Functional & Integrative Genomics (2023) 23:254
https://doi.org/10.1007/s10142-023-01179-w

ORIGINAL ARTICLE

# Exploring and targeting potential druggable antimicrobial resistance targets ArgS, SecY, and MurA in *Staphylococcus sciuri* with TCM inhibitors through a subtractive genomics strategy

**Aafareen Khan[1] · Saman Sohail[1] · Seerat Yaseen[2] · Sareen Fatima[3] · Ayesha Wisal[1] · Sufyan Ahmed[2] · Mahrukh Nasir[4] · Muhammad Irfan[4] · Asad Karim[4] · Zarrin Basharat[5] · Yasmin Khan[4] · Muhammad Aurongzeb[6] · Syed Kashif Raza[7] · Mohammad Y. Alshahrani[8] · Carlos M. Morel[9] · Syed S. Hassan[4,9]**

Received: 20 March 2023 / Revised: 14 July 2023 / Accepted: 14 July 2023 / Published online: 26 July 2023


## Abstract

*Staphylococcus sciuri* (also currently *Mammaliicoccus sciuri*) are anaerobic facultative and non-motile bacteria that cause significant human pathogenesis such as endocarditis, wound infections, peritonitis, UTI, and septic shock. Methicillin-resistant *S. sciuri* (MRSS) strains also infects animals that include healthy broilers, cattle, dogs, and pigs. The emergence of MRSS strains thereby poses a serious health threat and thrives the scientific community towards novel treatment options. Herein, we investigated the druggable genome of *S. sciuri* by employing subtractive genomics that resulted in seven genes/proteins where only three of them were predicted as final targets. Further mining the literature showed that the ArgS (WP_058610923), SecY (WP_058611897), and MurA (WP_058612677) are involved in the multi-drug resistance phenomenon. After constructing and verifying the 3D protein homology models, a screening process was carried out using a library of Traditional Chinese Medicine compounds (consisting of 36,043 compounds). The molecular docking and simulation studies revealed the physicochemical stability parameters of the docked TCM inhibitors in the druggable cavities of each protein target by identifying their druggability potential and maximum hydrogen bonding interactions. The simulated receptor-ligand complexes showed the conformational changes and stability index of the secondary structure elements. The root mean square deviation (RMSD) graph showed fluctuations due to structural changes in the helix-coil-helix and beta-turn-beta changes at specific points where the pattern of the RMSD and root mean square fluctuation (RMSF) (< 1.0 Å) support any major domain shifts within the structural framework of the protein–ligand complex and placement of ligand was well complemented within the binding site. The β-factor values demonstrated instability at few points while the radius of gyration for structural compactness as a time function for the 100-ns simulation of protein–ligand complexes showed favorable average values and denoted the stability of all complexes. It is assumed that such findings might facilitate researchers to robustly discover and develop effective therapeutics against *S. sciuri* alongside other enteric infections.

**Keywords** *Staphylococcus sciuri* · Druggable targets · TCM library · Subtractive genomics · Molecular docking · MD simulation

## Introduction

*Staphylococcus sciuri* was found by Kloos et al. (1976) after the genus *Staphylococcus* was defined in 1880. *Staphylococci* are Gram-positive, anaerobic facultative, catalase-positive but oxidase-negative non-motile bacteria with peptidoglycan and teichoic acid in their cell walls (Zimmerman and Kloos 1976, Kloos 1980, Kloos et al. 1997). It was reported as a common bacterium thriving in a wide variety of environments and was originally thought to be a commensal bacteria found in healthy or ill farms and wild animals, but it has also been found in hospitalized individuals. There

Aafareen Khan, Saman Sohail, Seerat Yaseen, Sareen Fatima, Ayesha Wisal, Sufyan Ahmed, Mahrukh Nasir, Muhammad Irfan, Asad Karim, Zarrin Basharat, Yasmin Khan, Muhammad Aurongzeb, Syed Kashif Raza, Mohammad Y. Alshahrani, Carlos M. Morel, and Syed S. Hassan contributed equally to this work.

✉ Carlos M. Morel
carlos.morel@cdts.fiocruz.br; cmmorel@gmail.com

Extended author information available on the last page of the article

are reports of MRSS incidence in healthy broilers showing that *S. sciuri* may be a source of virulence and resistance genes, also depicting the clonal nature of the methicillin-resistant strains. Other studies investigated the prevalence of methicillin-resistant coagulase-negative *Staphylococci* (MRCoNS), while rarer studies were reported about MRSS employing selective isolation (Zimmerman and Kloos 1976, Kloos et al. 1997, Dakic et al. 2005). The bacterium causes significant pathogenesis in humans such as endocarditis, wound infections, peritonitis, septic shock, and urinary tract infections; the pathogenicity in animal population, however, is little understood, though antimicrobial resistance is prevalent. The selected pressure of frequent and non-specific use of antimicrobials for preventative, therapeutic, or growth booster purposes, mostly in pigs and poultry, has enhanced resistance and has been seen primarily in domestic animals. This resistance has spread to other wild species that share the same habitats and resources as domestic animals (Hedin and Widerström 1998). These species collect resistance and virulence genes circulating in a particular environment because of their toughness, prolificacy, and dispersal capacities (Kolawole and Shittu 1997). Because these genes are usually encoded in mobile genomic components, this is made easier (plasmids, chromosomal cassettes, transposons). These components may easily be transported horizontally across microorganisms, regardless of whether the receivers are pathogenic or non-pathogenic (Horii et al. 2001). The ancestral *Staphylococcal* species are likely to be *Staphylococcus sciuri*. It is often found on the skin and mucous membranes of warm-blooded animals in the environment and people (Torres et al. 2020) and currently is linked to mastitis in dairy cattle (Rahman et al. 2005, Nemeghaire et al. 2014b), dermatitis in dogs (Hauschild and Wójcik 2007), and exudative epidermis in piglets (Hauschild and Schwarz 2003).

It has been discovered that *S. sciuri* possesses a close homolog of the methicillin-resistance gene mecA seen in *S. aureus* (Wu et al. 2001). The *S. sciuri* group comprises five species including *S. sciuri* (three subspecies), *S. lentus*, *S. vitulinus*, *S. fleurettii*, and *S. stepanovicii* that have been isolated from diseases caused in both animals and humans (Hauschild and Wójcik 2007, Nemeghaire et al. 2014a). *S. sciuri*, the genus's original bacterium, and its closely related species were shown to carry the possible evolutionary progenitor of numerous resistance genes that might serve as a reservoir for *S. aureus* resistance and virulence genes (Wu et al. 1996), which include mecA gene complex (A to E) and eight cassette chromosome recombinase (ccr) gene complexes (ccrA1B1, ccrA2B2, ccrA3B3, ccrA4B4, ccrC1, ccrA5B3, ccrA1B6, and ccrA1B3) (Katayama et al. 2001).

There have been recent findings of multi-drug resistant (MDR) *Staphylococci* such as *S. sciuri* carrying multiple resistant genes towards commonly available β-lactams and other antibiotics, in Africa (Adesoji et al. 2020, Egyir et al. 2022), North and South America (Meservey et al. 2020, Salazar-Ardiles et al. 2020, Saraiva et al. 2021, de Carvalho et al. 2022, Santos et al. 2022), Korean Peninsula (Kim et al. 2019), Asia (Zhang et al. 2022, Boonchuay et al. 2023), Europe (Paterson 2020, Gómez-Sanz et al. 2021, Rey Pérez et al. 2021), Middle East (Khazandi et al. 2018, Al-Hayawi 2022), and other parts of the world, which is an increasing public health concern for treatment of life-threatening infections.

To find novel therapy and other prophylaxis options towards drug, vaccine, and diagnostic biomarkers, the primary phase in all protocols is the target identification in the post-genomics era. This can be achieved through various experimental as well as commonly used in silico approaches including pangenomics, subtractive genomics, structure-based drug designing (SBDD), comparative genomics, genome mining for metabolic pathway reconstruction, network pharmacology–based analyses, and reverse vaccinology, among other recently established computer-aided techniques (Ibrahim et al. 2017, Dalal et al. 2019, Singhal and Mohanty 2019, Dhankhar et al. 2020, Singh, Dhankhar et al. 2022, Zhang et al. 2023, Aregbesola et al. 2021). These methods have widely been used for the identification of protein-based therapeutic and vaccine targets in common and XDR, MDR, and other Pan-drug-resistant pathogenic microorganisms including viruses, bacteria, parasites, and fungus (Hughes 2002, Somani et al. 2019, Khan et al. 2021a, Khan et al. 2022b, Zhang et al. 2023).

The emergence of antibiotic-resistant pathogens due to excessive and unnecessary medications causes their immediate control a challenging assignment, hence using integrated OMICS strategies including but not limited to transcriptomics, metabolomics, and proteomics, among others, for disease target and regulator/inhibitor discovery as well as disease origin and prevalence in a variety of infections to expedite the process with minimized expenses (Shouxiang, Xiaojuan et al. 2021, Lvqin, Xuefeng et al. 2021, Linhui, Yutao et al. 2022, Yuan, Zhang et al. 2022, Dindhoria et al. 2022, Laamarti et al. 2022, Deng et al. 2022). The advantages include reduced time, cost-effectiveness, robustness, labor, and reproducibility to fabricate broad-spectrum therapeutic candidates (Hassan et al. 2014, Radusky et al. 2015, Basharat et al. 2021, Aurongzeb et al. 2022, Irfan et al. 2023).

## Methodology, databases, and approaches

### Genome selection of Staphylococcus sciuri and prediction of the core genome

The *S. sciuri* was selected on the basis of broad-spectrum host pathogenesis, and their complete genomes available at the start of this work were retrieved from the Joint Genome Institute-Genome Online Database (JGI-GOLD) (https://

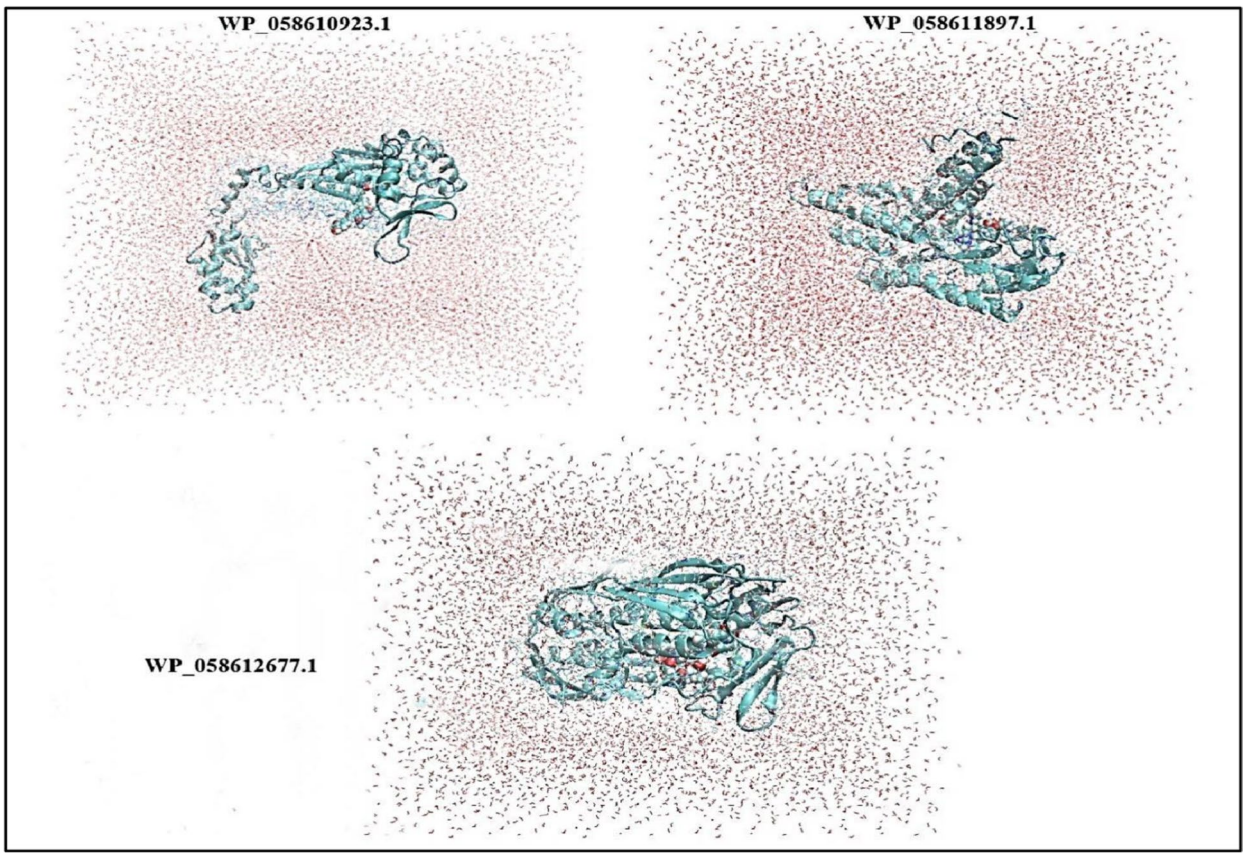


**Fig. 1** Solvation box surrounding the docked protein

gold.jgi.doe.gov/) where the genome data and other statistics are readily available for analyses. This database provides an open source of comprehensive access to information regarding metagenome sequencing projects and their associated metadata around the world (Mukherjee et al. 2017). To construct the core genome of *S. sciuri*, a high-throughput server called the Pathosystems Resource Integration Center (PATRIC, https://www.patricbrc.org/) was used to predict the core genome by randomly choosing one strain FDAARGOS_285 as the reference genome and the remaining ten strains were compared with this reference strain (Wattam et al. 2017).

### Non-host homologous, essential genome, and interactome prediction

After the prediction of core genome, the resultant data file was subjected to NCBI-BLASTp (https://www.ncbi.nlm.nih.gov/) (*e*-value = 0.0001, bit score = 200, and identity = 25%) against the human genome for filtering pathogen non-host homologs (Altschul et al. 1990). To identify conserved essential targets of *S. sciuri*, the set of core conserved proteins was submitted to the DEG (Database of Essential Genes) (http://tubic.tju.edu.cn/deg/) and CEG (Clusters of Essential Genes) (http://cefg.uestc.cn/ceg) servers using the default parameters. Essential genes in a bacterium constitute a minimal genome, forming a set of functional modules, which play key roles in the emerging field of synthetic biology and contain all the essential genes currently available (Luo et al. 2014, Liu et al. 2020). The STRING server (https://string-db.org/), utilized for the prediction of protein–protein interactome, serves as a biological database and web resource in molecular biology, encompassing both known and predicted protein–protein interactions (Szklarczyk et al. 2019).

### Comparative subcellular localization

The genes/proteins that were selected as non-redundant, essential, and human-non-homologous in the previous step were further analyzed for subcellular localization. This step is important to classify the proteins constituting the secretome and exoproteome of the pathogen. The exoproteome and secretome are considered as an excellent source for vaccine candidates. Subcellular localization of proteins was performed based on a comparative approach using two online subcellular localization tools: PSORTB (https://www.psort.org/psortb/) and CellO2GO (http://cello.life.nctu.edu.tw/cello2go/). The protein sequences in FASTA

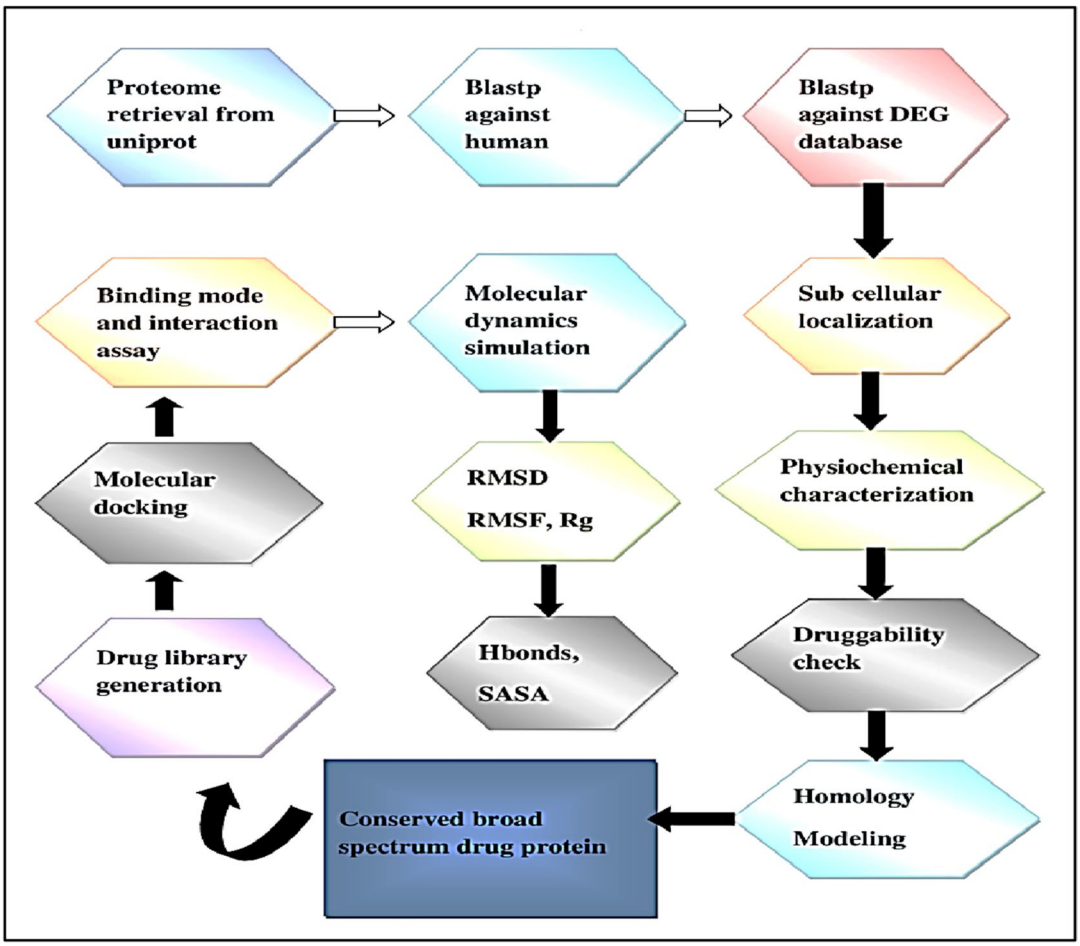


**Fig. 2** A schematic block diagram of different steps and tools employed in the subtractive genomics approach for mining druggable targets in *Staphylococcus sciuri* and identification of novel TCM inhibitors

format with organism type set to bacteria and gram stain set to negative were submitted. For bacteria, protein subcellular localization prediction (SCL) is the accurate tool to assign a possible localization site to a protein by using support vector machines (SVMs). Furthermore, it assigns the five subcellular locations, i.e., periplasm, cytoplasm, extracellular, inner outer membrane, and Gram-negative bacteria. In contrast, CELLO2GO considers SVM functionality at two levels; based on sequence-derived molecular descriptors followed by the probability of the subcellular location. Subcellular allocation and functional evaluation of a target protein is vital for proper drug design process and identification of a precise biological process (Yu et al. 2010, Yu et al. 2014).

## Identification of biological pathways and biological function

For the identification of different pathways involved in metabolism, the Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.genome.jp/kegg/pathway.html) was used to check which proteins are involved in unique or multiple pathways. KEGG is a collection of databases dealing with genomes, diseases, drugs, and biological pathways (Kanehisa et al. 2017). KAAS (KEGG Automatic Annotation Server) was used to filter the essential proteins for metabolic pathway analysis. In order to find out the biological/molecular function of proteins, UniProt (https://www.uniprot.org/) was consulted that is a freely accessible database of protein sequence and functional information. It contains a large amount of information about the biological function of proteins (Pundir et al. 2017). Furthermore, the molecular weight of each potential target was determined using the ProtParam software (https://web.expasy.org/protparam/) that helped in computational determination of the molecular weight of the subjected proteins. Virulent proteins were prioritized on the basis of molecular weight (Wilkins et al. 1999).

**Table 1** Total number of genes/proteins screened in each step of subtractive genomic/proteomic approach for druggable targets in *Staphylococcus sciuri.* The core genome drastically reduced after host homology and gene essentiality analyses and then to only three as the final pathogen targets

| S. no | Subtractive genomics hierarchy — tools/software | Total genes/proteins |
|---|---|---|
| 1 | Core Genome — PATRIC | 1784 |
| 2 | Non-Homologous/Essential Genome — NCBI/DEG/CEG | 170/35 |
| 3 | Interactome Prediction — STRING | 10 |
| 4 | Molecular Metabolic Pathways — KEGG | 07 |
| 5 | 3D Modelome — SWISS-MODEL | 03 |
| 6 | Structure Validation & Energy Minimization — UCSF Chimera | 03 |
| 7 | Druggability, VS & Docking Analyses — DoGSiteScorer/MOE (v2016) | 03 |
| 8 | Molecular Dynamics Simulation — AMBER12 | 03 |

## Drug target selection and 3D structure modeling

The potential druggable bacterial protein target has been identified from the reference strain of *S. sciuri* FDAARGOS_285. In the absence of their complete 3D structures, the possibility of comparative homology modeling was considered by evaluating the template availability of all 3 targets. Structural templates that showed at least 30% identity with > 90% query coverage were accepted. This assessment was carried out via comparison of the protein sequence against the structural resource RCSB-PDB, through the use of BLASTp functionality supported by NCBI. All these steps allowed to choose the current target for CADD analysis (Zhao et al. 2020). The three-dimensional structures of MurA, SecY, and ArgS were unavailable; thereby, as aforementioned, comparative modeling approaches were selected for 3D model prediction by selecting the filtered sequence throughout the pipeline to dig out the desired sequence from the reference genome of *Staphylococcus sciuri*. The template of MurA (PDB ID: 2AQ9) was selected for its modeling from *Escherichia coli* (Williams et al. 2007), with sequence identity = 91% and query coverage = 98%. The structural model of the target protein was constructed using the SWISS-MODEL that is an online homology-based web server (Colovos and Yeates 1993, Rufino et al. 1997).

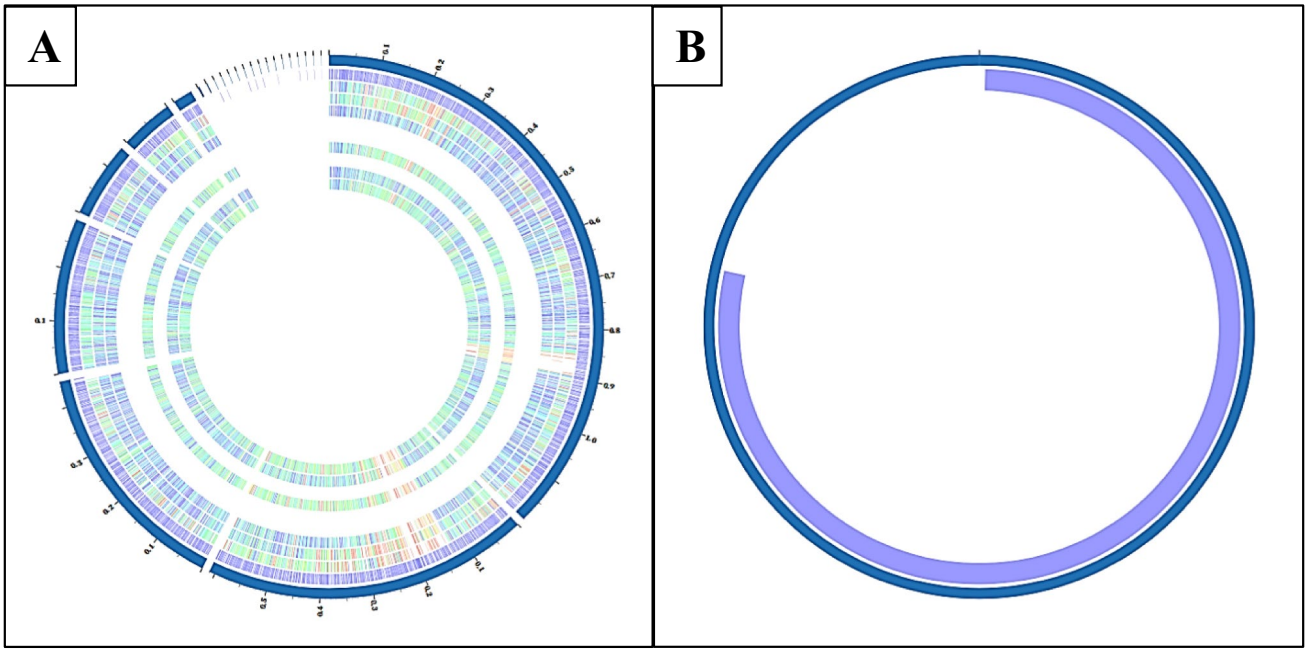


**Fig. 3** Circular comparative genome representation of *S. sciuri* genomes generated through the PATRIC Server. **A** FDAARGOS_285 vs nine strains (1/9). **B** FDAARGOS_285 vs one strain 1/1. Different colors and their intensities show the presence or absence of different genes, genic islets, genomic island, or other genetic materials among different strains of *S. sciuri*

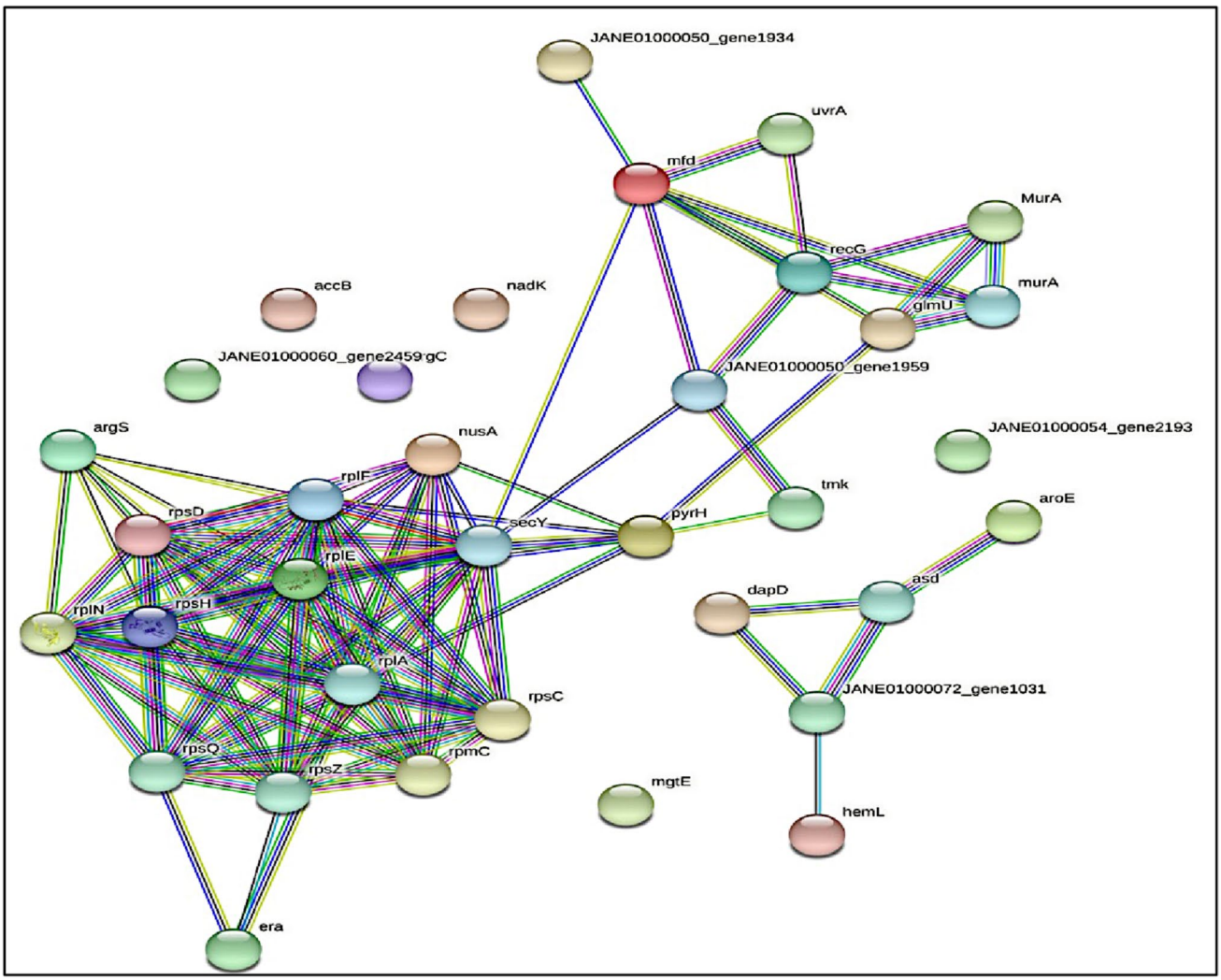


**Fig. 4** STRING analysis for protein–protein interactions. The different nodes in the network represent the proteins while the network edges represent specific and meaningful protein–protein associations. The network is a scalable vector graphic [SVG]; interactive. The different node colors show the different level of interactions whereas the edge colors show their known, predicted, and other interactions. The colored nodes show the query proteins and first shell of interactors, the white nodes represent second shell of interactors, empty nodes represent proteins of unknown 3D structure, and filled nodes represent that some 3D structure is known or predicted. The edges indicate both functional and physical protein associations, whereas line color indicates the type of interaction evidence and the line thickness indicates the strength of data support. Among the known interactions, those in cyan are from curated databases and those in purple are experimentally determined. In predicted interactions, those in green are from gene neighborhood analyses, those in red are gene fusion events, and those in blue are from gene co-occurrence. The other remaining interactions are olive = text-mining, black = co-expression, navy blue = protein homology

## Structure validation and energy minimization

Different online servers were utilized, namely, ERRAT (Colovos and Yeates 1993, Kumari et al. 2023), PDBSum (Rufino et al. 1997, Kumari and Dalal 2022), and ProCheck (Laskowski et al. 1993, Dhankhar et al. 2020, Singh, Dhankhar et al. 2022), to measure the quality of the modeled structure. The quality check measurements play a vital role in enhancing the 3D structure qualities, thereby improving the accuracy of drug-target interactions and increasing the efficacy of the drug.

The selected models of target protein targets, i.e., ArgS, MurA, and SecY, were then subjected to energy minimization to improve their quality. A powerful visualization tool UCSF Chimera was used to analyze the structures and to minimize energy. Gasteiger charges were assigned to proteins, and structural constraints were removed by 1500 rounds of minimization runs (750 steepest descent followed

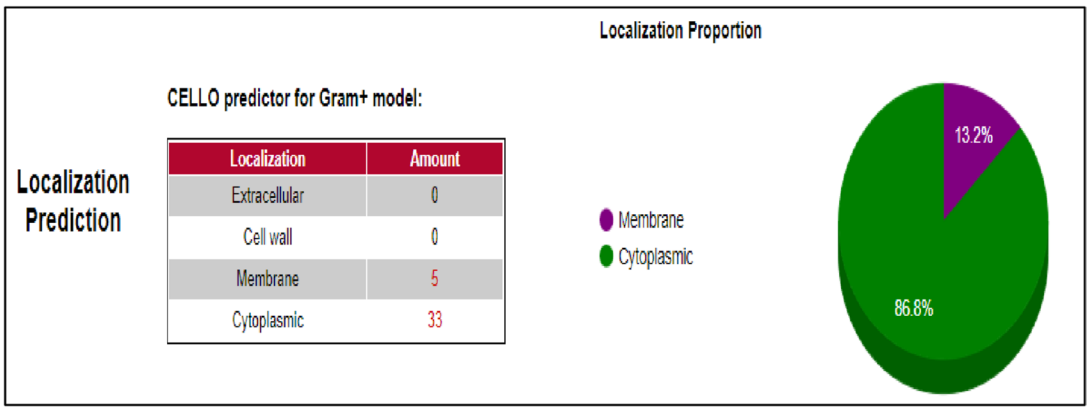


**Fig. 5** Comparative cellular localization prediction using PSORTb and Cello2go web servers. The relative abundance of the predicted targets as membrane and cytoplasmic proteins is shown. In most cases the membrane-bound small molecular structures/proteins are antigenic in nature and are regarded as good adjuvant/vaccine candidates, whereas the cytoplasmic proteins are generally considered as good drug targets for inhibiting vital metabolic cellular processes

by 750 conjugate gradients) with a step size of 0.02 Å, under ff03.rl force field. Protein targets having undergone energy minimization were evaluated through the validation process and then used for docking studies (Weiner and Kollman 1981, Malik, Dalal et al. 2019).

## Druggability, virtual screening, and docking analyses

The information obtained from 3D structures and druggability analyses are important features for prioritizing and authenticating putative pathogen targets. For druggability analyses, the final list of essential non-host and host homologous protein targets was subjected to DoGSiteScorer (https://bio.tools/dogsitescorer) in PDB format. DoGSiteScorer is an automated pocket detection and analysis tool for calculating the druggability of protein cavities (Volkamer et al. 2012). For efficient inhibition, the proper active cavity in the protein three-dimensional structure of molecule binding must be examined. The appropriate active site is categorized based on buriedness, size, shape, and the hydrophobic consideration of the specific site (Pettersen et al. 2004). The active sites of ArgS, MurA, and SecY were determined from different literature sources and was also affirmed manually in the target sequences through sequence alignment. In MOE (v2016) (https://www.chemcomp.com/index.htm) (Molecular Operating Environment), virtual screening (VS), docking, and visualization

**Table 2** Identification of biological pathways using the KEGG (Kyoto Encyclopedia of Genes and Genomes). The table describes the vital pathways of the seven putative targets and are tabulated as gene/protein names, protein functions, and the respective metabolic pathways in which they play key biological role/s

| Gene | Protein name | Protein function | KEGG pathways |
|---|---|---|---|
| RpIA | Ribose-5-phosphate isomerase A | Ribose-5-phosphate isomerase activity | Ribosome |
| Mfd | Transcription-repair-coupling factor | ATP binding–damaged DNA binding<br>DNA binding, DNA translocase activity, helicase activity, RNA polymerase core enzyme binding | Nucleotide excision repair |
| ArgS | Arginine-tRNA ligase | Arginine-tRNA ligase activity, ATP binding | Aminoacyl-tRNA biosynthesis |
| MurA | UDP-N-acetylglucosamine 1-carboxyvinyl transferase 1 | UDP-N-acetylglucosamine 1-carboxyvinyl-transferase activity | Amino sugar and nucleotide sugar metabolism, peptidoglycan biosynthesis, metabolic pathways<br>Biosynthesis of nucleotide sugars |
| SecY | Translocase subunit secY | Protein transmembrane transporter activity, signal sequence binding | Quorum sensing, protein export, bacterial secretion system |
| pPrH | Uridylate kinase | Allosteric enzyme | Pyrimidine metabolism, metabolic pathways, biosynthesis of cofactors |
| RecG | ATP dependent DNA helicase RecG | ATP binding, DNA binding, DNA helicase activity, hydrolase activity | Homologous recombination |

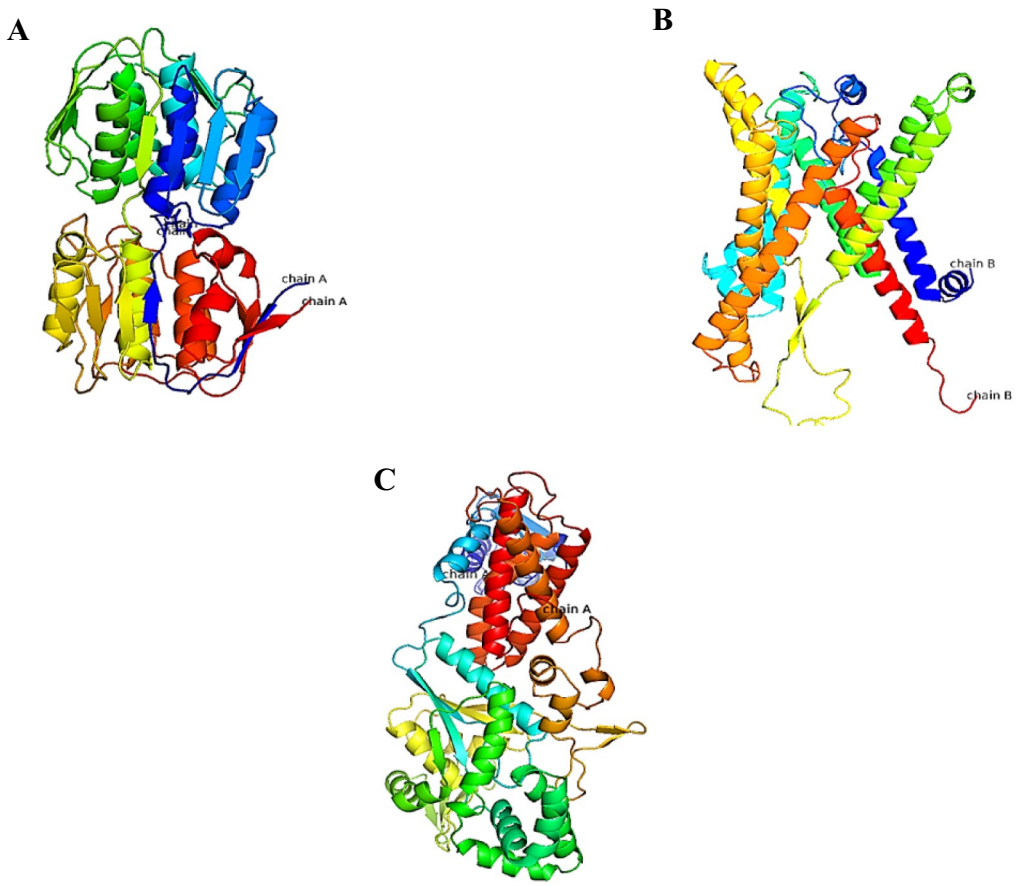


**Fig. 6** 3D structures generated through SWISS-MODEL. **A** WP_058612677, **B** WP_058611897, **C** WP_058610923

were performed following a slightly modified protocol adapted by Muneeba et al. (2023) and Hassan et al. (2022; Christoph, Sabine et al. 2015, Syed, Rida et al. 2022, Muneeba, Syed et al. 2023). The grid for molecular docking was centered around the previously selected active site residues/interface residues of the protein according to the protocol modified from Dalal et al. (Dalal et al. 2021). The 2D depiction of some of the residues interacting through H-bond with the corresponding ligands are shown as well (Fig. 10). The molecular docking strategy was divided into three major steps: active site identification, ligand preparation, and molecular docking. The docking procedure was done out using reduced protein and ligand molecules.

## Molecular dynamics simulation

Molecular docking simulations were used to investigate the behavior of docked proteins. The Assisted Model Building with Energy Refinement program (AMBER) was employed for this aim, and several modules were used for analysis (Weiner and Kollman 1981, Salomon-Ferrer et al. 2013). The details of biomolecule simulation were broken down into five stages, which are shown below.

### System preparation

For the simulation of docked proteins, the AMBER12, module tLEaP, was used, which is an unavoidable part of the system setup that provides an interface for preparing primary coordinates and topology files. The protein was solvated with a three-point transferable intermolecular potential (TIP3P) water box with 8.0 and force fields ff03.rl, GAFF, and ff99SB (Fig. 1). To ensure the accuracy of bonds in docked complexes, angles, and atom kinds, a docked protein system was employed. After preparing the starting files, the simulation procedure began.

### Minimization, heating, equilibration, and production

Minimization is usually done to eliminate undesirable confrontations. At a cutoff value of 8, the steepest descent technique and 1000 steps for conjugate gradient were used. After 10 ps of minimization heating using the Langevin dynamics method for temperature control, 100 ps of equilibration at a constant temperature of 300 K is required before the production run begins. During equilibration, the total energy of the system remains constant, while the kinetic and potential energies vary. The

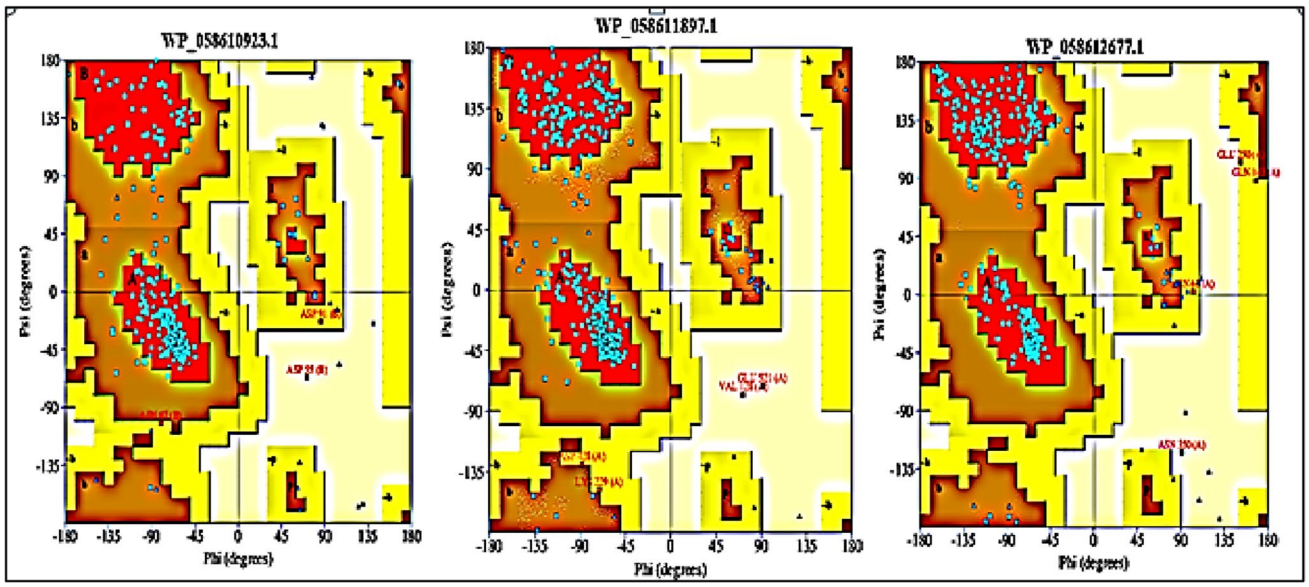


**Fig. 7** Ramachandran plot representing Psi and Phi angles of the selected models and showing the % amino acid residues of the 3D-modeled structures in four different quadrants of the Ramachandran plot

**Table 3** Stereo-chemical properties of the predicted final targets. The table shows the values in percentage of the amino acid residues of the 3D-modeled structures in different quadrants of the Ramachandran plot and the *Z*-scores as a measure of their respective qualities

| Accession number | Most favored regions [A, B, L] | Allowed regions [a, b, l, p] | Generously allowed regions [~a, ~b, ~l, ~p] | Disallowed regions | *Z*-score |
|---|---|---|---|---|---|
| WP_058610923.1 ArgS | 90.18% | 9.28% | 0.58% | 0.38% | −1.74 |
| WP_058612677.1 SecY | 91.78% | 7.58% | 0.68% | 0.68% | −9.47 |
| WP_058611897.1 MurA | 91.38% | 7.68% | 0.84% | 0.36% | −1.01 |

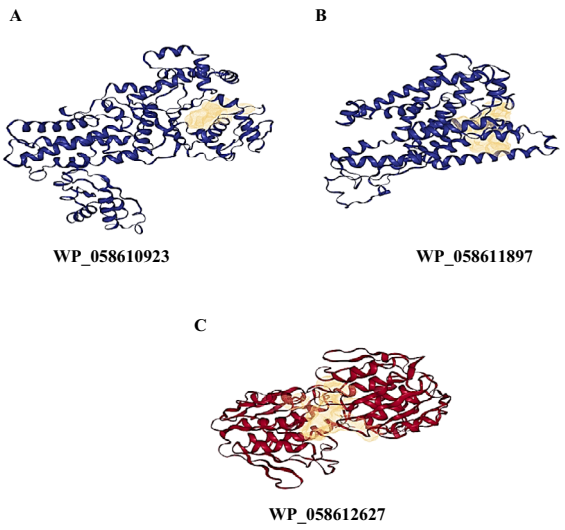


**Fig. 8** Identification of druggable pockets of the top three predicted targets using Protein + of the DoGSiteScorer

manufacturing run for the docked complex was completed in 100 ns (ns), followed by equilibration. Periodic boundary conditions were simulated in the simulation box using a canonical ensemble. To keep the temperature constant, the Berendsen

**Table 4** Pocket detection and protein druggability score for WP_058610923.1. The surface topology of the receptor macromolecule/s in terms of different physicochemical descriptors such as volume, surface area, and drug scores, etc., among others, determines the druggability of a pocket via DoGSiteScorer. A pocket with a drug score close to 1 is considered highly druggable pocket

| Name | Volume Å$^3$ | Surface Å$^2$ | Drug score | Simple score |
|---|---|---|---|---|
| POCKET_2 | 625.92 | 859.24 | **0.84** | 0.42 |
| POCKET_1 | 691.42 | 951.23 | **0.82** | 0.52 |
| POCKET_3 | 175.04 | 315.05 | 0.49 | 0.01 |
| POCKET_5 | 168.71 | 311.44 | 0.39 | 0.01 |
| POCKET_4 | 174.49 | 300.28 | 0.36 | 0.0 |

The "bold values" represent the highest scores of the druggable pockets as predicted by the DoGSiteScorer

**Table 5** Pocket detection and protein druggability score for WP_058611897.1. The surface topology of the receptor macromolecule/s in terms of different physicochemical descriptors such as volume, surface area, and drug scores, among others, determines the druggability of a pocket via DoGSiteScorer. A pocket with a drug score close to 1 is considered a highly druggable pocket

| Name | Volume Å$^3$ | Surface Å$^2$ | Drug score | Simple score |
|---|---|---|---|---|
| POCKET_2 | 840.66 | 1078.58 | **0.83** | 0.58 |
| POCKET_1 | 1676.67 | 2285.72 | **0.82** | 0.68 |
| POCKET_3 | 145.17 | 372.07 | 0.36 | 0.0 |
| POCKET_5 | 135.39 | 360.44 | 0.22 | 0.0 |
| POCKET_4 | 136.21 | 278.31 | 0.21 | 0.0 |

The "bold values" represent the highest scores of the druggable pockets as predicted by the DoGSiteScorer

**Table 6** Pocket detection and protein druggability score for WP_058612677.1. The surface topology of the receptor macromolecule/s in terms of different physicochemical descriptors such as volume, surface area, and drug scores, among others, determines the druggability of a pocket via DoGSiteScorer. A pocket with a drug score close to 1 is considered highly druggable pocket

| Name | Volume Å$^3$ | Surface Å$^2$ | Drug score | Simple score |
|---|---|---|---|---|
| POCKET_1 | 1577.92 | 1701.62 | **0.81** | 0.62 |
| POCKET_2 | 342.4 | 661.62 | **0.63** | 0.16 |
| POCKET_3 | 233.92 | 153.73 | 0.59 | 0.0 |
| POCKET_4 | 159.1 | 192.0 | 0.45 | 0.0 |
| POCKET_5 | 144.96 | 279.96 | 0.26 | 0.0 |

The "bold values" represent the highest scores of the druggable pockets as predicted by the DoGSiteScorer

coupling integration procedure was applied (Berendsen et al. 1984).

### Simulation trajectory analysis

The PTRAJ (Process Trajectory) module of AMBER12 was used to create output files for analysis and to compute four properties, namely, root mean square deviation (RMSD), root mean square fluctuation (RMSF), the radius of gyration (Rg), and their β-factor, and graphical representations were examined in XMgrace (https://plasma-gate.weizmann.ac.il/Grace/) (Vaught 1996).

### Root mean square deviation

The coordinates of alpha carbon (C) are commonly thought to indicate an amino acid's location in three-dimensional space. RMSD is a metric that allows you to compare the relative locations of protein C atoms by computing their averaged distance over a period of time (Kuzmanic and Zagrovic 2010). It is written mathematically as

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_i \mathrm{d}^2{}_i}$$

where $N$ is the number of compared atoms and $d_i$ is the distance between the $i$th pair of atoms.

### Root mean square fluctuation

RMSF is used to determine the backbone atoms of the docked target (N, C, and C). It is the root mean square of the averaged distance between an atom and its average geometric location in a particular set of dynamics, and it may be read as the set of atom positions recorded over a specific time scale. The RMSF is calculated using the following equation:

$$\mathrm{RMSF} = \sqrt{\sum Tkk\left(\frac{xi(tk) - x}{T}\right)}$$

where $T$ represents the time interval, $x_i$ represents the position of an atom at a particular time, and $x$ represents the averaged position of the atom.

### β-Factor

The term-factor, which is closely related to the RMSF, assesses the spatial displacement of atoms around their mean locations as a result of local vibrational and thermal motions (Kuzmanic and Zagrovic 2010). They may be equivalent in terms of RMSF since they measure fluctuations:

$$\beta - \mathrm{Factor} = \mathrm{RMSF}^2\left(\frac{8\pi 2}{3}\right)$$

### Radius of gyration

The radius of gyration is used to assess the overall packing quality and density of a structure. It is a physical characteristic that may be estimated experimentally, most commonly via small-angle X-ray scattering (SAXA). The following equation was used to quantify the compactness of a macromolecular system:

$$R_{\mathrm{g}} = \sum_i^N = {}_1 m_i \left(r_i - r_{\mathrm{cm}}\right)^2 / \sum_i^N = 1^{mi}$$

where $N$ is the total number of atoms, $mi$ denotes the mass of atom I, $r_i$ denotes the position vector of an atom I, and $r_{\mathrm{cm}}$ denotes the molecule's center of mass. Figure 2 represents an overview of all steps that have been followed in this work, whereas Table 1 shows the number of proteins/genes screened in each step of subtractive genomic/proteomic approach.

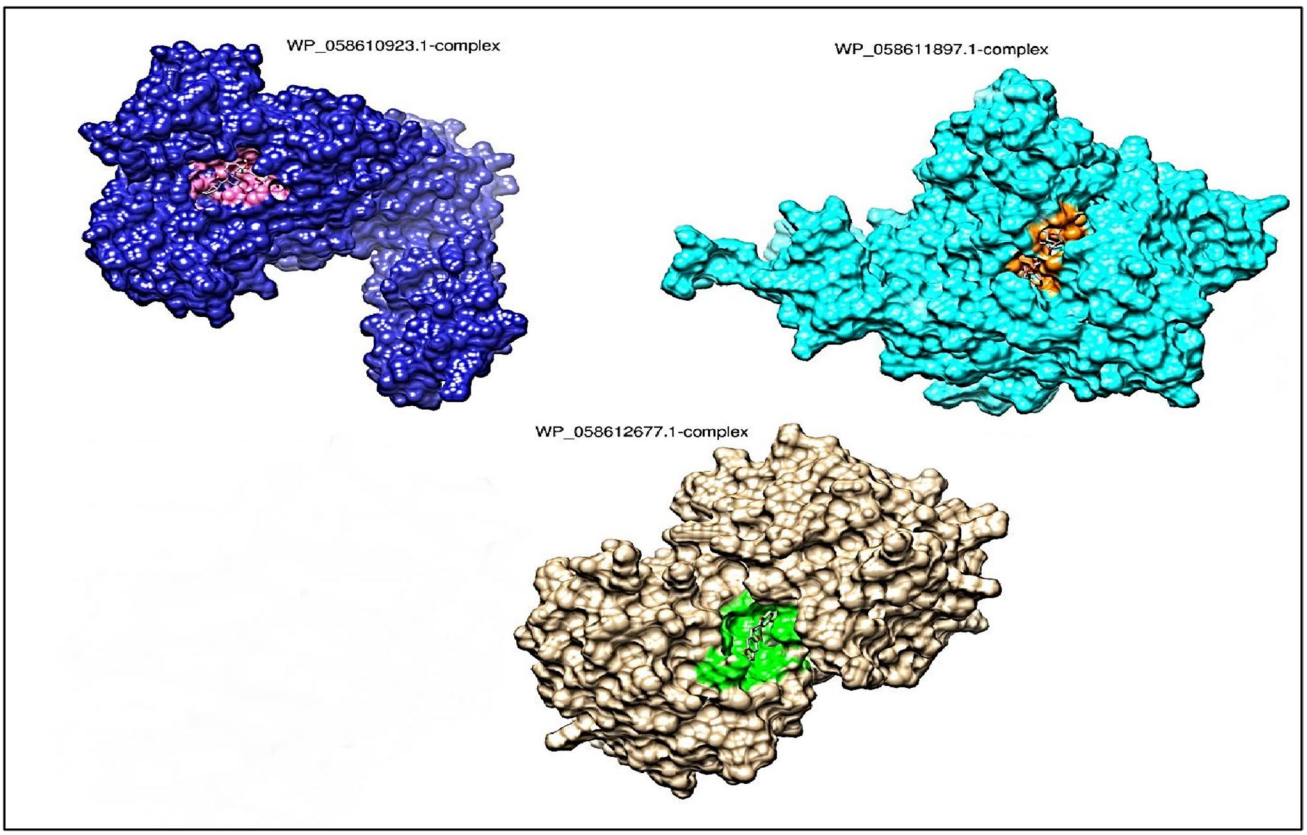


**Fig. 9** Depicting 3D graphics of all the three docked complexes with the best inhibitor shown in the binding cavity. The ligand-receptor complex was generated via the UCSF CHIMERA tool

## Results and discussion

### Data retrieval

In the present study, eleven (11) out of one hundred and twelve (112) strains of *Staphylococcus sciuri* were included that have been reported to be completely sequenced until 2021. All the sequence data is available at the National Center for Biotechnology Information (NCBI) https://www.ncbi.nlm.nih.gov/genome/ for downloading and downstream analyses. This study emphasizes on exploration of the genomes of the selected strains. A reference strain of *S. sciuri* (FDAARGOS_285) was randomly selected for further comparative analysis.

### Prediction of core and non-host homologous genes/proteins

For the construction of the core genome, the Pathosystems Resource Integration Center (PATRIC) was used. Among the 11 strains of *S. sciuri*, one strain FDAARGOS_285 was taken as a reference and the rest of strains were compared to it (Fig. 3). The core genome file contained 1784 genes that were then submitted to NCBI-BLASTp (*E*-value = 0.0001, bit score 100, and identity 25%) against the human genome for filtering pathogen-specific non-host homologs. Among these gene sequences, considering the human genome as the host, we found 170 non-host homologous proteins. This step is important to avoid cross-reactivity and binding of the drugs to undesired host protein sites.

### Analyses of essential genes and protein–protein interaction

The non-host homologous 170 core proteins were then subjected to BLASTp against essential proteins present in DEG (http://tubic.tju.edu.cn/deg/). The file was then subjected to NCBI BLASTp by using the Perl script with the threshold *E*-value = $10e^{-4}$, bit score = 100, and sequence identity = ≥ 30% against prokaryotes, eukaryotes, and archaea by which druggable targets were reduced to 35 potential targets. The STRING (https://string-db.org/)

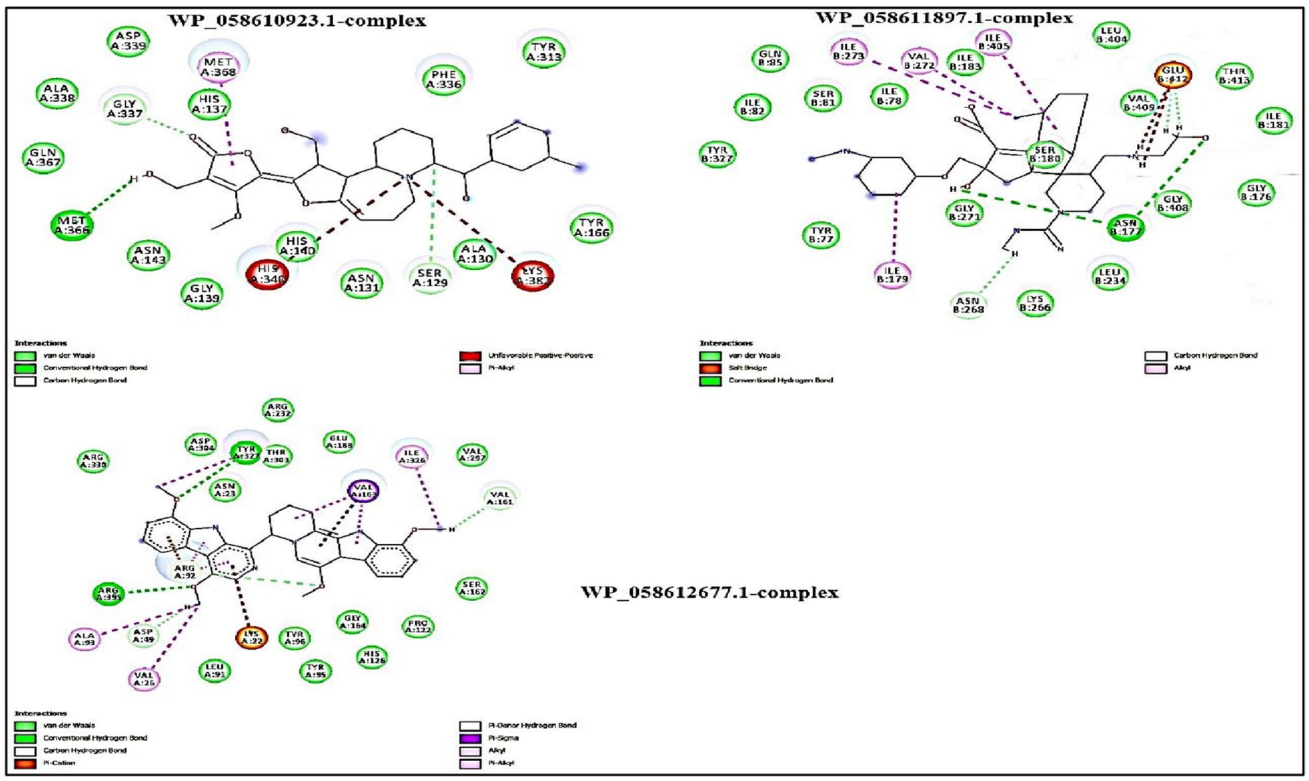


**Fig. 10** 2D depiction of the ligand-receptor complexes of the final protein targets representing H-bond with the corresponding best inhibitor. The dotted lines show the H-bond interactions b/w the inhibitor and the amino acid residues of the target protein. The different colors correspond to the chemical nature of interactions and amino acids

database is used to determine the inter relation between proteins, which is essential for the proper functioning and gives a detailed knowledge about protein involved in single or multiple pathways. Out of 35, 10 proteins showed multiple interactions. Thus, selecting them for drugs would account for more specificity and accurate results (Fig. 4).

**Table 7** Docking results of inhibitors with corresponding binding affinities via H-bond within the ArgS binding site. The S-score (docking score) of the MOE software manifest the thermodynamic stability of the ligand-receptor complex system

| TCM code | Docking score | Interactions |
|---|---|---|
| C25H28O6 | **−7.9** | **Met366, Gly337, Ser129, Met368** |
| C24H32O6 | −7.1 | Gly337, Asn143 |
| C14H28O4 | −7.8 | Phe336, Asn143, |
| C23H28O8 | −7.9 | Tyr313, Tyr166 |
| C15H20O3 | −7.6 | Ala130, Ser129 |

The "bold values" represent the highest scores of the druggable pockets as predicted by the DoGSiteScorer

## Comparative subcellular localization and identification of biological pathways

The 35 druggable targets were then further proceeded for subcellular localization prediction. We have predicted the comparative subcellular localization of all proteins by using PSORTb (https://www.psort.org/psortb/) and Cello2go (http://cello.life.nctu.edu.tw/cello2go/). Out of the total 35 proteins, 33 were cytoplasmic proteins and

**Table 8** Docking results of inhibitors with corresponding binding affinities via H-bond within the SecY binding site. The S-score (docking score) of the MOE software manifest the thermodynamic stability of the ligand-receptor complex system

| TCM code | Docking score | Interactions |
|---|---|---|
| C30H51N5O5 | **−7.7** | **Asn177, Glu412** |
| C29H36O15 | −7.5 | Ile181, Glu412 |
| C29H32O7 | −7.6 | Ser81, Val409 |
| C20H24O9 | −7.2 | Ile181, Asn268 |
| C25H34O7 | −6.4 | Glu412, Val272 |

The "bold values" represent the highest scores of the druggable pockets as predicted by the DoGSiteScorer

**Table 9** Docking results of inhibitors with corresponding binding affinities via H-bond within the MurA binding site. The S-score (docking score) of the MOE software manifest the thermodynamic stability of the ligand-receptor complex system

| TCM code | Docking score | Interactions |
|---|---|---|
| C35H45NO15 | **−7.9** | **Tyr327, Arg395, Asp49, Val161** |
| C18H34O11 | −7.6 | Asp49, Ser326 |
| C25H32O6 | −7.3 | Ala93, Ser326 |
| C11H10O4 | −7.3 | Ala93, Tyr327 |
| C21H30O5 | −7.2 | Val26, Tyr327 |

The "bold values" represent the highest scores of the druggable pockets as predicted by the DoGSiteScorer

5 were membrane proteins. The results are mentioned below in Fig. 5 and Table 4.

These proteins were then further subjected to the KEGG database for pathway analysis. It was discovered that seven proteins were involved in multiple pathways. The determination of molecular pathways is essential and a very important step because it tells us the estimate and extent to which a protein is necessary for a molecular pathway (Supplementary Fig. 1). Table 2 contains the functionally annotated 7 important non-host homologous proteins.

## Drug target selection and 3D structure modeling

Drug targets have been selected based on their mechanism of function and virulence check, molecular weight, pathway analysis, and druggability. Results inferred seven best targets responsible for resistance against antibiotics. Herein, further investigation showed that three targets, namely, ArgS (WP_058610923), MurA (WP_058611897), and SecY (WP_058612677), have more pathogenic responses according to literature support. The 3D structure of protein availability is the starting point for CADD analysis. Structures of protein (ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1) were generated from online servers like SWISS-MODEL. Models

**Table 10** Physicochemical molecular properties of top hits/ADME profile via SWISS ADME. The evaluation/features of top hist are important in computationally mining for potent inhibitors

| Formula | MW | RB | HBA | HBD | TPSA | BS | PA | SA | SC | LK | LV | VV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{25}H_{28}O_6$ | 424.49 | 10 | 6 | 0 | 71.06 | 0.55 | 0 | 3.43 | Poorly | −5.50 | 0 | 0 |
| $C_{30}H_{51}N_5O_5$ | 561.76 | 26 | 7 | 4 | 158.3 | 0.55 | 0 | 5.64 | Soluble | −5.05 | 1 | 2 |
| $C_{35}H_{45}NO_{15}$ | 719.73 | 19 | 15 | 1 | 189.68 | 0.17 | 0 | 6.79 | Moderately soluble | −9.72 | 2 | 2 |

*BS* bioavailability score, *PA* PAINS alerts, *SA* synthetic accessibility, *SC* Silicos-IT Class, *LK* log Kp (cm/s), *LV* Lipinski violations, *VV* Veber violations

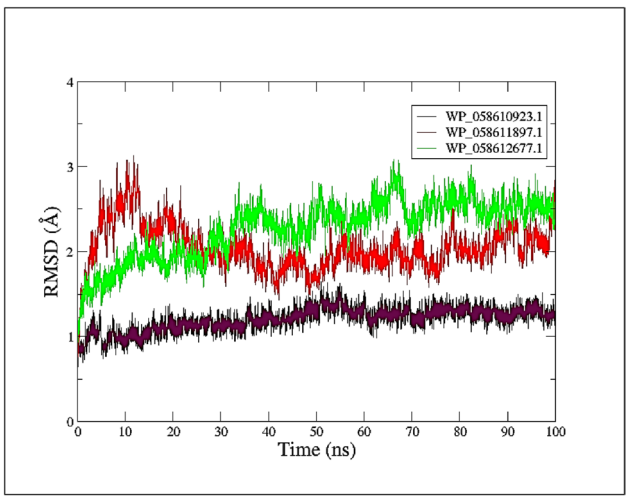


**Fig. 11** RMSD plot of simulated ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1) protein complex for the 100-ns simulation run

generated from SWISS-MODEL were selected for further analysis based on physicochemical properties and quality assessment measures. Structures generated through Swiss-Model are given below (Fig. 6).

Besides significant coverage, model 1 showed strong stereochemistry with no residue in the disallowed region and the lowest *Z*-score (Supplementary Fig. 2). Energy minimization was done to relax the structure and remove the steric clashes of the side chain. Ramachandran plots of the selective models showed that maximum residues are present in the most favored regions. Stereo-chemical properties of comparative homology modeled structure are given below (Fig. 7; Table 3).

## Validation of 3D models and druggability analysis

Only three cytoplasmic proteins were chosen as potential therapeutic targets out of a total of seven proteins based

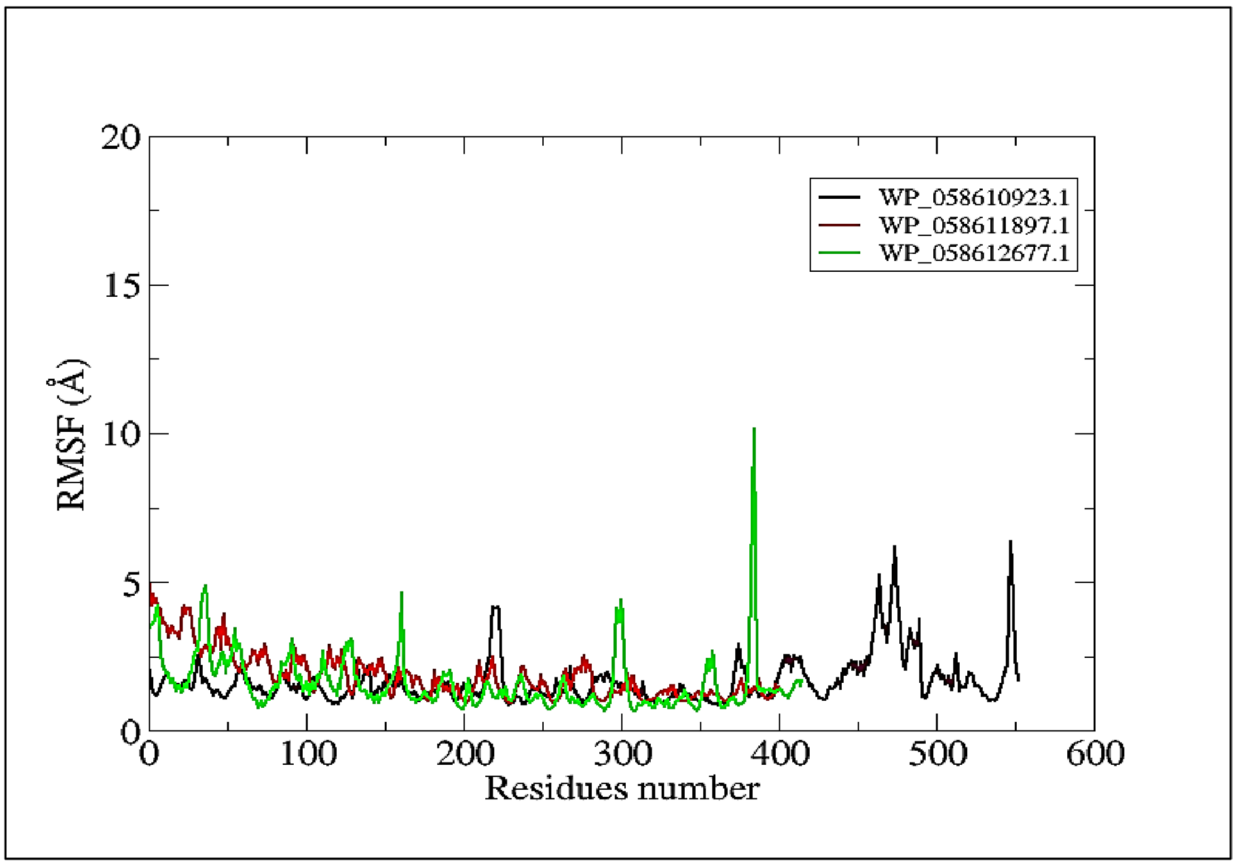


**Fig. 12** RMSF of simulated ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1) protein over the 100-ns simulation run

on their percentage identity of more than 25% and the pathways in which they are involved. The final list of essential non-host good-quality protein targets was subjected to DoGSiteScorer in PDB format. After that, Target Pathogen Database has been used in order to analyze druggability and other biochemical functions. Druggable pockets of final three targets are given below in Fig. 8 and Tables 4, 5, and 6.

## Molecular docking, inhibitor selection, and ADMET profiling

Active site information for the docking procedure is criterion-based. The following steps were involved for this procedure.

In the current study, the Traditional Chinese Medicine (TCM) library was used containing 36,043 compounds used as an inhibitor for docking into ArgS, MurA, and SecY active sites. The top hits of TCM were docked, and top five compounds were analyzed for each receptor.

A total of 36,043 ligands were docked into the active site of the target using the Molecular Operating Environment (MOE) software. For this, selected binding pocket orientation of the active compound was also identified.

Selected ligand molecules were docked into the active site of the target using MOE (Fig. 9). Corresponding hydrogen bonds and binding affinity were also calculated using MOE (Fig. 10). The highest score achieved for compound 1, compound 2, and compound 3 with binding affinities – 7.9, – 7.7, and – 7.9 kcal/mol against the target proteins, respectively. Docking scores and respective binding affinity for the top 5 compounds arranged in descending order are provided below against each protein target (Tables 7, 8, and 9). Detailed visualization analysis was carried out through MOE and the preferred orientation of the ligand binding.

In silico prediction of drug-likeness and ADMET profiling of drug candidates helps reduce the expense of synthesis, preclinical, and clinical research (Kar and Leszczynski 2020). Furthermore, molecular properties of

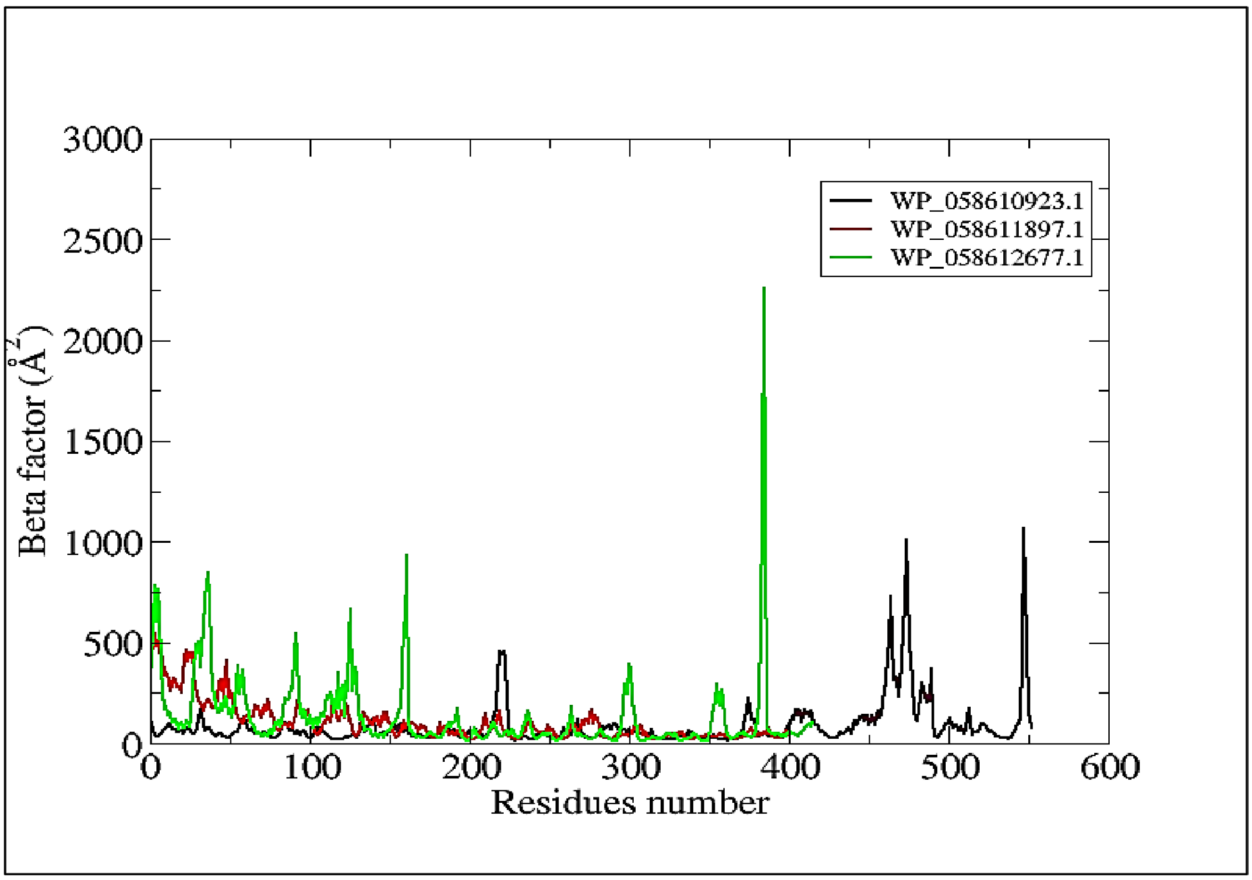


**Fig. 13** β-Factor graphs of simulated ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1) proteins over the 100-ns simulation run

top hits compounds were calculated using Swiss ADME (Table 10).

## Molecular dynamics simulation

The most fundamental element associated with the function of proteins is their conformational dynamics. Functional information of protein molecule in encrypted in its structure. To unravel its functional variability, a comprehensive understanding of the structure is needed. In the current study, MD simulation was performed to explore the conformational aspect of protein–ligand interactions and to evaluate the stability of the homology model and enzyme-inhibitor complex. Data reduction analyses like root mean square deviation (RMSD) and root mean square fluctuation (RMSF), the radius of gyration (Rg), and β-factor values were used to determine the conformational changes and stability index of secondary structure elements of the simulated complexes.

### Root mean square deviation

RMSD explains the backbone analysis and Cα atoms dynamics over the period of docked protein over the 100-ns time period, and it was observed at the 15-ns fluctuation, but the remaining graph of simulation stability was observed. The average RMSD value for docked protein was 1.17 Å. Figure 6 shows a maximum peak of 1.67 Å. Overall, the pattern of the RMSD graph does support any major domain shifts within the structural framework of the protein–ligand complex. The placement of ligand was well complemented within the binding site during simulation and does not destabilize the protein as shown in Fig. 11.

### Root mean square fluctuations

Structure flexibility and fluctuation of Cα residues over time are observed by the RMSF. The average RMSF of docked ArgS, MurA, and SecY proteins calculated from

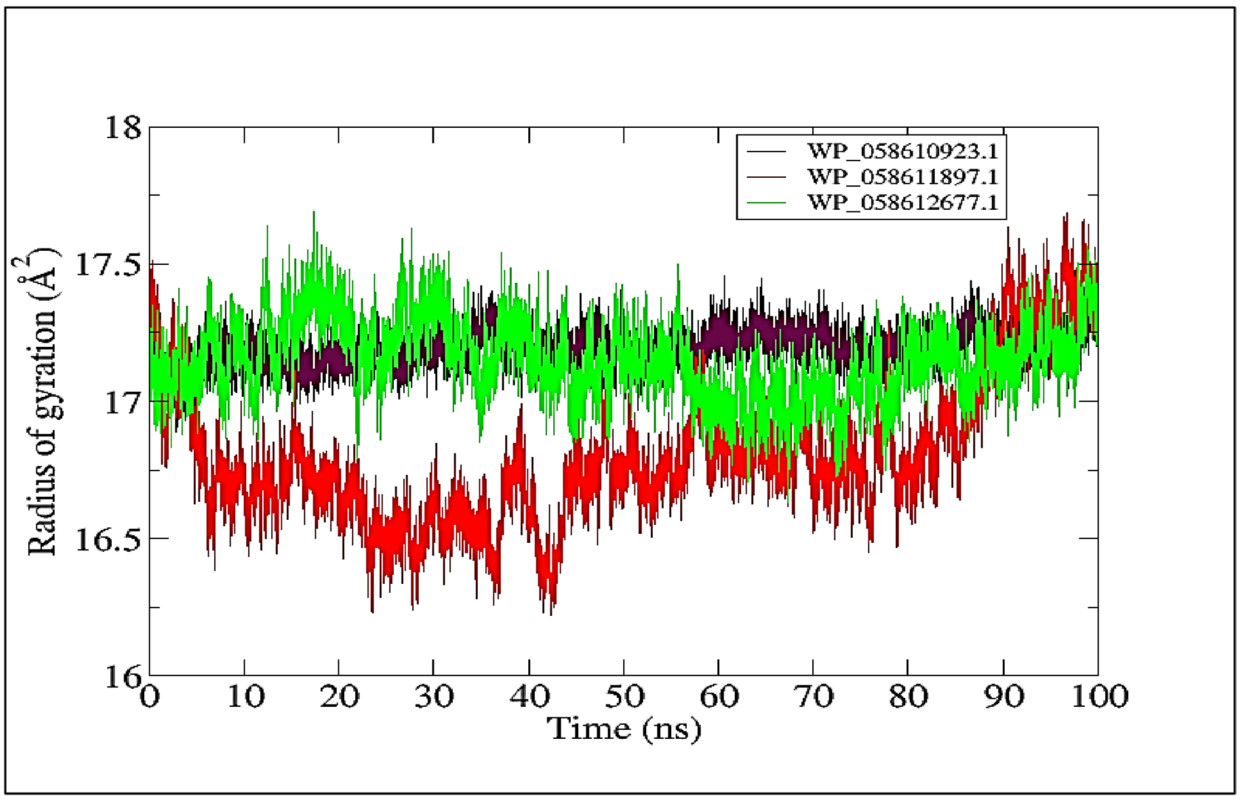


**Fig. 14** The radius of gyration of simulated proteins ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1) over the 100-ns simulation time period

100 ns was 1.3 Å with 2.4 and 2.7 Å while a maximum peak has been noticed at, while major fluctuations at 76,103, 203 to 263 and 336 residues then at the end of the graph for 518 and 560 residues were observed. That was mostly the loop region of the protein. Till the end of 100 ns, many fluctuations appeared in the graph of Fig. 12. One of the stability proofs of the protein in the simulation run was that the active site residue His125 had an RMSF value of less than 1.0 Å.

### β-Factor analysis

β-Factor explains the thermal stability and flexibility of the protein overtime. The quantity of β-factor is measured in RMSF. Therefore, its value on the level of localized atomic fluctuation collectively contributes to the global vibrational movement of the protein and its thermal stability. The average β-factor values for ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1) were calculated which are 86.7, 105.8, and 130.7 Å, respectively, demonstrating the higher instability from residue numbers 203 to 263, 502, and 599 of protein (Fig. 13).

### Radius of gyration

The radius of gyration was calculated to evaluate the structural compactness as a time function for the 100-ns simulation of protein–ligand complexes ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1). The average values of 17.3, 16.7, and 17.4 Å, respectively, for docked protein (Fig. 14) denoted the stability of the protein structure.

Here, by employing a subtractive genomics approach, we have reported some essential non-host homologous protein–based putative targets in principally an animal-associated bacterial species whose clinical relevance to humans is increasing day by day. These targets include **argS** (arginine-tRNA ligase) with arginine-tRNA ligase and ATP binding activities, and have an important role in the aminoacyl-tRNA biosynthesis pathway, a key player in protein synthesis; **murA** (UDP-N-acetylglucosamine 1-carboxyvinyl transferase 1), which is involved in amino sugar/nucleotide sugar metabolism and peptidoglycan biosynthesis; and finally, **secY** (translocase subunit secY) with protein transmembrane transporter and signal sequence binding activities. This target is involved

in multiple vital bacterial processes such as quorum sensing, protein export, and bacterial secretion system (Holden et al. 2004, Gill et al. 2005). In addition to the protocol followed here for therapeutics targets mining in *S. sciuri* and novel inhibitors, there are other well-established in silico computational approaches in the literature that have been used to identify other novel and potential antibacterial agents targeting other important protein-based targets in the Gram-positive *Staphylococcus* genus, with a main focus on *Staphylococcus aureus*. For example, The target FmtA is a core member of the *Staphylococcus aureus* cell wall stimulon, a factor that affects methicillin resistance in *S. aureus* strains, interacting with teichoic acids and shown to be localized to the cell division septum. FmtA, as part of the catalytic activity, hydrolyzes the ester bond between the backbone of teichoic acids and d-Ala, which are polyribitol-phosphate or polyglycerol-phosphate polymers found in the *S. aureus* cell envelope (Rahman et al. 2016). Recently, Vikram Dalal and his group have performed numerous biophysical, structural, and in silico studies to show the binding interaction and complex stabilities of newly identified inhibitors towards FmtA from *S. aureus*. However, the reported screened molecules need to be tested, modified, and experimentally validated to develop the effective antimicrobial compounds against *S. aureus* (Dalal et al. 2019, Dalal et al. 2021, Dalal et al. 2022, Singh, Dhankhar et al. 2022). Some other related in silico studies have reported potent inhibitors against GraR, a member of the two-component regulatory system GraR/GraS and is involved in resistance against cationic antimicrobial peptides (CAMPs) (Meehl et al. 2007, Dhankhar et al. 2020). Potential lead molecules were identified by performing a structure-based pharmacophore modeling against the lipophilic membrane (LLM) protein that regulates bacterial lysis rate and methicillin resistance level in *S. aureus* (Kumari and Dalal 2022). Similarly, two other individual studies by the same group have reported further novel inhibitors against the ribosome biogenesis GTP-binding protein (YsxC), a GTPase that interacts with 50S/30S subunits of the ribosome, and β′ subunit of RNA polymerase, and thereby play an important role in bacterial protein synthesis of *S. aureus* (Kumari et al. 2022, Kumari et al. 2023). FemC is another methicillin-resistance factor that regulates the synthesis of peptidoglycan in the Gram-positive *Staphylococcus aureus*. A set of natural product-like compounds from Selleckchem and Enamine databases were screened for inhibitor mining by taking into consideration the active site of the validated FemC model (Dalal and Kumari 2022). The methodology employed here and other in silico cloning and vaccine design studies, hereby, report potent protein-based targets and inhibitors that are required to be validated and may further be utilized to develop novel scaffolds for antimicrobials against *S. aureus* targets (Khan et al. 2021b, Khan et al. 2022a, Khan et al. 2022c).

## Conclusion

Research methodologies were adopted to identify the potential therapeutic candidates in the Gram-negative and MDR pathogen *S. sciuri*. Genome subtraction aids the identification of pathogen-specific potent drug targets involved in crucial metabolic pathways. Virtual screening and molecular docking were followed to mine the inhibitors from the TCM library. Molecular docking resulted in 1326 compounds as the top inhibitors against ArgS (WP_058610923.), MurA (WP_058611897.1), and SecY (WP_058612677.1). Furthermore, MD simulation confirmed that in the physiochemical environment, the drug-receptor complex attains stability due to structural rearrangements concerning time. Besides minor fluctuations, inside chain and loop movement stability of the inhibitor were observed. Structural stability observed in the docked complex after simulation studies confirms the prospective roles of the selected ligand as a lead compound. The ADMET profiling of the final three TCM compounds further paved a way for its practical feasibility whereas the predicted protein-based three targets could further aid, bridging the gap between the existing and novel pathogen targets. The literature survey of the predicted target proteins manifest that they play a pivotal role in bacterial survival, pathogenesis, and infection establishment. Synthesis of the cell-wall components/peptidoglycan biosynthesis is of utmost importance to retain structural integrity along with antibiotics resistance. On the other hand, protein biosynthesis, nucleotide metabolism, Quorum sensing, and the different types of bacterial secretion systems are always very attractive targets in any drug development procedures. These findings/outcomes of the current study could enhance pharmacological design to develop more potent, efficient, and specific drugs against MDR *S. sciuri*.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10142-023-01179-w.

**Acknowledgements** We would like to acknowledge the mutual research/academic collaboration of authors from different institutions including JRC Genome Research, ICCBS, University of Karachi, Karachi-Pakistan, Department of Chemistry, Islamia College Peshawar–Pakistan, Department of Chemistry and CDTS_Oswaldo Fiocruz, RJ—CAPES, Brazil (Finance Code 001).

**Author contribution** Conceptualization, Syed S Hassan and Aafareen Khan. Data curation, Syed S Hassan, Saman Sohail, and Aafareen Khan. Formal analysis, investigation, Zarrin Basharat. Methodology, Syed S Hassan, Saman Sohail, Muhammad Irfan. Project administration, Syed S Hassan, Mohammad Y. Alshahrani, and Carlos M. Morel. Resources, Carlos M. Morel. Software, Aafareen Khan and Saman Sohail. Supervision, Syed S Hassan. Validation, Asad Karim; Sareen

Fatima, Syed Kashif Raza. Visualization, Aafareen Khan, Saman Sohail, Zarrin Basharat, Mohammad Y. Alshahrani. Writing—original draft, Aafareen Khan, Saman Sohail, Ayesha Wisal. writing—review and editing, Seerat Yaseen, Mahrukh Nasir, Sufyan Ahmed, Muhammad Aurongzeb, Muhammad Irfan, Mohammad Y. Alshahrani and Carlos M. Morel. The final manuscript was reviewed and approved by all authors.

**Funding** The authors also express their appreciation to the Deanship of Scientific Research at King Khalid University, Saudi Arabia, (Research Group Program under funding grant number: RGP. 2/370/44).

**Data availability** Data is derived/used from the public domain (https://www.ncbi.nlm.nih.gov) for this work.

## Declarations

**Ethics approval** Not applicable.

**Competing interests** The authors declare no competing interests.

## References

Adesoji TO et al (2020) Antibiotic-resistant staphylococci from the wastewater treatment plant and grey-water samples in Obafemi Awolowo University, Ile-Ife, Nigeria. J Water Health 18(6):890–898

Afzal M, Hassan SS et al (2023) Genomic landscape of the emerging XDR Salmonella Typhi for mining druggable targets clpP, hisH, folP and gpmI and screening of novel TCM inhibitors, molecular docking and simulation analyses. BMC Microbiol 23:25

Al-Hayawi A (2022) The multiplex PCR assay detection of Staphylococcus sciuri antibiotic resistance, mecA gene, and the inhibitory effect of root exudate of Nigella sativa (black seeds) treated with magnetized water. J Med Life 15(2):228–233

Altschul SF et al (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410

Aregbesola OA et al (2021) Whole-genome sequencing, genome mining, metabolic reconstruction and evolution of pentachlorophenol and other xenobiotic degradation pathways in Bacillus tropicus strain AOA-CPS1. Funct Integr Genomics 21:171–193

Aurongzeb M et al (2022) Insights into genome evolution, pan-genome, and phylogenetic implication through mitochondrial genome sequence of Naegleria fowleri species. Sci Rep 12(1):1–13

Basharat Z et al (2021) Pan-genomics, drug candidate mining and ADMET profiling of natural product inhibitors screened against Yersinia pseudotuberculosis. Genomics 113(1):238–244

Berendsen HJ et al (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81(8):3684–3690

Boonchuay K et al (2023) Association of multilocus sequencing types and antimicrobial resistance profiles of methicillin-resistant Mammaliicoccus sciuri in animals in Southern Thailand. Vet World 16(2):291–295

Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2(9):1511–1519

Dakic I et al (2005) Isolation and molecular characterization of Staphylococcus sciuri in the hospital environment. J Clin Microbiol 43(6):2782–2785

Dalal V, Kumari R (2022) Screening and identification of natural product-like compounds as potential antibacterial agents targeting FemC of Staphylococcus aureus: an in-silico approach. ChemistrySelect 7(42):e202201728

Dalal V et al (2019) Repurposing an ancient protein core structure: structural studies on FmtA, a novel esterase of Staphylococcus aureus. J Mol Biol 431(17):3107–3123

Dalal V et al (2021) Structure-based identification of potential drugs against FmtA of Staphylococcus aureus: virtual screening, molecular dynamics, MM-GBSA, and QM/MM. Protein J 40(2):148–165

Dalal V et al (2022) Quantum mechanics/molecular mechanics studies on the catalytic mechanism of a novel esterase (FmtA) of Staphylococcus aureus. J Chem Inf Model 62(10):2409–2420

de Carvalho TP et al (2022) Mammaliicoccus (Staphylococcus) sciuri-induced suppurative meningoencephalitis and bacteremia in an infant western lowland gorilla (Gorilla gorilla gorilla). J Med Primatol 51(6):396–399

Deng J et al (2022) Comparative proteomic analyses of Tartary buckwheat (Fagopyrum tataricum) seeds at three stages of development. Funct Integr Genomics 22:1449–1458

Dhankhar P et al (2020) In-silico approach to identify novel potent inhibitors against GraR of S. aureus. Front Biosci (Landmark Ed) 25(7):1337–1360

Dindhoria K et al (2022) Computational approaches and challenges for identification and annotation of non-coding RNAs using RNA-Seq. Funct Integr Genomics 22:1105–1112

Egyir B et al (2022) Antimicrobial resistance and genomic analysis of staphylococci isolated from livestock and farm attendants in Northern Ghana. BMC Microbiol 22(1):180

Gill SR et al (2005) Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant Staphylococcus aureus strain and a biofilm-producing methicillin-resistant Staphylococcus epidermidis strain. J Bacteriol 187(7):2426–2438

Gómez-Sanz E et al (2021) The resistome and mobilome of multidrug-resistant Staphylococcus sciuri C2865 unveil a transferable trimethoprim resistance gene, designated dfrE, spread unnoticed. mSystems 6(4):e0051121

Hassan SS et al (2014) Proteome scale comparative modeling for conserved drug and vaccine targets identification in Corynebacterium pseudotuberculosis. BMC Genomics 15(7):1–19

Hassan SS et al (2022) Subtractive sequence analysis aided druggable targets mining in Burkholderia cepacia complex and finding inhibitors through bioinformatics approach. Mol Divers 26:1–25

Hauschild T, Schwarz S (2003) Differentiation of Staphylococcus sciuri strains isolated from free-living rodents and insectivores. J Vet Med Ser B 50(5):241–246

Hauschild T, Wójcik A (2007) Species distribution and properties of staphylococci from canine dermatitis. Res Vet Sci 82(1):1–6

He L et al (2021) Comparative transcriptome analysis reveals that deletion of CheY influences gene expressions of ABC transports and metabolism in Haemophilus parasuis. Funct Integr Genomics 21:695–707

Hedin G, Widerström M (1998) Endocarditis due to Staphylococcus sciuri. Eur J Clin Microbiol Infect Dis 17(9):673–675

Holden MT et al (2004) Complete genomes of two clinical Staphylococcus aureus strains: evidence for the rapid evolution of virulence and drug resistance. Proc Natl Acad Sci U S A 101(26):9786–9791

Horii T et al (2001) Intravenous catheter-related septic shock caused by Staphylococcus sciuri and Escherichia vulneris. Scand J Infect Dis 33(12):930–932

Hughes TR (2002) Yeast and drug discovery. Funct Integr Genomics 2(4–5):199–211

Ibrahim M et al (2017) Comparative pan genome analysis of oral Prevotella species implicated in periodontitis. Funct Integr Genomics 17(5):513–536

Irfan M, Tariq M, Basharat Z, Abid Khan RM, Jahanzaeb M, Shakeel M, Nisa ZU, Shahzad M, Jahanzaib M, Moin ST, Hassan SS, Khan IA (2023) Genomic analysis of Chryseobacterium indologenes and conformational dynamics of the selected DD-peptidase.

Res Microbiol 174(1–2):103990. https://doi.org/10.1016/j.resmic.2022.103990
Kanehisa M et al (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45(D1):D353–D361
Kar S, Leszczynski J (2020) Open access in silico tools to predict the ADMET profiling of drug candidates. Expert Opin Drug Discov 15(12):1473–1487
Katayama Y et al (2001) Genetic organization of the chromosome region surrounding mecA in clinical staphylococcal strains: role of IS 431-mediated mecI deletion in expression of resistance in mecA-carrying, low-level methicillin-resistant Staphylococcus haemolyticus. Antimicrob Agents Chemother 45(7):1955–1963
Khan T et al (2021a) Evaluation of the whole proteome of Achromobacter xylosoxidans to identify vaccine targets for mRNA and peptides-based vaccine designing against the emerging respiratory and lung cancer-causing bacteria. Front Med (lausanne) 8:825876
Khan T et al (2021b) A computational perspective on the dynamic behaviour of recurrent drug resistance mutations in the pncA gene from Mycobacterium tuberculosis. RSC Adv 11(4):2476–2486
Khan T et al (2022b) Towards specie-specific ensemble vaccine candidates against mammarenaviruses using optimized structural vaccinology pipeline and molecular modelling approaches. Microb Pathog 172:105793
Khan T et al (2022c) Subtractive proteomics assisted therapeutic targets mining and designing ensemble vaccine against Candida auris for immune response induction. Comput Biol Med 145:105462
Khan T, Khan A, Ansari JK, Najmi MH, Wei DQ, Muhammad K, Waheed Y (2022a) Potential immunogenic activity of computationally designed mRNA- and peptide-based prophylactic vaccines against MERS, SARS-CoV, and SARS-CoV-2: a reverse vaccinology approach. Molecules 27(7):2375. https://doi.org/10.3390/molecules27072375
Khazandi M et al (2018) Genomic characterization of coagulase-negative staphylococci including methicillin-resistant Staphylococcus sciuri causing bovine mastitis. Vet Microbiol 219:17–22
Kim SJ et al (2019) Antimicrobial resistance and genetic characterization of coagulase-negative staphylococci from bovine mastitis milk samples in Korea. J Dairy Sci 102(12):11439–11448
Kloos WE (1980) Natural populations of the genus Staphylococcus. Annu Rev Microbiol 34:559–592
Kloos W et al (1976) Preliminary studies on the characterization and distribution of Staphylococcus and Micrococcus species on animal skin. Appl Environ Microbiol 31(1):53–59
Kloos WE et al (1997) Ribotype delineation and description of Staphylococcus sciuri subspecies and their potential as reservoirs of methicillin resistance and staphylolytic enzyme genes. Int J Syst Evol Microbiol 47(2):313–323
Kolawole D, Shittu A (1997) Unusual recovery of animal staphylococci from septic wounds of hospital patients in Ile-Ife, Nigeria. Lett Appl Microbiol 24(2):87–90
Kumari R, Dalal V (2022) Identification of potential inhibitors for LLM of Staphylococcus aureus: structure-based pharmacophore modeling, molecular dynamics, and binding free energy studies. J Biomol Struct Dyn 40(20):9833–9847
Kumari R et al (2022) Structural-based virtual screening and identification of novel potent antimicrobial compounds against YsxC of Staphylococcus aureus. J Mol Struct 1255:132476
Kumari R et al (2023) Computational investigation of potent inhibitors against YsxC: structure-based pharmacophore modeling, molecular docking, molecular dynamics, and binding free energy. J Biomol Struct Dyn 41(3):930–941
Kuzmanic A, Zagrovic B (2010) Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. Biophys J 98(5):861–871
Laamarti M et al (2022) Genomic analysis of two Bacillus safensis isolated from Merzouga desert reveals desert adaptive and potential plant growth-promoting traits. Funct Integr Genomics 22:1173–1187
Laskowski RA et al (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26(2):283–291
Liu S, Wang SX, Liu W, Wang C, Zhang FZ, Ye YN, Wu CS, Zheng WX, Rao N, Guo FB (2020) CEG 2.0: an updated database of clusters of essential genes including eukaryotic organisms. Database (Oxford) 2020:baaa112. https://doi.org/10.1093/database/baaa112
Luo H et al (2014) DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. Nucleic Acids Res 42(D1):D574–D580
Malik A et al (2019) Structural insights into Entamoeba histolytica arginase and structure-based identification of novel non-amino acid based inhibitors as potential antiamoebic molecules. Febs J 286(20):4135–4155
Meehl M et al (2007) Interaction of the GraRS two-component system with the VraFG ABC transporter to support vancomycin-intermediate resistance in Staphylococcus aureus. Antimicrob Agents Chemother 51(8):2679–2689
Meservey A et al (2020) Staphylococcus sciuri peritonitis in a patient on peritoneal dialysis. Zoonoses Public Health 67(1):93–95
Mukherjee S, Stamatis D, Bertsch J, Ovchinnikova G, Verezemska O, Isbandi M, Thomas AD, Ali R, Sharma K, Kyrpides NC, Reddy TB (2017) Genomes online database (GOLD) v.6: data updates and feature enhancements. Nucleic Acids Res 45(D1):D446–D456. https://doi.org/10.1093/nar/gkw992
Nemeghaire S et al (2014) The ecological importance of the Staphylococcus sciuri species group as a reservoir for resistance and virulence genes. Vet Microbiol 171(3–4):342–356
Nemeghaire S et al (2014) Characterization of methicillin-resistant Staphylococcus sciuri isolates from industrially raised pigs, cattle and broiler chickens. J Antimicrob Chemother 69(11):2928–2934
Paterson GK (2020) Genomic epidemiology of methicillin-resistant Staphylococcus sciuri carrying a SCCmec-mecC hybrid element. Infect Genet Evol 79:104148
Pettersen EF et al (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25(13):1605–1612
Pundir S, Martin MJ, O'Donovan C (2017) UniProt protein knowledgebase. Methods Mol Biol 1558:41–55. https://doi.org/10.1007/978-1-4939-6783-4_2
Radusky LG et al (2015) An integrated structural proteomics approach along the druggable genome of Corynebacterium pseudotuberculosis species for putative druggable targets. BMC Genomics 16(5):1–8
Rahman MT et al (2005) Genetic analysis of mecA homologues in Staphylococcus sciuri strains derived from mastitis in dairy cattle. Microb Drug Resist 11(3):205–214
Rahman MM et al (2016) The Staphylococcus aureus methicillin resistance factor FmtA is a d-amino esterase that acts on teichoic acids. mBio 7(1):e02070-02015
Rey Pérez J, Zálama Rosa L, García Sánchez A, de Mendoza Hermoso, Salcedo J, Alonso Rodríguez JM, Cerrato Horrillo R, Zurita SG, Gil Molino M (2021) Multiple antimicrobial resistance in methicillin-resistant staphylococcus sciuri group isolates from wild ungulates in Spain. Antibiotics (Basel) 10(8):920. https://doi.org/10.3390/antibiotics10080920
Rufino SD et al (1997) Predicting the conformational class of short and medium size loops connecting regular secondary structures: application to comparative modelling. J Mol Biol 267(2):352–367
Salazar-Ardiles C, Caimanque T, Galetović A, Vilo C, Araya JE, Flores N, Gómez-Silva B (2020) Staphylococcus sciuri strain LCHXa is a Free-Living Lithium-Tolerant Bacterium Isolated from Salar

de Atacama, Chile. Microorganisms 8(5):668. https://doi.org/10.3390/microorganisms8050668
Salomon-Ferrer R et al (2013) An overview of the Amber biomolecular simulation package. Wiley Interdiscip Rev: Comput Mol Sci 3(2):198–210
Santos INM, Kurihara MNL, Santos FF, Valiatti TB, Silva JTPD, Pignatari ACC, Salles MJ (2022) Comparative phenotypic and genomic features of staphylococci from sonication fluid of orthopedic implant-associated infections with poor outcome. Microorganisms 10(6):1149. https://doi.org/10.3390/microorganisms10061149
Saraiva MMS et al (2021) Staphylococcus sciuri as a reservoir of mecA to Staphylococcus aureus in non-migratory seabirds from a remote oceanic island. Microb Drug Resist 27(4):553–561
Scholz C et al (2015) DOCKTITE - a highly versatile step-by-step workflow for covalent docking and virtual screening in the molecular operating environment. J Chem Inf Model 55:398–406
Singh V et al (2022) Drug-repurposing approach to combat Staphylococcus aureus: biomolecular and binding interaction study. ACS Omega 7(43):38448–38458
Singh V et al (2022) In-silico functional and structural annotation of hypothetical protein from Klebsiella pneumonia: a potential drug target. J Mol Graph Model 116:108262
Singhal K, Mohanty S (2019) Genome organisation and comparative genomics of four novel Wolbachia genome assemblies from Indian Drosophila host. Funct Integr Genomics 19(4):617–632
Somani D et al (2019) Transcriptomics analysis of propiconazole-treated Cochliobolus sativus reveals new putative azole targets in the plant pathogen. Funct Integr Genomics 19(3):453–465
Sun S et al (2021) A network-based approach to identify protein kinases critical for regulating srebf1 in lipid deposition causing obesity. Funct Integr Genomics 21:557–570
Szklarczyk D et al (2019) STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47(D1):D607–D613
Torres RT et al (2020) Wild boar as a reservoir of antimicrobial resistance. Sci Total Environ 717:135001
Vaught A (1996) Graphing with Gnuplot and Xmgr: two graphing packages available under linux. Linux J 1996(28es):7-es
Volkamer A et al (2012) DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. Bioinformatics 28(15):2074–2075
Wang L et al (2022) In silico development and experimental validation of a novel 7-gene signature based on PI3K pathway-related genes in bladder cancer. Funct Integr Genomics 22:797–811
Wattam AR et al (2017) Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res 45(D1):D535–D542
Weiner PK, Kollman PA (1981) AMBER: assisted model building with energy refinement. A general program for modeling molecules and their interactions. J Comput Chem 2(3):287–303
Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF (1999) Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112:531–52. https://doi.org/10.1385/1-59259-584-7:531
Williams JMG et al (2007) Autobiographical memory specificity and emotional disorder. Psychol Bull 133(1):122
Wu S et al (1996) Tracking the evolutionary origin of the methicillin resistance gene: cloning and sequencing of a homologue of mecA from a methicillin susceptible strain of Staphylococcus sciuri. Microb Drug Resist 2(4):435–441
Wu SW et al (2001) Recruitment of the mecA gene homologue of Staphylococcus sciuri into a resistance determinant and expression of the resistant phenotype in Staphylococcus aureus. J Bacteriol 183(8):2417–2424
Yu NY et al (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26(13):1608–1615
Yu C-S et al (2014) CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. PLoS One 9(6):e99368
Yuan WJ et al (2022) Comparative transcriptome analyses identify genes involved into the biosynthesis of forsythin and forsythoside A in Forsythia suspensa. Funct Integr Genomics 22:731–741
Zhang M et al (2022) Staphylococcus sciuri causes disease and pathological changes in hybrid sturgeon acipenser baerii × acipenser schrencki. Front Cell Infect Microbiol 12:1029692
Zhang X et al (2023) Identification of diagnostic molecules and potential therapeutic agents for atopic dermatitis by single-cell RNA sequencing combined with a systematic computing framework that integrates network pharmacology. Funct Integr Genomics 23(2):95
Zhao L et al (2020) Advancing computer-aided drug discovery (CADD) by big data and data-driven machine learning modeling. Drug Discov Today 25(9):1624–1638
Zimmerman RJ, Kloos WE (1976) Comparative zone electrophoresis of esterases of Staphylococcus species isolated from mammalian skin. Can J Microbiol 22(6):771–779

## Authors and Affiliations

**Aafareen Khan[1] · Saman Sohail[1] · Seerat Yaseen[2] · Sareen Fatima[3] · Ayesha Wisal[1] · Sufyan Ahmed[2] · Mahrukh Nasir[4] · Muhammad Irfan[4] · Asad Karim[4] · Zarrin Basharat[5] · Yasmin Khan[4] · Muhammad Aurongzeb[6] · Syed Kashif Raza[7] · Mohammad Y. Alshahrani[8] · Carlos M. Morel[9] · Syed S. Hassan[4,9]**

✉ Syed S. Hassan
hassanchemist83@gmail.com; hassanchemist83@iccs.edu

Aafareen Khan
aafreendarweshkhan@gmail.com

Saman Sohail
samansohail195@gmail.com

Seerat Yaseen
its_seerat93@hotmail.com

Sareen Fatima
sayreenkhan@gmail.com

Ayesha Wisal
ayeshawisal@gmail.com

Sufyan Ahmed
drsufyan76@yahoo.com

Mahrukh Nasir
drmahrukhnasir@gmail.com

Muhammad Irfan
mirfan046@gmail.com

Asad Karim
asad.karim.alvi@gmail.com

Zarrin Basharat
zarrin.iiui@gmail.com

Yasmin Khan
yasminkhanmpharm@gmail.com; yasminkhan@iccs.edu

Muhammad Aurongzeb
aurangzebku@gmail.com

Syed Kashif Raza
s_kashif_raza@outlook.com

Mohammad Y. Alshahrani
moyahya@kku.edu.sa

[1] Department of Chemistry, Islamia College Peshawar, Peshawar 25000, KP, Pakistan

[2] Abbasi Shaheed Hospital, Karachi Medical and Dental College, Karachi, Pakistan

[3] Department of Microbiology, University of Balochistan, Quetta, Balochistan, Pakistan

[4] Dr. Panjwani Center for Molecular Medicine, International Center for Chemical and Biological Sciences (ICCBS-PCMD), University of Karachi, Karachi 75270, Pakistan

[5] Alpha Genomics (Private) Limited, Islamabad 44710, Pakistan

[6] Faculty of Engineering Sciences & Technology, Hamdard University, Karachi 74600, Pakistan

[7] Faculty of Rehabilitation and Allied Health Sciences (FRAHS), Riphah International University, Faisalabad, Pakistan

[8] Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Khalid University, P.O. Box 61413, Abha 9088, Saudi Arabia

[9] Centre for Technological Development in Health (CDTS), Oswaldo Cruz Foundation (Fiocruz), Building "Expansão", 8Th Floor Room 814, Av. Brasil 4036 - Manguinhos, Rio de Janeiro, RJ 21040-361, Brazil