database [32] by calculating the volumes of atoms in macromolecules. PROCHECK examines the stereochemical qualities of models by generating Ramachandran plots and by comparing the folding states of protein models with the crystal structures in PDB database to validate them statistically. WHATCHECK, derived from some validation algorithms of WHATIF software [33], analyses many stereochemical parameters for models.

Results and discussion

In our previous study, we stated that performing self-shuffling with EG3 (Cel12A) gene has resulted in formation of novel enzyme, EG3_S2, with altered enzymatic and kinetic properties with respect to its native counterpart [2]. Here, we aim to elucidate the structural basis of these particular differentiations that basically determine the kinetic and enzymatic properties.

First, the generated protein models were examined as mentioned in method section and the details for model verifications are provided in Table S1. These results show that VERIFY 3D and PROVE gave a good result for only EG3_nat (pdb id: 1H8V). For ERRAT and PROCHECK results, the further steps from generation to equilibration are required. Of note, there were only a very little percentage of residues in disallowed region of Ramachandran plot for all models (see Table S1). Although none of them was perfect, all models were good enough for work based on WHATCHECK results. Since these analyses are mostly based on the statistical data derived from the good quality structures in PDB database, we obtain worse scores for EG3_S2 compare to EG3_nat. By taking the adequately good quality reports for the structures of EG3_nat and EG3_S2, especially Ramachandran plot analysis as the most sensitive method for validity of a model [34], we make a decision to carry on this study for further.

Upon self-shuffling approach, we obtain EG3_S2 enzyme with novel structural reorganization. For instance, there are reported 2 α -helices and 15 β -sheets in the crystal structures of EG3_nat [14] but 3 α -helices and 16 β -sheets in EG3_S2 enzyme [2]. As in line with expectations, these particular variations result in altered functionality and stability in EG3_S2 enzyme with respect to its counterpart [2], and now we report different RMSD patterns for EG3_S2 and EG3_nat enzymes in Figure 1. As displayed in Figure 1, more dynamic RMSD pattern has been captured for EG3_S2 rather than its native counterpart at all temperatures. Actually for EG3_S2 shuffled enzyme,

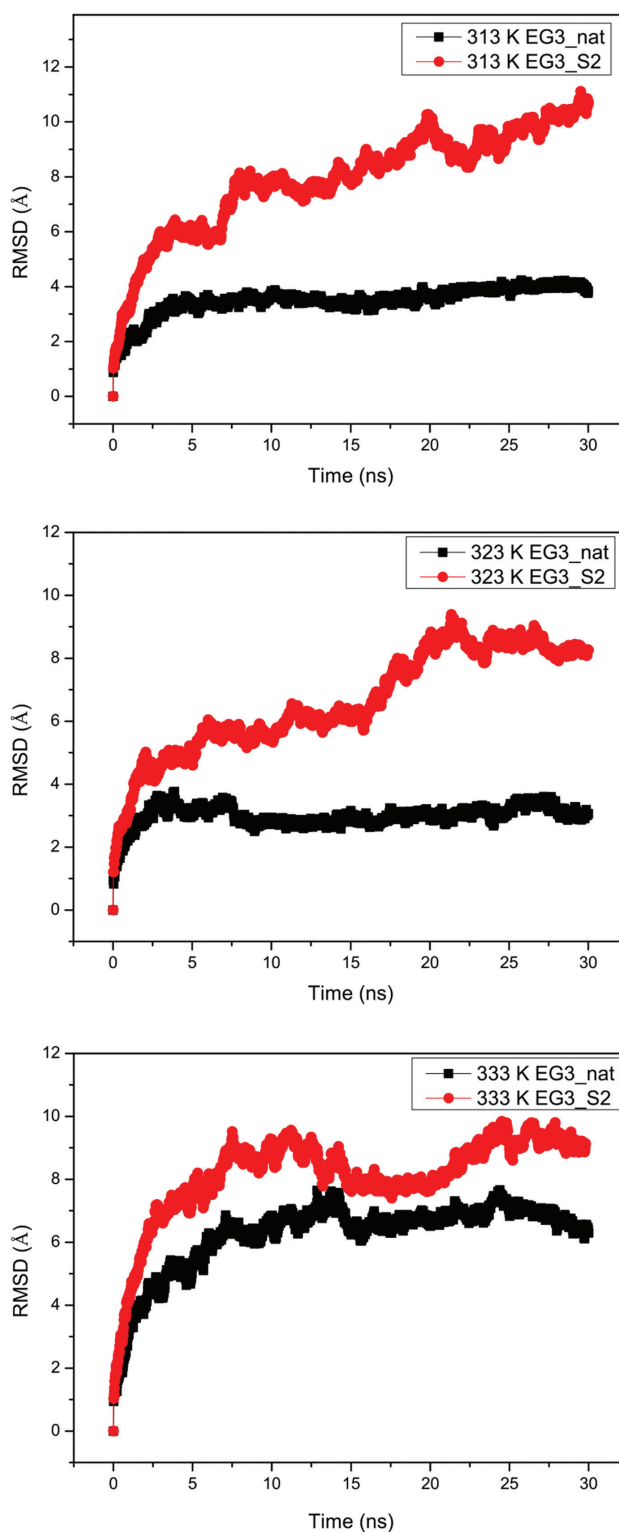


Figure 1: The RMSD patterns of the shuffled and native enzymes over MD trajectories at 313–333 K (10/K). In this figure, all RMSD results are reported as an average value over MD trajectories for all temperatures.

increase in running temperatures with 10 K intervals does not lead any drastic change in its RMSD pattern. This

fact proves us that EG3_S2 enzyme could easily adopt changes in temperature and hence more dynamic backbone response is captured along 30 ns. Even EG3_S2 and EG3_nat shared 65% identity in the level of amino acid sequence as we stated, the self-shuffling approach has resulted in formation of novel endoglucanase enzyme [2] and this particular novelty corresponds to the unique RMSD patterns (Figure 1). Since we have already known that thermal stability of EG3_S2 is better than its native counterpart based on our previous experimental study; this particular increase in RMSD pattern is not about the dissociation or unfolding of EG3_S2 but it is about the way of adopting temperature in a different way. Basically, the full or partial unfolding of enzyme structures are generally associated with an instant peak in RMSD patterns, which means ~ 2 – 3 or more fold increase in RMSD pattern within a very shorter time period ~ 0.02 – 0.08 ns [35], and RMSD patterns belonging to EG3_S2 could not be accounted as structural unfolding. Moreover, we saved the configurations of EG3_S2 along the MD trajectories, specifically at 10th ns, 20th ns and 30th ns, at different temperatures and we checked the structural integrity of EG3_S2 by performing TM-alignment. As provided in Supplementary Figures 7 and 8, the structural integrity of EG3_S2 are preserved along the trajectories at all running temperatures by obtaining TM-scores as being smaller than 5 Å. Herein, we conclude that the particular increase in RMSD pattern of EG3_S2 is not due to the loss of structural integrity as we previously suggested. As already mentioned in the method section, MD simulations have been carried out with the X-ray structure of EG3_nat with PDB ID: 1H8V [14], and 3D model of EG3_S2 generated by I-TASSER. Thus, we draw a conclusion such that running MD simulations in these conditions would be a reason for this particular difference that we clearly observe in RMSD patterns (Figure 1).

For native counterpart (EG3_nat) of shuffled enzyme, more flattered RMSD patterns are monitored along 30 ns (Figure 1). For both 313 K and 323 K, more or less same RMSD patterns are captured. Upon increase in temperature from 323 K to 333 K, RMSD values are almost doubled in a gradual manner.

Then, we focus on the RMSF patterns of EG3_S2 and EG3_nat enzymes along 30 ns. Here, we measure the RMSF values of C α -atoms in backbone rather than RMSF measurements including side-chain atoms. As in line with RMSD patterns of shuffled and native enzymes (Figure 1), the higher RMSF results, especially observed in the first 100 amino acids, are reported for EG3_S2 rather than native enzyme and it also supports the fact of more dynamic behavior of shuffled enzyme compared to EG3_nat. This particular RMSF pattern is also in line with the RMSD pattern displayed in Figure 1. Again for RMSF patterns, we suggest that running EG3_S2 as a modeled protein structure but EG3_nat as a crystal structure could lead this particular difference in the fluctuations of residues.

Since temperature intervals (10/K) are not so high; we do not observe a clear segregation in RMSF patterns of both shuffled and native enzymes but we clearly identify higher mobility in EG3_S2, accounted as its novel structural property. In Figure 2, the most interesting feature is captured in RMSF pattern of native enzyme as indicated with circle. Between Tyr76-Asn107 residues, the flexibilities increase almost up to ~ 6 -fold at 323 K compared to those at 313 K and 333 K. Here, it is crucial to point out that those RMSF values are back to those at 313 K upon increasing temperature from 323 K to 333 K, and it seems to be transient. In Figure S1, we provide the detailed structural information for Tyr76-Asn107 regions as this particular region is composed mainly β -sheet structures and is not covered the active site. Especially, Glu132 is so far from this particular region (see Figure S1).

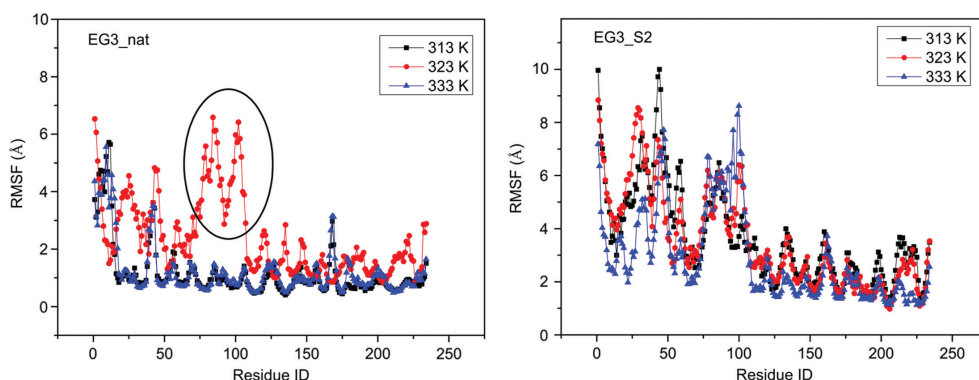


Figure 2: The RMSF patterns of the shuffled and native enzymes over MD trajectories at 313–333 K (10/K). In this figure, all RMSF results are reported as an average value over MD trajectories for all temperatures.

As stated before, the higher RMSF values are reported for EG3_S2 as its novel structural properties as being different than native one. Through the experimental studies, we have already revealed that there is an enzymatic activity of shuffled enzyme compared to its native counterpart [2]. By using modeled protein structure of the shuffled enzyme and crystal structure of native enzyme, we measured the flexibilities of particular residues, in close contact with catalytic ones, and concluded that there is a partial increase in those of EG3_S2 compared to native one [2]. This particular RMSF pattern also confirms the reported results. Previously, we explained this particular increase in EG3_S2 from the thermodynamic point of view as the flexibility of residues modulates the affinity and specificity of biomolecules through underlying energy landscapes [36]. This particular increase in flexibilities of residues results in a decrease in binding affinity of biomolecules since the binding process with increased flexibility might be associated with disorder in bound state of protein [37]. Indeed, binding affinity of protein is directly lowered via the contributions of increased flexibilities to the entropy of protein [38].

As a next step, we perform the multiple alignments of protein sequences belonging to shuffled and native sequences (Figure S2). The great variations in the amino acid level have been created between EG3_S2 and EG3_nat as provided in Figure S2. Of course, this particular variation has played a role in intramolecular interactions, e.g. salt bridge formations. Compared to EG3_S2, less salt bridge interactions are counted in EG3_nat and this fact could provide us a basis for improved thermal stability in EG3_S2 with respect to its native counterpart, which has been already demonstrated in our previous work [2]. Along 30 ns MD simulations, there exist only three salt bridge interactions in EG3_nat at all three running temperatures, e.g. Asp142-Lys139, Asp187-Lys189 and Asp197-Arg196. Upon increase in temperature from 323 K to 333 K, one more salt bridge interaction exists as Asp63-Lys2 in N-terminus of enzyme.

In Figure 3, we provide the patterns of salt bridge interactions in EG3_nat at 313 K, 323 K and 333 K by indicating the closeness of these particular interactions to catalytic Glu residues. Specifically, Asp142-Lys139 and Asp197-Arg196 interactions are more stable compared to Asp187-Lys189, which is actually more solvent exposed compared to Asp142-Lys139 and Asp197-Arg196. We only report fluctuations for Asp187-Lys189 interactions upon increasing temperature. With VMD, we search for residues in close contact with catalytic Glu132 and Glu216, ~ 5 Å, and conclude that there is no close contact residues to catalytic Glu, which is involved in any salt-bridge interactions.

As clearly displayed in Figure 3, all these listed salt bridge interactions are not closely located to catalytic site, Glu132 and Glu216. As previously mentioned, one more salt bridge interaction exists with increase in temperature to 333 K, Asp63-Lys2, in N-terminus of EG3_nat. Due to increase in provided kinetic energy to system, the conformational changes have been experienced in enzyme structure and hence residues having a potential to form salt bridge interactions, Asp63-Lys2, come into close contact. As in line with expectation, this particular Asp63-Lys2 interaction is not so stable over 30 ns trajectory (Figure S3). It becomes so strong in specific time periods, e.g. 0th–3rd ns, 5th–21th and 24th–25th ns time periods in the range of ~ 3 Å. But, the higher fluctuated peaks are introduced over the rest of trajectories as being weaker, in the range of ~ 8 –10 Å, and it displays a tendency to become weaker through the end of trajectory (Figure S3). In fact, the location of salt bridge interactions contributes to stability of enzyme structures in different manners, e.g. salt bridge interactions on surface contribute less than 1 kcal/mol but those on buried part of enzyme structure contribute more than 4 kcal/mol to stability of enzyme structures [39, 40]. If we consider both location and strength pattern of Asp63-Lys2 interaction, we could draw a conclusion that Asp63-Lys2 interaction does not contribute more to the stability of EG3_nat at 333 K.

As mentioned above, the self-shuffling approach of EG3 gene leads to generation of EG3_S2 with novel structural properties (Figure 1). The distributions of existed salt bridge interactions in EG3_S2 are also confirmed this fact as almost ~ 10 -fold more salt bridge interactions are existed at EG3_S2 on comparison to native structure for each temperature. Among these many, we just focus on Glu77-Lys76, Glu77-Lys84, Glu243-Arg242, Glu278-Arg253 and Glu335-Lys331, which are consistently existed at all temperatures. As displayed in Figure 4, the patterns of all these interactions are dynamic instead of being so stable, for instance Asp197-Arg196 interaction of EG3_nat (Figure 3). These particular patterns observed in salt bridge interactions of EG3_S2 could also verify its RMSD pattern, see Figure 1. Only for Glu278-Arg253 interaction of EG3_S2, we monitor the losing of this interaction depending on increasing temperature.

As being different than EG3_nat, both catalytic Glu residues involve in salt bridge formations; Glu224-Lys263 at 333 K, Glu224-Arg237 at 313 K and 323 K, and Glu308-Arg237 at 323 K and 333 K. In Figure 5, the evolutions of these salt bridge interactions are displayed by indicating their closeness to catalytic site of EG3_S2. Upon providing more kinetic energy to system (333 K), the conformational changes have been undergone by EG3_S2 enzyme

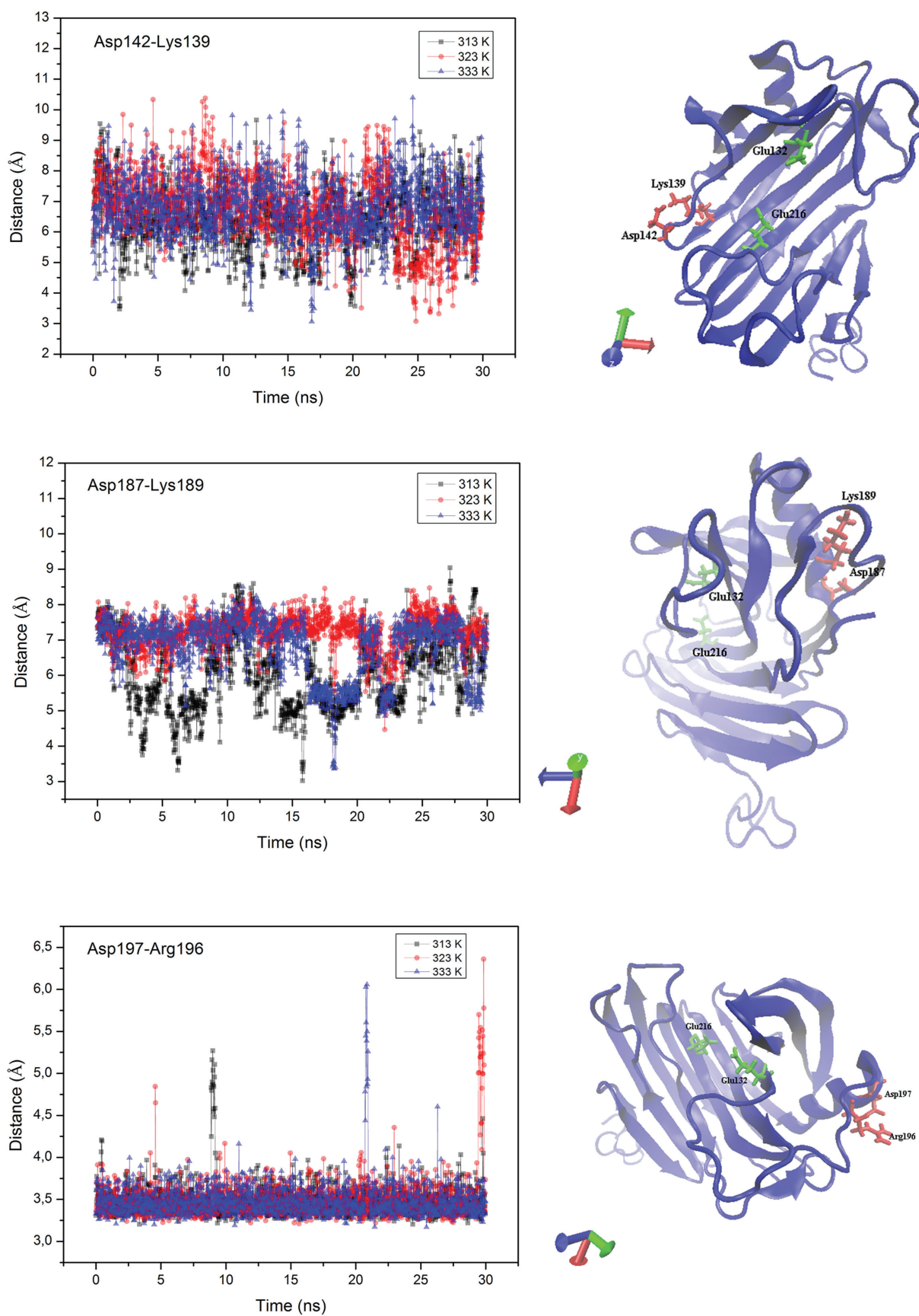


Figure 3: The evolutions of salt bridge interactions in the native enzyme along 30 ns MD trajectories. The evolutions of salt bridge interactions are reported as an average value of repeat simulations over 30 ns MD trajectories.

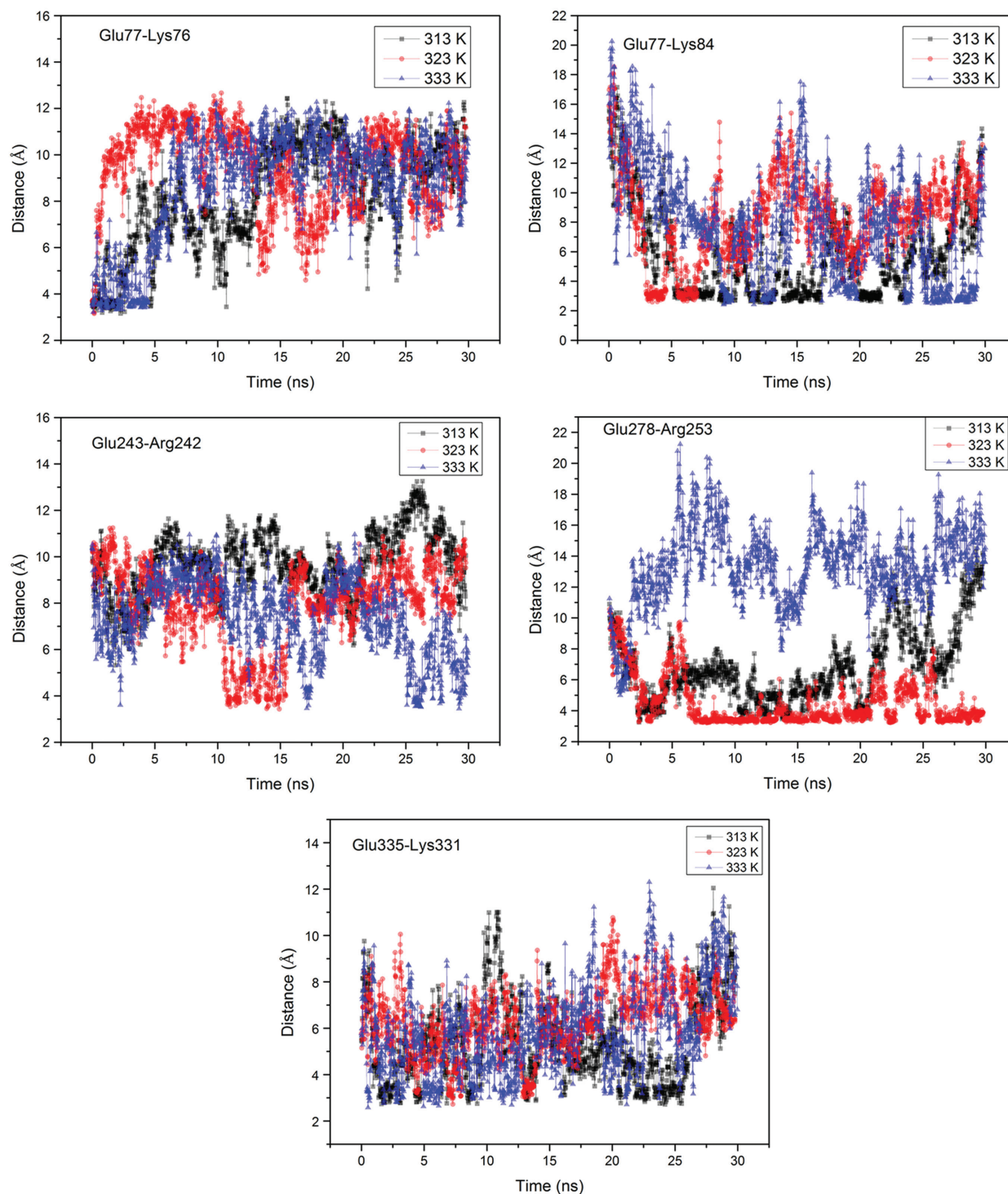


Figure 4: The evolutions of the salt bridge interactions in the shuffled enzyme along 30 ns MD trajectories.

The evolutions of salt bridge interactions are reported as an average value of repeat simulations over 30 ns MD trajectories.

and it results in movement of Lys263 close to Glu224 that leads to the formations of salt bridge interaction in middle-strength (Figure 5). Up to 5th ns of MD trajectory, this particular Glu224-Lys263 interaction is not so strong but then it reaches to ~4 Å average distance between

5th and 25th ns time periods. The similar patterns in evolution of Glu308-Arg237 and Glu224-Arg237 interactions are also reported in Figure 5. During the first 10th ns of MD trajectory, Glu308-Arg237 (323 K and 333 K) and Glu224-Arg237 (313 K and 323 K) interactions are not

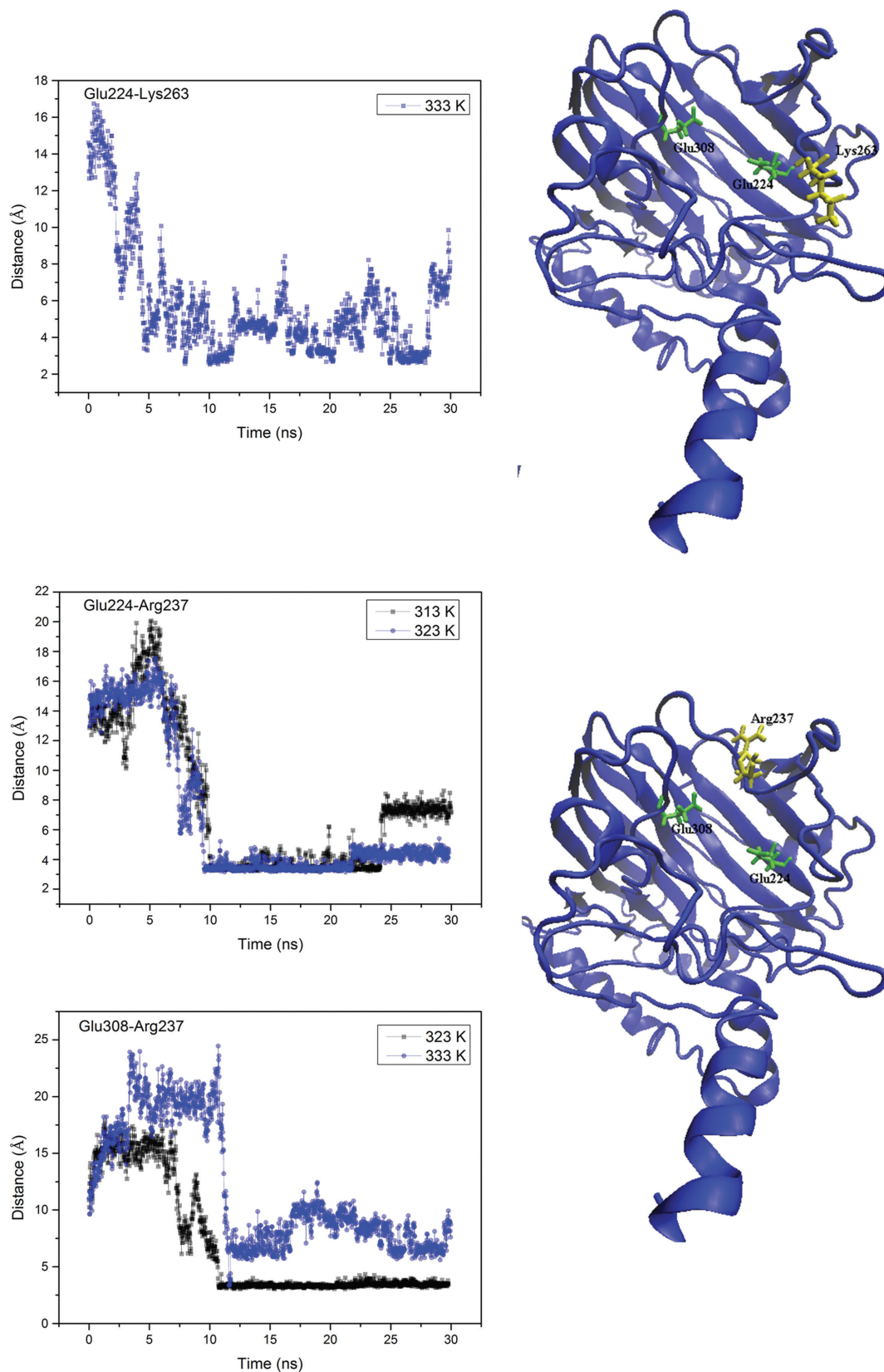


Figure 5: The evolution of salt bridge interactions formed by catalytic Glu residues in the shuffled enzyme along 30 ns MD trajectories. The evolutions of salt bridge interactions are reported as an average value of repeat simulations over 30 ns MD trajectories.

so strong but then these particular residues come closer and the stabilized and strong salt bridge interactions are established up to the end of whole trajectory and to 25th ns, respectively. For Glu308-Arg237 and Glu224-Arg237 interactions, we observe temperature-dependent loosening in thee through the whole trajectory in Glu308-Arg237 and from 10th ns to the end of trajectory in Glu224-Arg237. To sum up, these particular interactions existed in active site of EG3_S2 have demonstrated us that the active site of EG3_S2 becomes compacts through the enhancement of intramolecular interactions, even increase in reaction temperature. Actually, these particular interactions could help the preservations of catalytic residues in a proper manner required for effective chemical reaction, and thus we report EG3_S2 in our previous work as a better catalysis than its native counterpart with altered k_{cat}/k_m ratio, ~6.5-fold [2].

In fact, the adaptation of enzymes to higher temperatures is a complex issue to define, and it has been related to the Coulomb interactions, e.g. ion pairs, salt bridges and hydrogen bonds, and a set of cooperative effects, e.g. the minimization of surface energy, the hydration of nonpolar surface groups, the burying of hydrophobic residues, the optimization of core packing, weak protein-protein and protein-solvent interactions [41–43]. Reporting higher thermal stability for EG3_S2 than native in our previous work suggests that these Coulomb interactions and listed cooperative effects are arranged in a better way in EG3_S2 than its native counterpart. The differences in RMSD patterns of EG3_S2 and EG3_nat, and the evolution of existed salt bridge interactions have confirmed this fact.

Lastly, we report the radius of gyration (R_g) measurement results belonging to EG3_S2 and EG3_nat enzymes along 30 ns MD trajectories. Through R_g measurement, we implicitly assess the compactness of protein structure as the lowest R_g value corresponds to the tightest protein conformation. The results in Figure 6 tell us that EG3_nat displays the tightest conformation even increase in provided kinetic energy of system. These particular R_g patterns of EG3_S2 and EG3_nat are in line with RMSD results (Figure 1) that EG3_S2 displays more dynamic pattern, accounted as a unique structural property, with respect to its native counterpart. Here, we suggest that the tendency of EG3_S2 to deviate from its centered mass is actually due to the variations in amino acid sequences between EG3_S2 and EG3_nat (Figure S2), and their 3D distributions instead of being partially unfold. The level of variations in amino acid sequences is too high to draw a conclusion for unfolding process by taking R_g results into an account.

In sum up, we describe the differences in structural properties of EG3_S2 and EG3_nat enzymes along 30 ns

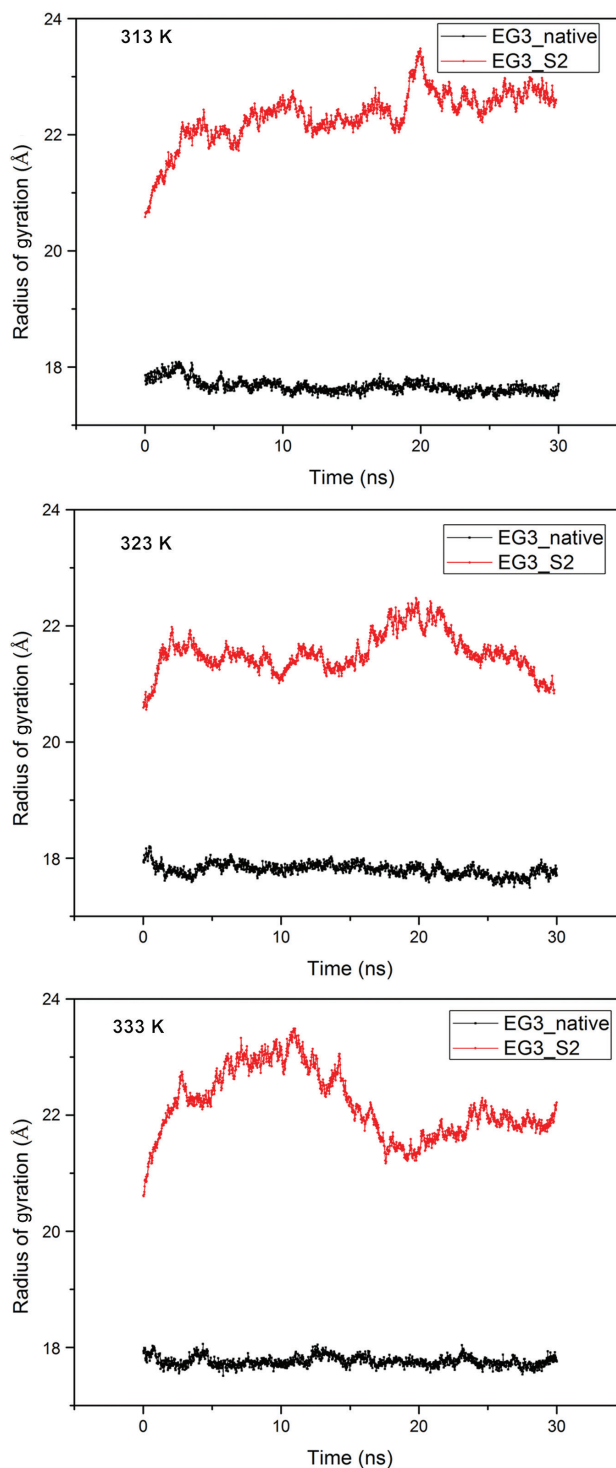


Figure 6: The radius of gyration measurement of the shuffled and native enzymes along 30 ns MD trajectories. In this figure, the radius of gyration results are reported as an average value of repeat simulations over 30 ns MD trajectories.

MD simulations. We have already known from our experimental study that self-shuffling approach leads to the formation of novel endoglucanase enzyme, EG3_S2, in terms

of enzymatic and kinetic properties. MD simulations help us to clearly elucidate how EG3_S2 is a better catalysis than its native counterpart from the structural point of view. Once again, this study demonstrates us how computational approaches are strong to support experimental results by filling the gaps from other perspectives.

Acknowledgment: The Scientific and Technological Research Council of Turkey supported this work with grant number 112T901.

Conflict of interest: The authors declare that they have no competing interest.

References

- Kurtzman AL, Govindarajan S, Vahle K, Jones JT, Heinrichs V, Patten PA. Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. *Curr Opin Biotechnol* 2001;12:361–70.
- Yenenler A, Sezerman OU. Design and characterizations of two novel cellulases through single-gene shuffling of Cel12A (EG3) gene from *Trichoderma reesei*. *Protein Eng Des Sel* 2016;29:219–29.
- Stemmer WP. Molecular breedings of genes, pathways and genomes by DNA shuffling. *Biotechnol Bioprocess Eng* 2002;7:121–9.
- Jackel C, Kast P, Hilvert D. Protein design by directed evolution. *Annu Rev Biophys* 2008;37:153–73.
- Song JK, Rhee S. Simultaneous enhancement of thermostability and catalytic activity of phospholipase A(1) by evolutionary molecular engineering. *Appl Environ Microbiol* 2000;66:890–4.
- Akbulut N, Ozturk MT, Pijning T, Ozturk SI, Gumusel F. Improved activity and thermostability of *Bacillus pumilus* lipase by directed evolution. *J Biotechnol* 2013;164:123–9.
- Cramer A, Whitehorn EA, Tate E, Stemmer WP. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol* 1996;14:315–9.
- Yano T, Oue S, Kagamiyama H. Directed evolution of an aspartate aminotransferase with new substrate specificities. *Proc Natl Acad Sci USA* 1998;95:5511–5.
- Chockalingam K, Chen Z, Katzenellenbogen JA, Zhao H. Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc Natl Acad Sci USA* 2005;102:5691–6.
- Wilson DB. Cellulases and biofuels. *Curr Opin Biotechnol* 2009;20:295–9.
- Bajpai P. Application of enzymes in the pulp and paper industry. *Biotechnol Prog* 1999;15:147–57.
- Bayer EA, Morag E, Lamed R. The cellulosome – a treasure-trove for biotechnology. *Trends Biotechnol* 1994;12:379–86.
- Ulker A, Sprey B. Characterization of an unglycosylated low molecular weight 1,4-beta-glucan-glucanohydrolase of *Trichoderma reesei*. *FEMS Microbiol Lett* 1990;57:215–9.
- Sandgren M, Shaw A, Ropp TH, Wu S, Bott R, Cameron AD, et al. The X-ray crystal structure of the *Trichoderma reesei* family 12 endoglucanase 3, Cel12A, at 1.9 Å resolution. *J Mol Biol* 2001;308:295–310.
- Nidetzky B, Steiner W, Hayn M, Claeysens M. Cellulose hydrolysis by the cellulases from *Trichoderma reesei*: a new model for synergistic interaction. *Biochem J* 1994;298:705–10.
- Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 2008;9:40.
- Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* 2015;43:174–81.
- Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* 1996;14:33–8, 27–8.
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem* 2005;26:1781–802.
- Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. *J Comput Chem* 2009;30:1545–614.
- Mackereel AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 1998;102:3586–616.
- Mackereel AD, Feig M, Brooks CL. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 2004;126:698–9.
- Jorgensen WL, Chandrasekhar J, Madura JD. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–35.
- Darden T, York D, Pedersen L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J Chem Phys* 1993;98:10089.
- Essman U, Perera L, Berkowitz M, Darden T, Lee H, Pedersen GG. A smooth particle mesh Ewald method. *J Chem Phys* 1995;103:8577–93.
- Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. *Nature* 1992;356:83–5.
- Bowie JU, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science (New York, N.Y.)* 1991;253:164–70.
- Colovos CT, Yeates O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 1993;2:1511–9.
- Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol* 1996;264:121–36.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993;26:283–91.
- Hooft RW, Vriend G, Sander C, Abola EE. Errors in protein structures. *Nature* 1996;381:272.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res* 2000;28:235–42.
- Vriend G. WHAT IF: a molecular modeling and drug design program. *J Mol Graph* 1990;8:52–6.
- Kleywegt GJ, Jones TA. Phi/psi-chology: Ramachandran revisited. *Structure* 1996;4:1395–400.
- Yenenler A, Venturini A, Burdugoglu HC, Sezerman OU. Investigating the structural properties of the active conformation BTL2 of a lipase from *Geobacillus thermocatenulatus* in toluene using molecular dynamic simulations and engineering BTL2 via in-silico mutation. *J Mol Model* 2018;24:229.

36. Chu X, Wang J. Specificity and affinity quantification of flexible recognition from underlying energy landscape topography. *PLoS Comput Biol* 2014;10:1003782.
37. Spolar RS, Record MT. Coupling of local folding to site-specific binding of proteins to DNA. *Science* 1994;263:777–84.
38. Grunberg R, Nilges M, Leckner J. Flexibility and conformational entropy in protein-protein binding. *Structure* 2006;14:683–93.
39. Marti DN, Bosshard HR. Electrostatic interactions in leucine zippers: thermodynamic analysis of the contributions of Glu and His residues and the effect of mutating salt bridges. *J Mol Biol* 2003;330:621–37.
40. Anderson DE, Becktel WJ, Dahlquist FW. pH-induced denaturation of proteins: a single salt bridge contributes 3–5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* 1990;29:2403–8.
41. Jaenicke R, Bohm G. The stability of proteins in extreme environments. *Curr Opin Struct Biol* 1998;8:738–48.
42. Kirino H, Aoki M, Aoshima M, Hayashi Y, Ohba M, Yamagishi A, et al. Hydrophobic interaction at the subunit interface contributes to the thermostability of 3-isopropylmalate dehydrogenase from an extreme thermophile, *Thermus thermophilus*. *Eur J Biochem* 1994;220:275–81.
43. Tanaka S, Igarashi S, Ferri S, Sode K. Increasing stability of water-soluble PQQ glucose dehydrogenase by increasing hydrophobic interaction at dimeric interface. *BMC Biochem* 2005;6:1.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/tjb-2018-0180>).