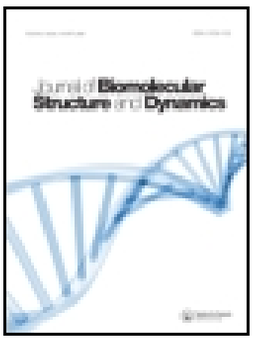


Taylor & Francis
Taylor & Francis Group

Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: http://www.tandfonline.com/loi/tbsd20

Computational simulations assessment of mutations impact on streptokinase (SK) from a group G streptococci with enhanced activity - insights into the functional roles of structural dynamics flexibility of SK and stabilization of SK-μplasmin catalytic complex

Faegheh Kazemi, Seyed Shahriar Arab, Nasir Mohajel, Malihe Keramati, Niloofar Niknam, Mohammad Mehdi Aslani & Farzin Roohvand

To cite this article: Faegheh Kazemi, Seyed Shahriar Arab, Nasir Mohajel, Malihe Keramati, Niloofar Niknam, Mohammad Mehdi Aslani & Farzin Roohvand (2018): Computational simulations assessment of mutations impact on streptokinase (SK) from a group G streptococci with enhanced activity - insights into the functional roles of structural dynamics flexibility of SK and stabilization of SK-μplasmin catalytic complex, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2018.1472668

To link to this article: https://doi.org/10.1080/07391102.2018.1472668

Publisher: Taylor & Francis

Journal: *Journal of Biomolecular Structure and Dynamics*

DOI: http://doi.org/10.1080/07391102.2018.1472668

Computational simulations assessment of mutations impact on streptokinase (SK) from a group G *streptococci* with enhanced activity - insights into the functional roles of structural dynamics flexibility of SK and stabilization of SK-µplasmin catalytic complex

Running head: Insights into the mutations impact on the structure of a SK variant with enhanced activity by computational simulations

Faegheh Kazemi[1,4], Seyed Shahriar Arab[2], Nasir Mohajel[1], Malihe Keramati[3], Niloofar Niknam[2], Mohammad Mehdi Aslani[4], Farzin Roohvand[1]*

1: Virology department, Pasteur Institute of Iran, Tehran, Iran.
2: Biophysics Department, School of Biological Sciences, Tarbiat Modares University (TMU), Tehran, Iran.
3. Nano-biotechnology department, Pasteur Institute of Iran, Tehran, Iran.
4. Microbiology department, Pasteur Institute of Iran, Tehran, Iran.

***Corresponding Author:**
Farzin Roohvand (Ph.D),
Role: Supervisor - design and supervision of the project, proof reading and submission of the manuscript.

Virology department, Pasteur Institute of Iran (IPI),No. 69,Pasteur Ave, Tehran, Iran.

Post Code**:** 1316943551

E-mail: farzin.roohvand2@gmail.com

rfarzin@pasteur.ac.ir

Tel/Fax: + 98 21 66496682,

ORCID ID: 0000-0002-9597-3555

Faegheh Kazemi

Role: Ph.D. candidate - performing all the computation studies, analyzing the whole data, writing the manuscript draft.

ORCID ID: 0000-0001-7968-0456

Tel: 00982166405535

E-mail: faeghehkazemi1@gmail.com

Dr Seyed Shahriar Arab

Role: Supervisor - MD simulations results and analyzing, assistance in drafting the manuscript.

ORCID ID: 0000-0002-2739-1617

Tel: 00982182883494

E-mail: sh.arab@modares.ac.ir

Dr Nasir Mohajel:

Role: Advisor - RINs results and analyzing, assistance in drafting the manuscript.

ORCID ID: 0000-0001-8883-4429

Tel: 00982164112250

E.mail: n_mohajel@pasteur.ac.ir

Dr Malihe Keramati:

Role: Advisor - developing the idea, assistance in drafting the manuscript.

ORCID ID: 0000-0002-7196-8751

Tel: 00982164112168

E-mail: keramati.malihe@gmail.com

Niloofar Niknam:

Role: Assistance in performing MD simulations and analyzing the data.

ORCID ID: 0000-0001-9835-7021

Tel: 00989178382245

E-mail: nniknam82@gmail.com

Dr Mohammad Mehdi Aslani:

Role: Supervisor- Initiation of the project.

ORCID ID:0000-0002-3013-6940

Tel: 00982164112248

E-mail: mmaslani@yahoo.com

Abstract:

Streptokinase (SK), a plasminogen activator (PA) that converts inactive plasminogen (Pg) to plasmin (Pm), is a protein secreted by groups A, C and G *streptococci* (GAS, GCS and GGS, respectively), with high sequence divergence and functional heterogeneity. While roles of some residual changes in altered SK functionality are shown, the underlying structural mechanisms are less known. Herein, using computational approaches, we analyzed the conformational basis for the increased activity of SK from a GGS (SKG132) isolate with four natural residual substitutions (Ile33Phe, Arg45Gln, Asn228Lys, Phe287Ile) compared to the standard GCS (SKC). Using the crystal structure of SK-Pm catalytic complex as main template SKC.µPm catalytic complex was modeled through homology modeling process and validated by several online validation servers. Subsequently, SKG132.µPm structure was constructed by altering the corresponding residual substitutions. Results of three independent MD simulations showed increased RMSF values for SKG132.µPm, indicating the enhanced structural flexibility compared to SKC.µPm, specially in 170 and 250 loops and three regions: R1 (149-161), R2 (182-215) and R3 (224-229). In parallel, the average number of Hydrogen bonds in 170 loop, R2 and R3 (especially for Asn228Lys) of SKG132 compared to that of the SKC was decreased. Accordingly, residue interaction networks (RINs) analyses indicated that Asn228Lys might induce more level of structural flexibility by generation of free Lys256, while Phe287Ile and Ile33Phe enhanced the stabilization of the SKG132.µPm catalytic complex. These results

denoted the potential role of the optimal dynamic state and stabilized catalytic complex for increased PA potencies of SK as a thrombolytic drug.

Keywords:

Activity; Molecular dynamics simulation; Plasminogen; Residue interaction network; Streptokinase

Abbreviations:

- Cα - C-alpha
- RMSD - Root mean square deviation
- RMSF - Root mean square fluctuation
- Hbonds - Hydrogen bonds
- RINs - Residue interaction networks
- RING - Residue interaction network generator
- VDW - van der Waals
- MD simulation – Molecular dynamics simulation
- Pg – Plasminogen
- Pm – Plasmin
- μPm – microplasmin
- SK – Streptokinase
- PAs - Plasminogen activators
- tPAs - Tissue type-PAs
- K – kringle
- R – Region
- PDB - Protein Data Bank

1. Introduction

Thrombotic disorders are resulted from blood clots during a pathologic blood coagulation process. In such conditions, the usual medical intervention is treatment with plasminogen activators (PAs) like streptokinase (SK) (Roohvand, 2017; Zhalyalov, Panteleev, Gracheva, Ataullakhanov, & Shibeko, 2017). PAs convert the endogenous human plasminogen (Pg: the inactive proenzyme) to plasmin (Pm: the active serine protease), which cleaves the fibrin

network of the blood clot (Klegerman, 2017). Recent crystallographic data on Pg indicated five kringle domains (K1-K5) and a serine protease catalytic, so called microplasmin (μPm) domain along this 791 amino acids protein (Law et al., 2012).

SK has no intrinsic enzymatic activity and its mechanism of Pg activation relies on a series of protein-protein interactions that ends up with the formation of the active SK.Pm complex (Boxrud, Verhamme, & Bock, 2004; Verhamme and Bock, 2014).

SK is a 47 kDa protein consisting of a single chain, 414 amino acids composed of three structurally autonomous sequential domains (SKα 1-146, SKβ 147- 290 and SKγ 291-414) connected by two flexible coil regions (Parrado et al., 1996). Elucidation of the crystal structure of the SK.Pm complex (Wang, Lin, Loy, Tang, & Zhang, 1998) and various mechanistic studies by deletion and/or peptide walking, permitted the identification of several functional regions or exosites across the domains of the SK (Aneja, Datt, Yadav, & Sahni, 2013) including: *i*) The residues 1-59 of SKα (specially Asp41–His48 region) that play roles in binding and activation of the Pg (Kim, Lee, Kim, Kim, & Byun, 2000; Bean, Verhamme, & Bock, 2005; Sazonova, Robinson, Gladysheva, Castellino, & Reed, 2004); *ii*) The Val158-Arg219 region and the 250 loop (including the pair of Lys 256, Lys 257) of SKβ for SK-Pm complex formation (Sazonova, et al., 2004) and Pg recognition and processing (Chaudhary et al., 1999; Tharp et al., 2009), respectively and *iii*) The coiled region of SKγ (Leu314-Ala342) that play roles in Pg activation by stabilizing the SK.μPm catalytic complex (Wu et al., 2001). In addition, hydrophobic interactions between SK and μPm might play significant role on generation of a stable complex which is essential for PA activity (Wu, et al., 2001). Apart from functional regions and exosites, several functionally important, critical single residues (so called "hot spots") in SK structure are identified (Yadav et al., 2011). For example, substitutions like V19F, V35E and S44K in the

SKα-domain contributed to the conformational activation of the SK-Pm complex (Kim et al., 2000). Conversely, alanine substitution of Pro177 and Lys180 (Aneja, Datt, Singh, Kumar, & Sahni, 2009) as well as charge-reversal mutations of two Lys residues at the tip of 250 loop (K256E, K257E) showed a drastically decline on PA properties of SK (Chaudhary, et al., 1999). Interestingly, E253R substitution showed additive/synergistic influence on Pg activation, apparently due to the facilitation of the Pg substrate recognition and processing by SK (Tharp, et al., 2009). Application of the gained structure-function information for rational engineering of SK has resulted to invention of superior SKs with improved Pg activation properties (Arabi et al., 2011; Keramati, Arabi Mianroodi, Memarnejadian et al., 2013a). However, besides such mutation and mechanistic studies, the presence of heterogeneous enzyme families with evolutionary substituted/mutated residues towards altered functionality/activity, might be a natural source for identification of the hot spots and functional mechanisms. This is exactly the case with SKs secreted by various beta-hemolytic *streptococci* (*βHS*) groups A, C and G. Indeed, SKs are heterologous proteins even within the isolates of the same group, showing different protein sequences and PA potencies (McArthur et al., 2008). Of note, differences in PA potencies of the *streptococcal* culture supernatants (SCS) have been traditionally used to identify superior producer strains (Keramati, Roohvand, Aslani, Motevalli, & Memarnejadian, 2013b; Keramati, Roohvand, et al., 2013c). Accordingly, various assays were developed to correlate the heterogeneity of SK protein sequences and their PA potencies in defined categories (Keramati et al., 2012; Zhang, Mayfield, Ploplis, & Castellino, 2014). Recently, kinetic studies on SK from a group G *streptococci* (SKG132) isolate with four natural substitution/mutation (Ile33Phe, Arg45Gln, Asn228Lys and Phe287Ile) indicated around two times increased PA potencies compared to that of the standard SK (SKC) (Keramati, Aslani, Khatami, & Roohvand, 2017).

Although some or all of these substitutions might have potentially contributed as hot spots to the increased activity of SKG132, but the exact role of each residue substitution and the functional mechanism was not known.

Computing power and theoretical modeling of protein-ligand has emerged as a valuable tool in understanding biological functions and enzyme mechanisms (Broomhead and Soliman, 2017). Availability of molecular dynamics (MD) simulation algorithms for accurate *in-silico* analyses in recent years, provided relevant three dimensional (3D) structural-computational models of enzymatic systems (Culka, Gisdon, & Ullmann, 2017). MD simulation also computes the trajectory of a system including coordinates and velocities of atoms using "Newtonian equations of motion" which provides atomic level information on biological macromolecules (Dror, Dirks, Grossman, Xu, & Shaw, 2012). Therefore it is widely employed in structure-based drug discovery projects (L. Chen et al., 2012) and studies on the impact of mutations on the protein structure at the atomic level (Krishnamoorthy, Gajendrarao, Olivotto, & Yacoub, 2017; Sutthibutpong, Rattanarojpong, & Khunrae, 2017; Thirumal et al., 2017; Zhao, Yang, Xiang, Gao, & Zeng, 2017). In addition, availability of other complementary tools like residue interaction networks (RINs) which use the MD simulation data, provide the detailed analyzes of the residue interactions in the enzyme complex at the atomic level (Piovesan, Minervini, & Tosatto, 2016). Indeed, the computer-generated predictions have currently changed the classic view of proteins from simple static macromolecules to more dynamic structures and have proposed new strategies for protein engineering by optimizing the dynamic state of several enzyme structures (Craveur et al., 2015; McAuley and Timson, 2017).

Despite the availability of the crystal structure of SK.μPm catalytic complex (Wang, et al., 1998) for MD simulation applications, only few studies addressed the computational approaches to

investigate the mechanism and the exact role of SK residues for Pg activation. Results of these prior studies implied the potential importance of some functional loops and flexible regions of SK (170 and 250 loops) as well as several single residues (Asp173, Asp174, and Arg176) for PA potencies (Yadav, Datt, Singh, & Sahni, 2008; Aneja, et al., 2013), but the underlying mechanisms and the role of the SK critical residues in the SK-Pm complex are not fully understood.

In the present study, using computational approaches, we attempted to gain insights about the atomic basis of the increased activity of SKG132 and the role of the natural residual substitutions (Ile33Phe, Arg45Gln, Asn228Lys, Phe287Ile) to that of the standard SK (SKC) which further our knowledge on structural-functional enzymatic properties of this protein. To the best of our knowledge, this is the first MD simulation and RINs study on structure of SK.μPm complex.

2. Material and methods

Construction and refinement of the predicted models

SK sequence from *Streptococcus equisimilis*, group C, ATCC 9542 (SKC-2; protein_id: AAC60418.1) which is known as a standard strain for production of SK (Arabi, et al., 2011) was selected as the main sequence for analyses (hereafter: SKC). This SK was already used for kinetic analyses comparing the activity of SKG132, too (Keramati, et al., 2017). To construct the structure of the model, multiple template approach was used in the homology modeling process. To this end, the sequence of SKC was downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/protein/AAC60418.1) and BLASTP search was performed against the Protein Data Bank (PDB; https://www.rcsb.org) to find suitable templates. Due to the availability of SK.μPm crystal structure (i.e.: SK-Pm catalytic complex; residues 1 to 372 of SK and 1 to 250 of

μPm) with PDB code of 1BML at 2.9 Å resolution (Wang, et al., 1998), this PDB was used as the main template for molecular modeling and construction of the SKC.μPm structure. Coordinate of this complex structure shares more than 87% sequence identity with SKC (https://www.ebi.ac.uk/Tools/psa/). However, due to the lack of the first 11 N-terminal amino acids and some other residues in the β domain (residues 175-181 and 252-262) of SK in the SK.μPm crystal structure (1BML), two other available PDBs of SK (PDB codes: 1L4Z at 2.8 Å resolution with 81.1% identity with SKC) (Wakeham et al., 2002) and (1QQR at 2.3 Å resolution with 95.7% identity with SKC) (Wang, Tang, Hunter, & Zhang, 1999) were also used as the inputs in the MODELLER v.9.13 (Webb and Sali, 2014) to fill-in the missing residues in 1BML. The best constructed model was selected based on the lowest DOPE (Discrete Optimized Protein Energy score which is a statistical potential optimized for model assessment) calculated based on the standard MODELLER energy function. Residual substitutions for the SKG132 (Ile33Phe, Arg45Gln, Asn228Lys, Phe287Ile) were performed using YASARA v.11.11.2 (Krieger and Vriend, 2014). Reliability of the predicted models was evaluated using several online validation servers including: RAMPAGE (http://mordred.bioc.cam.ac.uk/rapper/rampage.php) (Lovell et al., 2003), MolProbity (http://molprobity.biochem.duke.edu/) (Chen et al., 2010), ProSA (https://prosa.services.came.sbg.ac.at/prosa.php) (Wiederstein and Sippl, 2007) and Verify-3D (http://servicesn.mbi.ucla.edu/Verify3d/) (Luthy, Bowie & Eisenberg, 1992). Visualization of the modeled structures was achieved by the Pymol program (DeLano, 2002).

Molecular Dynamics (MD) simulation

The dynamic behavior of the modeled structures (SKG132.μPm and SKC.μPm) were studied by MD simulation via GROMACS v. 5.0.7 program package (Abraham et al., 2015), using the GROMOS96 53a6 force field with 2-fs time steps during 100 nanoseconds (ns) for three

repeated times (Oostenbrink, Villa, Mark, & Van Gunsteren, 2004). To this end, each modeled structure (SKG132.µPm and SKC.µPm) was centered in a cubic box and SPC water molecules were added with a minimum spacing of 10.0 Å from the protein surface atoms and simulation box edge. To eliminate boundary effect complications, three-dimensional periodic boundary conditions (PBCs) were considered. To neutralize simulation system charges, water molecules were replaced with Na^+ counter ions and energy minimization was executed through the steepest descent integrator. The maximum force on each system was set to 100 kJ/mol. The equilibration was performed in two steps, nvt and npt, for 2 ns long each and with time step of 2 femtosecond (fs). The electrostatic interactions were calculated by the particle mesh Ewald (PME) method (Oostenbrink, et al., 2004) with a cutoff value of 12 Å. LINCS algorithm was applied to constrain all bond lengths (Hess, Bekker, Berendsen, & Fraaije, 1997) and neighbor searching was performed every 10 steps. Temperature was kept constant at 300 K using v-rescale method for both steps (Bussi, Donadio, & Parrinello, 2007). Berendnsen algorithm was used for pressure coupling and was kept at 1 bar in npt step (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984). Coupling time for both pressure and temperature was equal to 2 ps. All parameters were kept similar to npt step except that Parrinello-Rahman was used as pressure coupling algorithm (Parrinello and Rahman, 1981). Production runs with 2-fs time steps during 100 nanoseconds (ns) were repeated three times to increase the accuracy of simulation.

Trajectory analyses

In order to investigate variation of the structural features for SKG132.µPm compared to that of SKC.µPm, several parameters such as: root mean square deviation (RMSD) of C-alpha (Cα) atoms (Cα-RMSD), root mean square fluctuation (RMSF) of Cα atoms (Cα-RMSF) as well as occupancy and number of hydrogen bonds (Hbonds) were calculated by gmx rms, gmx rmsf and

gmx hbond gromacs built-in tools, respectively. Number of Hbonds were determined while a donor-acceptor distance was less than 0.35 nanometer (nm) and a donor-hydrogen-acceptor angle was 30 degrees.

Analysis of the residue interaction networks (RINs)

To construct the residue interaction networks (RINs), the average structure derived from the last 50 ns trajectory of each protein complex structure (SKC.μPm and SKG132.μPm) was submitted to the Residue Interaction Network Generator (RING) 2.0.1 web server (Piovesan, et al., 2016). The non-covalent interactions, including van der Waals (VDW), Hbond, salt bridge, pi-pi and pi-cation were calculated according to the RING manual (Piovesan, et al., 2016). Residue interaction maps of the protein complex structures were developed using the RINalyzer plug-in of Cytoscape platform (Shannon et al., 2003) and finally visualized using RING-Viz script for Pymol (Shannon, et al., 2003).

3. Results and discussion

Structure validation of the predicted models

The best constructed model was selected based on the lowest DOPE score (-31236.52). The DOPE score profile was calculated/plotted using assess_dope function of MODELLER (Fig. S1). As shown in figure S1, the superposed model profile on the template profile gave an idea about the quality of input alignment which were almost similar to each other. To validate the quality of the modeled structures, several online validation methods like Ramachandran plot, ProSA and Verify-3D plot were used. Ramachandran maps were generated using RAMPAGE Ramachandran plot assessment (Lovell, et al., 2003). Results of the Ramachandran plot for the best predicted model and the template indicated that 90%, 7.3% and 2.7% of all the residues

were in the favored, allowed and outlier regions for the modeled structure, respectively (Fig. S2A and Table 1) and 81%, 13.6% and 5.4% for the template, respectively (Fig. S2B and Table 1). These results showed that only very few residues could be found in disallowed regions, indicating (validating) the quality of the predicted structures (Figures S2A & S2B and Table 1). In addition, evaluation by MolProbity also indicated that 88.38% and 3.78% of all residues in the best predicted model as well as 78.8% and 7.91% of all residues in the template were in Ramachandran favored and Ramachandran outliers regions, respectively (results are not shown). The stereochemical quality of the modeled structures was also assessed using Verify-3D tool. In this tool, residues with a scoring function over 0.2 in the 3D profile are considered to be reliable. Results showed that, 80.65% and 79.25% and of the residues in the modeled structure (Fig. S3A) and template (Fig. S3B) respectively, had a score over 0.2, indicating (validating) the quality of the predicted structures. (Table 1). To further validate the quality of the predicted structures, ProSA Z-score assessments was also considered in our study. The ProSA Z-score value for the modeled (Fig. S4A and the template (Fig. S4B) and structures was -8.16 and -8.52, respectively, indicating a perfect correlation between the two structures (Table 1). Accordingly, the Energy plot also indicated that the most of the residues were with negative values for the model (Fig. S5A) and template (Fig. S5B). Collectively, all the validation assessments indicated the proper quality of the predicted model.

The position of Table 1

Protein structure drift and flexibility

Cα-RMSD was calculated as an indicator of the structure drift for SKC.μPm and SKG132.μPm with respect to the initial structure. As shown in Figure 1, Cα-RMSD values of all three runs reached to a plateau after 30 ns for SKC.μPm (Fig. 1A) and after 50ns for SKG132.μPm (Fig. 1B). Therefore, throughout the study, RMSF, Hbonds and RINs were analyzed by trajectories extracted from the last 50 ns of the simulations for both SKC.μPm and SKG132.μPm.

To investigate the effect of the four residue substitutions (Ile33Phe, Arg45Gln, Asn228Lys and Phe287Ile) on flexibility of the SKG132.μPm compared to that of the SKC.μPm, Cα-RMSF analysis was performed. The average RMSF values of three independent runs (run 1, run 2 and run 3) for SKC.μPm and SKG132.μPm are presented in Figure 1C (error bars reflect the observed fluctuations observed for the three independent runs). Based on RMSF data (Fig. 1C), flexibility of some loops and regions in SKG132.μPm compared to SKC.μPm structure was significantly increased. This increased flexibility was concentrated in the loops of SK part in the SKG132.μPm, comprising the 170 and 250 loops of SKβ and three regions (Rs) of 149-161 (R1), 182-215 (R2) and 224-229 (R3) with highest values corresponding to 170 loop and R2 (Fig. 1C-right, Fig. S6, Table 2). In agreement with our finding, a recent computational docking and MD simulation study also indicated the importance of the single residual substitutions in cytochrome P450 (CYP) 2B6 proteins for enhanced flexibility and substrate recognition (Kobayashi et al., 2014). More recently, the phenomenon of the increased conformational flexibility and its relation to increased enzyme activity through enhancement of the substrate binding is also reported for

Aerococcus viridans lactate oxidase (Stoisser, Brunsteiner, Wilson, & Nidetzky, 2016). Interestingly, our results for increased Cα-RMSF values of SKG132 (Fig. 1C-right, SK part and Fig. S6) are in accordance with several prior studies on the importance of 170 and 250 loops of SK for substrate recognition and processing. In this context, deletion or alanine substitution of Pro177 and Lys180 of 170 loop (Aneja, et al., 2009) of SK were previously shown for their significant contribution to the declined activities. Indeed, prior studies had implied the potential importance of the structural configuration, flexibility and/or surface accessibility of 170 loop (Aneja, et al., 2013) for efficient SK-substrate interactions and processing. Accordingly, the importance of 250 loop of SK for substrate interaction was also previously demonstrated by Alanine-scanning and charge-reversal mutations studies (Chaudhary, et al., 1999; Tharp, et al., 2009). Of note, our results for Cα-RMSF values (Fig. 1C and Fig. S6), additionally indicated the enhanced flexibility of the three specified SK regions (R1-R3) which are spatially located around 250 (R1-R3) and 170 (R2) loops (Fig. S7). Therefore, due to the critical role of 170 and 250 loops of SK for substrate Pg activation, the increased flexibility of these loops (and presence of regions R1 to R3 with further increased flexibility, as well) in SKG132 part of SKG132.µPm complex compared to SKC.µPm (findings of the present computational study) might further improve the exposure of these loops during substrate Pg activation in a synergistic manner. These increased flexibilities might have totally resulted to the higher Pg activation of SKG132 compared to SKC, as previously reported by experimental analyses (Keramati, et al., 2017). As expected however and contrary to the SK part of the SKG132.µPm complex, the Cα RMSF values of most residues in the µPm part of SKG132.µPm complex were lower than 3 Å (Fig. 1C-left, µPm part) indicating the stable nature of µPm part of the complex, particularly in the active sites (residues 58-63, 105 and 194-205) which might be required for an efficient and active site

formation by binding of the two components (Axelsson, 1995, http://www.uniprot.org/uniprot/P00747).

The position of Figure 1

The position of Table 2

Analysis of the hydrogen bonds

As shown in Table 3 and Figures S8 to S10, the average number of Hbonds in 170 loop, R2 (182-215) and R3 (224-229) of SKG132.µPm were significantly decreased compared to that of the SKC.µPm in all three runs. The decrease in the number of Hbonds was in accordance with increased Cα RMSF values (Table 2, Fig. 1C-right, Fig. S6) for enhanced structural flexibility in 170 loop, R2 and R3 of SKG132.µPm complex compared to that of SKC.µPm. Of particular note, the Hbond in the Val224-Lys256 residue pairs of R3 in SKC.µPm with average occupancy rate of 82% for all three runs (77.9% for run1; Fig. 2) was absent in the SKG132.µPm. This absence resulted to the generation of free (unbond) Lys256 residue in SK part of SKG132.µPm. Indeed, disruption of this Hbond in SKG132.µPm might significantly contribute to the increased flexibility of R3 while, generation of free Lys256 residue might also induce configurational changes in 250 loop (for better exposure of this loop) to enhance SK activity as previously suggested (Chaudhary, et al., 1999; Tharp, et al., 2009).

In accordance with our observation, results of a prior study indicated that alanine and charge reversal substitutions of Lys256 at the tip of 250 loop showed a drastically decline of Pg

activation by SK, implying the importance of this residue for substrate processing and recognition (Chaudhary, et al., 1999). Importance of the alterations in the number of Hbonds due to the residue substitution, for the enzyme activity is also recently reported for anaplastic lymphoma kinase (ALK)-tyrosine kinase through free energy calculation studies that are in agreement with our findings (Sun, Li, Li, & Hou, 2013). Therefore, the Asn228Lys mutation in R3 of SKG132 may have a critical role as a hot-spot in the enhancement of the flexibility and thereby the reported increase in the activity of SKG132 (Keramati, et al., 2017).

The position of Figure 2

The position of Table 3

To further explore the consequences of the residual substitutions at the atomic level, RINs analyses of all four substituted amino acids (Ile33Phe, Arg45Gln, Asn228Lys, Phe287Ile) in SKG132.µPm complex were carried out. In general, RINs results indicated the altered residue interactions for all of the four substituted residues of SKG132.µPm compared to that of SKC.µPm. (Fig. 3, Table 4). The overall surface expanded geometry of the SKG132.µPm compared to SKC.µPm complex with the specified locations of the altered residue interactions are provided in Figure S11. As shown in Figures 3A and S11, the residue Ile33 forms an Hbond with Gly24 in the SK part of the SKC.µPm complex. However, Ile33Phe substitution in SKG132.µPm complex resulted to formation of two new VDW interactions with Pro101 and Thr102 residues, besides retaining the same Hbond with Gly24. Of note, Pro101 and Thr102 residues surround the Asp105 in the active site (residues 58-63, 105 and 194-205) of the µPm part in the SKG132.µPm complex (Fig. 3A). Therefore, Ile33Phe substitution in SK part of the

SKG132.μPm complex might be considered as a potentially important structural/functional substitution. Because, these new interactions might have directly resulted to the formation of a more stable complex and/or altered active site geometry to enhance the activity of SKG132 compared to SKC. In fact, it was already well documented that formation of more stable complex between SK-μPm partners, plays a central role in Pg activation (Wu, et al., 2001). Moreover, our finding is in accordance with the results of a recent study on alteration of the mycobacterium tuberculosis isocitrate lyase enzymatic activity, apparently due to the altered conformational topology and geometry of the active site that was consequence of a single residue substitution (L418A) (Shukla, Shukla, Sonkar, & Tripathi, 2017).

As shown in Figures 3B and S11, Arg45Gln substitution resulted to the formation of new interactions for Gln45 with Arg10 (via Hbond and VDW) and Val13 (via VDW) in the SK part of the SKG132.μPm complex. Our analyses however, could not attribute any significant role for this substitution in structural changes of SKG132.μPm compared to SKC.μPm.

As shown in Figures 3C and S11, Asn228 in the SKC.μPm complex interacts with Lys61 via VDW. However, Asn228Lys residue substitution in the SKG132.μPm complex resulted to the presence of the positively charged residue around Lys228 that disrupted all of the interactions around this residue, including interactions with Lys61 (which is present in SKC.μPm complex). Therefore, in accordance with results of Hbond evaluations, RINs analysis also indicated that generation of free (unbound) Lys228 in the SKG132.μPm complex might play an important role in increased flexibility of R3 to induce conformational changes in favor of higher Pg activation and thus reported increased activity of SKG132 compared to SKC (Keramati, et al., 2017).

As shown in Figures 3D and S11, Phe287 in the SKC.μPm complex has no any interaction with several other residues of the complex. However, Phe287Ile residue substitution resulted to

formation of new external interactions with Phe42 residue (VDW) in the μPm part of the SKG132.μPm complex and with Tyr217, Leu277, Lys282 (via VDW) and Pro283 (via VDW and Hbond) in the SK part of SKG132.μPm complex. Formation of a new external interaction of SK with its μPm partner in the SKG132.μPm complex may have a direct role in the formation of a more stable complex with increased Pg activation properties. Of particular note, as shown in Figure S12, the negatively charged glutamic acid (Glu91) in the μPm part of the SKG132.μPm complex is located juxtaposition of the Ile287 in the SK part of this complex. This observation might suggest that alteration of Phe287 to a positively charged residue might induce a salt-bridge interaction with the juxta-positioned Glu91, resulting to the formation of still a more stable complex with increased Pg activation properties (Fig. S12). Overall, the RINs result for Phe287Ile substitution, in accordance with our findings for Ile33Phe substitution and prior reports on the importance of the stable complex formation (SK-μPm partners) for Pg activation (Wu, et al., 2001) might be considered as another potential atomic basis for the increased activity of SKG132 compared to SKC, besides the enhanced flexibility of the critical regions in SK structure, shown by MD simulation and Hbond analysis.

The position of Figure 3.

The position of Table 4.

From the classic point of view, proteins were inflexible (static) macromolecules with defined rigid secondary structures, less rigid random coils and active sites generally found at the protein core. The emerging evidences and recently accumulated data however, suggested that the structural flexibility/plasticity and dynamic/mobile properties of residues, especially distant from

the active site, might be critical in shaping biological functions such as enzyme catalysis and activity regulation in some proteins (Estarellas et al., 2016; Nevin Gerek, Kumar, & Banu Ozkan, 2013). In fact, flexibility might be essential for functional properties of many proteins to interact properly with their partner substrates, ligands, nucleic acids or other proteins to form complex structures (Craveur, et al., 2015). Therefore, findings of the present study, consistent with recent reports, suggested a probable correlation between altered residues and increased flexibility of the critical regions of SKG132 distant from the active site, besides enhanced stable complex formation (SK-μPm partners) around active site in the Pm part of the SKG132.μPm complex. These alterations might have contributed, individually or collectively, to the increased PA potency of the SKG132 compared to that of the SKC. Our data supports the recent proposition that beside an optimal structure, an optimal dynamic state might be also required by proteins for their optimum functionality (McAuley and Timson, 2017).

Conclusions:

In the present study, using computational approaches, we attempted to gain insights about the conformational basis of the increased activity of SK from a recently isolated group G *streptococci* (SKG132) with four natural residual substitutions (Ile33Phe, Arg45Gln, Asn228Lys, Phe287Ile) compared to that of the standard group C (SKC). Results of the MD simulation for SKG132 indicated the enhanced flexibility of two important loops (170 and 250) and regions (R1-R3) which might induce better exposure of these functionally important areas in the SK molecule. Hydrogen bond analyses indicated the average decrease of the number of Hbonds in 170 loop and R2-R3 of SKG132 compared to that of the SKC, especially for Asn228Lys substitution in R3, supporting the MD simulation data for the potential role of the increased flexibility in the critical regions for the increased activity of SKG. Finally, RINs

analysis indicated that Asn228Lys substitution which is located in the R3 (224-229) might induce the highest level of structural changes (and thus the most profound contribution in the increased Pg activation) in SKG132.µPm compared to SKC.µPm due to the increased flexibility of R3 through generation of free (unbound) Lys256 (in 250 loop) and thus better exposure of 250 loop for substrate interaction. Therefore, Asn228Lys substitution might be considered as the most important residue substitution and a hotspot in the SKG132. On the other hand, due to the important role of Phe287Ile substitution in the stability of the complex formation and potential role of the Ile33Phe substitution in both stability of the complex and alteration of the active site geometry, these substitutions might be considered as other important substitutions after Asn228Lys for the increased activity of the SKG132 compared to SKC. RINs analyses also suggested that the alteration of Phe287 to a positively charged residue might induce a salt-bridge interaction with the juxta-positioned Glu91, resulting the formation of a more stable complex with increased Pg activation properties. Finally, Arg45Gln with minor induced structural changes in SKG132.µPm complex might be considered as an insignificant substitution for altered Pg activation of SKG132. These results may have important implications for rational design of the SK as a fibrinolytic drug with increased plasminogen activation properties for the final aim of the thrombolytic therapy, in specific and denote the importance of the optimal dynamic state in protein engineering approaches, in general

Acknowledgments:

This study was financially supported by Pasteur Institute of Iran in partial fulfilment of the Ph.D thesis of Mrs. Faegheh Kazemi in medical biotechnology program.

Conflict of interest

The authors have no conflicts of interest.

References:

Abraham, M., Van Der Spoel, D., Lindahl, E., Hess, B., van Buuren, A., Apol, E., . . . Feenstra, K. (2015). GROMACS User Manual version 5.1.

Aneja, R., Datt, M., Singh, B., Kumar, S., & Sahni, G. (2009). Identification of a new exosite involved in catalytic turnover by the streptokinase-plasmin activator complex during human plasminogen activation. *Journal of Biological Chemistry, 284*(47), 32642-32650.

Aneja, R., Datt, M., Yadav, S., & Sahni, G. (2013). Multiple Exosites Distributed across the Three Domains of Streptokinase Co-Operate to Generate High Catalytic Rates in the Streptokinase–Plasmin Activator Complex. *Biochemistry, 52*(49), 8957-8968.

Arabi, R., Roohvand, F., Norouzian, D., Sardari, S., Aghasadeghi, M. R., Khanahmad, H., . . . Motevalli, F. (2011). A comparative study on the activity and antigenicity of truncated and full-length forms of streptokinase. *Pol J Microbiol, 60*(3), 243-251.

Axelsson, F. (1995). Plasminogen: Product monograph, Available fromL URL:http: www. Chromogenix.Se/.

Bean, R. R., Verhamme, I. M., & Bock, P. E. (2005). Role of the streptokinase α-domain in the interactions of streptokinase with plasminogen and plasmin. *Journal of Biological Chemistry, 280*(9), 7504-7510.

Berendsen, H. J., Postma, J. v., van Gunsteren, W. F., DiNola, A., & Haak, J. (1984). Molecular dynamics with coupling to an external bath. *The Journal of chemical physics, 81*(8), pp. 3684-3690.

Boxrud, P. D., Verhamme, I. M., & Bock, P. E. (2004). Resolution of conformational activation in the kinetic mechanism of plasminogen activation by streptokinase. *Journal of Biological Chemistry, 279*(35), 36633-36641.

Broomhead, N. K., & Soliman, M. E. (2017). Can We Rely on Computational Predictions To Correctly Identify Ligand Binding Sites on Novel Protein Drug Targets? Assessment of Binding Site Prediction Methods and a Protocol for Validation of Predicted Binding Sites. *Cell biochemistry and biophysics*, 75(1), 15-23.

Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of chemical physics, 126*(1), p 014101.

Chaudhary, A., Vasudha, S., Rajagopal, K., Komath, S. S., Garg, N., Yadav, M., . . . Sahni, G. (1999). Function of the central domain of streptokinase in substrate plasminogen docking and processing revealed by site-directed mutagenesis. *Protein Science, 8*(12), 2791-2805.

Chen, L., K Morrow, J., T Tran, H., S Phatak, S., Du-Cuny, L., & Zhang, S. (2012). From laptop to benchtop DeLano, W. L. (2002). The PyMOL molecular graphics system to bedside: structure-based drug design on protein targets. *Current pharmaceutical design, 18*(9), 1217-1239.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., . . . Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography, 66*(1), pp. 12-21.

Craveur, P., Joseph, A. P., Esque, J., Narwani, T. J., Noël, F., Shinada, N., . . . Bertrand, O. (2015). Protein flexibility in the light of structural alphabets. *Frontiers in molecular biosciences, 2,* 20.

Culka, M., Gisdon, F. J., & Ullmann, G. M. (2017). Computational Biochemistry—Enzyme Mechanisms Explored. *Advances in Protein Chemistry and Structural Biology,* 109, 77-112.

Dror, R. O., Dirks, R. M., Grossman, J., Xu, H., & Shaw, D. E. (2012). Biomolecular simulation: a computational microscope for molecular biology. *Annual review of biophysics, 41*, 429-452.

Estarellas, C., Capece, L., Seira, C., Bidon-Chanal, A., Estrin, D., & Luque, F. (2016). Chapter Three-Structural Plasticity in Globins: Role of Protein Dynamics in Defining Ligand Migration Pathways. *Advances in protein chemistry and structural biology, 105*, "pp. 59-80".

Hess, B., Bekker, H., Berendsen, H. J., & Fraaije, J. G. (1997). LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry, 18*(12), 1463-1472.

Keramati, M., Arabi Mianroodi, R., Memarnejadian, A., Mirzaie, A., Sazvari, S., Mehdi Aslani, M., . . . Roohvand, F. (2013a). Towards a superior streptokinase for fibrinolytic therapy of vascular thrombosis. *Cardiovascular & Hematological Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Cardiovascular & Hematological Agents), 11*(3), 218-229.

Keramati, M., Aslani, M. M., Khatami, S., & Roohvand, F. (2017). Sequence and kinetic analyses of streptokinase from two group G streptococci with high fibrin-dependent plasminogen activities and the identification of novel altered amino acids as potential hot spots. *Biotechnology Letters, 39*(6), 889-895.

Keramati, M., Roohvand, F., Aslani, M. M., Motevalli, F., & Memarnejadian, A. (2013b). Pitfalls in screening streptococci for retrieving superior streptokinase (SK) genes: no activity correlation for streptococcal culture supernatant and recombinant SK. *Journal of industrial microbiology & biotechnology, 40*(1), 151-158.

Keramati, M., Roohvand, F., Aslani, M. M., Khatami, S., Aghasadeghi, M., Sadat, M., . . . Motevalli, F. (2013c). Screening, Cloning and Expression of Active Streptokinase from an Iranian Isolate of S. equisimilis Group C in E. coli. *Iranian journal of basic medical sciences, 16*(4), 620.

Keramati, M., Roohvand, F., Eslaminejad, Z., Mirzaie, A., Nikbin, V. S., & Aslani, M. M. (2012). PCR/RFLP-based allelic variants of streptokinase and their plasminogen activation potencies. *FEMS microbiology letters, 335*(2), 79-85.

Kim, D. M., Lee, S. J., Kim, I. C., Kim, S. T., & Byun, S. M. (2000). Asp41-His48 region of streptokinase is important in binding to a substrate plasminogen. *Thrombosis research, 99*(1), 93-98.

Klegerman, M. E. (2017). Translational initiatives in thrombolytic therapy. *Frontiers of medicine*, *11*(1), 1-19.

Kobayashi, K., Takahashi, O., Hiratsuka, M., Yamaotsu, N., Hirono, S., Watanabe, Y., & Oda, A. (2014). Evaluation of influence of single nucleotide polymorphisms in cytochrome P450 2B6 on substrate recognition using computational docking and molecular dynamics simulation. *PloS one, 9*(5), e96789.

Krieger, E., & Vriend, G. (2014). YASARA View—molecular graphics for all devices—from smartphones to workstations. *Bioinformatics, 30*(20), 2981-2982.

Krishnamoorthy, N., Gajendrarao, P., Olivotto, I., & Yacoub, M. (2017). Impact of disease-causing mutations on inter-domain interactions in cMyBP-C: a steered molecular dynamics study. *Journal of Biomolecular Structure and Dynamics, 35*(9), pp. 1916-1922.

Law, R. H., Caradoc-Davies, T., Cowieson, N., Horvath, A. J., Quek, A. J., Encarnacao, J. A., . . . Lu, B. G. (2012). The X-ray crystal structure of full-length human plasminogen. *Cell reports, 1*(3), 185-190.

Lovell, S. C., Davis, I. W., Arendall, W. B., De Bakker, P. I., Word, J. M., Prisant, M. G., . . . Richardson, D. C. (2003). Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics, 50*(3), pp. 437-450.

Luthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature, 356*(6364), pp. 83-85.

McArthur, J. D., McKay, F. C., Ramachandran, V., Shyam, P., Cork, A. J., Sanderson-Smith, M. L., . . . Ranson, M. (2008). Allelic variants of streptokinase from Streptococcus pyogenes display functional differences in plasminogen activation. *The FASEB Journal, 22*(9), 3146-3153.

McAuley, M., & Timson, D. J. (2017). Modulating mobility: a paradigm for protein engineering? *Applied biochemistry and biotechnology, 181*(1), 83-90.

Nevin Gerek, Z., Kumar, S., & Banu Ozkan, S. (2013). Structural dynamics flexibility informs function and evolution at a proteome scale. *Evolutionary applications, 6*(3), 423-433.

Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of computational chemistry, 25*(13), 1656-1676.

Parrado, J., Conejero-Lara, F., Smith, R. A., Marshall, J. M., Ponting, C. P., & Dobson, C. M. (1996). The domain organization of streptokinase: nuclear magnetic resonance, circular dichroism, and functional characterization of proteolytic fragments. *Protein Science, 5*(4), 693-704.

Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics, 52*(12), 7182-7190.

Piovesan, D., Minervini, G., & Tosatto, S. C. (2016). The RING 2.0 web server for high quality residue interaction networks. *Nucleic acids research, 44*(W1), W367-W374.

Roohvand, F. (2017). Streptokinase for Treatment of Thrombotic Disorders: The End? Or the End of the Beginning? *Iranian biomedical journal*, *22*(3), 140-141.

Sazonova, I. Y., Robinson, B. R., Gladysheva, I. P., Castellino, F. J., & Reed, G. L. (2004). α domain deletion converts streptokinase into a fibrin-dependent plasminogen activator through mechanisms akin to staphylokinase and tissue plasminogen activator. *Journal of Biological Chemistry, 279*(24), 24994-25001.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research, 13*(11), 2498-2504.

Shukla, H., Shukla, R., Sonkar, A., & Tripathi, T. (2017). Alterations in conformational topology and interaction dynamics caused by L418A mutation leads to activity loss of Mycobacterium tuberculosis isocitrate lyase. *Biochemical and biophysical research communications, 490*(2), 276-282.

Stoisser, T., Brunsteiner, M., Wilson, D. K., & Nidetzky, B. (2016). Conformational flexibility related to enzyme activity: evidence for a dynamic active-site gatekeeper function of Tyr215 in Aerococcus viridans lactate oxidase. *Scientific Reports, 6*(27892).

Sun, H., Li, Y., Li, D., & Hou, T. (2013). Insight into crizotinib resistance mechanisms caused by three mutations in ALK tyrosine kinase using free energy calculation approaches. *Journal of chemical information and modeling, 53*(9), 2376-2389.

Sutthibutpong, T., Rattanarojpong, T., & Khunrae, P. (2017). Effects of helix and fingertip mutations on the thermostability of xyn11A investigated by molecular dynamics simulations and enzyme activity assays. *Journal of Biomolecular Structure and Dynamics*, 10.1080/07391102.2017.1404934.

Tharp, A. C., Laha, M., Panizzi, P., Thompson, M. W., Fuentes-Prior, P., & Bock, P. E. (2009). Plasminogen Substrate Recognition by the Streptokinase-Plasminogen Catalytic Complex Is Facilitated by Arg253, Lys256, and Lys257 in the Streptokinase β-Domain and Kringle 5 of the Substrate. *Journal of Biological Chemistry, 284*(29), 19511-19521.

Thirumal, D. K., George, C. P. D., Sneha, P., Tayubi, I., Siva, R., Chakraborty, C., & Magesh, R. (2017). Influence of V54M mutation in giant muscle protein titin: a computational screening and molecular dynamics approach. *Journal of biomolecular structure & dynamics, 35*(5), pp. 917-928.

Verhamme, I. M., & Bock, P. E. (2014). Rapid binding of plasminogen to streptokinase in a catalytic complex reveals a three-step mechanism. *Journal of Biological Chemistry, 289*(40), 28006-28018.

Wakeham, N., Terzyan, S., Zhai, P., Loy, J., Tang, J., & Zhang, X. (2002). Effects of deletion of streptokinase residues 48-59 on plasminogen activation. *Protein engineering, 15*(9), 753-761.

Wang, X., Lin, X., Loy, J. A., Tang, J., & Zhang, X. C. (1998). Crystal structure of the catalytic domain of human plasmin complexed with streptokinase. *Science, 281*(5383), 1662-1665.

Wang, X., Tang, J., Hunter, B., & Zhang, X. C. (1999). Crystal structure of streptokinase β-domain. *FEBS letters, 459*(1), 85-89.

Webb, B., & Sali, A. (2014). Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics, 47,* 5.6.1-32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25199792

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research, 35*(suppl_2), pp. W407-W410.

Wu, D.-H., Shi, G.-Y., Chuang, W.-J., Hsu, J.-M., Young, K.-C., Chang, C.-W., & Wu, H.-L. (2001). Coiled Coil Region of Streptokinase γ-Domain Is Essential for Plasminogen Activation. *Journal of Biological Chemistry, 276*(18), 15025-15033.

Yadav, S., Aneja, R., Kumar, P., Datt, M., Sinha, S., & Sahni, G. (2011). Identification through Combinatorial Random and Rational Mutagenesis of a Substrate-interacting Exosite in the γ Domain of Streptokinase. *Journal of Biological Chemistry, 286*(8), 6458-6469.

Yadav, S., Datt, M., Singh, B., & Sahni, G. (2008). Role of the 88–97 loop in plasminogen activation by streptokinase probed through site-specific mutagenesis. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1784*(9), 1310-1318.

Zhalyalov, A. S., Panteleev, M. A., Gracheva, M. A., Ataullakhanov, F. I., & Shibeko, A. M. (2017). Co-ordinated spatial propagation of blood plasma clotting and fibrinolytic fronts. *PloS one, 12*(7), e0180668.

Zhang, Y., Mayfield, J. A., Ploplis, V. A., & Castellino, F. J. (2014). The β-domain of cluster 2b streptokinase is a major determinant for the regulation of its plasminogen activation activity by cellular plasminogen receptors. *Biochemical and biophysical research communications, 444*(4), 595-598.

Zhao, F.-L., Yang, G.-H., Xiang, S., Gao, D.-D., & Zeng, C. (2017). In silico analysis of the effect of mutation on epidermal growth factor receptor in non-small-cell lung carcinoma: from mutational analysis to drug designing. *Journal of Biomolecular Structure and Dynamics, 35*(2), pp. 427-434.

Table 1: Comparison of the model validation parameters for the model and template structures

Validation parameters		**Template**	**Model**
Ramachandran plot	Favored	81%	90%
	Allowed	13.6%	7.3%
	Outliers	5.4%	2.7%

Verify-3D	79.25%	80.65%
ProSA Z-score	-8.52	-8.16

Table 2. The root mean square fluctuation (RMSF) of Cα atoms (Cα-RMSF) values for various loops and regions of SK part of SKC132.µPm and SKC.µPm structures (run1, run2, run3). Cα-RMSF values were calculated by gmx rmsf-gromacs built-in tools ± (Standard deviation).

Loops and regions	**Average RMSF of related loops and regions in**					
	Run 1		**Run 2**		**Run 3**	
	SKC.µPm	**SKG132.µPm**	**SKC.µPm**	**SKG132.µPm**	**SKC.µPm**	**SKG132.µPm**
170 loop	0.23±0.03	0.33±0.04	0.17±0.03	0.31±0.06	0.17±0.02	0.27±0.03
250 loop	0.14±0.01	0.24±0.05	0.1±0.008	0.17±0.01	0.11±0.01	0.16±0.02
149-161 (R1)	0.21±0.04	0.33±0.04	0.16±0.04	0.25±0.03	0.16±0.01	0.25±0.03
182-215 (R2)	0.20±0.03	0.35±0.10	0.15±0.03	0.31±0.07	0.16±0.02	0.28±0.08
224-229 (R3)	0.21±0.04	0.30±0.08	0.14±0.02	0.25±0.06	0.14±0.02	0.29±0.07

Table 3. Comparison of the average number of hydrogen bonds (Hbonds) in 170 loop, R2 (182-215) and R3 (224-229) of SKG132.µPm and SKC.µPm in all three runs (run 1, 2, 3). Hbonds were calculated when a donor-acceptor distance was less than 0.35 nm and a donor-hydrogen-acceptor angle was 30 degrees. Hbonds were calculated by gmx hbond-gromacs built-in tools. All calculations corresponded to the last 50 ns of MD simulations ± (Standard deviation).

Loops and regions	Average number of Hbonds					
	Run 1		Run 2		Run 3	
	SKC.µPm	SKG132.µPm	SKC.µPm	SKG132.µPm	SKC.µPm	SKG132.µPm
170 loop	10.31±2.72	6.31±2.06	10.9±2.93	6.27±2.18	11.35±2.83	6.14±2.38
182-215 (R2)	41.75±3.95	34.87±3.08	42.89±4.25	35.09±3.01	41.88±3.96	33.13±3.06
224-229 (R3)	17.7±2.94	12.3±2.67	15.97±3.36	11.64±2.78	15.34±3.36	10.95±2.74

Substitution site	Residue interactions before and after substitution		
	SKC.µPm	SKG132.µPm	
Ile33Phe	(Hbond) Gly24	(Hbond) Gly24 (VDW) Pro101 (VDW) Thr102	
Arg45Gln	None	Arg10 (Hbond and VDW)	

Table 4. Results of the residue interaction networks (RINs) analysis for the substituted amino acids (Ile33Phe, Arg45Gln, Asn228Lys, Phe287Ile) in SKG132.µPm complex (i.e.: after substitution) compared to SKC.µPm (i.e.: before substitution). RINs data correspond to the outputs of the Run 1 MD simulation.

		& (VDW) Val13
Asn228Lys	(VDW) Lys61	None
Phe287Ile	None	(VDW) Phe42* (VDW) Tyr217 (VDW)Leu277 (VDW)Lys282 (VDW and Hbond) Pro283

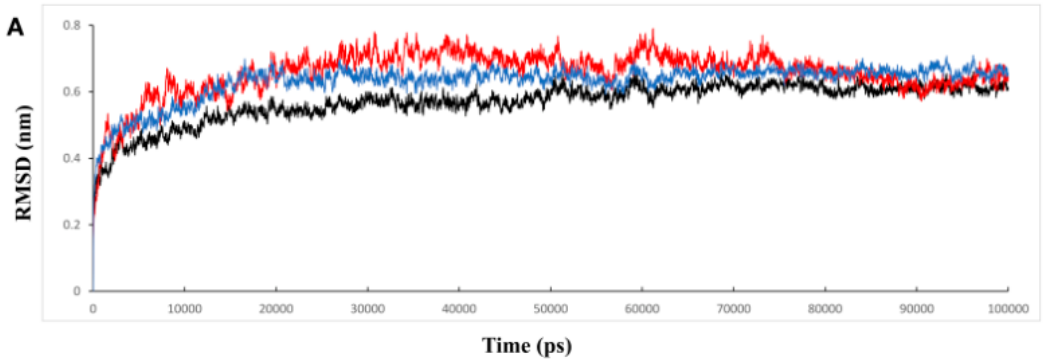
Figure legends

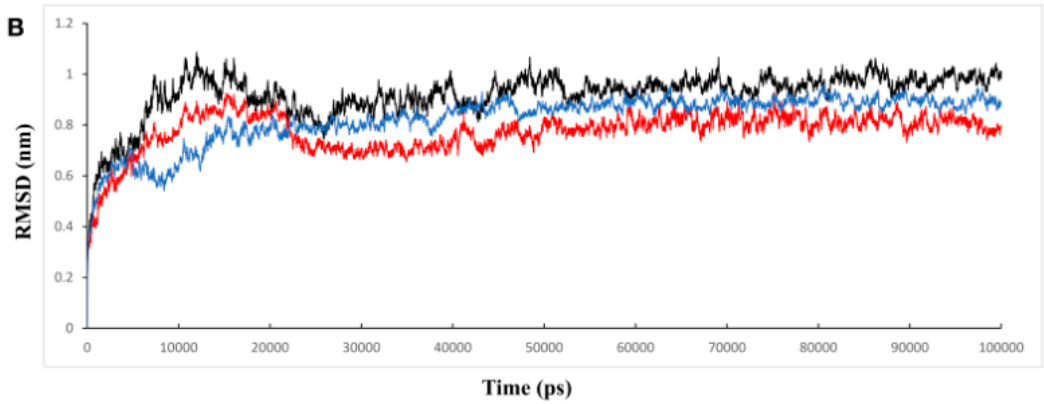
Figure 1. Structural drift and flexibility analyses of SKC.µPm and SKG132.µPm. The root mean square deviation (RMSD) of C-alpha (Cα) atoms (Cα-RMSD) were calculated with respect to the initial structure as a function of simulation time for (A) SKC.µPm; and (B) SKG132.µPm. The three different colors denote the tree different runs: run1 (black), run2 (red) and run3 (blue). Cα-RMSD values reached to a plateau after 30 ns for SKC.µPm and after 50ns for SKG132.µPm for all three runs. Thus in all analyses the trajectories were extracted from the last 50 ns of the MD simulation. (C) The comparison of average root mean square fluctuation (RMSF) of Cα atoms (Cα-RMSF) of all three runs (runs 1, 2 and 3) for both SKG132.µPm and SKC.µPm with error bars. Cα-RMSF values generally indicated increased flexibility of some regions and loops in SKG132.µPm (red) compared to SKC.µPm (black) (C-right). Increase in flexibility was concentrated in the loops of SK part in the SKG132.µPm, including the 170 and 250 loops of SKβ and three regions (Rs) including: 149-161 (R1), 182-215 (R2) and 224-229 (R3), with highest values corresponding to the 170 loop and R2.

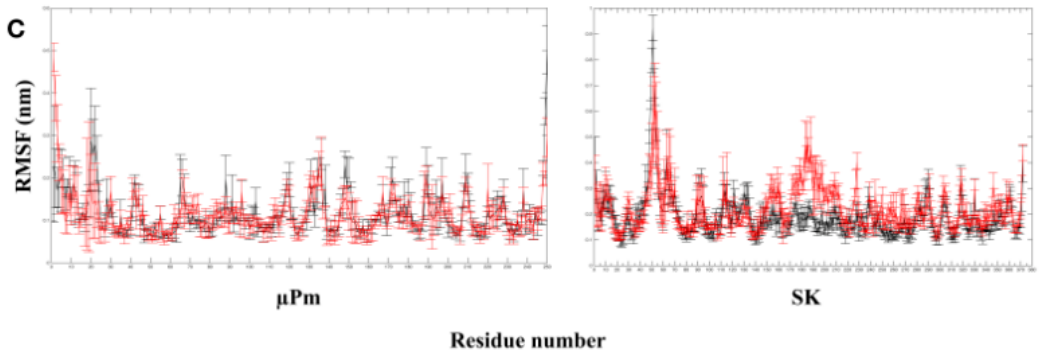
Figure 2. Analysis of the hydrogen bond occupancy in SKC.µPm (run1). Occupancy was calculated when a donor-acceptor distance was less than 0.35 nm and a donor-hydrogen-acceptor angle was 30 degrees. Hbond occupancy in Val224-Lys256 residue pairs of R3 in SKC.µPm was 77.9%, while it was absent in the SKG132.µPm. Calculations were corresponded to the last 50 ns of MD simulations.

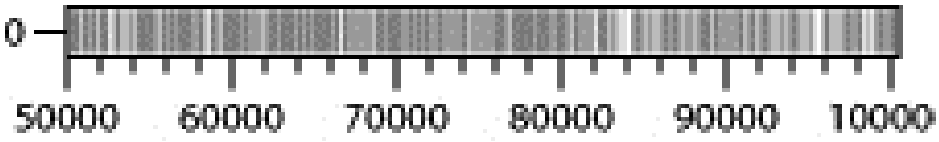
Figure 3. Residue interaction networks (RINs) analysis for SKC.µPm (left) and SKG132.µPm (right) complexes. SK parts of the complexes appear in light and µPm in dark colors. (A) Ile33 formed an Hbond with Gly24 in the SK part of the SKC.µPm complex (Left). Ile33Phe substitution, besides retaining the same Hbond with Gly24 in SKG132.µPm complex, also induced the formation of two new VDW interactions with Pro101 and Thr102 residues (that surround the Asp105 in the active site), in the µPm part of the SKG132.µPm complex (right). (B) There was no interaction between Arg45 and other residues of SKC.µPm (left). Arg45Gln substitution induced the formation of two new interactions for Gln45 with Arg10 (via Hbond and VDW) and Val13 (via VDW) in the SK part of the SKG132.µPm complex (right). (C) Asn228 in the SKC.µPm complex interacts with Lys61 via VDW (left). Asn228Lys substitution resulted to the loss of interaction of this residue in SKG132.µPm (right). (D) Phe287 has no any interaction with other residues of the complex (left), but Phe287Ile residue substitution resulted to formation of a new external interactions with Phe42 residue (VDW) in the µPm part of the SKG132.µPm complex and new internal interactions with Tyr217, Leu277, Lys282 (via VDW) and Pro283 (via VDW and Hbond) in the SK part of SKG132.µPm complex. Residue interactions were

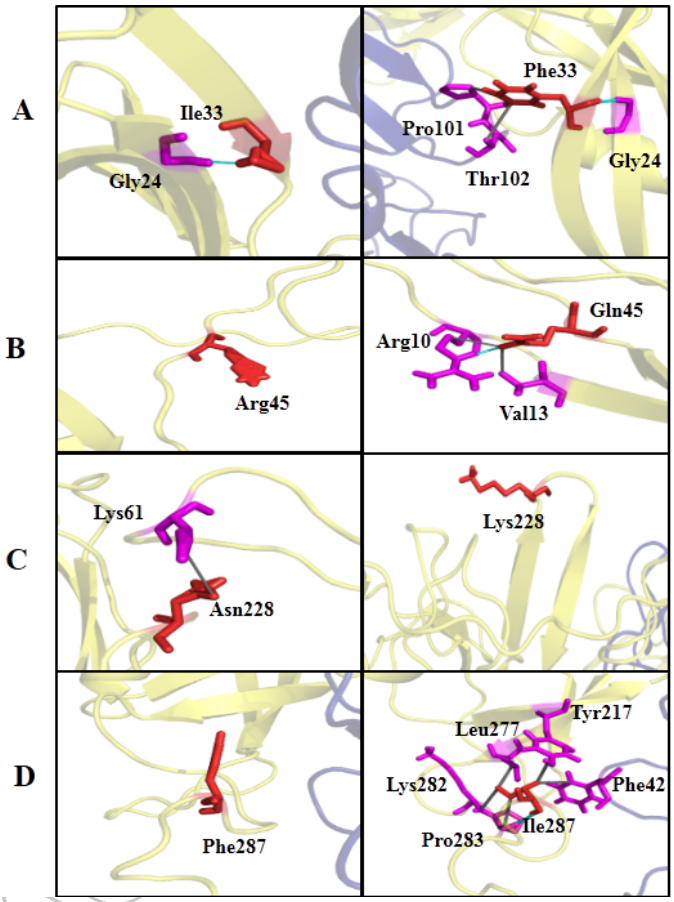
developed by the RINalyzer plug-in of Cytoscape platform and visualized by RING-Viz script for Pymol application from the last 50 ns trajectory of each protein complex structure. VDW and Hbond denote the van der Waals and hydrogen bonds interactions, respectively. An enlarged surface representation of the SKC.µPm and SKG132.µPm complex with specified interactions of Figures 3A to 3D is presented in supplementary Figure 11. RINs data correspond to the outputs of the run 1 MD simulation.











Supplementary materials:

Figure S1: Analysis of DOPE score profile (calculated/plotted using assess_dope function of MODELLER software). The Blue and red lines show the energy profile of template and the modeled structures, respectively. As shown, the superposed profiles of the template and modeled structures were almost similar.

Figure S2: Ramachandran plot for the modelled and template structures of SK. Ramachandran maps were generated by Rampage validation program. Results indicated that 90%, 7.3% and 2.7% of residues in the modeled structure (A) and 81%, 13.6% and 5.4% of residues in the template structure (B) were in the favored, allowed and outlier regions, respectively

Figure S3: Evaluation of the stereochemical quality of the modeled structures by Verify-3D tool. As shown, 80.65% and 79.25% of the residues in the modeled structure (A) and template (B) respectively, had a score over 0.2, validating the quality of the predicted structures.

Figure S4: Evaluation of the quality of the modelled and template structures of SK using the ProSA server. The plot shows z-score value of -8.16 for modeled structure (A) and z-score value of -8.52 for template structure (B) indicating a perfect correlation between the two structures.

Figure S5: Energy plots of modelled and template structures obtained by ProSA server. As shown, most of the residues were with negative values for the model (A) and template (B).

Figure S6: Analysis of the structural flexibility. Root mean square fluctuation (RMSF) of Cα atoms (Cα-RMSF) or both SKC.μPm (A) and SKG132.μPm (B) were calculated. The three different colors denote the tree different runs: run1 (black), run2 (red) and run3 (blue). Values generally indicated increased flexibility of some regions and loops in SK part of SKG132.μPm in all three runs (B-right) compared to SKC.μPm. (Aright). Increase in the flexibility was concentrated in the loops of SK part in the SKG132.μPm (B-right), including the 170 and 250 loops of SKβ and three regions (Rs) including: 149-161 (R1), 182-215 (R2) and 224-229 (R3), with highest values corresponding to 170 loop and R2.

Figure S7: Cartoon representation of the flexible regions (R1-R3) around 170 and 250 loops. 170 and 250 loops are colored in pink and magenta, respectively. Three flexible regions including R1 (149-161), R2 (182-215) and R3 (224-229) are colored in green, blue and yellow, respectively. Three enhanced flexible regions (R1-R3) are spatially located around 250 (R1-R3) and 170 (R2) loops. The structure is the average PDB structure of SKG132.μPlm which retrieved from the last 50 ns of MD simulations.

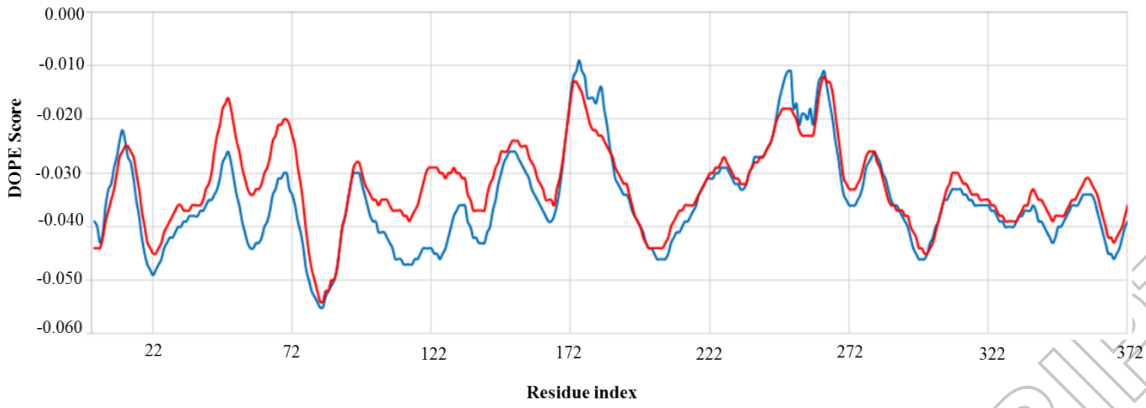
Figure S8: Analysis of the hydrogen bonds (Hbonds) for 170 loop of SKG132.µPm and SKC.µPm in three independent runs (A, B & C). Number of Hbonds were calculated when a donor-acceptor distance was less than 0.35 nm and a donor-hydrogen-acceptor angle was 30 degrees. The average number of Hbonds of SKG132.µPm (red) were significantly decreased compared to SKC.µPm (black), (A) run1 (6.31 versus 10.31 Hbonds), (B) run2 (6.27 versus 10.9 Hbonds), (C) R3 (6.14 versus 11.35 Hbonds). All calculations corresponded to the last 50 ns of MD simulations

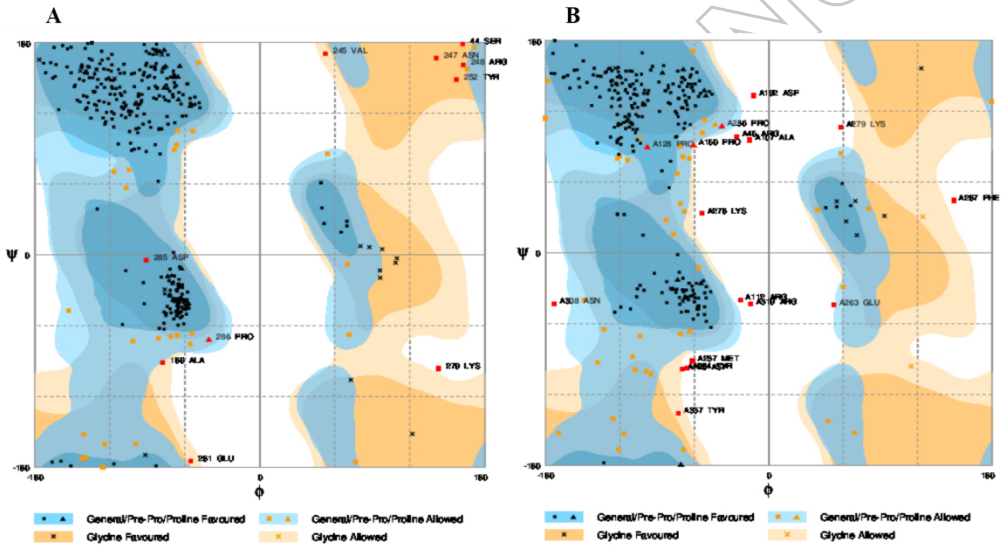
Figure S9: Analysis of the hydrogen bonds (Hbonds) for R2 (182-215) of SKG132.µPm and SKC.µPm in three independent runs (A, B & C). Nnumber of Hbonds were calculated when a donor-acceptor distance was less than 0.35 nm and a donor-hydrogen-acceptor angle was 30 degrees. The average number of Hbonds of SKG132.µPm (red) were significantly decreased compared to SKC.µPm (black), (A) run1 (34.87 versus 41.75 Hbonds) (B) run2 (35.09 versus 42.89 Hbonds), (C) R3 (33.13 versus 41.88 Hbonds). All calculations corresponded to the last 50 ns of MD simulations.

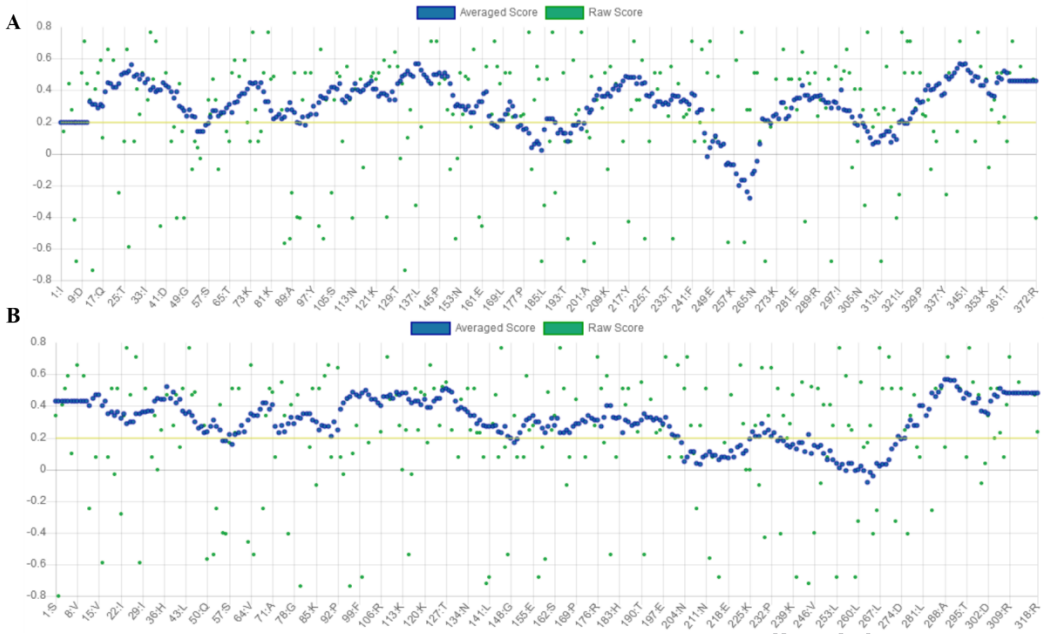
Figure S10: Analysis of the hydrogen bonds (Hbonds) for R3 (224-229) of SKG132.µPm and SKC.µPm in three independent runs (A, B & C). Number of Hbonds were calculated when a donor-acceptor distance was less than 0.35 nm and a donor-hydrogen-acceptor angle was 30 degrees. The average number of Hbonds of SKG132.µPm (red) were significantly decreased compared to SKC.µPm (black), (A) run1 (12.3 versus 17.7 Hbonds), (B) run2 (11.64 versus 15.97 Hbonds), (C) run3 (10.95 versus 15.34 Hbonds). All calculations corresponded to the last 50 ns of MD simulations.

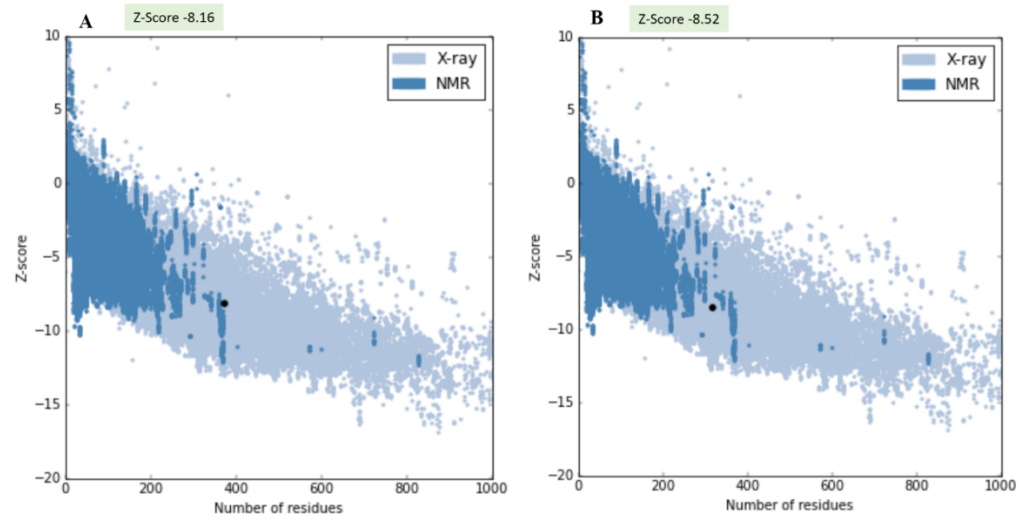
Figure S11: Enlarged view of the Residue interaction networks (RINs) analysis for SKC.µPm (up) and SKG132.µPm (down) complexes. SK parts of the complexes appear in yellow and µPm in blue colors. Ile33 formed Hbond interaction with Gly24 in the SK part of the SKC.µPm complex (A-up). Ile33Phe substitution, besides retaining the formation of Hbond with Gly24, also induced the formation of two new VDW interactions with Pro101 and Thr102 residues (surrounding the active site, Asp105) in the µPm part of the SKG132.µPm complex (A-down). There was no interaction between Arg45 and other residues of SKC.µPm (B-up). Arg45Gln substitution resulted to the formation of new interactions for Gln45 with Arg10 (via Hbond and VDW) and Val13 (via VDW) in the SK part of the SKG132.µPm complex (B-down). Asn228 in the SKC.µPm complex interacted with Lys61 via VDW (C-up). Asn228Lys substitution resulted to the loss of interaction of this residue in SKG132.µPm (C-down). Phe287 had no interaction with other residues of the complex (D-up), while, Phe287Ile residue substitution resulted to formation of a new external interaction (VDW) with Phe42 residue in the µPm part of the SKG132.µPm complex and several internal interactions with residues of the SK part of the complex including: Tyr217, Leu277, Lys282 (via VDW) and Pro283 (via VDW and Hbond) (D-down). Residue interactions were developed by the RINalyzer plug-in of Cytoscape platform and visualized by RING-Viz script for Pymol application from the last 50 ns trajectory of each protein complex structure. VDW and Hbond denote the van der Waals and hydrogen bonds interactions, respectively. VDW and Hbond interactions depicted as gray and blue lines, respectively. The RINs analyses are based on the MD simulation outputs of run1.

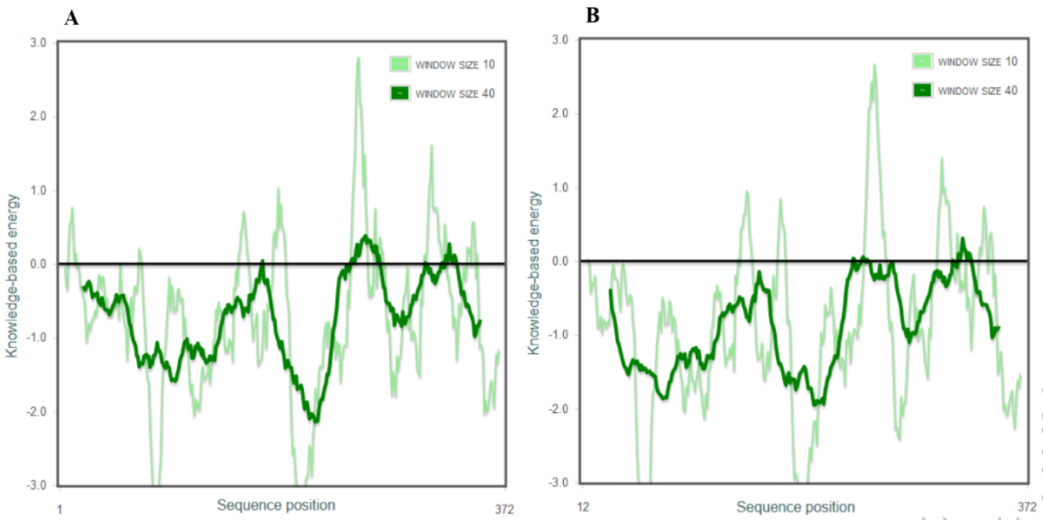
Figure S12: Enlarged residue interaction networks (RINs) analysis for Ile287 in SKG132.µPm complex. Phe287Ile substitution formed a new external interaction (VDW) with Phe42 residue in the µPm part of the SKG132.µPm complex. The figure shows the position of Glu91 in µPm part of SKG132.µPm complex juxtaposition of the Ile287. SK part of the complex appear in yellow and µPm in blue colors. VDW denote the van der Waals and depicted as gray line.

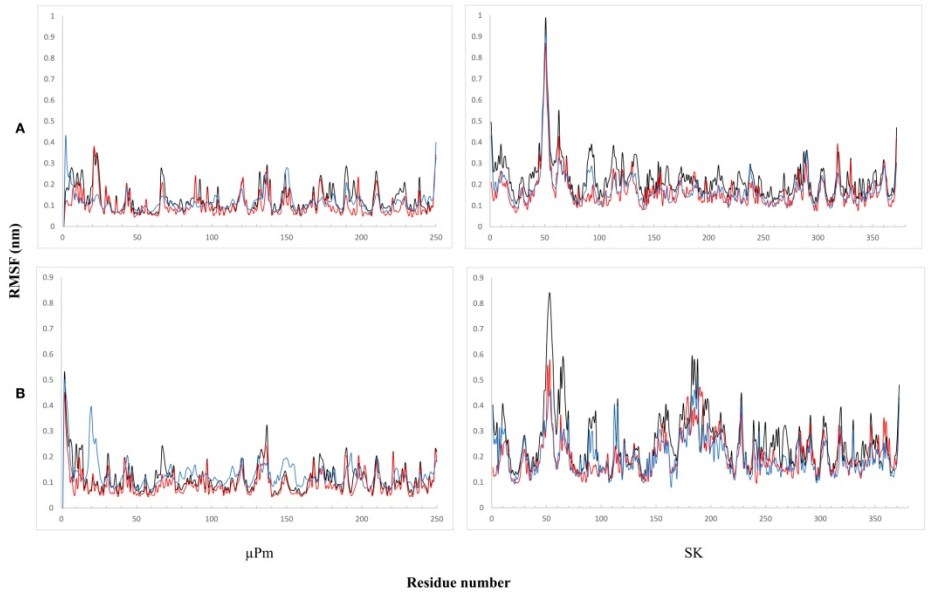


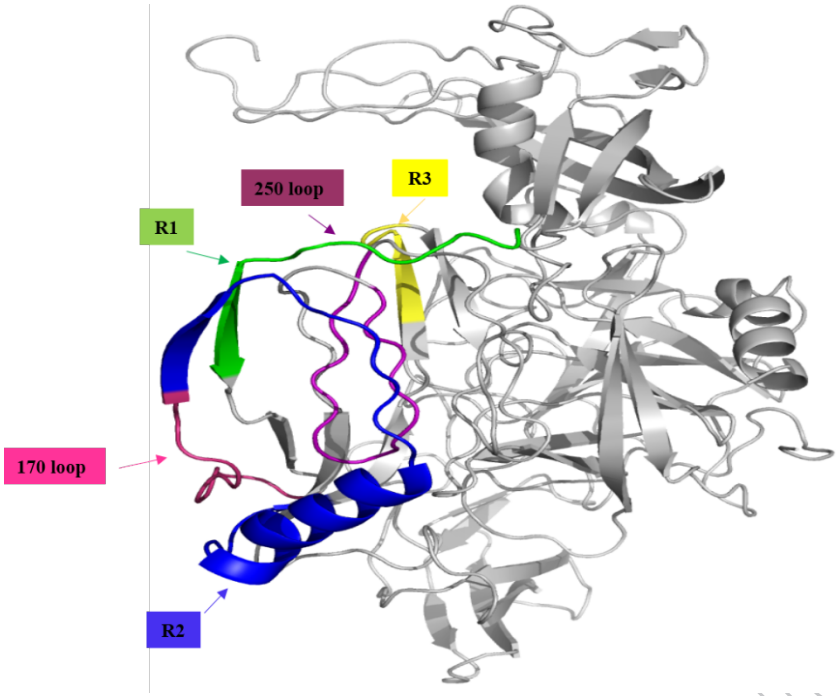


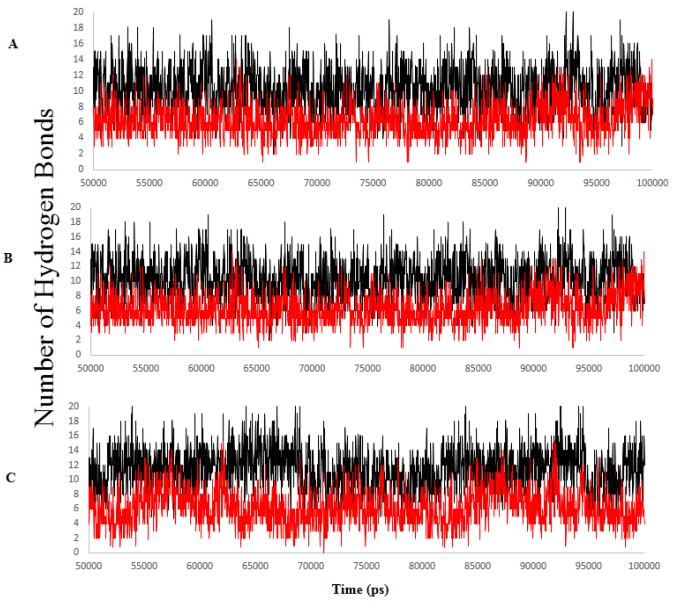

A
Averaged Score
Raw Score
B
Averaged Score
Raw Score

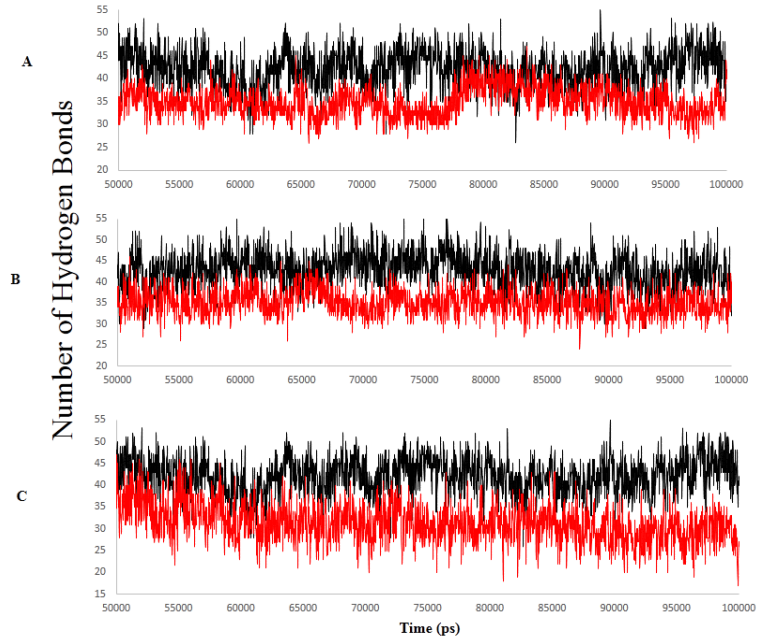

A
Z-Score -8.16
X-ray
NMR
Z-score
Number of residues
B
Z-Score -8.52
X-ray
NMR
Z-score
Number of residues

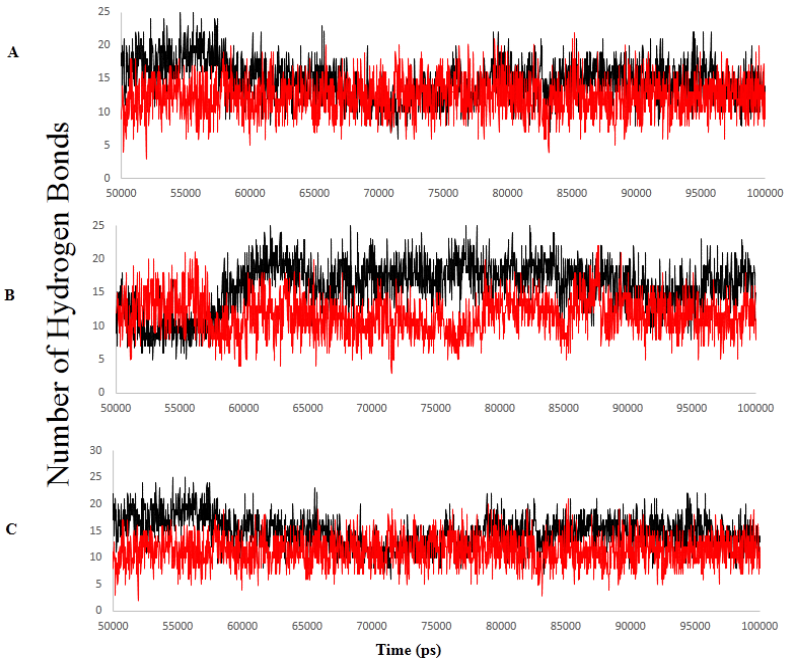

A
B
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
1
372
12


A
B
RMSF (nm)
μPm
SK
Residue number


250 loop
R3
R1
170 loop
R2


A
B
C
Number of Hydrogen Bonds
Time (ps)


A
B
C
Number of Hydrogen Bonds
Time (ps)


A
B
C
Number of Hydrogen Bonds
Time (ps)