Molecular Docking and MD Simulation Analysis of L-asparaginase of *Streptomyces iranensis, Streptomyces himalayensis and Streptomyces griseus*

Achyutuni Venkata Naga Tejaswini[1], Ramesh Malothu[1*]

[1]Centre of Microbiology, School of Biotechnology, Institute of Science and Technology, Jawaharlal Nehru Technological University Kakinada, Kakinada 533003, Andhra Pradesh, India.
[*]Corresponding author: ramesh_biotech@jntuk.edu.in

Abstract

The present study focuses on structural, molecular docking and simulation analysis of L-asparaginase protein belonging to *Streptomyces iranensis* (*S. iranensis*) (WP_044568095), *Streptomyces himalayensis* (*S. himalayensis*) (WP_181864635) and *Streptomyces griseus* (*S. griseus*) (WP_037680047). The three protein sequences retrieved from NCBI are also assessed for characteristic, secondary structure and localization studies. The analysis with BLAST tool predicted WP_044568095 to be 40.18% and 39.16% identical with D. chrysanthemi (P06608) and E. coli L-asparaginase II (P00805). Similarly WP_181864635 is identified to be 38.39% and 37.58% identical with P06608, P00805 respectively while no resemblance is observed with WP_037680047.1. The three *Streptomyces* proteins on subsequent studies by I-TASSER and GalaxyRefine are found to have 85.9%, 84.2% and 79.2% residues in the favoured zone for WP_044568095, WP_181864635 and WP_037680047 on validation by PROCHECK. Further, acceptable findings are obtained from studies by ERRAT and ProSA. Molecular docking performed by CB-Dock and PyRx resulted in similar predictions about the active site amino acids forming hydrogen bond with L-asparagine. WP_044568095 docked with Asn resulted in lowest energy of - 4.6 kcal/mol while - 4.4 kcal/mol is obtained for WP_181864635, WP_037680047 evaluated by AutoDock Vina of PyRx software. The docked structures of WP_044568095 and WP_181864635 under controlled conditions by CABS-Flex 2.0 are found stable in comparison with WP_037680047. As L-asparaginase possess high therapeutic application in various malignancies treatment, the evaluated proteins of *S. iranensis*, *S. himalayensis* and *S. griseus* are worth analysing further for in vitro enzyme activity and are considered to have promising therapeutic importance.

Keywords: AutoDock Vina, Cabs Flex, CB-Dock, I-TASSER, *Streptomyces*.

Introduction

The genus *Streptomyces* possesses a wide assortment of bioactive compounds producing pathways. The *Streptomyces* species (sp) are considered valuable due to their potentiality of producing plethora of secondary metabolites (1-3). The research studies describe that *Streptomyces* sp are vital for synthesis of antibiotics (4, 5) and many enzymes. *Streptomyces* sp are described as ubiquitous (6), are also reported to be a proven source for the significant anticancer compound L-asparaginase (asparaginase) production.

The enzyme asparaginase finds application in conversion of the L-asparagine (Asn), crucial for the cellular functioning of a cancer cell into aspartic acid and ammonia. The decrease and exhaustion of the key amino acid impacts the functioning and proliferation of the tumour cells ultimately resulting in apoptosis (7). Asparaginase targets the oncogenic factors and is recognized as the most essential component in the treatment of acute lymphoblastic leukemia (ALL) (8). The clinical outcomes of chemotherapy involving asparaginase in conjunction with other medicines resulted in full remission in nearly 90% of patients (9). Therapy involving asparaginase is also recommended for acute myeloid leukaemia (10), lymphosarcoma (11), chronic myeloid leukaemia (12) as well as for aggressive NK-cell leukaemia (13), according to the scientific studies. In the present day, asparaginases derived from *Escherichia coli* (*E. coli*) and *Dickeya chrysanthemi* (*Erwinia chrysanthemi)* (*D. chrysanthemi*) are currently found effective and accessible for the treatment of ALL (14). The research and medical history regarding the significant enzyme states that the chemotherapeutic compound is also related with development of allergic responses in certain individuals (15, 16). The in silico analysis of asparaginase from the new species is highly important to discover the characteristics of the protein and its affinity towards Asn.

The present study aims for computational analysis of L-asparaginase proteins of *Streptomyces iranensis* (*S. iranensis*), *Streptomyces himalayensis* (*S. himalayensis*) and *Streptomyces griseus* (*S. griseus*). According to Netzker *et al* (17), rapamycin which finds application as an immunosuppressant, anticancer as well as anti-inflammatory agent (18) is also produced by *S. iranensis* (17). *S. griseus* characterized by Krainsky (19) is described as the largest genus of the Actinobacteria (20) is also known for the production of antibiotic Streptomicin (21), asparaginase (22, 23), Frigocyclinone (24), Bafilomycin (25), trypsin as a as source of insulin synthesis (26), enzymes protease *(27),*

Chitinase (28), Lipase (29) and pronase (30).

Thus the sequence analysis, structural prediction, virtual screening and simulation methods of the three different *Streptomyces* species provides valuable insights on the protein characteristics and affinity towards Asn.

Materials and methods

Retrieval of Streptomyces protein sequences

Asparaginase protein sequences belonging to *Streptomyces* species namely *S. iranensis* (accession number WP_044568095), *S. himalayensis* (WP_181864635) and *S. griseus* (WP_037680047) are obtained from the non-redundant (nr) database of the National Center for Biotechnology Information database (NCBI) (31). The proteins WP_044568095, WP_181864635 and WP_037680047 are searched against UniProtKB (32) by BLAST (33) tool to identify asparaginase proteins exhibiting identity above 35% with proteins of other genera specifically with *E. coli* and *D. chrysanthemi*. The retrieved protein sequences and WP_044568095, WP_181864635, WP_037680047 are aligned via ClustalW (34) and further visualized by a phylogenetic tree of MEGA (35). The domain categorization of the three query proteins is further obtained by submission to Pfam database (36).

Interpretation of various characteristics of proteins

The asparaginase protein sequences WP_044568095, WP_181864635 and WP_037680047 are assessed by ExPASy ProtParam server (37) to determine the properties such as molecular weight (Mwt), formula, isoelectric point (pI), factors concerning protein stability and thermal stability, Extinction coefficient (EC), charge, half-life studies (HL) and hydropathicity index (37).

Evaluation of secondary and tertiary structures and validation of asparaginase proteins

The proteins WP_044568095, WP_181864635 and WP_037680047 are subjected to determination of secondary structural (2D) elements by PSIPRED (38) and SOPMA (39). The three dimensional formation (3D) of protein is worth evaluating to interpret the protein function (40). The 3D structure of the Streptomyces asparaginase proteins are evaluated by submitting the sequences to the I-TASSER server (41) which are eventually assayed upon a confidence score (C-score) calculated between -5 to 2 (42). The C-score is formulated considering the alignments of the threading templates, thus it attributes to the quality of each protein model along with the associated TM-score and RMSD values (42). Proceeding further, the modelled proteins are subjected to required structural refinements by GalaxyRefine (43). The process of refinement proceeds by necessary side chain changes to the side chains of the protein followed by structural relaxation (43). The quality assessment of the protein structures is based on the Φ and Ψ statistics of the Ramachandran plot by PROCHECK (44) followed by ERRAT (45) and further by ProSA (46) studies.

Docking and simulation studies of the three Asparaginase proteins

The resulted protein structures are reviewed for significant ligand binding efficiency via an automated tool CB-Dock (47) and also by a user-input application PyRx (48), both the computational methods involve the application of AutoDock Vina (49). The refined and validated protein models in pdb format and the ligand Asn (Pubchem ID 6267) in SDF format are submitted to both the docking applications. The energy minimization of the ligand is executed by Open Babel (50), an integrated tool of PyRx (48). The relevant dimension suitable for the receptor-ligand interaction is an important step in drug design (51). The best protein-ligand docked structures and the interacting residues of all three species namely *S. iranensis*, *S. himalayensis* and *S. griseus* are visualized by Discovery studio (52) and Ligplot+ software (53). The molecular dynamic (MD) simulations of the docked *S. iranensis* and *S. himalayensis* proteins are executed by CABS Flex 2.0 (54). The methodology of evaluation of the stability of the docked protein is performed by maintaining the simulation criteria default with 50 cycles and 50 trajectory frames at temperature 1.4 for duration of 10 ns.

Results and Discussion

Retrieval of asparaginase proteins of Streptomyces species

The asparaginase protein sequences WP_044568095, WP_181864635 and WP_037680047 obtained from the nr database of the NCBI (31) are found to possess 379 amino acids (AA), 377 AA, 338AA as per the initial sequence information. The BLASTp results concerning the proteins WP_044568095, WP_181864635 are described in moderate similarity with asparaginases of *E. coli, D. chrysanthemi* and other taxa. The *S. iranensis* (WP_044568095) and *S. himalayensis* (WP_181864635) proteins are identified to have 40.18% and 38.39% similarity with L-asparaginase of D. chrysanthemi (P06608), respectively. Similarly, WP_044568095 and WP_181864635 proteins are also found to be E. coli L-asparaginase II (P00805) with 39.16%, 37.58% identical respectively. Contrastingly no resemblance is determined between the *S. griseus* protein WP_037680047.1 and asparaginase proteins belonging to other genera, although the similarity is observed with the various *Streptomyces* species. The search for proteins representing identity > 90% to WP_037680047.1 resulted only with *Streptomyces* sp. BK161 (A0A4R7HBL9), the remaining identical proteins of *Streptomyces* are categorized as having less than 90% similarity. The phylogenetic analysis of

the retrieved protein sequences and WP_044568095, WP_181864635, WP_037680047 is determined through MEGA software (35) by primarily aligning the sequences via ClustalW (34) and further prediction by neighbor-joining method (NJ) (55) (Fig. 1). The Pfam studies revealed that the three proteins belonging to the same genera are identified to be members of two different asparaginase family. *S. iranensis* (WP_044568095) and *S. himalayensis* (WP_181864635) belong to an N-terminal PF00710 family consisting asparaginase proteins and a glutaminase/asparaginase C-terminal domain, PF17763. Studies with *S. griseus* (WP_037680047) described that the protein belongs to PF06089 described to be Asparaginase II.

Computation of physicochemical features of the proteins

The elucidation of the asparaginase protein sequences by the ExPASy ProtParam online tool is
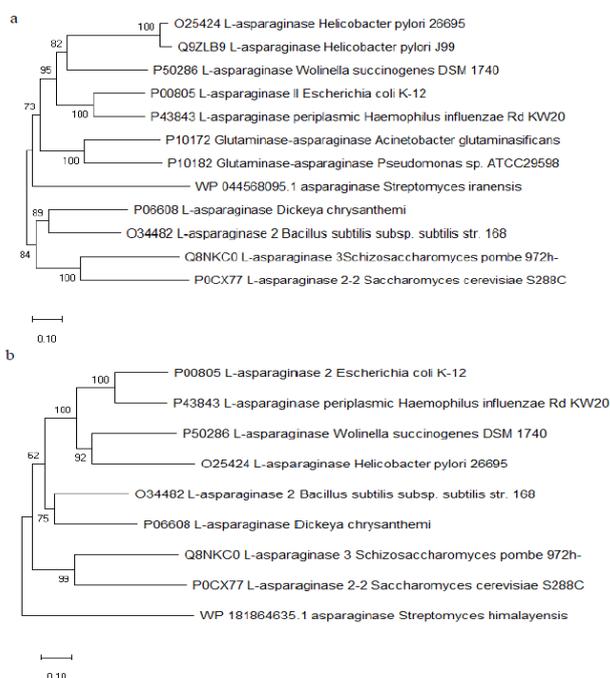
Figure 1: Phylogenetic relationship of (a) WP_044568095 (b) WP_181864635 with asparaginases of different genera determined by NJ method.

detailed in the Table 1. WP_044568095 possessing 379 AA is discovered to have a higher Mwt of 40182.58 kDa than the Mwt of WP_181864635 and WP_037680047 calculated to be 39282.38 kDa and 35196.13kDa accordingly.

Similarly, the determination of the charges of the AA residues of the three asparaginase proteins infers that WP_044568095 possessed 41 positively charged, 33 negatively charged residues. Contrastingly WP_181864635 and WP_037680047 indicated the presence of highly negatively charged residues with positive to negative residue ratios to be 38:40, 33:45

respectively (Table 1). The computation of pI and HL of a protein provides insights into the protein stability (37, 56, 57). The pI values of the three proteins varied from 5.0 to 9.5, with WP_181864635 and WP_037680047 being and acidic, WP_044568095 protein is found to be basic. Concerning the HL studies, methionine is identified to be the N-terminal residue of all the three asparaginase proteins with HL ranging between 10 – 30 hours involving *E. coli* and yeast in vivo and in vitro studies of mammalian reticulocytes. The computed instability indexes of the three asparaginase proteins (Table 1) are impressively stable with values less than 40, above which impacts the protein stability (58). In addition to II, the aliphatic index (AI) is also pivotal as a high AI value pertains to increased stability of the protein in varied temperatures (59). The

Table 1: Physicochemical parameters of WP_044568095, WP_181864635, and WP_037680047 computed by ExPASy ProtParam tool. Molecular wt: Molecular weight, pI: isoelectric point, II: instability index, AI: aliphatic index, Gravy: grand average of hydropathicity

Analysis criteria WP_044568095		Sequence details		
		(WP_181864635)	(WP_037680047)	
Molecular wt		40182.58	39282.38	35196.13
Formula		$C_{1763}H_{2839}N_{521}O_{537}S_8$	$C_{1714}H_{2778}N_{486}O_{550}S_9$	$C_{1538}H_{2496}N_{458}O_{467}S_{10}$
pI		9.54	5.86	5.31
positive AA		41	38	33
negative AA		33	40	45
Ec value	CC	34045	35535	18700
	CR	33920	35410	18450
II		27.67	18.59	33.33
AI		86.28	86.53	102.6
Gravy		-0.147	-0.044	0.098

estimated AI value defined by the presence of the aliphatic side chains (59), detailed that WP_037680047 have a higher AI value of 102.6 in comparison to the remaining two asparaginase proteins. The AI of WP_044568095 and WP_181864635 is computed to be nearly similar with 86.28 and 86.53 values accordingly inferring all the three proteins to be stable. Determining the Extinction coefficient (EC) of the protein is essential in measuring the quantity of light being absorbed in a specific wavelength, ideally referring to 280nm (37). The EC values of the three asparaginase proteins ranging from 18700 - 34045 are measured on the amount of tyrosine, tryptophan and cystine residues present in a protein (37, 60). The GRAVY index indicated that the proteins WP_044568095, WP_181864635 are categorized as hydrophilic and

WP_037680047 is predicted to be a membrane protein

Molecular docking and MD simulation analysis of L-asparaginase of three different *Streptomyces* species

with hydrophobic characteristics (61).

Determination of secondary and tertiary structures and quality assessment

The evaluated secondary structure by SOPMA (39) infers that all three asparaginase proteins WP_044568095, WP_181864635, WP_037680047 represented high occurrence of coils, followed by α-helices, strands, ß-turns indicating a folded structural organization. However, the percentage of each secondary element varied and is tabulated in Table 2. Simultaneously the secondary structure interpreted by the PSIPRED server (38) is graphically depicted in Fig 2.

Table 2: Secondary structural elements of WP_044568095, WP_181864635, WP_037680047 proteins predicted by SOPMA (listed in percentage)

Structural elements	Sequence details		
	WP_044568095	WP_181864635	WP_037680047
α-helices	34.83%	33.69%	37.87%
Strand	19.79%	21.22%	11.24%
ß-turn	5.01%	5.04%	11.83%

| Coil | 40.37% | 40.50% | 39.05% |
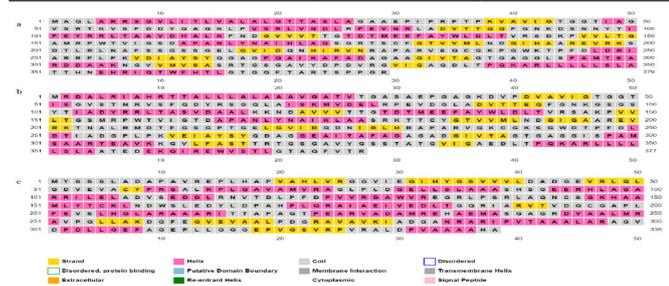
Figure 2: Pictorial representation of the secondary structural elements of Asparaginase protein (a) WP_044568095 (b) WP_181864635 (c) WP_037680047 by PSIPRED server.

The 3D structural elucidation of WP_044568095, WP_181864635, WP_037680047 determined by I-TASSER (41) is observed to have varied C-scores. The structures with C-scores -0.49 concerning *S. iranensis* (WP_044568095) protein, -0.51 related to *S. himalayensis* (WP_181864635) and -3.79 of *S. griseus* (WP_037680047) are chosen for further computational studies (Fig. 3). The TM-score and RMSD of WP_037680047 are estimated to be 0.30±0.10 and 15.9±3.2Å.

In a similar process, the TM-score and RMSD of WP_044568095 and WP_181864635 are also deduced which unexpectedly found to be similar correlating to a TM-score of 0.65±0.13 and RMSD of 7.8±4.4Å accordingly.

The prediction of the LOMETS server (62) based on the Z-scores listed many asparaginase proteins in alignment with *S. iranensis* (WP_044568095) and *S. himalayensis* proteins (WP_181864635). The data comprises asparaginase proteins of *Erwinia carotovora* (PDB id 2JK0A), *Erwinia chrysanthemi* (1HFJA) and a GatDE protein of *Methanothermobacter* species (2D6F) among the templates displaying alignment with both WP_044568095 and WP_181864635. In addition, asparaginases of diverse bacterial taxonomy such as *Wolinella succinogenes* (5K3OA), *Helicobacter pylori* (2WLTA) are identified to have high alignment with WP_181864635. The LOMETS server has also described alignment between *S. iranensis* (WP_044568095) and the asparaginase protein of *E. chrysanthemi* (1O7JA). Contrastingly, the templates listed for *S. griseus* (WP_037680047) have a diverse range of proteins namely a penicillin-binding protein (2WAE), *Yersinia pseudotuberculosis* (6RWB), esterase of *Alcaligenes* sp. (1QLW). Thus it is observed that the proteins of *S. iranensis* and *S. himalayensis* are noteworthy for further validation depending upon the similarity with a diverse range of asparaginases. Also, an attempt to understand the asparaginase of *S. griseus* is performed by further analysis.

The refinement of the proteins by GalaxyRefine indicated the proteins WP_044568095, WP_181864635 and WP_037680047 to have high Rama favoured values of 92.6, 92.3, and 87.5 respectively. Further in-depth analysis by PROCHECK indicated that *S. iranensis* and *S. himalayensis* belonging to PF00710 comprises a majority of residues greater than 80% in the favoured zone (Fig. 3).

The protein WP_044568095 is evaluated to have 85.9% (274 AA) favoured, 10.7% (34 AA) additionally allowed and 0.6% (2 AA) generously allowed and 2.8% (9AA) of disallowed residues. The protein WP_181864635 displayed 84.2% residues (271 AA) in the favoured and 11.2% (36 AA), 0.9% (3 AA), 3.7% (12 AA) of residues in the additionally, generously and disallowed regions accordingly. WP_037680047 protein is calculated to have 79.2%, 14.8%, 2.8%, 3.2% residues equaling to 224 AA in favoured, 42 AA in additionally allowed, 8 AA in generously allowed residues and 9 AA in disallowed regions respectively (Fig. 3) (Table 3).

The elucidation of the z-scores based on ProSA analysis (46) is measured between -7 to -9 characterizing the asparaginase proteins to be present within the score range of the native proteins in PDB. The protein models WP_044568095, WP_181864635 and WP_037680047 are found to have a z-score of -9.06, -8.71 and -7.82 consecutively. The proteins WP_044568095, WP_181864635 are termed to be of higher quality than WP_037680047 based upon the concept of energy plot as a function of the AA residues (46). The energy plot of WP_044568095 and WP_181864635 displayed minute peaks in the N-terminal part of the protein with a maximum

of AAs below the baseline with negative residual charges (Fig. 4). Although few exceptional positive peaks
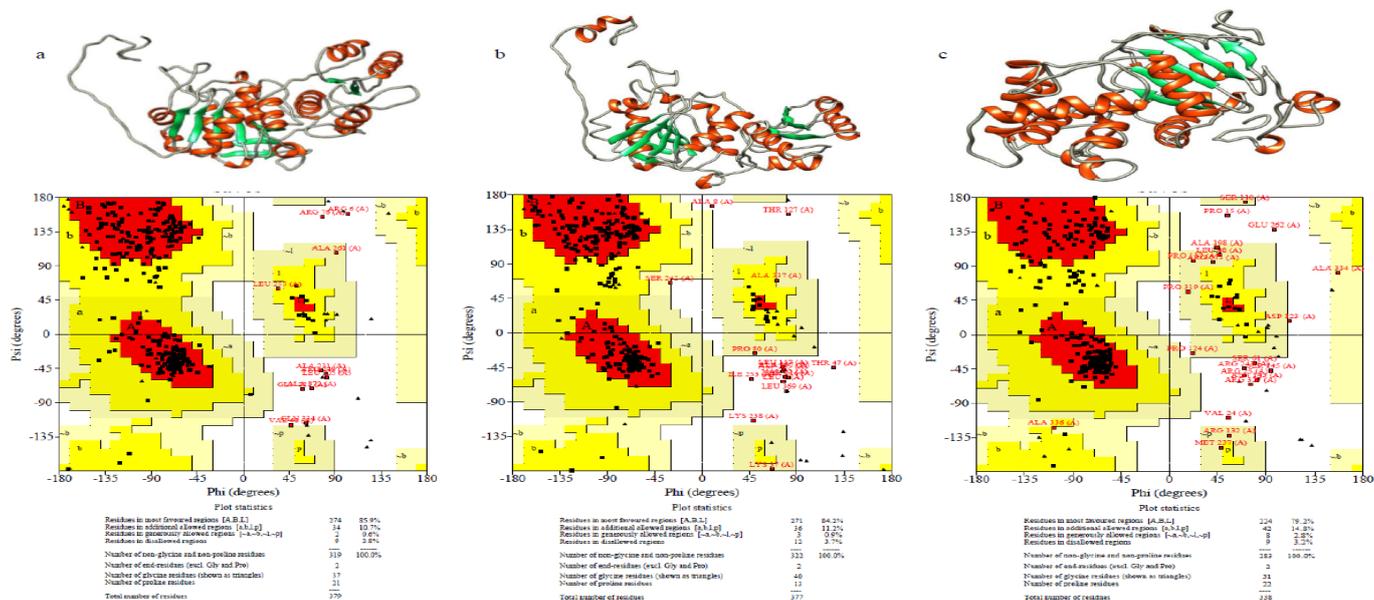
Figure 3: The 3D Structural determination (upper panel) and Ramachandran plot analysis (lower panel) of (a) WP_044568095 (b) WP_181864635 (c) WP_037680047.

concerning the energy plot of WP_037680047, the protein model can be still considered acceptable as the z-score is within the range of proteins (Fig. 4).

Table 3: Structural assessment of WP_044568095, WP_181864635, WP_037680047 asparaginase proteins by PROCHECK and ProSA

Protein	Protein structural assessment criteria				ProSA z-score
	Favoured	additionally allowed	generously allowed	disallowed	
WP_044568095	85.90%	10.70%	0.60%	2.80%	-9.06
WP_181864635	84.20%	11.20%	0.90%	3.70%	-8.71
WP_037680047	79.20%	14.80%	2.80%	3.20%	-7.82

Proceeding further, refined proteins WP_044568095, WP_181864635 and WP_037680047 are reported to have an ERRAT quality score of 79.213, 89.049 and 71.733 sequentially (Fig. 5). Thus the proteins are termed to have good quality based on the non-bonded interactions with scores greater than 50.

Molecular docking

Evaluation of the macromolecule-ligand interaction by computational approach prior to drug

Figure 4: Quality assessment of (a) WP_044568095 (b) WP_181864635 (c) WP_037680047 by ProSA z-score (left) and energy plot (right) of the modelled proteins accordingly.

Molecular docking and MD simulation analysis of L-asparaginase of three different *Streptomyces* species
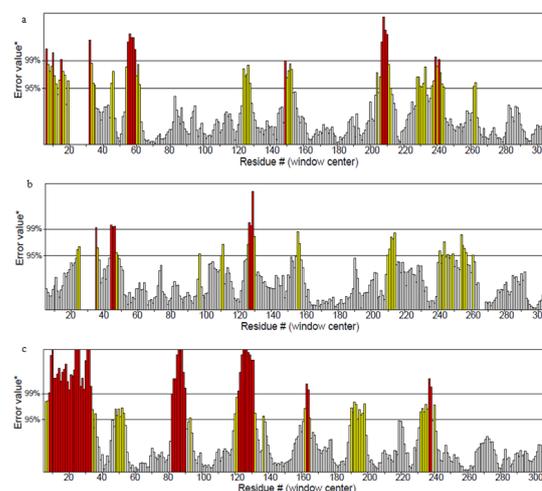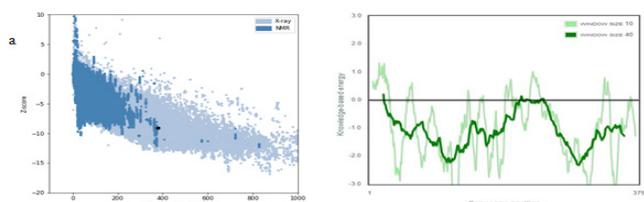
development is essential as it allows for the assessment of protein function. Reliable docked poses and dock scores between the receptor and

Figure 5: ERRAT Quality assessment (a) WP_044568095 (b) WP_181864635 (c) WP_037680047. The ERRAT scores for proteins represented with a, b, c in the image are calculated to be 79.213, 89.049 and 71.733 respectively.

respective ligand Asn is obtained by performing molecular docking methodology with CB-Dock and PyRx. As both the distinct docking applications are based on AutoDock Vina, the derived interpretations are also found similar thus confirming the results to be accurate.

Protein-ligand interactions analysed by PyRx

The AutoDock Vina application, embedded in PyRx software calculated the docked protein-ligand complex of *S. iranensis* to have the lowest possible energy of -4.6 kcal/mol while *S. himalayensis* and *S. griseus* are determined to have -4.4 kcal/mol. The docking of *S. iranensis* asparaginase protein (WP_044568095) with the ligand is performed within the specified grid dimensions 57.91 x 56.61 x 49.96 Å. The 2D representation illustrated by the Discovery studio for the protein WP_044568095 detailed Ser69, Asp84, Val85 and Thr87 forming prominent hydrogen bonds with the ligand (Fig. 6). The ligand is also encircled with Van der Waal interactions by Ile32, Pro35, Lys37, Val residues 68 and 72, Thr86 (Table 4).

The docked complex of *S. himalayensis* is determined within the specified Vina search space of 64.82 x 53.57 x 50.57 Å (x, y, z coordinates). It is
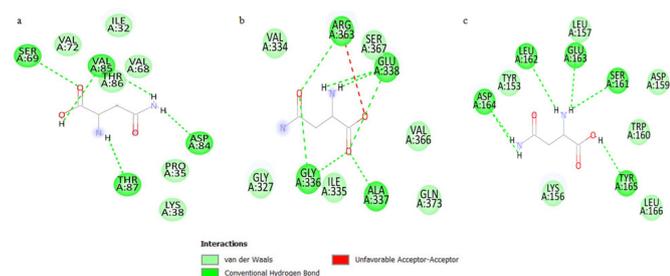
observed that the residues Gly336, Ala337, Glu338, Arg363 of *S. himalayensis* (WP_181864635) interacted with Asn by hydrogen bond formation (Table 4) (Fig. 6).

Figure 6: A 2D representation of the docked structure (a) WP_044568095, (b) WP_181864635, (c) WP_037680047 with Asn by Discovery studio.

The AA GLY 336 and GLU338 are identified important residues in the active site as GLY 336 formed hydrogen bonds with two oxygen atoms and GLU338 bonded with two hydrogen and a single oxygen atom of the ligand. Gly327, Val334, Ile335, Val366, Ser367,

Gln373 exhibited Van der Waals attractions with Asn. The AA Arg363 in addition to the stable hydrogen bond also formed an unfavourable acceptor-acceptor bond with an oxygen atom of the ligand.

The *S. griseus* active site is predicted in the x, y, z coordinates of 47.33 x 57.62 x 56.60 Å. The protein interaction exhibited more hydrogen bonds in comparison with WP_044568095, WP_181864635. The WP_037680047 residues between 161 to 165 namely, Ser161, Leu162, Glu163, Asp164, Tyr 165 formed a hydrogen bond with Asn (Fig. 6). In addition, few residues between 153 to 166 displayed Van der Waals bond formation with ligand (Table 4).

Protein-ligand interactions analysed by CB dock

The predictions of CB-Dock regarding the docked poses and the residues involved in the active site are similar to the predictions of PyRx. However, the binding affinity slightly varied. The binding affinity of WP_044568095 and WP_037680047 with Asn is calculated to be -4.4 kcal/mol while -4.1 kcal/
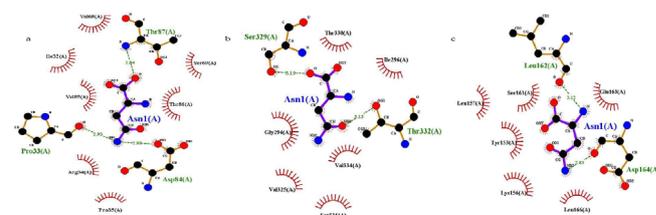
mol is identified between the WP_181864635 and the ligand Asn. The predicted docked structures on further visualization by Ligplot+ desktop application depicted hydrogen and hydrophobic interactions of WP_044568095, WP_181864635, WP_037680047 with Asn are listed in Table 4 and illustrated in Fig. 7

Figure 7: Ligplot+ analysis of (a) WP_044568095, (b) WP_181864635, (c) WP_037680047 with Asn. Green dotted lines with respective bond length mentioned represent hydrogen bond between protein and Asn, the surrounding AA in semi circle indicates hydrophobic interactions.

Analysis of the docked structures

As the current study involves molecular docking by two significant applications based on Autodock Vina, the analysis of the best docked protein structure will emphasize confidence in the prediction of active sites and binding possibilities. Pro33, Asp84 and Thr87 of WP_044568095 bonded with Asn via hydrogen bond with bond length measuring 2.93 Å, 2.88 Å and 2.84 Å accordingly. The AA Asp84, Thr87 forming hydrogen bond and Ile32, Pro35, Val68, Thr86 featuring hydrophobic interactions are also identified by PyRx studies along with few additional AA residues. The AA residues in the active site of WP_181864635 discovered by both the software's based on Autodock Vina are found

Table 4: The active site residues predicted by PyRx and CB-Dock. Both the docking methods predicted similar residues in active site for WP_044568095, WP_181864635 and WP_037680047 modelled proteins

Protein	PyRx		CB-Dock	
	Type of interaction	AA involved	Type of interaction	AA involved
WP_044568095	hydrogen bond	Ser69, Asp84, Val85, Thr87	hydrogen bond	Pro33, Asp84, Thr87
	van der Waals	Ile32, Pro35, Lys37, Val68, Val72, Thr86	hydrophobic interaction	Ile32, Arg34, Pro35, Val68, Ser69, Val85, Thr86
WP_181864635	hydrogen bond	Gly336, Ala337, Glu338, Arg363	hydrogen bond	Ser329, Thr332
	van der Waals	Gly327, Val334, Ile335, Val366, Ser367, Gln373	hydrophobic interaction	Gly294, Ile296, Ser316, Val325, Thr330, Val334
WP_037680047	hydrogen bond	Ser161, Leu162, Glu163, Asp164, Tyr165	hydrogen bond	Leu162, Asp164
	van der Waals	Tyr153, Lys156 Leu157, Asp159, Trp160, Leu166	hydrophobic interaction	Tyr153, Lys156, Leu157, Ser161, Glu163, Leu166

to vary however, the predicted residues of the active sites are considered accurate as both the applications have identified residues between 325 and 375 forming hydrogen and hydrophobic bonds. The active sites of WP_037680047 for the ligand Asn are identified between AA 150-170. Leu162, Asp164 of WP_037680047 where a hydrogen bond measuring the bond length of 3.12 Å, 2.83 Å respectively with Asn. The residues exhibiting hydrophobic interaction constituting Tyr153, Lys156, Leu157 and Leu166 are also found similar by both docking methodologies. The docked protein-ligand structures with a strong affinity to Asn analysed by the two docking applications revealed the docked proteins to be accurate with the majority of the active site residues being identical.

MD Simulation analysis

The analysis of the bound receptor by MD simulation studies provides insights into structural rearrangement, efficiency and predicts a reliable protein model [63]. The MD simulations for WP_044568095, WP_181864635 and WP_037680047 yielded ten protein-ligand complexes, with the first model being chosen due to its enhanced stability.

The simulation studies of *S. iranensis* evaluated by CABS-Flex 2.0 detailed the root mean square fluctuation (RMSF) graph to have varied fluctuations ranging between 0.1 and 4.2 Å (Fig. 8). The N-terminal residues between 8 and 35 are recorded to have significant variations with AA 29 indicating RMSF of 4.250. The computed high RMSF values imply higher AA flexibility and vice-versa [64]. WP_181864635 exhibited high RMSF values fluctuating between 4.13 to 9.74 Å in the N- terminal region with the highest amplitude where

by AA 34 (Fig. 8).

The remaining residues showed moderate fluctuations with RMSF values varied between 0.1 to 3.0 Å. The residues of WP_044568095 and WP_181864635 forming hydrogen bonds with Asn indicated fluctuations below 0.1810 Å and 0.7190 Å respectively. This infers that *S. iranensis* and *S. himalayensis* proteins bound with the ligand Asn are identified to be stable and the
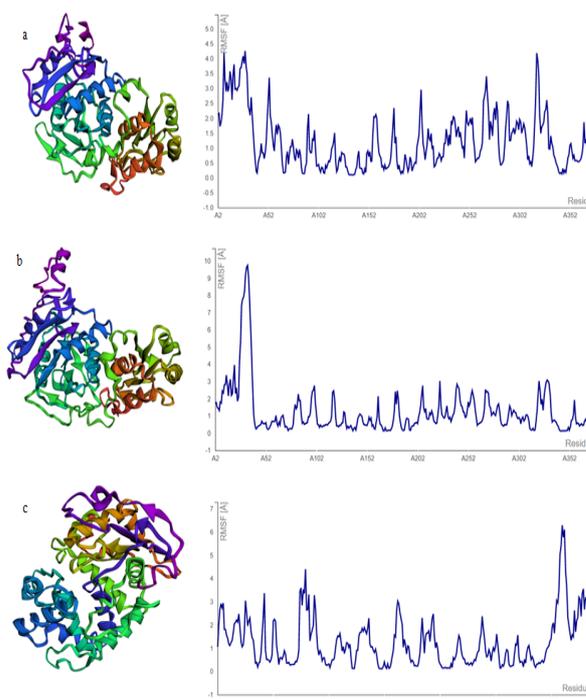
hydrogen bond interaction is observed to be rigid. The simulation analysis of WP_037680047, depicted

residual RMSF values to be within 1.06-6.29 Å, with a higher RMSF value observed at residual position 312. The active site residues of WP_037680047 interacting via hydrogen bond with the ligand has fluctuations ranging from 1.4860 Å - 3.0270 Å.

Figure 8: CABS-Flex 2.0 simulation analysis of (a) WP_044568095 (b) WP_181864635 (c) WP_037680047 asparaginase proteins. The protein model after subjecting to simulation (left) and the corresponding RMSF plot of the residues is also displayed (left).

It is observed that determined RMSF values of WP_037680047 active site are slightly higher than WP_044568095, WP_181864635.

Conclusion

The broad therapeutic use of asparaginase emphasizes the importance of research on the proteins from distinct sources, since they are extremely important in the treating ALL, acute and chronic myeloid leukemia and lymphosarcoma. The application of computational techniques to analyse the various criteria of asparaginase proteins from three different species resulted in a better knowledge of the similar and unique properties among the proteins from the same *Streptomyces* genus. Primarily, WP_044568095 and WP_181864635 displayed moderate similarity with asparaginases of D. chrysanthemi (P06608) and E. coli L-asparaginase II (P00805). The three modeled *S. iranensis* (WP_044568095), *S. himalayensis* (WP_181864635) and *S. griseus* (WP_037680047) proteins exhibited affinity towards Asn, necessitated for asparaginase activity. In addition, the study also reports that the MD simulations of the docked structures to be stable. The findings conclude that the assessed proteins are worth further research for in vitro enzyme activity and are considered to have potential therapeutic value.

Acknowledgements

Conflict of interest

The authors declare no conflicts of interest.

References

1. Lee, L.H., Zainal, N., Azman, A.S., Eng, S.K., Goh, B.H., Yin, W.F., Ab Mutalib, N.S. and Chan, K.G. (2014). Diversity and antimicrobial activities of actinobacteria isolated from tropical mangrove sediments in Malaysia. The scientific world journal, 2014.

2. Ser, H.L., Palanisamy, U.D,, Yin, W.F., Abd Malek, S.N., Chan, K.G., Goh, B.H. and Lee, L.H. (2015). Presence of antioxidative agent, Pyrrolo (1, 2-a] pyrazine-1, 4-dione, hexahydro-in newly isolated Streptomyces mangrovisoli sp. nov. Frontiers in microbiology, 20: 854.

3. Tan, L.T., Ser, H.L., Yin, W.F., Chan, K.G., Lee, L.H. and Goh, B.H. (2015). Investigation of antioxidative and anticancer potentials of Streptomyces sp. MUM256 isolated from Malaysia mangrove soil. Frontiers in microbiology, 26: 1316.

4. Glasby, J.S. (1993). Encyclopedia of Antibiotics (3rd edition), John Wiley and Sons, Chichester, pp. 1-507.

5. Lee, L.H., Zainal, N., Azman, A.S., Eng, S.K., Ab Mutalib, N.S., Yin, W.F. and Chan, K.G. (2014). Streptomyces pluripotens sp. nov., a bacteriocin-producing streptomycete that inhibits meticillin-resistant Staphylococcus aureus. International Journal of Systematic and Evolutionary Microbiology, 64: 3297-3306.

6. Selvakumar, J.N., Chandrasekaran, S.D. and Vaithilingam M. (2015). Bio prospecting of marine-derived Streptomyces spectabilis VITJS10 and exploring its cytotoxicity against human liver cancer cell lines. Pharmacognosy magazine, 11: S469.

7. Ueno, T., Ohtawa, K., Mitsui, K., Kodera, Y., Hiroto, M., Matsushima, A., Inada, Y. and Nishimura, H. (1997). Cell cycle arrest and apoptosis of leukemia cells induced by L-asparaginase. Leukemia, 11: 1858-61.

8. Avramis, V.I. and Tiwari, P.N. (2006). Asparaginase (native ASNase or pegylated ASNase) in the treatment of acute lymphoblastic leukemia. International journal of nanomedicine, 1: 241.

9. Piątkowska-Jakubas, B., KrawczykKuliś, M., Giebel, S., AdamczykCioch, M., Czyż, A., Lech-Marańda, E., Paluszewska, M., Pałynyczko, G., Piszcz, J. and Hołowiecki, J. (2008). Use of Lasparaginase in acute lymphoblastic leukemia: recommendations of the Polish Adult Leukemia Group. Polskie Archiwum Medycyny Wewnętrznej= Polish Archives of Internal Medicine, 118.

10. Ahmed, T., Holwerda, S., Klepin, H.D., Isom, S., Ellis, L.R., Lyerly, S., Manuel, M., Dralle, S., Berenzon, D., Powell, B.L. and Pardee, T.S. (2015). High dose cytarabine, mitoxantrone and l-asparaginase (HAMA) salvage for relapsed or refractory acute

myeloid leukemia (AML) in the elderly. Leukemia research, 39: 945-9.

11. Margolin. J.F. (1997) Acute lymphoblastic leukemia. Principle and Practice of Pediatric Oncology, 409-62.

12. Song, P., Ye, L., Fan, J., Li, Y., Zeng, X., Wang, Z., Wang, S., Zhang, G., Yang, P., Cao, Z. and Ju, D. (2015). Asparaginase induces apoptosis and cytoprotective autophagy in chronic myeloid leukemia cells. Oncotarget, 6: 3861.

13. Jaccard, A., Petit, B., Girault, S., Suarez, F., Gressin, R., Zini, JM., Coiteux, V., Larroche, C., Devidas, A., Thieblemont, C. and Gaulard, P. (2009). L-asparaginase-based treatment of 15 western patients with extranodal NK/T-cell lymphoma and leukemia and a review of the literature. Annals of oncology, 20: 110-6.

14. Ehsanipour, E.A., Sheng, X., Behan, J.W., Wang, X., Butturini, A., Avramis V.I. and Mittelma, S.D. (2013). Adipocytes cause leukemia cell resistance to L-Asparaginase via release of glutamine. Cancer Res, 73: 2998–6.

15. Mitchell, L., Hoogendoorn, H., Giles, A., Vegh, P. and Andrew, M. (1994). Increased endogenous thrombin generation in children with acute lymphoblastic leukemia: risk of thrombotic complications in L'Asparaginase-induced antithrombin III deficiency. Blood, 83: 386-91.

16. Rizzari, C., Zucchetti, M., Conter, V., Diomede, L., Bruno, A., Gavazzi, L., Paganini, M., Sparano, P., Nigro, LL., Arico, M. and Milani, M. (2000). L-asparagine depletion and L-asparaginase activity in children with acute lymphoblastic leukemia receiving im or iv Erwinia C. or E. coli L-asparaginase as first exposure. Annals of oncology, 11: 189-93.

17. Netzker, T., Schroeckh, V., Gregory, M.A., Flak, M., Krespach, M.K., Leadlay, P.F. and Brakhage, A.A. (2016). An efficient method to generate gene deletion mutants of the rapamycin-producing bacterium Streptomyces iranensis HM 35. Applied and environmental microbiology, 82: 3481-92.

18. Li, J., Kim, S.G. and Blenis, J. (2014). Rapamycin: one drug., many effects. Cell metabolism, 19: 373-9.

19. Krainsky, A. (1914). Die Aktinomyceten und ihre Bedeutung in der Natur., Centr. Bakteriol. Parasitenk. Abt. II, 42: 649-88.

20. Kampfer, P. (2006). The family Streptomycetaceae, part I: taxonomy. The prokaryotes, 3: 538-604.

21. Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M., Ikeda, H., Yamashita, A., Hattori, M. and Horinouchi, S. (2008). Genome sequence of the streptomycin-producing microorganism Streptomyces griseus IFO 13350. Journal of bacteriology, 190: 4050-60.

22. Dejong, P.J. (1972). L-Asparaginase production by Streptomyces griseus. Applied microbiology, 23: 1163-4.

23. Meena, B., Anburajan, L., Sathish, T., Raghavan, R.V., Dharani, G., Vinithkumar, N.V. and Kirubagaran, R. (2015). L-Asparaginase from Streptomyces griseus NIOT-VKMA29: optimization of process variables using factorial designs and molecular characterization of l-asparaginase gene. Scientific reports, 5: 1-12.

24. Bruntner, C., Binder, T., Pathom-aree, W., Goodfellow, M., Bull, A.T., Potterat, O., Puder, C., Hörer, S., Schmid, A., Bolek, W. and Wagner, K. (2005). Frigocyclinone., a novel angucyclinone antibiotic produced by a Streptomyces griseus strain from Antarctica. The Journal of antibiotics, 58: 346-9.

25. Zhang, D.J., Wei, G., Wang, Y., Si, C.C., Tian, L., Tao, L.M. and Li, Y.G. (2011). Bafilomycin K., a new antifungal macrolide from Streptomyces flavotricini Y12-26. The Journal of antibiotics, 64: 391-3.

26. Zhang, Y., Liang, Q., Zhang, C., Zhang, J., Du, G. and Kang, Z. (2020). Improving production of Streptomyces griseus trypsin for enzymatic processing of insulin precursor. Microbial cell factories, 19: 1-1.

27. Sidhu, S.S., Kalmar, G.B., Willis, L.G. and Borgford, T.J. (1994). Streptomyces griseus protease C. A novel enzyme of the chymotrypsin superfamily. Journal of Biological Chemistry, 269: 20167-71.

28. Berger, L.R., Reynold, D.M. (1958). The chitinase system of a strain of Streptomyces griseus. Biochimica et biophysica acta, 29: 522-34.

29. Vishnupriya, B., Sundaramoorthi, C., Kalaivani, M. and Selvam K. (2010). Production of lipase from Streptomyces griseus and evaluation of Bioparameters. Int J Chem Tech Res, 2: 1380-3.

30. Narahashi, Y. and Yanagita, M. (1967). Studies on Proteolytic Enzymes (Pronase) of Streptomyces griseus K-1: I. Nature and Properties of the Proteolytic Enzyme System. The Journal of Biochemistry, 62: 633-41.

31. NCBI Resource Coordinators. (2018). Database

resources of the National Center for Biotechnology Information. Nucleic Acids Research, 46: 8–13.

32. UniProt Consortium. (2019). UniProt: a worldwide hub of protein knowledge. Nucleic acids research, 47: D506-15.

33. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman DJ. (1990). Basic local alignment search tool. Journal of molecular biology, 215: 403-10.

34. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting., position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22: 4673-80.

35. Kumar, S., Stecher, G. and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular biology and evolution, 33: 1870-4.

36. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer EL., Tosatto, S.C., Paladin, L., Raj, S., Richardson, L.J. and Finn, R.D. (2021). Pfam: The protein families database in 2021. Nucleic Acids Research, 49: 412–19.

37. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, MR., Appel, RD. and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook, 571-607.

38. Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. Journal of molecular biology, 292(2): 195-202.

39. Geourjon, C. and Deleage, G. (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics, 11: 681-4.

40. de Lima Corrêa, L., Borguesan, B., Krause, M.J. and Dorn, M. (2018). Three-dimensional protein structure prediction based on memetic algorithms. Computers & Operations Research, 91: 160-77.

41. Yang, J. and Zhang, Y. (2015). I-TASSER server: New development for protein structure and function -predictions. Nucleic Acids Res, 43: 174–81.

42. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. Nature methods, 12: 7-8.

43. Heo L., Park H. and Seok C. (2013). GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic Acids Res, 41: 384-8.

44. Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr, 26: 283–91.

45. Colovos, C., Yeates, T.O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Sci, 2: 1511–9.

46. Wiederstein, M. and Sippl, M.J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res, 35: 407-10.

47. Liu, Y., Grimm, M., Dai, W.T., Hou, M.C., Xiao, Z.X. and Cao, Y. (2020). CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. Acta Pharmacologica Sinica, 41: 138-44.

48. Dallakyan, S. and Olson, A.J. (2015). Small-molecule library screening by docking with PyRx. Methods Mol Biol, 1263: 243–50.

49. Trott, O. and Olson, A.J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function., efficient optimization., and multithreading. J Comput Chem, 31: 455-61.

50. O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T. and Hutchison GR. (2011). Open Babel: an open chemical toolbox. J Cheminform, 3: 1-4.

51. Alberg, D.G. and Schreiber, S.L. (1993). Structure-based design of a cyclophilin-calcineurin bridging ligand. Science, 262: 248-50.

52. BIOVIA. (2016). Dassault Systemes. Discovery Studio Modeling Environment v2021. San Diego: Dassault Systemes.

53. Laskowski, R.A., Swindells, M.B. (2011). LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. Journal of Chemical Information and Modeling, 51: 2778-2786.

54. Kuriata, A., Gierut, A.M., Oleniecki, T., Ciemny, M.P., Kolinski, A., Kurcinski, M., Kmiecik, S. (2018). CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. Nucleic acids research, 46: W338-43.

55. Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Bio Evol, 4: 406-25.

56. Ciechanover, A. and Schwartz, A.L. (1989). How are substrates recognized by the ubiquitin-mediated proteolytic system. Trends in biochemical sciences, 14: 483-8.

57. Varshavsky, A. (1997). The N-end rule pathway of protein degradation. Genes to cells, 2: 13-28.

58. Guruprasad, K., Reddy, B.B. and Pandit, M.W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Engineering, Design and Selection, 4: 155-61.

59. Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. The Journal of Biochemistry, 88: 1895-8.

60. Gill, S.C., Von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Analytical biochemistry, 182: 319-26.

61. Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of molecular biology, 157: 105-32.

62. Wu, S. and Zhang, Y. (2007). LOMETS: a local meta-threading-server for protein structure prediction. Nucleic acids research, 35: 3375-82.

63. Hospital, A., Goñi, J.R., Orozco, M. and Gelpí, J.L. (2015). Molecular dynamics simulations: advances and applications. Advances and applications in bioinformatics and chemistry, 8: 37.

Kmiecik, S., Gront, D., Kolinski, M., Wieteska, L., Dawid, A.E. and Kolinski, A. (2016). Coarse-grained protein models and their applications. Chemical reviews, 116: 7898-936.