Research in Pharmaceutical Sciences, August 2019; 14(4): 359-368
Received: March 2019
Accepted: May 2019

School of Pharmacy & Pharmaceutical Sciences
Isfahan University of Medical Sciences

Original Article

Improving the solubility, activity, and stability of reteplase using *in silico* design of new variants

Hooria Seyedhosseini Ghaheh[1], Mohamad Reza Ganjalikhany[2], Parichehreh Yaghmaei[1], Morteza Pourfarzam[3], and Hamid Mir Mohammad Sadeghi[4,*]

[1]Department of Biology, Science and Research Branch, Islamic Azad University, Tehran, I.R. Iran.
[2]Department of Biology, Faculty of Sciences, University of Isfahan, Isfahan, I.R. Iran.
[3]Department of Clinical Biochemistry, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, I.R. Iran.
[4]Department of Pharmaceutical Biotechnology, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, I.R. Iran.

Abstract

Reteplase (recombinant plasminogen activator, r-PA) is a thrombolytic agent recombined from tissue-type plasminogen activator (t-PA), which has several prominent features such as strong thrombolytic ability and *E. coli* expressibility. Despite these outstanding features, it demonstrates reduced fibrin binding affinity, reduced stimulation of protease activity, and lower solubility, hence higher aggregation propensity, compared to t-PA. The present study was devoted to design r-PA variants with comparable structural stability, enhanced biological activity, and high solubility. For this purpose, computational molecular modeling techniques were utilized. The supercharging technique was applied for r-PA to designing new species of the protein. Based on the results from *in silico* evaluation of selected mutations in comparison to the wild-type r-PA, the designed supercharged mutant (S7 variant) exhibited augmented stability, decreased solvation energy, as well as enhanced binding affinity to fibrin. The data also implied increased plasminogen cleavage activity of the new variant. These findings have implications to therapies which involve removal of intravascular blood clots, including the treatment of acute myocardial infarction.

Keywords: *In silico* design; Reteplase; Supercharging; Thrombolysis; Tissue-type plasminogen activator.

INTRODUCTION

Reteplase (recombinant plasminogen activator, r-PA) is a recombinant form of human tissue plasminogen activator (t-PA) and is the first third-generation thrombolytic drug. This protein can be used for removal of clots in blood vessels by converting plasminogen to plasmin which triggers fibrinolysis in plasma. Therefore, it is utilized in the treatment of acute myocardial infarction a major cause of mortality worldwide.

Structurally, r-PA consists of two domains, including serine protease and kringle-2, with a total of 355 amino acids and molecular weight of 39 kDa (1). The active site of r-PA, the same as other serine protease enzymes, consists of the classic triplet, Ser, Asp, and His, acting as the nucleophile agent, the acidic agent and the catalytic place, respectively. Equivalent residues in the active site of r-PA are Ser-306, Asp-199, and His-150 (2). The catalytic domain also contains four specificity pockets, S1-S4, conferring the remarkable specificity to the enzyme in its plasminogen cleavage. S4 is the largest in size, with its Trp-346 residue being the largest group and composing the entrance of the active site (3). The kringle-2 domain is implicated to contain a lysine-binding site which is used for binding to fibrin, an essential cofactor for t-PA-mediated plasminogen activation (4).

Compared to t-PA, r-PA presents several outstanding features making it suitable as a thrombolytic drug.

*Corresponding author: H. Mir Mohammad Sadeghi
Tel: +98 3137927059, Fax: +98-3136680011
Email: h_sadeghi@pharm.mui.ac.ir

Access this article online

Website: http://rps.mui.ac.ir

DOI: 10.4103/1735-5362.263560

Since the carbohydrate side chains on the three potential N-glycosylation sites of r-PA are not required for its function, it can be expressed in *Escherichia coli* in non-glycosylated form (5). Due to the deletion of finger domain, epidermal growth factor, and kringle-1 regions as well as the carbohydrate side chains, the plasma half-life of r-PA increases when compared to full length t-PA. It has also other prominent properties such as strong thrombolytic ability, reduced hepatic clearance, and low molecular weight in comparison with t-PA. However, the stimulation of protease activity by the kringle-2 domain in the presence of fibrin is known to be lower in r-PA than in t-PA. Its binding affinity to fibrin is also lower than that of t-PA (about 5 folds) due to the deletion of the fibronectin finger region (6). Another limitation, particularly emerging in reteplase production in large quantities, is its relatively poor solubility which leads to its misfolding into inclusion bodies during expression/purification (3). Therefore, it is imperative to find strategies to simultaneously tackle these multiple challenges in usage of r-PA as a therapeutic agent.

In general, protein aggregation is a well-known culprit in pharmaceutical biotechnology. In this regard, increasing the net charge of protein surface (supercharging) has been suggested to help resolve this challenge by shifting the isoelectric point of proteins and reinforcing polar interactions with water molecules (7). The present study aimed to take this approach, presuming its efficacy in preserving or improving the potency of the enzyme in plasminogen activation and fibrin binding (8). Because of their contribution to the cost-effective development of protein drugs, the use of computational techniques, such as *in silico* design, molecular dynamic (MD) simulations, molecular docking and physicochemical predictions are proposed to investigate the protein modifications to enhance their desired features (9). Here, these methods are utilized to design supercharged mutants, with the objective of examining the stability, solubility, and activity of designed r-PA variants in comparison to currently used r-PA form.

MATERIALS AND METHODS

Homology modeling

The three dimensional (3D) structure of r-PA has not yet been determined experimentally, hence we used homology modeling protocol to predict its structure. This method was performed by Modeller v9.18 software using the crystal structure of t-PA catalytic domain (PDB ID: 1BDA, with 100.0% sequence identity to that domain in our r-PA sequence) and kringle-2 domain (PDB ID: 1PML, with 100.0% sequence identity to that domain in our r-PA sequence) as templates. Among 1000 generated models, we chose the best model in terms of the lowest discrete optimized protein energy (DOPE) score, and the quantitative evaluation of selected models was carried out using PROCHECK web tool (10) and VADAR algorithm (11). In order to protect the r-PA activity, the highly conserved and functional residues to be involved in the proper folding and function of r-PA were identified by multiple alignments and excluded from mutagenesis.

In silico engineering of supercharged reteplase

For increasing the net charge of the protein, we used the Supercharge tool of Rosie server implemented in Rosetta-Commons (12). Supercharged r-PA was generated by considering the isoelectric pH (pI) of the protein, which was calculated using PDB2PQR version 2.1.1 (http://agave.wustl.edu/pdb2pqr).
Simultaneously, net charge of the protein surface was increased as suggested by AvNAPSA method. To avoid manipulation of important functional residues in the active site of the enzyme, server-suggested mutations with up to 7.5 Å distance from the active site of the enzyme were excluded, using the Swiss PDB Viewer program version 4.1.0, because sequences in the proximity of the active site residues are highly conserved in the protease family (13). Suggested mutations from each method were introduced

to r-PA structure by Rosetta Backrub server (http://kortemmelab.ucsf.edu/backrub).

In the next step, the stability of the designed mutants was evaluated by PoPMuSiC and Eris servers. The PoPMuSiC program introduces the possible single-site mutations in a protein structure, predicting the resulting folding free energy changes ($\Delta\Delta G_{computed} = \Delta G_{mutant} - \Delta G_{wild\text{-}type}$) by utilizing database-derived potentials (14). For more accuracy, we also used the Eris server to calculate the change of the protein stability induced by mutations utilizing the recently developed Medusa modeling suite (15). In order to improve the stability prediction accuracy, the backbone structure of the wild-type protein was allowed to be relaxed before prediction, by choosing pre-relaxation option.

Based on modifications suggested by the Supercharge tool of Rosie server, and after excluding mutations in vicinity of the active site, the net charge of r-PA surface was increased from +1e to +5e (in mutation N113R, F288R, E212K, and I149K) and to +7e (in mutation N113R, A284R, E291R, E212K, and F288R). The +7e supercharged r-PA suggested by Rosetta program was termed S7 variant. The net charge of the protein surface was also increased by AvNAPSA method, from +1e to +5e (in mutation Q82K, D270K, and E191K) and to +7e (in mutation E191K, E213K, Q267K and D270K).

Molecular dynamic simulations and analyses

Molecular dynamic simulations were carried out by Gromacs (version 4.5) software package using the g43a1 force filed (16). The system was solvated with a simple point charge (SPC216) model of water. The particle mesh Ewald method was adopted for handling the long-range electrostatic interactions at cutoff distance of 10 Å (17). A linear constraint solver (LINCS) algorithm was applied to constrain the bonds involving the hydrogen atom. The 3D structure of wild type or mutant r-PA models were solvated in a solvation box with 10 Å distance between the edges of the box and the protein surface. For this purpose, a triclinic box, applying 3D periodic boundary conditions was selected. The correct number of negative ions (Cl^-) was added for neutralizing the positively-charged system. At first the system was energy minimized, using steepest descent and then conjugate gradient algorithms, to remove steric hindrance due to added hydrogen atoms. Position restraint procedure was conducted in association with NVT (for 500 ps) and NPT (for 1000 ps) ensembles in two phases, to stabilize pressure and temperature. MD simulation was performed for 50 ns. Time steps of NVT, NPT, and MD simulations were 0.002 ps. To calculate the Λ parameter, MD simulations were also implemented at three temperatures (i.e. 300 K, 350 K, and 400 K). Overall, fifteen molecular dynamics simulations of r-PA were performed. Root mean squared deviation (RMSD), root mean squared fluctuation (RMSF), radius of gyration (Rg), average solvent accessible surface area (ASA), and number of hydrogen bonds were calculated via Gromacs analysis tools. The $F_{overall}$ quantity which is known as a measure of overall flexibility of protein was calculated using RMSF values based on following formula, where '*i*' is the residue number (18):

$$F_{overall} = \left[\frac{1}{n}\sum_{i=1}^{n} rmsf_i^2\right]^{1/2} \quad \ldots\ldots (1)$$

The stability was also computed and compared between variants using the Λ parameter which is an inverse measure of order/rigidity in macromolecular structures (19), and is calculated as follows:

$$\Lambda = \frac{d\ln(1-S^2 \times 0.89)}{d\ln T} \qquad (2)$$

where, S known as the order parameter, describes the orientational fluctuations of the backbone NH vectors and is obtained in squared form (S^2) from output of MD simulations in low, medium and high temperatures. Linear regression is used to extract Λ values, and the quality of the fit is measured by the regression coefficient R^2 (19).

Based on the simulation output, solvation/hydration energy (ΔG_{sol}) of proteins was also calculated by using Adaptive Poisson-Boltzmann solver (APBS) method via the non-linearized Poisson Boltzmann

equation. The solvation energy was calculated as the sum of polar and nonpolar energies obtained by the program (20).

Molecular docking of reteplase to fibrin

For enhancing the activation of plasminogen, t-PA has shown high specificity to fibrin over fibrinogen. Because of the subtle difference between the two proteins, it has been hypothesized that the conversion of fibrinogen to fibrin causes a conformational change, exposing binding sites for t-PA (21). To perform docking, we obtained the crystal structure of fibrinogen from PDB (ID: 3GHG) and removed its β and γ chains in order to have the α-chain alone with all residues exposed. Docking of α-chain to native and mutant r-PA models was carried out by HADDOCK (http://haddock.science.uu.nl/services/HADDOCK/haddockserver-easy.html), a popular program which follows a true data-driven strategy (22). As the input to HADDOCK, fibrin residues 148-160 (KRLEVDIDIKIRS) and r-PA residues 37-127 (the Kringle-2 domain) which lied on the protein surface (according to their ASA values), were defined as active residues.

RESULTS

Structure of wild type and supercharged variants of reteplase

Ramachandran plot statistics of the best homology model of r-PA i.e. that with the lowest DOPE score indicated that 80.1%, 16.0%, 2.7%, and 1.2% of residues are located in the most favored regions, additional allowed areas, generously allowed areas and disallowed areas, respectively (Fig. 1A). Analysis of secondary structures using VADAR algorithms represented 7% helix, 36% β strands, 56% coil, and 27% turn. The C-terminus of r-PA has a high content of β-sheets, while the N-terminus is highly coiled and is presumed to be disordered. All modified residues locate on sequence regions with low conservation, which ensures the protein function not to be disrupted (Fig. 1B).

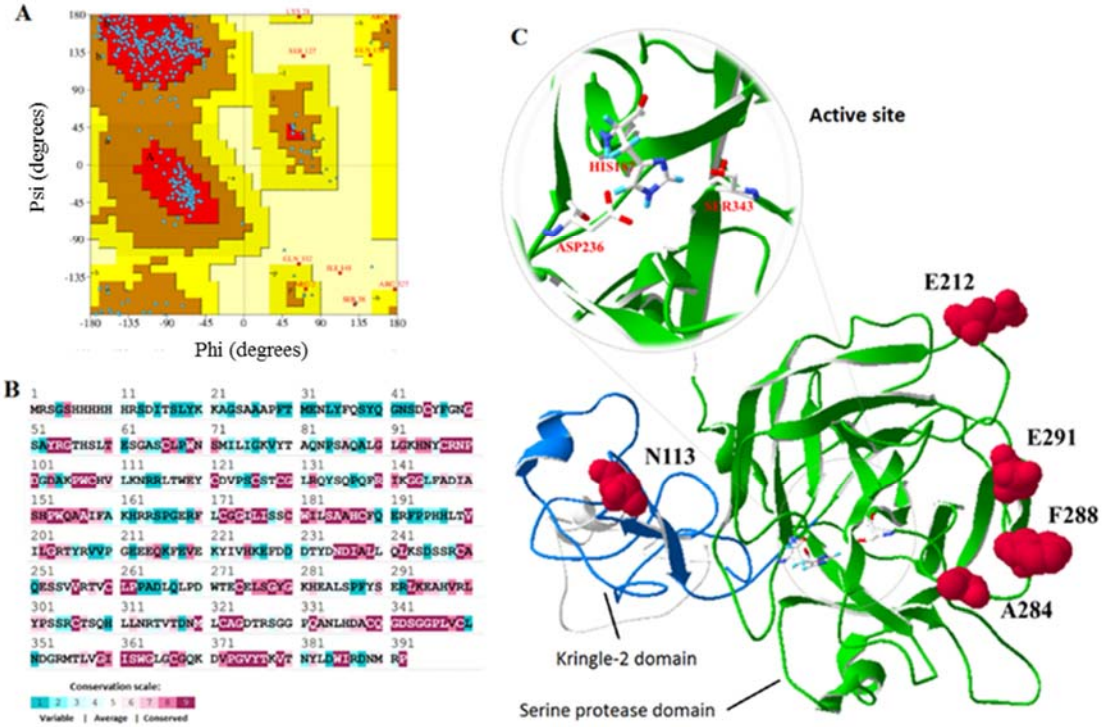

Fig. 1. Validation and display of the designed r-PA variant. (A) The Ramachandran plot of the best homology model of wild type r-PA; (B) amino acid sequence of wild type r-PA, color-coded by level of conservation; and (C) overall view of the Reteplase structure and its active site (magnified). Residues subject to mutation in S7 supercharged variant are represented by red spheres. Image rendered by PyMol (www.pymol.org). r-PA, Reteplase.

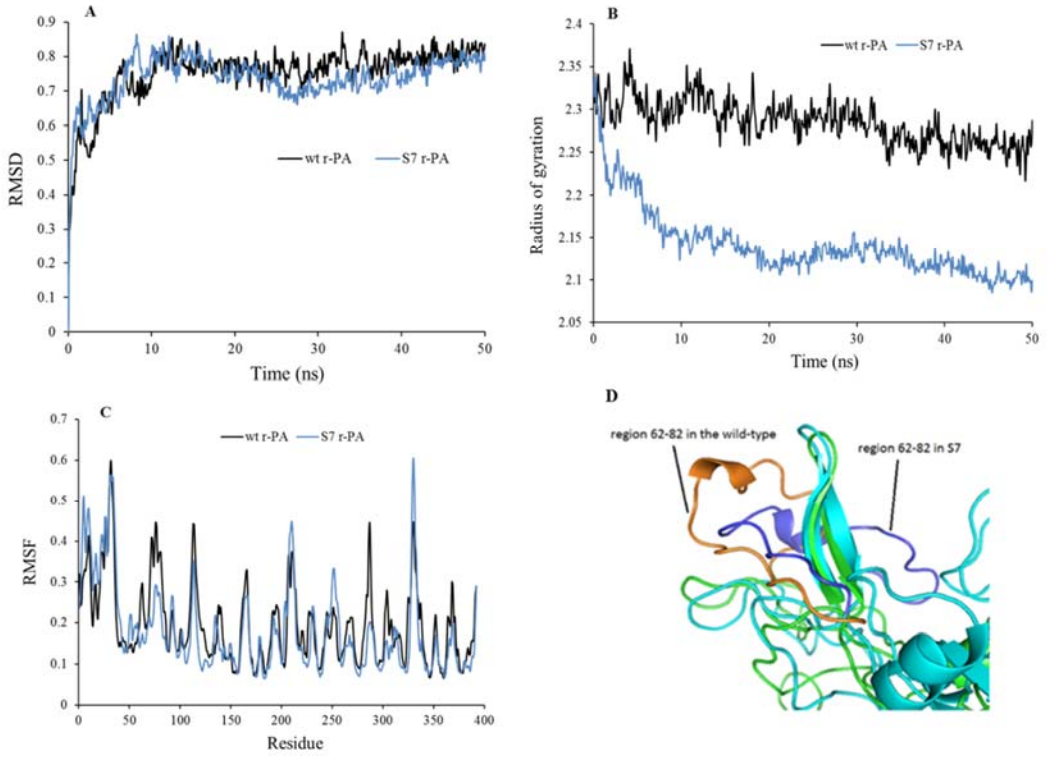

Fig. 2. Dynamics of r-PA S7 variant. (A) RMSD plot, (B) the radius of gyration plot, and (C) RMSF plot for the molecular dynamic simulation of wild type and supercharged variants. (D) Superimposed kringle-2 domains of the two variants, illustrating the flexible area of this domain. Image rendered by PyMol (www.pymol.org). r-PA, Reteplase; RMSD, root mean squared deviation; RMSF, root mean squared fluctuation; wt, wild type.

Molecular dynamics

RMSD plots drawn for +5e-AvNAPSA, +7e-AvNAPSA, +5e-Rosetta, and +7e-Rosetta supercharged variants in 300, 350, and 400 K confirmed higher/comparable stability of the selected variant S7. RMSD plots of wild type and S7 r-PA variants (in 300 K) are compared in Fig. 2A, indicating equilibration of the system for both variants achieved after about 15 ns followed by comparable fluctuations for both systems until the end of simulation time. Rg as a measure of the structural compactness was demonstrated in Fig. 2B. As can be seen, S7 variant is more compact, which can result from enhanced hydration due to the stronger interaction of charged groups with water molecules, compared to neutral/polar ones in the wild type r-PA. To investigate whether this compact state affects the enzyme's active site, average accessible surface area of the active site residues from the last 25 ns of the simulation were calculated and showed almost unchanged ASA of the active site for the S7 variant compared to that of wild type (5.317 ± 0.3 nm^2 *vs* 4.799 ± 0.17 nm^2, respectively). The same comparison for Trp-383 (equivalent to Trp-346 of the S4 specificity pocket) also showed that the exposed state of this residue is preserved in S7 compared to wild type (0.21 ± 0.06 nm^2 and 0.155 ± 0.1 nm^2, respectively).

Local structural fluctuations illustrated by RMSF plots demonstrated no remarkable difference between wild type and S7 variants in the serine protease domain, while a small region in kringle-2 (residues 62-82) with a coil-helix-coil structure adjacent to the N113R mutation site is conferred some degree of rigidity in the S7 variant compared to the wild type r-PA (Fig. 2C and D). Computation of $F_{overall}$ for the second half of the simulation time indicated lower overall flexibility of S7 than wild type ($F_{overall}$ of 0.35 and 0.38, respectively),

however high fluctuations limiting to the region 62-82 shows that the observed overall rigidity actually occurs in a local scale, rather than globally. This region does not contribute to fibrin binding by the kringle-2 domain. Active site residues maintain their stable status in both variants, with all last 25-ns RMSF values being less than 1 Å. Among mutation sites of S7, the only remarkable change compared to wild type was seen in 288 position, where the mutation caused reduced fluctuation of the corresponding loop, probably because of higher hydration and hydrogen bonds with water molecules of the Arg residue in S7 compared to Phe in the wild type. The hydrogen bonds inside the protein structure and those between protein and solvent were plotted in Fig. 3 and showed that the r-PA variants maintain their overall hydrogen-bonding pattern after introducing the mutations.

Stability of designed reteplase variants

Stability check with PoPMuSiC and Eris servers showed negative values for the folding free energy (ΔΔG) for all supercharged variants. The Λ parameter computed for wild type and the four supercharged variants indicated higher stability for the S7 variant (Table 1). Average number of intramolecular hydrogen bonds calculated for the last 25 ns of molecular dynamic simulations also implies comparable degree of stability for all variants (Table 2). Counted number of salt bridges in the last frame of simulation (9 for S7, *vs* 5 for wild type) is an additional indication of enhanced enthalpic stability for the S7 variant of r-PA.

Solubility of designed reteplase variants

Calculated solvation energy of mutants showed the most negative free energy of solvation for the S7 variant, indicating highly significantly increased solubility of this species compared to all other variants (Table 1). The number of protein-solvent H-bonds calculated for the last 25 ns of MD simulation was unchanged/higher in S7, partly explaining the calculated solvation free energy. In fact, introduced substitutions in the surface of r-PA accompanied remarkable expansion of hydrophilic and significant shrinkage of hydrophobic surface area of the protein (Table 2).

Based on the stability and solubility results, the S7 mutant species was selected for further analyses and functional investigations. The candidate sites of mutation as suggested for this variant are illustrated in the constructed 3D model of r-PA in Fig. 1C.

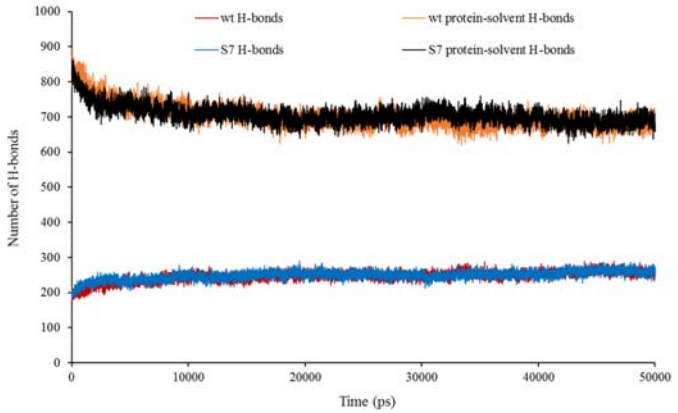

Fig. 3. Number of protein-protein and protein-solvent hydrogen bonds during the time of simulation.

Table 1. Calculated lambda parameter, its model fit measure, and free energy of solvation for reteplase variants.

Variants	Λ (K^{-1})	R^2	$\Delta G_{solvation}$ (kJ/mol)
Wild type	0.0071	0.9932	-19268.0398 ± 170.21
+5e r-PA by AvNAPSA	0.009	0.9761	-19381.8048 ± 138.776
+5e r-PA by Rosetta	0.0088	0.9908	-22206.7260 ± 197.707
+7e r-PA by AvNAPSA	0.0084	0.9896	-19426.4235 ± 199.274
+7e r-PA by Rosetta (S7)	0.0065	0.9188	-25050.2398 ± 150.872

Table 2. Average number of hydrogen bonds and solvent accessible surface area during the last 25 ns of molecular dunamic simulation for reteplase variants.

Variants	Intramolecular (protein-protein)	Intermolecular (protein-solvent)	Hydrophobic Surface	Hydrophilic surface
Wild type	254.14 ± 9.90	672.77 ± 17.28	122.97 ± 2.02	29.85 ± 1.05
+5e r-PA by AvNAPSA	267.1 ± 9.30	671.26 ± 16.13	125.67 ± 1.78	30.80 ± 1.15
+5e r-PA by Rosetta	244.47 ± 8.69	716.87 ± 16.98	121.02 ± 1.73	31.83 ± 1.13
+7e r-PA by AvNAPSA	252.95 ± 9.09	680.42 ± 18.23	124.97 ± 2.01	31.27 ± 1.17
+7e r-PA by Rosetta (S7)	254.03 ± 9.86	697.32 ± 18.27	118.69 ± 1.89	33.43 ± 1.08

Table 3. Parameters computed for protein-protein docking between fibrin and reteplase variants.

Parameters	Fibrin-wild type r-PA complex	Fibrin-S7 r-PA complex
HADDOCK score	-77.0 ± 10.5	-103.8 ± 9.1
Cluster size	8	4
RMSD from the overall lowest-energy structure	31.3 ± 0.7	0.8 ± 0.5
Van der Waals energy	-58.6 ± 1.5	-63.2 ± 5.2
Electrostatic energy	-209.7 ± 20.6	-415.6 ± 43.8
Desolvation energy	2.3 ± 11.0	26.9 ± 7.5
Restraints violation energy	213.0 ± 95.83	156.6 ± 64.24
Buried surface area	1543.9 ± 33.5	2196.8 ± 170.8
Z-Score	-1.5	-2.8

r-PA, Reteplase; RMSD, root mean squared deviation.

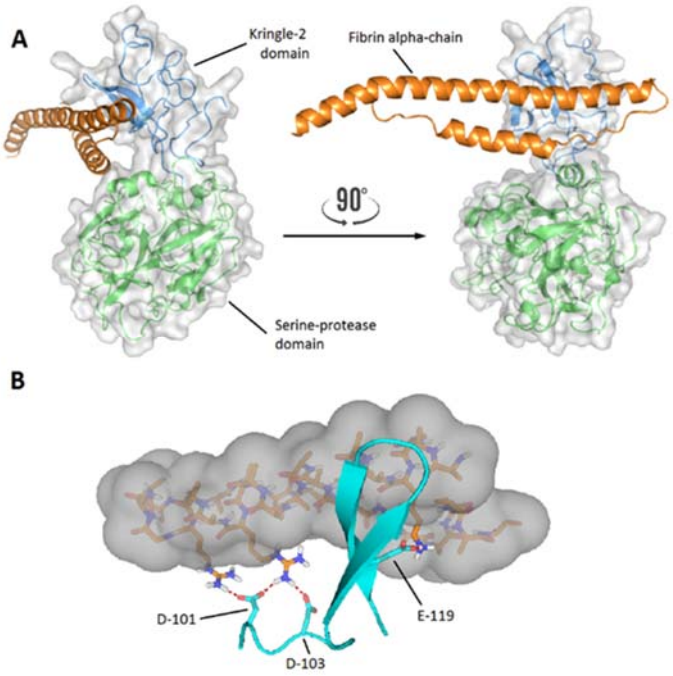

Fig. 4. Interaction of reteplase S7 with fibrin. (A) The complex of reteplase S7 and the α-chain of fibrin, (B) highly-conserved negatively-charged residues of kringle-2 domain (cyan ribbon) interacting with Lys/Arg residues from fibrin (surface-represented backbone). Images rendered by PyMol (www.pymol.org).

Docking results

For fibrin interaction to each of the mutant and native forms of r-PA, HADDOCK clustered several structures, outputting the docking scores, and energies (Table 3). The docking indicated that the interface of α-chain of fibrin and lysine-binding site of kringle-2 is energetically the most favorable site for the binding (Fig. 4A). Interestingly, the docking score as well as both electrostatic and van der Waals components of the binding energy for the fibrin/S7 complex showed more favorable values than the fibrin/ wild type r-PA, while electrostatics energies play

the major role. The ionic bonds involve Lys/Arg residues from fibrin and expectedly Asp/Glu ones from r-PA. The kringle-2 domain contains only three conserved negatively-charged residues (Fig. 1B), all of which participate in the columbic interactions with fibrin α-chain (Fig. 4B).

DISCUSSION

Designing of globular proteins with more solubility is a substantially challenging task. This is why the researches have so far been conducted to improve the activity and solubility/aggregation resistance of r-PA, as well as its production yield (1,5,23) have gained limited success. Because of the high potential of computational tools in the area of protein engineering, we adopted these design methods to investigate whether the promising supercharge approach can decrease the aggregation propensity of r-PA by enhancing its solubility, as well as to examine its capability to retain/improve the enzyme biological function and stability.

Common strategies to improve solubility during protein expression include the use of weak promoters, low IPTG concentration, low temperature (24), modified growth media (25), co-expression with molecular chaperons (26), fusion with solubility enhancing tags (27), and use of various expression hosts (28). Unfortunately, these experimental studies have had low yield of protein solubility and require denaturation/refolding processes, hence being cost-intensive and time-consuming. In this study, increasing the net charge of protein surface, as a feature against aggregation, was hypothesized as an approach to this end. This method was used in the field of aggregation resistance of proteins. The study by Der *et al*. on variants of green fluorescent protein showed the supercharge protocol as an effective approach for charge-based improvement for refolding (29). There have been also several other successful applications of the supercharging technique to enhance solubility and refolding yield (30,31). Applying the same technique in this study showed the significant effect of designed mutations on the free energy of solvation for the engineered r-PA species. This is probably due to mutation of suggested sites to arginine, an effective residue in preventing of aggregation, improving the solubilization, and increasing the purification yield of proteins (32,33).

The supercharging approach has shown to have a remarkable positive contribution to the fibrin binding activity of the enzyme. We examined the binding energy of designed mutant r-PA variants to fibrin, due to the importance of their binding for formation of a ternary complex of fibrin, plasminogen and the activator. Study of the fibrin binding through molecular docking revealed more favorable interaction of the supercharged species with fibrin, compared to the wild type, which may have a positive implication regarding the function of the designed supercharged r-PA enzyme. The increased flexibility of the region 62-82, which is probably induced by the adjacent N113R substitution, may provide explanation for this result, as the specificity of protein-ligand interactions is now widely accepted to result largely from variable loops, connecting secondary structure elements, especially those near the binding sites (34).

Small increase of the solvent-accessible surface of the active site for the S7 variant against that of wild type can provide better structural accommodation in the active site for plasminogen cleavage activity. In addition, retained exposed state of the key Trp residue of the S4 specificity pocket, as well as the preserved flexibility of the active site residues despite the more compact structure of S7, imply to preserved or even higher plasminogen cleavage activity of the serine protease domain of r-PA. Furthermore, the activity may be enhanced by the reinforced surface charge. In fact, a few clusters of water molecules (presumably bound to charged groups on the enzyme surface) have been claimed to be required for enzymatic activity, but they are stripped by H_2O molecules in the bulk water, thereby lowering the catalytic function (8). Accordingly, supercharging could result in higher activity through increasing the surface hydration, especially considering that four out

of five suggested mutations in the S7 variant locate on the serine protease domain.

The supercharge technique can also augment the protein stability as shown by other works. Miklos *et al*. applied this technique to structural design of thermo-resistant antibodies, and reported this method is useful in enhancing the resistance to thermal inactivation due to induced aggregation, as well as in reinforcing the antibody affinity to antigen (35). Stability analyses in our study, using the wild type as reference, showed comparable or higher thermodynamic (ΔG) and thermal stability of the supercharged r-PA. In addition, the number of intermolecular and protein-solvent hydrogen bonds did not significantly change for S7 compared to the wild type, which demonstrated preserved stability of the whole system.

CONCLUSION

By insightful utilization of homology modeling, MD simulations, and molecular docking, a model for designing new r-PA variants capable of avoiding aggregation and improving function was developed, with the aim of yielding a better recombinant protein production on overexpression in *E. coli*. Among generated r-PA supercharged species, the S7 variant was highly capable of improving both solubility and activity as well as maintaining sustainability *in silico*. The observations for this variant, i.e. preserved stability, supercharging-mediated intramolecular, and solvent-induced structural changes which result in improved plasminogen cleavage activity, enhanced binding affinity to fibrin, as well as the significant augmentation of r-PA solubility, have positive implications to design of fibrinolytic protein drug of r-PA, with the purpose of treating acute myocardial infarction. *In vitro* assessment of stability, solubility, and activity for new r-PA variants is under way in our laboratory.

ACKNOWLEDGEMENTS

The authors thank Dr. Karim Mahnam, the faculty member of Shahrekord University, Shahrekord, I.R. Iran, for his masterly technical guidance and Dr. Fatemeh Shafiee, the faculty member of Isfahan University of Medical Sciences, Isfahan, I.R. Iran, for her helpful review of the manuscript.

REFERENCES

1. Shafiee F, Moazen F, Rabbani M, Mir Mohammad Sadeghi H. Expression and activity evaluation of Reteplase in Escherichia coli TOP10. J Param Sci. 2015;6(3):58-64.
2. Kumar A, Pulicherla KK, Ram KS, Rao KR. Evolutionary trend of thrombolytics. Int J Bio Sci Bio Technol. 2010;2(4):51-68.
3. Renatus M, Bode W, Huber R, Stürzebecher J, Prasa D, Fischer S, *et al*. Structural mapping of the active site specificity determinants of human tissue-type plasminogen activator. Implications for the design of low molecular weight substrates and inhibitors. J Biol Chem. 1997;272(35):21713-21719.
4. Hudson NE. Biophysical mechanisms mediating fibrin fiber lysis. BioMed Res Int. 2017;(2017). Article ID 2748340, 17 pages.
5. Mandi N, Sundaram KR, Tandra SK, Bandyopadhyay S, Padmanabhan S. Asn^{12} and Asn^{278}: critical residues for *in vitro* biological activity of reteplase. Adv Hematol. 2010;(2010). Article ID 172484, 9 pages.
6. Kohnert U, Rudolph R, Verheijen JH, Weening-Verhoeff EJ, Stern A, Opitz U, *et al*. Biochemical properties of the kringle 2 and protease domains are maintained in the refolded t-PA deletion variant BM 06.022. Protein Eng. 1992;5(1):93-100.
7. Lawrence MS, Phillips KJ, Liu DR. Supercharging proteins can impart unusual resilience. J Am Chem Soc. 2007;129(33):10110-10112.
8. Adamczak M, Krishna SH. Strategies for improving enzymes for efficient biocatalysis. Food Technol Biotechnol. 2004;42(4):251-264.
9. Chang CC, Song J, Tey BT, Ramanan RN. Bioinformatics approaches for improved recombinant protein production in *Escherichia coli*: protein solubility prediction. Brief Bioinform. 2013;15(6):953-962.
10. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993;26(2):283-291.
11. Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, *et al*. VADAR: a web server for quantitative evaluation of protein structure quality. Nucleic Acids Res. 2003;31(13):3316-3319.
12. Moretti R, Lyskov S, Das R, Meiler J, Gray JJ. Web-accessible molecular modeling with Rosetta: the Rosetta online server that includes everyone (ROSIE). Protein Sci. 2018;27(1):259-268.

13. Brenner S. The molecular evolution of genes and proteins: a tale of two serines. Nature. 1988;334(6182):528-530.
14. Dehouck Y, Kwasigroch JM, Gilis D, Rooman M. PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. BMC Bioinformatics. 2011;12(1):151-162.
15. Yin S, Ding F, Dokholyan NV. Eris: an automated estimator of protein stability. Nat Methods. 2007;4(6):466-467.
16. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, *et al*. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013;29(7):845-854.
17. Darden T, York D, Pedersen L. Particle mesh Ewald: An N.log (N) method for Ewald sums in large systems. J Chem Phys. 1993;98(12): 10089-10092.
18. Basu S, Sen S. Do homologous thermophilic-mesophilic proteins exhibit similar structures and dynamics at optimal growth temperatures? A molecular dynamics simulation study. J Chem Inf Model. 2013;53(2):423-434.
19. Zeiske T, Stafford KA, Palmer AG. Thermostability of enzymes from molecular dynamics simulations. J Chem Theory Comput. 2016;12(6):2489-2492.
20. Cerutti DS, Jain T, McCammon JA. CIRSE: A solvation energy estimator compatible with flexible protein docking and design applications. Protein Sci. 2006;15(7):1579-1596.
21. Medved L, Nieuwenhuizen W. Molecular mechanisms of initiation of fibrinolysis by fibrin. Thromb Haemost. 2003;89(03):409-419.
22. Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc. 2003;125(7):1731-1737.
23. Khodabakhsh F, Dehghani Z, Zia MF, Rabbani M, Sadeghi HM. Cloning and expression of functional reteplase in *Escherichia coli* top10. Avicenna J Med Biotechnol. 2013;5(3):168-175.
24. Ghosh S, Rasheedi S, Rahim SS, Banerjee S, Choudhary RK, Chakhaiyar P, *et al.* Method for enhancing solubility of the expressed recombinant proteins in *Escherichia coli*. Biotechniques. 2004;37(3):418-423.
25. Prasad S, Khadatare PB, Roy I. Effect of chemical chaperones in improving the solubility of recombinant proteins in Escherichia coli. Appl Environ Microbiol. 2011;77(13):4603-469.
26. Trésaugues L, Collinet B, Minard P, Henckes G, Aufrère R, Blondeau K, *et al*. Refolding strategies from inclusion bodies in a structural genomics project. J Struct Funct Genomics. 2004;5(3): 195-204.
27. Costa S, Almeida A, Castro A, Domingues L. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. Front Microbiol. 2014;5:63-82.
28. Aghaabdollahian S, Rabbani M, Ghaedi K, Sadeghi HM. Molecular cloning of reteplase and its expression in *E. coli* using tac promoter. Adv Biomed Res. 2014;3:190.
29. Der BS, Kluwe C, Miklos AE, Jacak R, Lyskov S, Gray JJ, *et al*. Alternative computational protocols for supercharging protein surfaces for reversible unfolding and retention of stability. PLoS One. 2013;8(5):e64363.
30. Simeonov P, Berger-Hoffmann R, Hoffmann R, Sträter N, Zuchner T. Surface supercharged human enteropeptidase light chain shows improved solubility and refolding yield. Protein Eng Des Sel. 2010;24(3):261-268.
31. Mosavi LK, Peng ZY. Structure-based substitutions for increased solubility of a designed protein. Protein Eng. 2003;16(10):739-745.
32. Tsumoto K, Umetsu M, Kumagai I, Ejima D, Philo JS, Arakawa T. Role of arginine in protein refolding, solubilization, and purification. Biotechnol Prog. 2004;20(5):1301-1308.
33. Das U, Hariprasad G, Ethayathulla AS, Manral P, Das TK, Pasha S, *et al.* Inhibition of protein aggregation: supramolecular assemblies of arginine hold the key. PLoS One. 2007;2(11):e1176.
34. Shehu A, Kavraki LE. Modeling structures and motions of loops in protein molecules. Entropy. 2012;14(2):252-290.
35. Miklos AE, Kluwe C, Der BS, Pai S, Sircar A, Hughes RA, *et al*. Structure-based design of supercharged, highly thermoresistant antibodies. Chem Biol. 2012;19(4):449-455.