ORIGINAL RESEARCH ARTICLE

CrossMark

Subtractive Genomics, Molecular Docking and Molecular Dynamics Simulation Revealed LpxC as a Potential Drug Target Against Multi-Drug Resistant *Klebsiella pneumoniae*

Sajjad Ahmad[1] · Afifa Navid[1] · Amina Saleem Akhtar[1] · Syed Sikander Azam[1] · Abdul Wadood[2] · Horacio Pérez-Sánchez[3]

Abstract

The emergence and dissemination of pan drug resistant clones of *Klebsiella pneumoniae* are great threat to public health. In this regard new therapeutic targets must be highlighted to pave the path for novel drug discovery and development. Subtractive proteomic pipeline brought forth UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine deacetylase (LpxC), a Zn^{+2} dependent cytoplasmic metalloprotein and catalyze the rate limiting deacetylation step of lipid A biosynthesis pathway. Primary sequence analysis followed by 3-dimensional (3-D) structure elucidation of the protein led to the detection of *K. pneumoniae* LpxC (KpLpxC) topology distinct from its orthologous counterparts in other bacterial species. Molecular docking study of the protein recognized receptor antagonist compound 106, a uridine-based LpxC inhibitory compound, as a ligand best able to fit the binding pocket with a Gold Score of 67.53. Molecular dynamics simulation of docked KpLpxC revealed an alternate binding pattern of ligand in the active site. The ligand tail exhibited preferred binding to the domain I residues as opposed to the substrate binding hydrophobic channel of subdomain II, usually targeted by inhibitory compounds. Comparison with the undocked KpLpxC system demonstrated ligand induced high conformational changes in the hydrophobic channel of subdomain II in KpLpxC. Hence, ligand exerted its inhibitory potential by rendering the channel unstable for substrate binding.

Keywords *K. pneumoniae* HS11286 · LpxC · Subtractive proteomics · Homology modeling · Molecular docking · MD simulations

✉ Syed Sikander Azam
 ssazam@qau.edu.pk

1. National Center for Bioinformatics (NCB), Quaid-i-Azam University, Islamabad 45320, Pakistan

2. Department of Biochemistry, Abdul Wali Khan University-Mardan, Shankar Campus, Mardan, Khyber Pukhtoonkhwa, Pakistan

3. Structural Bioinformatics and High Performance Computing Research Group (BIO-HPC), Universidad Católica San Antonio de Murcia (UCAM), Murcia, Spain

 ∅ Springer

1 Introduction

Klebsiella pneumoniae is one of the most prevalent and clinically significant pathogens of humans, associated with pneumonia, respiratory diseases, urinary tract infections and blood infections in neonates, elderly and immune-compromised patients [1–3]. In recent years, *K. pneumoniae* has gained global attention as one of the most predominant and troublesome nosocomial pathogen due to rapid dissemination of resistance against all major classes of antibiotics [4–6]. More worrisome are carbapenemase producing strains with large-scale endemics in health-care facilities world-wide. The advent of resilient strains and inefficacy of frontline antibiotics has necessitated the application of highly efficient computational paradigm of drug design in order to equip modern medicine with novel anti-*K. pneumoniae* drugs [7].

Trending application of bioinformatics expertise has immensely accelerated the standard procedure of drug

design [8–10]. In silico drug target mining via subtractive proteomics lays the groundwork in this regard and has been favoured in numerous researches concerning pathogenic bacteria [11–17]. It is a whole genome comparative approach in which microbial proteome is extensively traversed through a series of bioinformatics tools in order to mine therapeutically important, unique and druggable targets [18]. Application of subtractive proteomics, also known as differential proteomics mining, has been successful in bringing attention to the previously unexplored bacterial biomolecules. Such unique targets can be focused upon by therapeutic drugs in an attempt to achieve a more positive outcome. Moreover, use of computer aided-drug designing (CADD) methods such as three-dimensional (3D) structural elucidation of drug targets, molecular docking and molecular dynamics simulations analysis contribute substantially towards the identification of potent antibacterial compounds and determination of time-dependent nature of drug–target interactions [19–22].

Current research is driven by the necessity to address the incessantly increasing health-risk posed by life threatening *K. pneumoniae* subsp. *pneumoniae* HS11286 (KPHS). Application of in silico subtractive proteome mining and characterization of druggable proteome yields essential, pathogen-specific drug target in the shape of *K. pneumoniae* LpxC (KpLpxC), a zinc-dependent metalloprotease, namely UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine deacetylase. It is a hydrolase enzyme and performs rate-limiting deacetylation of its substrate in Lipid A biosynthetic pathway [23]. Lipid A biosynthesis is pivotal to the structural integrity of external lipopolysaccharide (LPS) membrane in Gram-negative bacteria [23, 24]. Administering potent inhibitory drugs against KpLpxC deteriorates outer protective LPS membrane of bacteria and compromises its defence mechanism, thereby providing a focal target for therapeutic intervention against bacterial pathogen [25]. The scope of the current work is to identify a potent receptor antagonist against pathogenic target and assess the native configuration of drug target complex. Insights from this study will be conducive to pathogen-specific drug designing protocols, which in turn can be extrapolated to improve the existing healthcare systems.

2 Materials and Methods

Figure 1 outlines the methodology applied in current study that will be detailed in the following section.

2.1 Differential Proteome Mining

The complete proteome of KPHS was retrieved from UniprotKB and subjected to differential proteome mining
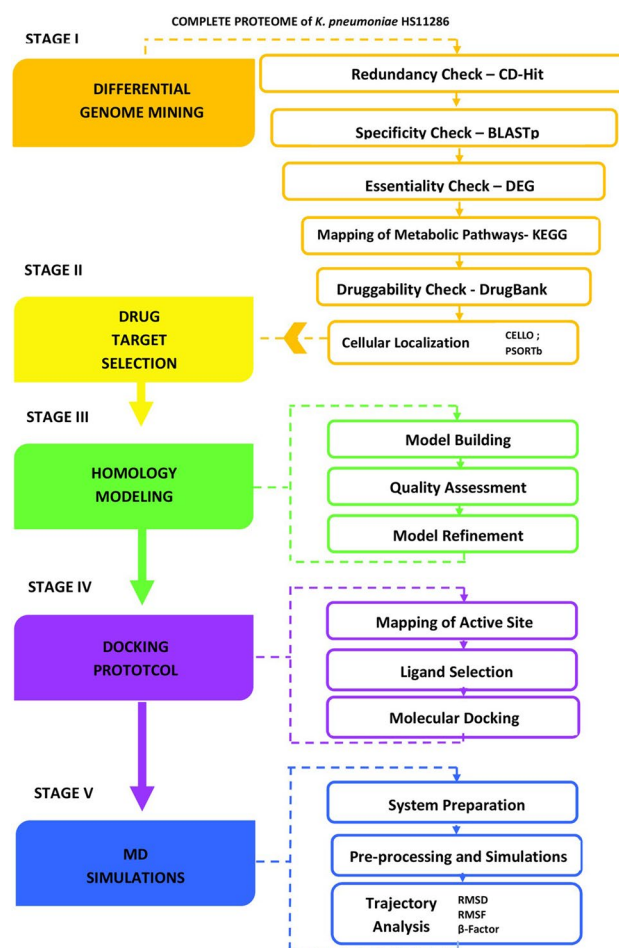

Fig. 1 A flowchart of the integrative methodology applied in the study

pipeline [26]. UniProtKB is a comprehensive repository of proteins with both manually annotated protein sequences found under UniProtKB/Swiss-Prot and automatically annotated proteins present in UniProtKB/TrEMBL [27]. The proteome data set was then passed through a Cluster Database at High Identity with Tolerance (CD-Hit) filter in order to remove paralogous sequences (http://weizhong-lab.used.edu/cdhit_suit/cgi-bin/index.cgi). At redundancy check, pair-wise comparison of sequences is carried out on the basis of short sequence matching computed from identity of the sequences, followed by the removal of redundant sequences from the dataset at a similarity threshold of 0.6 (>60%) [28, 29]. For specificity check, non-redundant protein sequences were queried at the NCBI BLASTp server (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to screen out human homologous pathogen sequences [30]. Similarity search was performed against the human proteome (TaxID: 9606) at the Refseq database with a default threshold E value of 10^{-3} [31]. To perform essentiality check, bacterial essential proteome was identified by sifting the set

of non-homologous KPHS sequences through DEG (Database of essential genes) database (http://tubic.tju.edu.cn/deg/), with a cut-off E value set at 10^{-10}, bit score > 100 and sequence identity ≥ 30% [32]. The DEG database contains information related to gene dataset essential for the survival of organism and contains 43 representatives' prokaryotic and eukaryotic genomes [33]. In order to annotate identified essential genes set the services of KEGG Automatic Annotation Server (KAAS) (http://www.genome.jp/kegg/kaas/) was utilized [34]. KAAS KPHS proteome represented by three-letter code 'kpm' was used as the reference set and bi-directional hit (BDH) method was selected for Kyoto Encyclopaedia of Genes and Genomes (KEGG) orthology (KO) assignment. For unique pathway identification, manual mapping of metabolic pathways was performed by consulting the KEGG pathway database [35]. Drugability check was addressed by performing a drugability assessment of pathogen-specific essential proteins at DrugBank [36]. The default parameter of E value 10^{-5} was used to guide the similarity search for druggable homologs of screened candidate sequences. Last, CELLO and PSORTb were employed to prioritize druggable KPHS target; proteins on the basis of subcellular localization [37, 38].

2.2 Drug Target Selection

Guided by the final proteome subtractive step, the choice of a prospective drug target was further restricted to proteins with the cytoplasmic site of action. Function-related attributes for the 'sifted' cytoplasmic drug targets were acquired from UniprotKB (Entry codes G8VWJ8, G8W2U2, G8VY79, G8VY78, and G8VWP3). Prospects for structure-based drug design modules were assessed by querying template availability for each target. For this purpose, amino acid sequences of target proteins were subjected to similarity search at BLASTp server against RCSB Protein Data Bank (PDB). Templates with sequence identity ≥ 30 and query coverage ≥ 80% were considered suitable templates.

2.3 Homology Modelling

The KpLpxC 3D structure was modelled via MODEL-LERv.9.12 using *E. coli* LpxC (PDB Id: 4MQY, Chain A) as a template [39]. The experimentally derived structure of the template was in complex form with a bound LPC-138 inhibitor and also included the functionally important Zn^{+2} ion as cofactor. All non-standard molecules with the exception of Zn^{+2} were removed from the template structure. MODEL-LERv.9.12 was employed for command-based generation of Zn^{+2} incorporated KpLpxC structure [40]. Zn^{+2} inclusion was achieved by manually inserting Zn^{+2} as a separate chain in KpLpxC, in the template-target sequence alignment at the preliminary stage of model building. Co-ordinates for the placement of the Zn^{+2} cofactor in KpLpxC were derived from the template structure and adjusted to suitably fit the structural environment of the model. For a comparative analysis, web-based tools i.e. I (Iterative)-TASSER [41], ModWeb [42], 3D-JIGSAW [43], SWISSMODEL [44], and EsyPred3D [45] were also utilized to model 3D structures of target protein.

An assortment of structure validation tools: PROCHECK [46], ERRAT [47], Verify-3D [48], and (protein structure analysis) ProSA [49] were used to evaluate stereochemical properties and model quality of generated models [50]. Root mean square deviations (RMSD) of generated models were calculated individually by superimposing upon the reference structure at UCSF Chimera [51]. For further structural refinement, the model with best stereochemistry was subjected to energy optimization at UCSF Chimera. The structure was assigned Gasteiger–Huckel charges and a total of 1500 energy minimization steps were performed by applying Tripos force-field (TFF) in order to remove residual steric clashes.

2.4 Docking Protocol

Literature references to significantly similar known homologous structures of KpLpxC, belonging to *Escherichia coli* (template, PDB ID: 4MQY: A) [39], *Yersinia enterocolitica* (PDB ID: NZK: A) [52] and *Pseudomonas aeruginosa* (PDB ID: 2VES: A) [53] were scrutinized to determine the primary ligand-binding residues. These multiple sequences were aligned with KpLpxC using clustal omega (clustalo) [54] in order to manually map the corresponding active site residues in target protein. Selection of the ligands was largely guided by the information gathered from Brenda [55]. The database was explored by querying Enzyme Commission (EC) no. 3.5.1.108 for the enzymatic target protein in order to acquire a list of compounds with established inhibitory activity against the protein [53, 56–66]. In addition, some recently explored classes of inhibitory agents were also taken into consideration [67–70]. The scientifically accurate structures of selected ligands were drawn and their respective 3D atomic coordinate files (.pdb) were generated through ChemDraw Ultra 8.0 application of the integrated suite, ChemOffice 2004 [71]. The final structures then underwent energy minimization at Chimera using ff03r.1 force field. GOLD (Genetic Optimization for Ligand Docking) program was employed to predict the structure of protein–ligand docked complexes [72]. Default parameters were opted for exploring the binding orientation of the ligands within the 10 Å radius of the user-defined active site residue, which was specified to be Phe121:N. Genetic algorithm was used to explore

optimal docking solutions, of which 10 best scoring solutions were retained. GOLD Score function was utilized to calculate and score the 'fitness' of ligands into binding pocket. Additionally, molecular docking was performed using AutoDock Vina [73] in order to obtain binding affinities of ligands for the receptor protein. Lamarckian Genetic Algorithm was employed to perform rigid receptor dockings, with grid box centered at the ligand, having dimensions $190.771 \times 107.552 \times 0.762$. Docking results were visualized for bonded–non-bonded interactions using Visual Molecular Dynamics (VMD) [74], Ligplot [75], and Discovery Studio (DS) Visualizer 3.5 software [76].

2.5 MD Simulations

MD Simulations were performed for both undocked KpLpxC and docked KpLpxC complex by using different modules of the Amber suite [77]. Basic simulation environment for both docked and undocked systems were prepared by Leap module of Amber 12. In both systems, ff03r.1 force field was used to define KpLpxC. For docked system, ligand force field parameter file defining the topology and connectivity of ligand was created by employing GAFF (general amber force field) in Antechamber program of Amber 12. A previously derived [78] Zn^{+2} parameter file was used as a template to define Zn^{+2} force field parameters for KpLpxC, which included non-bonded metal coordination (van der Waals interactions) and charge-based terms. The parameter file was loaded in Xleap, the graphical interface to the Amber 12 module, and a charge of $+2.0$ was assigned to Zn atom. The generated Zn^{+2} library was then added to both undocked and docked KpLpxC systems. Solvated environment was simulated using TIP3P implicit water model by adding 9 and 10 net neutralization sodium ions (Na^+) to the undocked and docked systems, respectively. A solvent box of $65.386\,\text{Å} \times 81.019\,\text{Å} \times 71.778\,\text{Å}$ for both systems, comprised of 8549 water molecules, with a cut off value of distance of 8 Å box margins was employed.

MD simulations for both systems were performed using Sander module of Amber 10. To remove possible steric clashes, the protein was subjected to a total of 5000 minimization steps, comprising of 2500 conjugate gradient and 2500 steepest descent gradient runs. Heating of the systems from 0 to 300 K at 1 atm was performed for 10 ps. Langevin Dynamics were applied for controlling the temperature of the system. Systems were equilibrated under similar conditions for 100 ps at constant temperature (300 K) and pressure. Periodic boundary conditions were applied and Shake Algorithm was used to constrain the calculation of hydrogen bonds. Simulations of 12 and 30 ns were performed for undocked and docked systems, respectively,

and trajectories for the respective systems were saved for a time-step of 2 ps. Command-based Ptraj module of Amber10 was used to calculate physical properties of the systems and Xmgrace was employed for graphical analysis of the output files.

3 Results and Discussion

3.1 Differential Proteome Mining

Differential proteome mining resulted in progressive 'subtraction' of proteins from complete pathogenic proteome in search of druggable targets. The graph in Fig. 2 presents the overall outcome of this strategy by outlining the number of KPHS proteins that were brought forth in each step. The basic requirement of the procedure was the availability of the complete proteome which can serve as an input file for the pipeline. In this respect complete proteome of *K. pneumoniae* HS11286 was retrieved from UniProtKB, a comprehensive repository of proteins sequences. The protein data were then subjected to CD-Hit redundancy analysis to remove paralogous sequences and led to the identification of 153 paralogous sequences sharing $> 60\%$ sequence identity, leaving behind 5575 non redundant protein sequences. The specificity checked against the human proteome using BLASTp provided no significant hits for 5141 non-redundant bacterial proteins, allowing selection of pathogen-specific proteins with no orthologous counterparts in human hosts. Essentiality checked in DEG database substantially subtracted the protein set wherein only 1475 out of 5141 proteins were identified as essential
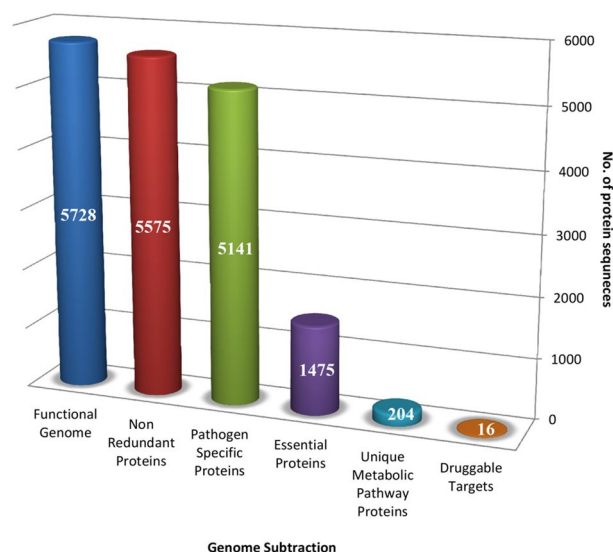

Fig. 2 Graph representing number of KPHS genes screened in progressive stages of differential proteome mining

bacterial proteins and were forwarded along the differential proteome channel. KAAS functional annotation identified the involvement of minimal KPHS proteome set in 128 essential metabolic pathways. Manual comparison of the human host and pathogenic *K. pneumoniae* HS11286 pathways lists led to the identification of 37 pathways that were exclusive to the pathogen. Out of 1475, 204 essential proteins were employed by these unique metabolic pathways. Distribution of 204 pathogen-specific, essential proteins in 37 unique metabolic pathways of *K. pneumoniae* HS11286 is illustrated in Fig. SI-1. These unique metabolic pathway proteins when subjected to druggability check yielded 16 essentially druggable targets. These targets were then checked for their preferred sites of action within the cell and 5 out of 16 putative targets were unanimously categorized as cytoplasmic proteins and were further scrutinized as prospective drug targets.

Proteome subtractive channel ensured a holistic sequence-based screening of *K. pneumoniae* HS11286 proteome for novel targets possessing principle qualifying features of an effective therapeutic agent. Targeting the 1475 essential, non-homologous genes brought forth by essentiality and specificity check provides a twofold advantage of annihilating the bacteria while minimizing the probability of cross reactivity in host [79]. Functional characterization of the essential gene set to 37 unique, pathogen-specific metabolic pathways further augmented the confidence in pathogen-specific drug target selection. Moreover, such essential pathways mark a critical consideration in targeted drug design approach [80]. The highest proportion of essential proteins belonged to pathways including: two-component system, phosphotransferase system, lipopolysaccharide biosynthesis, peptidoglycan synthesis and bacterial secretion system. Consolidating results in parallel differential proteomic studies provide a unanimous assessment of the importance of the pathway proteins in the survival of pathogen, hence providing promising targets for antibacterial drugs [18, 81]. Assessment of drug-binding potential or 'druggability' is the fundamental feature for the identification of a drug target. To this end, the next stage in proteome subtraction involved sequence to structure extrapolative comparisons with established drug targets for the presence of drug-binding sites in the targets. Proteins with significant homologs are considered essential, druggable therapeutic targets. Hence, the 16 non-redundant, pathogen-specific, 'essential', unique and 'druggable' putative therapeutic targets brought forth in current work are suitable drug targets (Table Sl-1) and can be pursued for biomedical research purposes.

3.2 Drug Target Selection

The observed functional characteristics of the five prospective 'drug' targets of cytoplasmic localization are listed in Table 1. Templates were available for all identified targets. UreA and UreB were subunits of Urease amidohydrolase hetrotrimer and were ruled out owing to their dependency on a third non-essential subunit, UreC, for their activation. Remaining three proteins MurF, PtsI and LpxC were further assessed on the basis of functional importance and availability of the best template. Analytical observations revealed LpxC as the most suitable drug target to be pursued in the current study.

The preference for candidate therapeutic targets with certain subcellular location is ascertained by the requirement of the research. While vaccine targets encompass membranous, excreted and transport proteins, it is the cytoplasm localized proteins that are preferred drug targets [18]. Adhering to the purpose of current study, the five cytoplasmic localized proteins were further scrutinized to determine their

Table 1 Functional features of identified cytoplasmic drug targets acquired from UniprotKB

Protein ID	Protein name	Biological function/EC no.	Length	Co-factor	Subunit structure	3D-structure available
G8VWJ8	UDP-*N*-acetylmuramoyl-tripeptide-D-alanyl-D-alanine ligase (murF)	Peptidoglycan biosynthesis [EC:6.3.2.10]	452	No	Monomer	No
G8W2U2	Phosphotransferase system, enzyme I (PtsI, ptsI)	Phosphotransferase system [EC:2.7.3.9]	575	Mg^{+2}	N/A	No
G8VY79	Urease subunit beta; Urea amidohydrolase subunit beta (UreB)	Urea catabolic process [EC 3.5.1.5]	106	Ni	Subunit of Hetrotrimer	No
G8VY78	Urease subunit gamma; Urea amidohydrolase subunit gamma (UreA)	Urea catabolic process [EC 3.5.1.5]	100	Ni	Subunit of Hetrotrimer	No
G8VWP3	UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine deacetylase (LpxC)	Lipo-polysaccharide biosynthesis [EC:3.5.1.108]	234	Zn^{+2}	Monomer	No

prospects as drug targets. Insights into the molecular functions in fundamental, pathogen-specific biological processes guide the selection of viable drug candidates [82]. Exclusive to Gram-negative bacteria, the outer membrane synthesizing lipopolysaccharide biosynthesis has garnered an active therapeutic interest from the bio- medicinal community [56]. In view of the fact that *K. pneumoniae* HS11286 belongs to the group of a Gram-negative bacteria, LpxC belonging to lipopolysaccharide biosynthetic pathway presented an attractive target. Interestingly, *K. pneumoniae* HS11286 LpxC (KpLpxC) was found to have a shorter sequence (234 aa) than its orthologous counterparts in other species. Analysis of structural implications of this sequence variation, corresponding behaviour of Zn^{+2} cofactor and effect on ligand binding was an additional factor that guided the selection of this target.

3.3 Homology Modelling

X-ray crystallographic structure of *E. coli* LpxC (EcLpxC) (PDB Id: 4MQY, Chain A) was identified as the most suitable template for modelling of *K. pneumoniae* HS11286 LpxC (KpLpxC) using MODELLER*v*.9.12 and other web-based software [39]. The structure of *E. coli* LpxC (EcLpxC) as a template was used because of its high sequence identity of 96% and query coverage of 100% with KpLpxC and as such, KpLpxC can be considered as high-quality model that is comparable with medium-resolutionNMR solution or low-resolution X-ray crystallography structure without any biasness. Parameters for the quality assessment of the generated models are listed in Table 2. Ensuing comparison of the model quality and stereochemistry guided the selection of the best model. None of the web-based servers was able to support the inclusion of Zn^{+2} cofactor within the models. Model generated via MODELLER, on the other hand, had the added benefit of Zn^{+2} incorporation in addition to overall exceptional stereochemistry. With best PROCHECK statistics (94.8% residues in the core favourable region), highest ERRAT quality (83.256) and RMSD of 0.246 Å, MODELLER model was deemed the most reliable and high-quality representative 3D structure of KpLpxC. The knowledge-based energy graph for the top modelled LpxC is provided in Fig. SI-2. The graph shows local model quality by plotting energy as a function of amino acid. As can be seen in the Figure there are two types of lines: thick and thin. The thick line represents the average energy over each 40 residue fragments while the thin line depicts window size of 10 residues and be seen in the background of the graph. In general, the positive values correspond to erroneous part of the input structure. As can be seen in the graph majority portion of the top modelled structure has negative value and hence can be considered as the most reliable model for the protein.

Table 2 Stereochemical evaluation parameters for KpLpxC structure models generated by MODELLER and web-based servers

Models	PROCHECK								ERRAT	Verify-3D	ProSA-web	Chimera
	No. of non-glycine and non-proline residues				No. of proline, glycine and end residues	Residue coverage	Bad contacts	G factor	Quality factor (%)	3D-1D scores> 0.2 (%)	Z Score	RMSD (Å)
	Most favoured region [A,B,L]	Additionally allowed region [a,b,l,p]	Generously allowed region [~a,~b,~l,~p]	Disallowed region								
MODELLER	199 94.8%	10 4.8%	1 0.5%	0 0.0%	24 + Zn	235	−0.4	1.2	83.256	93.62%	−5.57	0.246
SWISS MODEL	179 86.9%	26 12.6%	1 0.5%	0 0.0%	24	230	−0.4	1.6	90.686	98.70	−5.8	0.065
3D-JIGSAW	172 80.4%	37 17.3%	5 2.3%	0 0.0%	24	238	0.4	−1.6	59.009	82.43%	−5.69	0.601
EsyPred3D	195 94.2%	11 5.3%	1 0.5%	0 0.0%	24	231	−0.4	1.3	74.766	100%	−5.73	0.221
I-TASSER	180 85.7%	22 10.5%	5 2.4%	3 1.4%	24	234	0.1	0.2	26.484	90.21%	−5.81	0.478
ModWeb	197 95.6%	8 3.9%	1 0.5%	0 0.0%	23	229	−0.4	1.5	84.762	99.57	−5.9	0.250

Fig. 3 **a** 3D model of EcLpxC (template) in comparison with MODELLER-based 3D structure of KpLpxC (target), depicting lack of subdomain I and partial retention of Domain I in KpLpxC. **b** Superimposed structures of EcLpxC (white) and KpLpxC (aqua), highlighting conformational symmetry of Zn^{+2} tetrahedral geometry in the structural environment
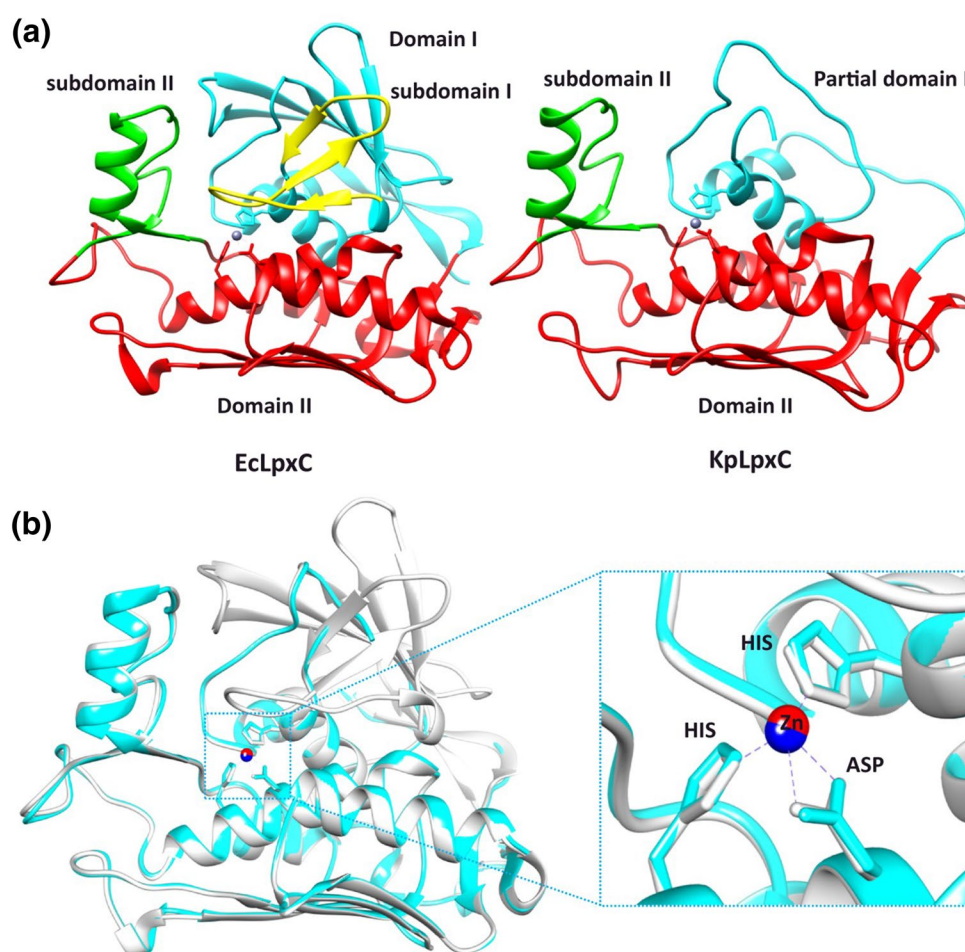
Visual inspection of 3D structural model of KpLpxC highlighted a key structural difference from the template structure (highlighted in Fig. 3a). The EcLpxC reference structure is characterized by a typical LpxC β–α–α–β sandwich topology, formed by terminal domains I and domain II. Each domain comprises of α+β topology wherein two α helices are packed between β sheets comprising of five β-strands. Three oppositely directed β-sheets form subdomain I, whereas subdomain II exhibit β–α–β topology [39]. The generated model, however, was devoid of the N-terminal β-sheet structure belonging to domain I and completely lacked subdomain I. The presence of two α helices at N terminus signified a partial retention of the domain I encompassing protein residues Met1 to Lys57. α+β assembly of domain II was found to be completely conserved, wherein residues Phe58–Arg119 were involved in the formation of α helices, whereas residues Asp148–Ala234 were responsible for modelling the 5-stranded β sheet. The intervening region, i.e. Thr120–VAL147 partook the formation of the sub-domain II structure comprising of β–α–β organization. A high degree of conformational symmetry between the target and reference structure was observed and the placement of Zn^{+2} in the 3D conformational space of KpLpxC model was in sync with that of the reference structure as shown in Fig. 3b.

Comparative model building by web-based models such as I-TASSER and 3D-JigSaw yielded models of low ERRAT quality and comparatively bad stereochemistry. Esypred-3D model also demonstrated relatively low ERRAT quality. ModWeb generated a model with stereochemistry and quality comparable to that of MODELLER model. However, LpxC is a zinc-dependent metalloenzyme, incorporation of Zn^{+2} ion in the structural model was a feature supported by MODELLER alone, thereby guiding the selection of the most structurally relevant as well a good quality model.

KpLpxC 3D structure acquired in this study differed from the typical β_1–α–α–β_2 geometry of LpxC protein in lacking the N-terminal β_1 sheet of domain I and complete sub-domain I structure. The absence of these N-terminal regions in KpLpxC can be rationalized by previous observations wherein a short amino acid sequence implicates that the essential protein is inherently devoid of this structural organization. The resultant distinctive KpLpxC topology

contained partially conserved domain I and completely conserved domain II and subdomain II. Ensuing concerns regarding the structural intactness of the binding pocket were answered by determination of maximum conservation in primary binding residues. Structure-based drug-binding potential of a drug target is primarily assessed by the architecture of the active site [83]. While studies emphasize the importance of deep hydrophobic binding pocket, bacterial targets tend to possess a largely polar surface area that is conducive to target–drug interactions [82]. In accordance with these observations, the binding site of KpLpxC surrounded by the hydrophobic residues of subdomain II. Ile32 and Phe121 at the base of the pocket was also hydrophobic in nature. A number of polar and charged residues such as Ser4, Thr5, Glu7, His8, Lys168, His167, Asp171 and His194 in the peripheral regions were indicative of strong ligand-binding interactions of the protein. Crystal structure elucidation studies of orthologous LpxC proteins similarly highlighted the involvement of corresponding residues in forming the primary binding pocket of the protein [84]. As the metal cofactor is instrumental to the catalytic activity of LpxC, the complete conservation Zn interacting residues centred within the binding pocket were an additional feature

that deemed the modelled structure suitable for structure based drug designing.

3.4 Docking Protocol

Multiple sequence alignment of KpLpxC against PaLpxC, YeLpxC and EcLpxC for active site prediction is illustrated in Fig. SI-3. Primary binding site of KpLpxC was identified at the interface of the main domains I and II and subdomain II. Catalytic Zn^{+2} triad of His8, His167, Asp171,the domain I residues Ser4, Thr5, Glu7, Ile32, subdomain II residues Thr120, Phe121 and domain II residues Lys168, His194 were identified as the conserved core active site residues.

The chemical structure, GOLD scores and AutoDock Vina acquired binding affinities (kJ mol^{-1}) of 249 docked compounds can be seen in Table SI-2. GOLD fitness scores of the ligand poses spanned the range of 28.35–67.29. Binding affinities of the inhibitor compounds obtained from AutoDock Vina ranged from -8.6 to -4.0 kJ mol^{-1}. A ranking of the docked poses brought forth plausibly potent inhibitors against KpLpxC. With highest GOLD fitness score of 67.29 and binding affinity of -7.9 kJ mol^{-1}, binding

Fig. 4 **a** Surface view of KpLpxC complexed with compound 106. **b** Docked pose of compound 106 in the active site of KpLpxC. **c** Compound 106 (green) in close contact with active site residues (yellow), highlighting sixfold coordination contact of Zn^{+2}
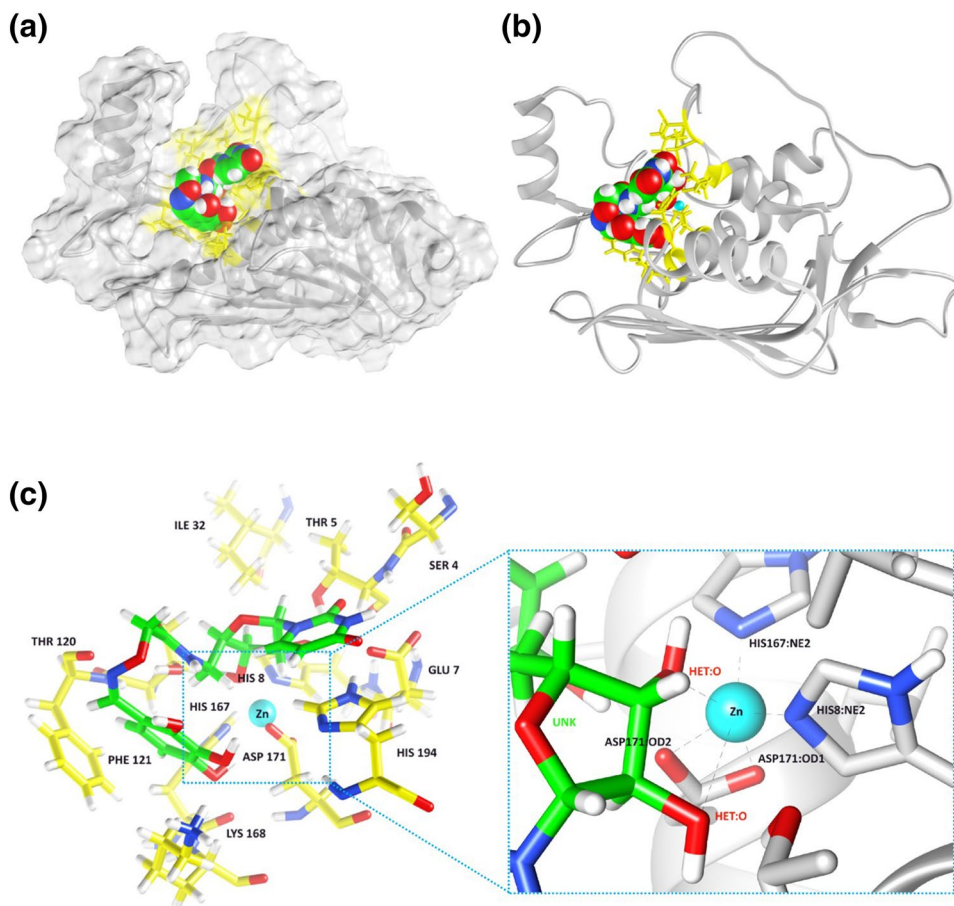
orientation of compound 106 exhibited best complementary binding in the active site of KpLpxC. Molecular surface view of the best docked complex is presented in Fig. 4a and the binding pose of compound 106 in the active site of KpLpxC is depicted in Fig. 4b. Uridine moiety of the compound interacts with Zn^{+2} ion. This particular placement of ligand outlines the entryway to the hydrophobic passage of subdomain II. On the other side, the relaxed conformational constraints of the partial domain I provide an alternate channel for accommodating the acetyl tail of the inhibitor. The acetyl tail of the compound folds in on itself at the interface of the domains. Figure 4c provides a schematic representation of Zn^{+2} coordination with surrounding protein and ligands residues, highlighting hexagonal coordination geometry of the cofactor in docked complex.

Molecular interaction analysis of compound 106-KpLpxC complex by VMD showed that in addition to Zn^{+2} binding residues His8, His167 and Asp171, Thr5, Glu7, Thr120, Phe121, Lys 168 and His194 potentially engaged the bound ligand in an elaborate network of hydrogen bonding at distances ranging from 1.5 to 3.5 Å. The identified bonds are listed in Table 3. Glu7:OE2 additionally formed ionic bonds with ligand N atoms.

The 2D image of Ligplot in Fig. 5a also reveals an extensive coordination between ligand and KpLpxC characterized by multiple hydrogen and polar bonds. His8, His167, Asp171, Thr5 and Lys168 exhibited hydrogen bond formation with ligand atoms at distances of 2.45, 2.77, 2.52, 3.05 and 3.18 Å, respectively. 2D depiction of molecular interactions obtained from DS Visualizer (Fig. 5b) illustrates two hydrogen bonds of Thr5 with the phenyl ring in the hydrophobic tail of the inhibitor. Glu7 and Thr120 contributed their side chain electronegative O atoms for the formation of hydrogen bonds with ligand H atoms. Pi–Pi interactions were observed between the aromatic rings of His194 and ligand acetyl chain. Positively charged Lys168 also showed Pi interaction with the uridine unit of the ligand. Additionally, Lys168 and Asp191 formed hydrogen bonds with ligand's terminal O and H atoms, respectively. The surrounding residues including Ser4, Ile32 and Gly122 participated in the non-bonded van der Waals interactions within the complex. Studies elucidating binding mechanisms of potent substrate like LpxC inhibitors have stated a binding of the hydroxamate moiety in the core active site; with a preferred adjustment of their long acetyl tails in the hydrophobic channel of the protein [23, 85].

The best docked compound belonged to a different class of uridine-based ligands designed with the intention of exploring alternate ligand binding dynamics. A relevant study conducted by Barb et al. proposed a binding mechanism for these ligands wherein the ligand binds to the peripheral region of subdomain II and induces

Table 3 H-bond interactions between KpLpxC and compound 106 identified by using VMD

| Protein residue	Ligand residue	Distance (Å)
HIS8:NE2	UNK1:H	2.74
HIS8:NE2	UNK1:H	2.99
HIS8:HE1	UNK1:O	2.93
HIS8:HD2	UNK1:O	3.64
GLU7:HB2	UNK1:O	3.30
GLU7:OE2	UNK1:H	2.28
GLU7:OE2	UNK1:H	2.98
THR5:HG1	UNK1:O	1.86
THR5:HG1	UNK1:O	3.48
THR5:OG1	UNK1:H	3.06
THR5:OG1	UNK1:H	3.35
THR5:OG1	UNK1:H	3.44
THR5:HB	UNK1:O	2.86
THR5:HB	UNK1:O	3.44
THR5:H	UNK1:O	2.42
THR5:HA	UNK1:O	2.46
THR5:HA	UNK1:O	2.77
THR120:HB	UNK1:O	2.82
THR120:OG1	UNK1:H	2.57
THR120:OG1	UNK1:H	1.79
THR120:OG1	UNK1:H	2.14
THR120:OG1	UNK1:H	3.34
THR120:OG1	UNK1:H	2.59
THR120:HG1	UNK1:O	2.96
THR120:HG1	UNK1:O	3.43
THR120:HG1	UNK1:N	3.67
THR120:HB	UNK1:O	2.82
PHE121:O	UNK1:H	2.69
PHE121:O	UNK1:H	2.97
PHE121:O	UNK1:H	3.43
HIS167:NE2	UNK1:H	2.55
HIS167:NE2	UNK1:H	3.50
HIS167:HD2	UNK1:O	2.26
HIS167:O	UNK1:H	3.29
LYS168:HA	UNK1:O	2.79
ASP171:OD2	UNK1:H	2.59
ASP171:OD2	UNK1:H	3.25
ASP171:OD2	UNK1:H	3.26
ASP171:OD2	UNK1:H	2.29
HIS194:HD2	UNK1:O	3.39
HIS194:HE1	UNK1:O	3.09
HIS194:HE2	UNK1:O	3.25
HIS194:HE2	UNK1:N	3.16
HIS194:HE2	UNK1:N	2.74
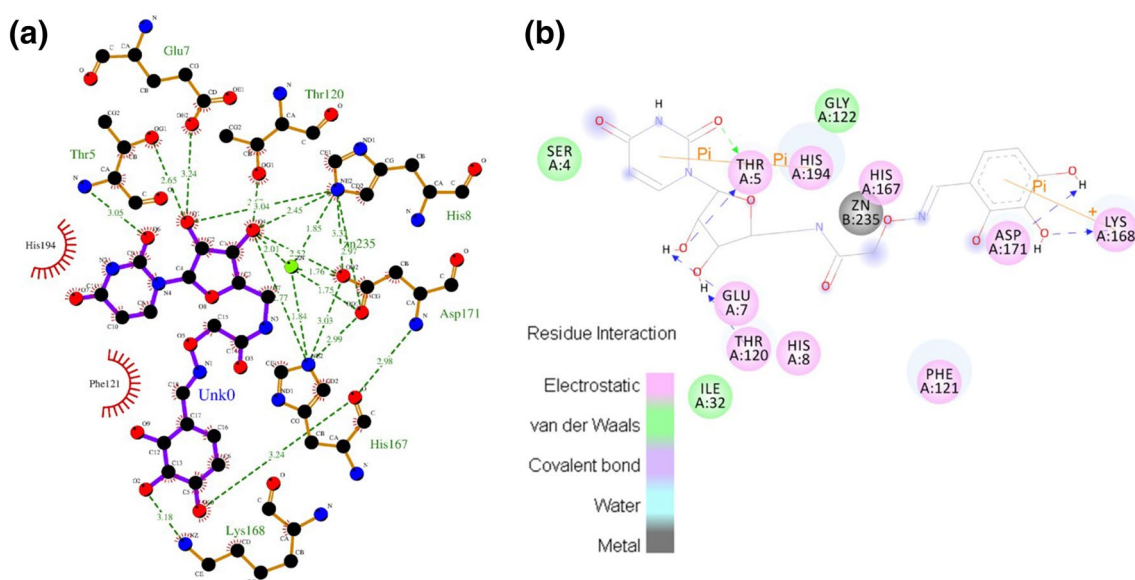
Fig. 5 **a** 2D Ligplot image of KpLpxC-compound 106 complex, representing an extensive network of hydrogen bonds and polar bonds between KpLpxC residues, Zn^{+2} (green) and ligand atoms. **b** 2D molecular interaction image of the complex depicted by DS Visualizer, highlighting Pi–Pi interactions in addition to multiple hydrogen bonds

destabilising effect on its hydrophobic sub-domain [56]. Current study witnessed binding preference of uridine-based ligand for the primary active site of KpLpxC; however, in accordance with the nature of uridine-based inhibitors, the ligand tail did not extend towards the hydrophobic channel but folded into the active site forming strong hydrogen bonds with polar residues. Only base residues of the hydrophobic channel (Thr120 and Phe121) formed hydrogen bonding interactions with the ligand. Thr5, Thr120 and Lys168 most commonly contributed to the intermolecular hydrogen bond network. In a recent study, Hale et al. have documented the significance of residues corresponding to Thr120 and Lys168 in terms of forming LpxC-ligand stabilizing interactions [84].

3.5 MD Simulations

Root mean square deviation (RMSD) graph for 12 ns simulation of undocked KpLpxC is represented in Fig. 6a. Backbone RMSD of Cα atoms showed a steady increase in initial 4 ns and was stabilized till 8th ns. A marked increase in RMSD was observed in 9th ns which continued uptill 12th ns. Overall, simulation run of 12 ns revealed an average root mean square deviation of 2.453 Å for undocked KpLpxC structure with maximum value of 3.579 Å observed in the 11th ns. Average root mean square fluctuation (RMSF) of the undocked system was 1.19 Å with maximum value of 5.415 Å observed for residue 233. Peaks indicated in the RMSF graph (Fig. 6b) represent residues with highest atomic oscillations whereby peaks 1, 2 and 4 correspond to loop regions and peak 3 and 5 belong to sub-domain II and C-terminal regions, respectively. β (Beta)-factor graph in Fig. 6c is in accordance with RMSF trends, whereby average value of β-factor for undocked KpLpxC was 52.458 $Å^2$ with a maximum value of 778.75 $Å^2$ corresponding to residue 233. Figure 7a represents undocked KpLpxC structure after 12 ns simulation run. The 12-ns structure has been superimposed with 1 ns structure in order to highlight the conformational variations over the course of simulation. Figure 7b highlights subdomain II as a region of conformational flexibility in undocked system. Additionally, superimposed RMSD, RMSF and β-factor graphs of undocked and docked KpLpxC systems after 12 ns simulations is illustrated in Fig. SI-4.

Tetrahedral coordination geometry of Zn^{+2} co-factor in undocked KpLpxC oscillated between fivefold and sixfold coordination after simulations. Figure 7c is representative of the sixfold coordination geometry of Zn^{+2} in undocked KpLpxC system observed after 1 ns. Loss of contacts with Zn^{+2} interacting triad (His8:NE2, His167:NE2 and Asp171:OD2) was observed. However, Zn maintained long-range native contact with Asp171:OD1. Hexagonal interaction geometry was completed by assuming additional contacts with Glu7:OE1, Glu7:OD2 and three water molecules (WAT 3814: O; WAT 3375: O; and WAT 4001: O). Figure 7d represents the fivefold coordination geometry of Zn in undocked KpLpxC system at the end of 12 ns

Fig. 6 **a** RMSD graph of undocked KpLpxC protein system over 12 ns MD simulations. **b** RMSF. **c** β-Factor
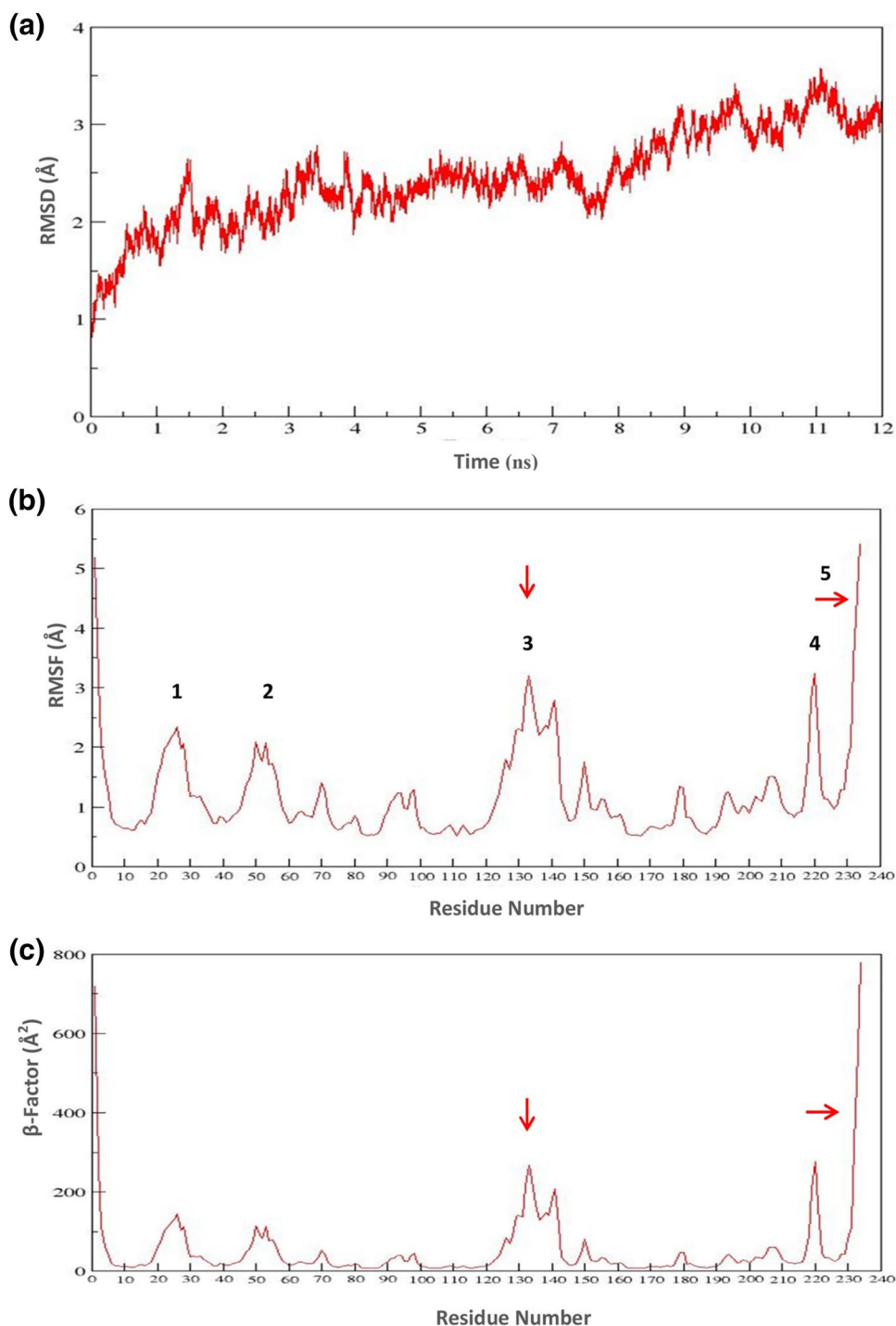
simulation. The change in coordination was assumed upon loss of contact with water molecule (WAT 4001: O) during 6th ns. Figure 8a depicts the movement of Zn^{+2} ion associated with its transitioning 6/5 fold coordination observed after 6th, 7th and 8th ns. Fivefold contacts were maintained after 8 ns. For every 150th frame over 12 ns simulation, average distances of Zn^{+2} metal from the aforementioned Zn^{+2} interacting residues are illustrated in Fig. 8b.

Average RMSD of the Cα atoms of KpLpxC in docked complex over the time period of 30 ns was 2.470 Å. Recurrent rising trends in the RMSD graph, for instance maximum RMSD value of 3.581 during 8th ns in Fig. 9a, can be attributed to the structural deviations exhibited by localized regions of the system. These regions of conformational changes corresponded to the connective helical structure between $β_1$ and $β_2$ of sub-domain II and the C-terminal

Fig. 7 **a** Superimposed image of undocked KpLpxC protein, obtained after 1st ns (green) and 12th ns (blue) MD simulations. **b** Substrate binding subdomain II representing conformational flexibility in the system. **c** Sixfold coordination geometry of Zn in undocked KpLpxC observed after 1 ns. **d** Fivefold coordination geometry of Zn in undocked KpLpxC system observed after 12 ns
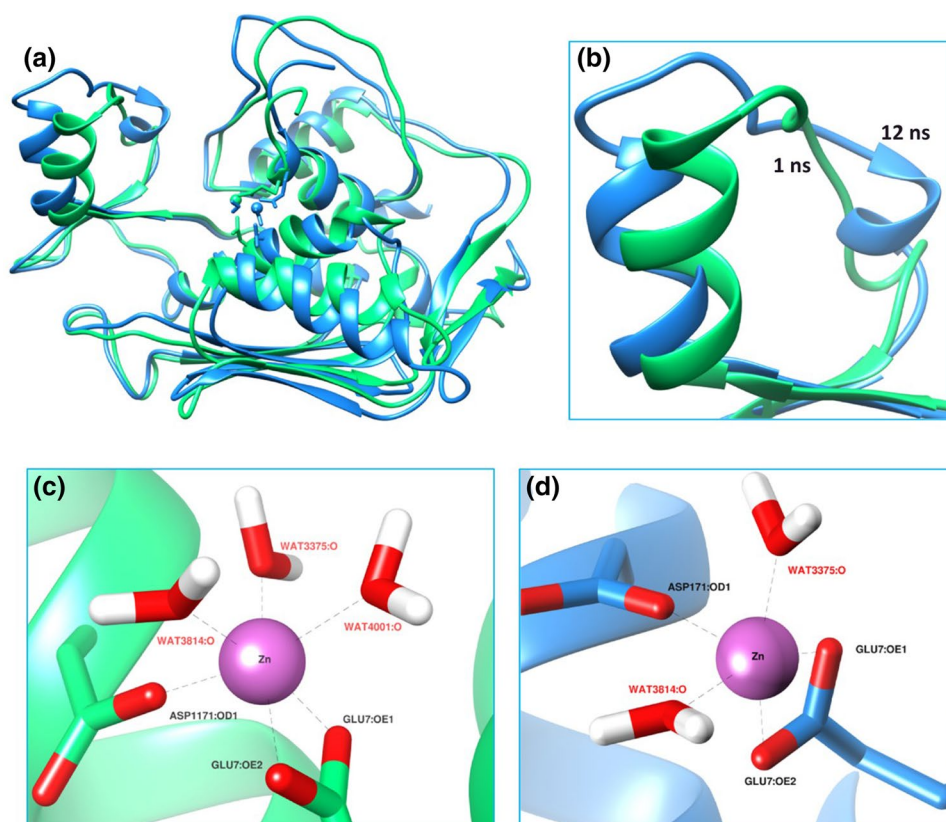

region of KpLpxC. At 30 ns, average RMSF for the docked complex was 1.231 Å. All the active site residues showed minimum fluctuations. High fluctuations were observed for the residues 131–139 with the highest value of 3.580 Å (residue 139). A sharp rise of RMSF in residues 232–234 marked the maximum value 6.590 Å. The peaks, indicated as 3 and 5 in Fig. 9b, also correspond to substrate binding subdomain II and C-terminus of KpLpxC, respectively, thereby correlating the structural deviations with fluctuations at atomic level. Peaks 1, 2 and 4 on the other hand, belonged to loop regions. With an average value of 59.088 Å², β-factors in Fig. 9c followed the same trend as that of RMSF with peak values corresponding to the regions of high fluctuations. Figure 10a depicts a snapshot of KpLpxC complex at the end of 30 simulation runs. The 30 complex has been superimposed with 1 ns complex in order to highlight the structural variations that may have occurred over the time period of 30 ns. Figure 10b highlights subdomain II as a region of high conformational variations. The change in ligand (compound 106) orientation within the active site of KpLpxC at the end of 30 ns is highlighted in Fig. 10c. The acetyl tail of uridine-based compound 106 exhibits a preference for N-terminal groove, instead of the conventional substrate binding hydrophobic channel of domain II. Figure 10d represents the sixfold coordination geometry of catalytic Zn after 30 ns simulation.

Interactions with Zn interacting triad (His8:NE2; His 167:NE2; Asp171:OD1 and OD2) were retained. Loss of contact with two ligand atoms was compensated by nearby residue atoms, Glu7:OE1 and WAT 3904:O.

For the docked system, the sixfold Zn metal coordination identified prior to simulations was maintained throughout the simulation. Figure 11a depicts the conservation of hexagonal coordination geometry by superimposing coordinates of complexes obtained after 1, 5, 10, 20 and 30 ns simulations. An average distance graph of Zn^{+2} interacting residues plotted for every 150th frame over 30 ns simulations is illustrated in Fig. 11b. Zn coordination dynamics observed in undocked and docked KpLpxC systems over 12 and 30 ns simulations, respectively can be seen in Table SI-3.

Capitalizing on the ability to simulate hydrated biological systems for proteins, MD simulations have been employed for systematic evaluation of the time-dependent behavior of pharmacologically important proteins in both docked and undocked systems [20]. Some LpxC proteins exhibit inherently variable structural conformations, whereas other orthologous proteins show ligand-dependent conformational flexibility [86]. In order to determine a potent inhibitor-target combination for LpxC, Lee et al. [86] have emphasized the need to explore the structural dynamics of the protein in solution form. Along the same tangent, time-dependent dynamics of KpLpxC

Fig. 8 **a** Transitional coordination geometry of Zn^{+2} in undocked KpLpxC, represented by superimposing 3D coordinates obtained at 1 ns (pink), 6 ns (purple), 7 ns (blue), 8 ns (green) and 12 ns (grey). Change from sixfold to fivefold geometry is accompanied by Zn^{+2} movement after 6 ns. Labelled residues are involved in 5/6 fold coordination with Zn^{+2}. Three interacting water molecules have not been represented. **b** Distance graph of Zn^{+2} interacting residues from Zn ion; plotted for every 150th frame of 12 ns simulations. HIS8:NE2, HIS167:NE2 and ASP171:OD2 are at maximum distance and do not interact with Zn^{+2}
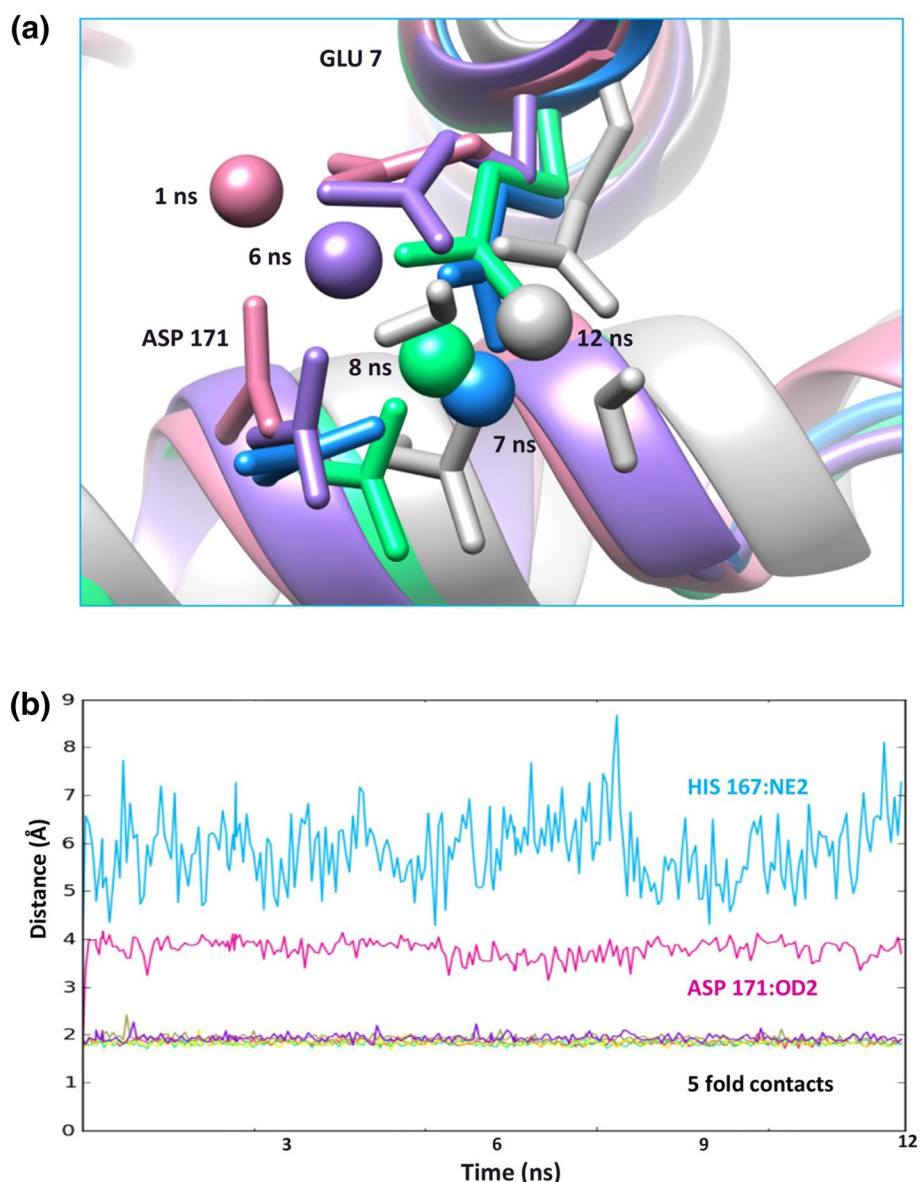

were observed in both docked and undocked systems. Simulation-derived RMSD trends in relation to RMSF and β-factors, observed in current work, illustrated intrinsic flexibility of KpLpxC hydrophobic subdomain II and C-terminal domain residues. The trend demonstrated by C-terminal residues concurs with inherent instability of terminal 3–4 residues; a phenomenon well documented in various LpxC structure-based studies [87]. Although a high degree of conformational flexibility has been associated with sub-domain II between different orthologous LpxC proteins, no local disorderliness of the region has been stated. A noticeably elevated residue fluctuations and disorderliness observed in ligand-bound KpLpxC complex indicated ligand-induced rise in conformational plasticity of these regions. Therefore, in light of these observations,

it can be reasoned that inherent flexibility of KpLpxC protein is subjected to marginal conformational changes upon ligand binding. Moreover, it can be reasoned that high degree of structural instability observed in the substrate binding domain of KpLpxC docked complex renders the channel in unbound fashion, hence inhibiting the associated function of the protein. In this regard, experimental investigations on KpLpxC topology and structural dynamics can further support these observations.

A consensus study on the behaviour of zinc ion in metalloproteins has demonstrated the innate preference of Zn^{+2} for five and sixfold coordination in catalytic enzymes [88]. In concurrence with this finding, the Zn^{+2} coordination geometry for undocked system transitioned between sixfold and fivefold contacts with an associated shift in the metal

Fig. 9 a RMSD graph of
compound 106 docked KpLpxC
protein system over 30 ns
MD simulations. **b** RMSF. **c**
β-Factor
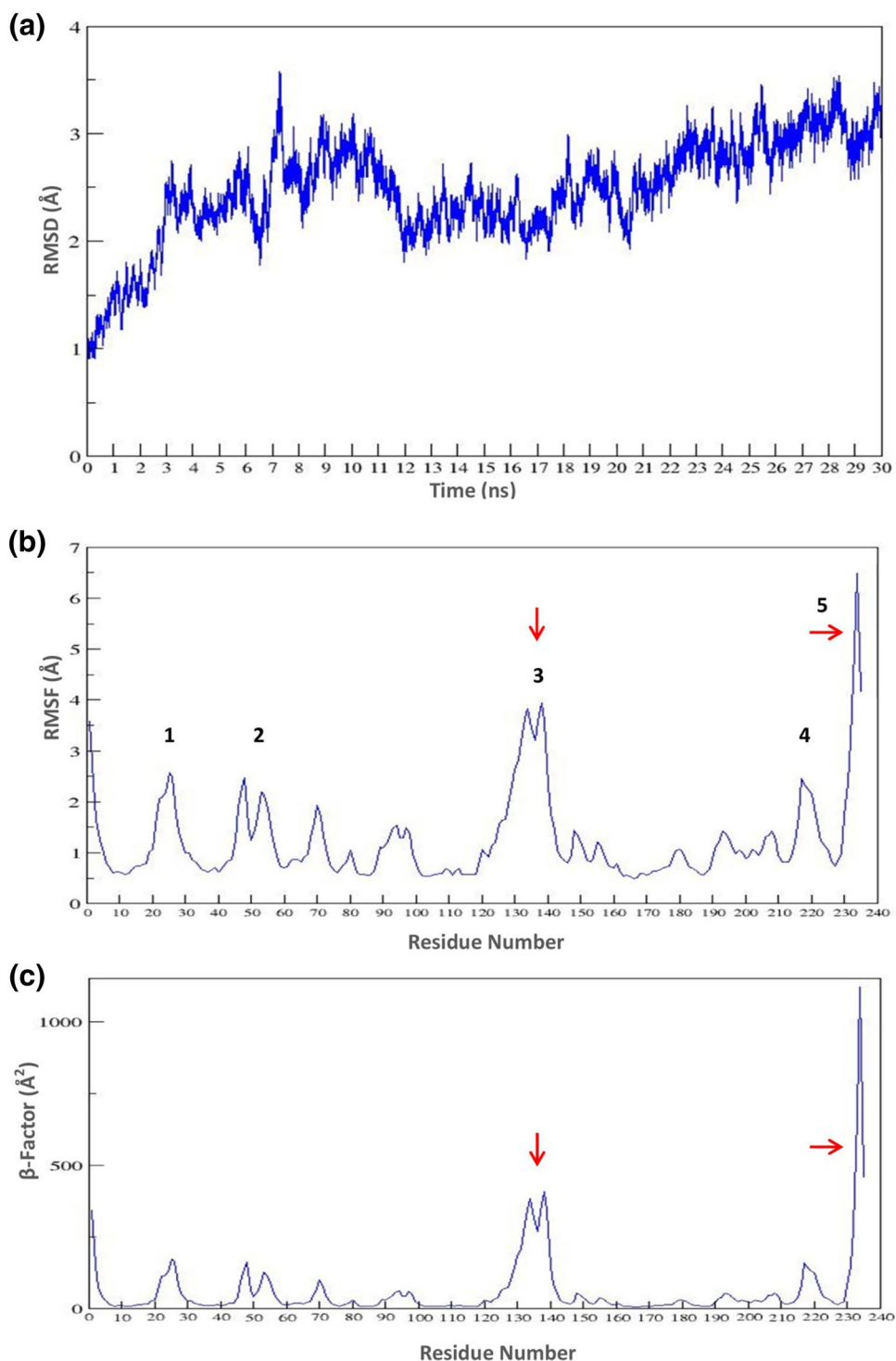
ion positioning. Similarly, Zn^{+2} demonstrated a preference
for hexagonal coordination contacts in docked system, both
prior to and after simulations. Moreover, zinc maintained
its characteristics contacts with the catalytic triad. Hence,
it can be concluded that sixfold coordination of Zn ion was
stabilized by the presence of ligand in the active site demon-
strating environment dependent behavior of KpLpxC.

4 Conclusions

Current work culminates into a successful realization of
outlined objectives by systematic in silico application of
drug design modules. The study characterized 16 putative
druggable compounds from whole proteome subtractive

Fig. 10 **a** Superimposed image of Compound 106-KpLpxC complex, obtained after 1st ns (red) and 30th ns (blue) MD simulations. **b** Substrate-binding subdomain II representing high conformational flexibility in the system. **c** Ligand poses after the time intervals of 1 ns (yellow) and 30 ns (green). **d** Sixfold coordination geometry of Zn in docked KpLpxC system, following 30 ns simulation
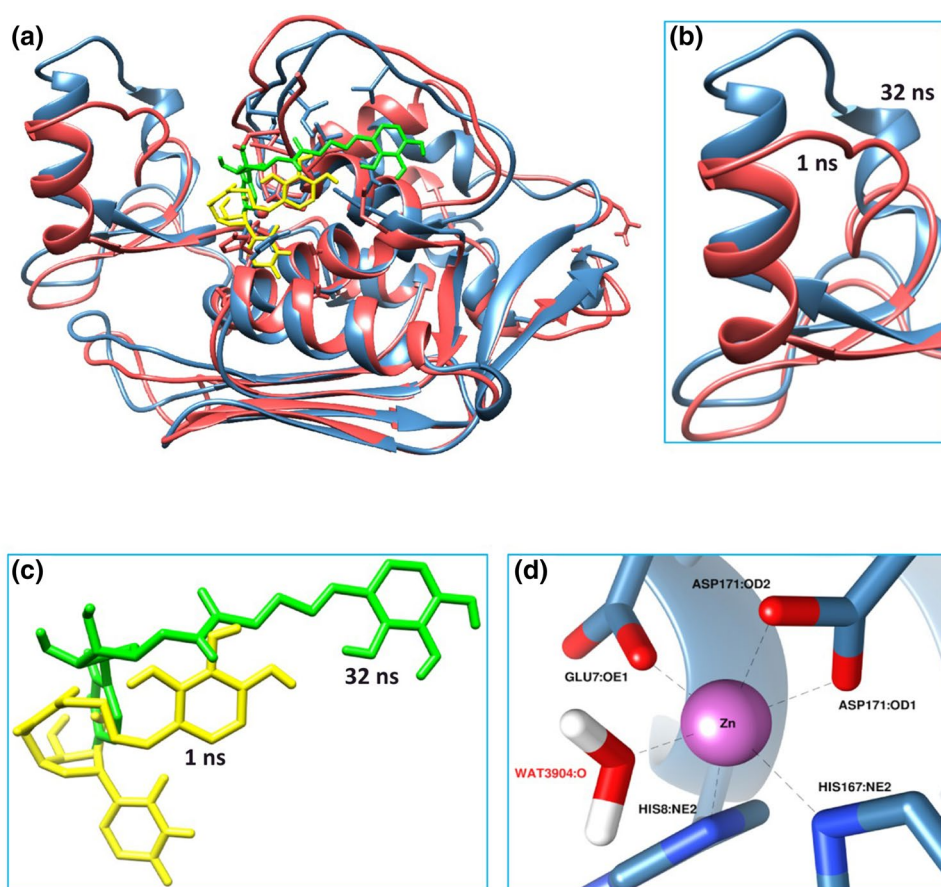
analysis of *K. pneumoniae* HS11286. LpxC, the selected Zn^{+2} dependent metalloprotein of lipopolysaccharide biosynthesis pathway claims an established pharmacological importance in antibacterial therapies targeting gram negative bacteria. The previously unexplored topology revealed by homology modelling of KpLpxC protein marked one of the key findings of the study. On the basis of 3D structure analysis, it was concurred that partial retention of N-terminal domain and complete lack of subdomain I is responsible for this unusual topology of KpLpxC. Molecular docking study concluded into the identification of uridine-based receptor antagonist (compound 106) as best potential inhibitory agent against KpLpxC that might be able to block or dampen a biological response by competitively binding and thus blocking KpLpxC activity. Insights from molecular docking and MD simulations led to the deduction that KpLpxC undergoes species-specific conformational changes which in combination with its unique topology, influences the binding pattern of uridine-based compound 106. Comparison of both docked and undocked systems revealed ligand-induced, marginally enhanced destabilization of KpLpxC subdomain II and C-terminal regions. Zn ion dynamics in both systems led to the deduction that zinc exhibits environment-dependent behavior; whereas it shows transitional behaviour in undocked KpLpxC, it prefers to maintain sixfold coordination in the presence of uridine moiety within the active site of KpLpxC. As the result of extensive dissection of molecular interactions in docked KpLpxC complex, it can be concluded that compounds mimicking the chemical structure of compound 106 can be employed to destabilize the substrate-binding site (subdomain II) of KpLpxC and induce target-specific inhibitory effects. These insightful findings pertaining to an unusual KpLpxC topology, ligand binding patterns and accompanied Zn^{+2} cofactor dynamics can enhance the existing pharmacological designs to develop more potent, specific and efficient drugs against MDR *K. pneumoniae*.

Fig. 11 **a** Conserved hexagonal coordination geometry of Zn^{+2} in docked KpLpxC complex, represented by superimposing 3D coordinates obtained at 1 ns (pink), 5 ns (purple), 10 ns (blue), 20 ns (green) and 30 ns (grey). Labelled residues (HIS 8:NE2, HIS 167:NE2, ASP 171:OD1, ASP 171:OD2, GLU 7:OE1 and WAT 3904:O) are involved in sixfold coordination with Zn. **b** Distance graph of Zn^{+2} interacting residues from Zn ion; plotted for every 150th frame of 30 ns simulations; HIS167:NE2 is at maximum distance from Zn ion
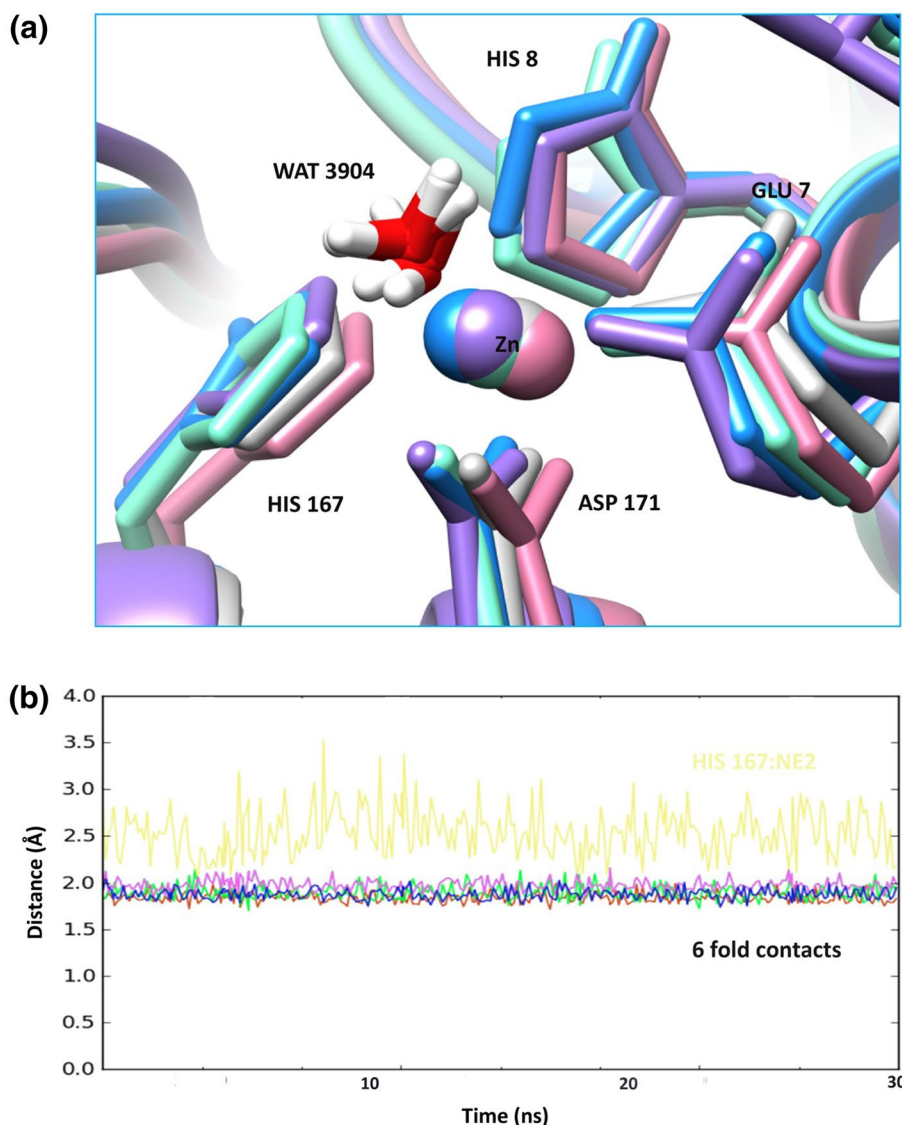
Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Paczosa MK, Mecsas J (2016) Klebsiella pneumoniae: going on the offense with a strong defense. Microbiol Mol Biol Rev 80:629–661
2. Band VI, Satola SW, Burd EM, Farley MM, Jacob JT, Weiss DS (2018) Carbapenem-resistant *Klebsiella pneumoniae* exhibiting clinically undetected colistin heteroresistance leads to treatment failure in a murine model of infection. mBio 9:02448-17
3. Farhadi T, Fakharian A, Ovchinnikov RS (2017) Virtual screening for potential inhibitors of CTX-M-15 protein of *Klebsiella pneumoniae*. Interdiscip Sci Comput Life Sci. https://doi.org/10.1007/s12539-017-0222-y
4. Nordmann P, Cuzon G, Naas T (2009) The real threat of *Klebsiella pneumoniae* carbapenemase-producing bacteria. Lancet Infect Dis 9:228–236
5. Center for Disease Control and Prevention (CDC) (2013) Antibiotic resistance threats in the United States. U.S Department of Health and Human Services
6. Mantzarlis K, Makris D, Manoulakas E, Karvouniaris M, Zakynthinos E (2013) Risk factors for the first episode of *Klebsiella pneumoniae* resistant to carbapenems infection in critically ill patients: a prospective study. BioMed Res Int 2013:850547. https://doi.org/10.1155/2013/850547
7. Tzouvelekis LS, Markogiannakis A, Psichogiou M, Tassios PT, Daikos GL (2012) Carbapenemases in *Klebsiella pneumoniae* and other Enterobacteriaceae: an evolving crisis of global dimensions. Clin Microbiol Rev 25:682–707

8. Phoebe Chen YP, Chen F (2008) Identifying targets for drug discovery using bioinformatics. Expert Opin Ther Target 12:383–389

9. Searls DB (2000) Using bioinformatics in gene and drug discovery. Drug Discov Today 5:135–143

10. Sliwoski G, Kothiwale S, Meiler J, Lowe EW (2014) Computational methods in drug discovery. Pharmacol Rev 66:334–395

11. Ahmad S, Raza S, Uddin R, Azam SS (2017) Binding mode analysis, dynamic simulation and binding free energy calculations of the MurF ligase from *Acinetobacter baumannii*. J Mol Graph Model 77:72–85

12. Butt AM, Nasrullah I, Tahir S, Tong Y (2012) Comparative genomics analysis of *Mycobacterium ulcerans* for the identification of putative essential genes and therapeutic candidates. PLoS One 7:e43080

13. Chawley P, Samal HB, Prava J, Suar M, Mahapatra RK (2014) Comparative genomics study for identification of drug and vaccine targets in Vibrio cholerae: MurA ligase as a case study. Genomics 103:83–93

14. Singh S, Singh DB, Singh A, Gautam B, Ram G, Dwivedi S, Ramteke PW (2016) An approach for identification of novel drug targets in *Streptococcus pyogenes* SF370 through pathway analysis. Interdisc Sci Comput Life Sci 8:388–394

15. Vijayalakshmi P, Nisha J, Rajalakshmi M (2014) Virtual screening of potential inhibitor against FtsZ protein from *Staphylococcus aureus*. Interdiscip Sci Comput Life Sci 6:331–339

16. Hosen MI, Tanmoy AM, Mahbuba DA, Salma U, Nazim M, Islam MT, Akhteruzzaman S (2014) Application of a subtractive genomics approach for in silico identification and characterization of novel drug targets in *Mycobacterium tuberculosis* F11. Interdiscip Sci Comput Life Sci 6:48–56

17. Uddin R, Siddiqui QN, Azam SS, Saima B, Wadood A (2018) Identification and characterization of potential druggable targets among hypothetical proteins of extensively drug resistant *Mycobacterium tuberculosis* (XDR KZN 605) through subtractive genomics approach. Eur J Pharm Sci 114:13–23

18. Wadood A, Jamal A, Riaz M, Khan A, Uddin R, Jelani M, Azam SS (2017) Subtractive genome analysis for in silico identification and characterization of novel drug targets in *Streptococcus pneumonia* strain JJA. Microb Pathog 115:194–198

19. Azam SS, Abbasi SW, Akhtar AS, Mirza ML (2014) Comparative modeling and molecular docking studies of d-alanine: d-alanine ligase: a target of antibacterial drugs. Med Chem Res 23:4108–4137

20. Azam SS, Abro A, Raza S, Saroosh A (2014) Structure and dynamics studies of sterol 24-C-methyltransferase with mechanism based inactivators for the disruption of ergosterol biosynthesis. Mol Biol Rep 41:4279–4293

21. Pérez-Castillo Y, Froeyen M, Cabrera-Pérez M, Nowé A (2011) Molecular dynamics and docking simulations as a proof of high flexibility in *E. coli* FabH and its relevance for accurate inhibitor modeling. J Comput Aided Mol Des 25:371–393

22. Sinko W, de Oliveira C, Williams S, Van Wynsberghe A, Durrant JD, Cao R, Oldfield E, McCammon JA (2011) Applying molecular dynamics simulations to identify rarely sampled ligand-bound conformational states of undecaprenyl pyrophosphate synthase, an antibacterial target. Chem Biol Drug Des 77:412–420

23. Barb AW, Zhou P (2008) Mechanism and inhibition of LpxC: an essential zinc-dependent deacetylase of bacterial lipid A synthesis. Curr Pharm Biotechnol 9:9–15

24. Yethon JA, Whitfield C (2001) Lipopolysaccharide as a target for the development of novel therapeutics in gram-negative bacteria. Curr Drug Targets Infect Disord 1:91–106

25. Mdluli KE, Witte PR, Kline T, Barb AW, Erwin AL, Mansfield BE, McClerren AL, Pirrung MC, Tumey LN, Warrener P, Raetz CR (2006) Molecular validation of LpxC as an antibacterial

26. drug target in *Pseudomonas aeruginosa*. Antimicrob Agents Chemother 50:2178–2184

26. Magrane M, UniProt C (2011) UniProt knowledgebase: a hub of integrated protein data. Database (oxford). https://doi.org/10.1093/database/bar009 **(bar 009)**

27. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A (2007) UniProtKB/Swiss Prot. Methods Mol Biol 406:89–112

28. Huang Y, Niu B, Gao Y, Fu L, Li W (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26:680–682

29. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659

30. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

31. Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35:61–65

32. Zhang R, Ou HY, Zhang CT (2004) DEG: a database of essential genes. Nucleic Acids Res 32:271–272

33. Zhang R, Lin Y (2009) DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. Nucleic Acids Res 37:D455-D458

34. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:182–185

35. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30

36. Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, Gautam B, Hassanali M (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res 36:D901-D906

37. Yu CS, Chen YC, Lu CH, Hwang JK (2006) Prediction of protein subcellular localization. Proteins Struct Funct Bioinform 64:643–651

38. Nancy YY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608–1615

39. Fiser A (2010) Template-based protein structure modeling. Methods Mol Biol 673:73–94

40. Šali, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815

41. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinform 9:40

42. Pieper U, Webb BM, Dong GQ, Schneidman-Duhovny D, Fan H, Kim SJ, Khuri N, Spill YG, Weinkam P, Hammel M, Tainer JA (2014) ModBase, a database of annotated comparative protein structure models and associated resources. Nucleic Acids Res 42:336–346

43. Bates PA, Kelley LA, MacCallum RM, Sternberg MJ (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins 45:39–46

44. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195–201

45. Lambert C, Leonard N, De Bolle X, Depiereux E (2002) ESyPred3D: prediction of proteins 3D structures. Bioinformatics 18:1250–1256

46. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291

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

48. Eisenberg D, Lüthy R, Bowie JU (1997) VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods Enzymol 277:396–404

49. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:407–410

50. Meng EC, Pettersen EF, Couch GS, Huang CC, Ferrin TE (2006) Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinform 7:339

51. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612

52. Cole KE, Gattis SG, Angell HD, Fierke CA, Christianson DW (2010) Structure of the metal-dependent deacetylase LpxC from *Yersinia enterocolitica* complexed with the potent inhibitor CHIR-090. Biochemistry 50:258–265

53. Mochalkin I, Knafels JD, Lightle S (2008) Crystal structure of LpxC from *Pseudomonas aeruginosa* complexed with the potent BB-78485 inhibitor. Protein Sci 17:450–457

54. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539

55. Schomburg I, Chang A, Placzek S, Söhngen C, Rother M, Lang M, Munaretto C, Ulas S, Stelzer M, Grote A, Scheer M (2013) BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic Acids Res 41:764–772

56. Barb AW, Leavy TM, Robins LI, Guan Z, Six DA, Zhou P, Bertozzi CR, Raetz CR (2009) Uridine-based inhibitors as new leads for antibiotics targeting *Escherichia coli* LpxC. Biochemistry 48:3068–3077

57. Clements JM, Coignard F, Johnson I, Chandler S, Palan S, Waller A, Wijkmans J, Hunter MG (2002) Antibacterial activities and characterization of novel inhibitors of LpxC. Antimicrob Agents Chemother 46:1793–1799

58. Coggins BE, McClerren AL, Jiang L, Li X, Rudolph J, Hindsgaul O, Raetz CR, Zhou P (2005) Refined solution structure of the LpxC–TU-514 complex and pK_a analysis of an active site histidine: insights into the mechanism and inhibitor design. Biochemistry 44:1114–1126

59. Gennadios HA, Whittington DA, Li X, Fierke CA, Christianson DW (2006) Mechanistic inferences from the binding of ligands to LpxC, a metal-dependent deacetylase. Biochemistry 45:7940–7948

60. Jackman JE, Raetz CR, Fierke CA (2001) Site-directed mutagenesis of the bacterial metalloamidase UDP-(3-*O*-acyl)-*N*-acetylglucosamine deacetylase (LpxC). Identification of the zinc binding site. Biochemistry 40:514–523

61. Kadam RU, Garg D, Chavan A, Roy N (2007) Evaluation of *Pseudomonas aeruginosa* deacetylase lpxc inhibitory activity of Dual PDE4–TNFα inhibitors: a multiscreening approach. J Chem Inf Model 47:1188–1195

62. Kadam RU, Garg D, Roy N (2008) Selective mapping of chemical space for *Pseudomonas aeruginosa* deacetylase LpxC inhibitory potential. Chem Biol Drug Des 71:45–56

63. Kline T, Andersen NH, Harwood EA, Bowman J, Malanda A, Endsley S, Erwin AL, Doyle M, Fong S, Harris AL, Mendelsohn B (2002) Potent, novel in vitro inhibitors of the *Pseudomonas aeruginosa* deacetylase LpxC. J Med Chem 45:3112–3129

64. Pirrung MC, Tumey LN, Raetz CR, Jackman JE, Snehalatha K, McClerren AL, Fierke CA, Gantt SL, Rusche KM (2002) Inhibition of the antibacterial target UDP-(3-*O*-acyl)-*N*-acetylglucosamine deacetylase (LpxC): isoxazoline zinc amidase inhibitors bearing diverse metal binding groups. J Med Chem 45:4359–4370

65. Pirrung MC, Tumey LN, McClerren AL, Raetz CR (2003) High-throughput catch-and-release synthesis of oxazoline hydroxamates. Structure–activity relationships in novel inhibitors of *Escherichia coli* LpxC: in vitro enzyme inhibition and antibacterial properties. J Am Chem Soc 125:1575–1586

66. Mansoor UF, Vitharana D, Reddy PA, Daubaras DL, McNicholas P, Orth P, Black T, Siddiqui MA (2011) Design and synthesis of potent Gram-negative specific LpxC inhibitors. Bioorg Med Chem Lett 21:1155–1161

67. Brown MF, Reilly U, Abramite JA, Arcari JT, Oliver R, Barham RA, Che Y, Chen JM, Collantes EM, Chung SW, Desbonnet C (2012) Potent inhibitors of LpxC for the treatment of Gram-negative infections. J Med Chem 55:914–923

68. McAllister LA, Montgomery JI, Abramite JA, Reilly U, Brown MF, Chen JM, Barham RA, Che Y, Chung SW, Menard CA, Mitton-Fry M (2012) Heterocyclic methylsulfone hydroxamic acid LpxC inhibitors as Gram-negative antibacterial agents. Bioorg Med Chem Lett 22:6832–6838

69. Montgomery JI, Brown MF, Reilly U, Price LM, Abramite JA, Arcari J, Barham R, Che Y, Chen JM, Chung SW, Collantes EM (2012) Pyridone methylsulfone hydroxamate LpxC inhibitors for the treatment of serious gram-negative infections. J Med Chem 55:1662–1670

70. Warmus JS, Quinn CL, Taylor C, Murphy ST, Johnson TA, Limberakis C, Ortwine D, Bronstein J, Pagano P, Knafels JD, Lightle S (2012) Structure based design of an in vivo active hydroxamic acid inhibitor of *P. aeruginosa* LpxC. Bioorg Med Chem Lett 22:2536–2543

71. Li Z, Wan H, Shi Y, Ouyang P (2004) Personal experience with four kinds of chemical structure drawing software: review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. J Chem Inf Comput Sci 44:1886–1890

72. Jones G, Willett P, Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 245:43–53

73. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461

74. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–38

75. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT a program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127–134

76. Accelrys Software Inc (2012) Discovery studio visualizer, Release 3.5. Accelrys Software Inc., San Diego

77. Weiner PK, Kollman PA (1981) AMBER: assisted model building with energy refinement. A general program for modeling molecules and their interactions. J Comput Chem 2:287–303

78. Yang W, Riley BT, Lei X, Porebski BT, Kass I, Buckle AM, McGowan S (2017) Generation of AMBER force field parameters for zinc centres of M1 and M17 family aminopeptidases. J Biomol Struct Dyn 28:1–10

79. Dickerson JE, Zhu A, Robertson DL, Hentges KE (2011) Defining the role of essential genes in human disease. PLoS One 6:e27368

80. Galperin MY, Koonin EV (1999) Searching for drug targets in microbial genomes. Curr Opin Biotechnol 10:571–578

81. Sakharkar KR, Sakharkar MK, Chow VT (2004) A novel genomics approach for the identification of drug targets in pathogens, with special reference to *Pseudomonas aeruginosa*. In Silico Biol 4:355–360

82. Bakheet TM, Doig AJ (2010) Properties and identification of antibiotic drug targets. BMC Bioinform 11:195

83. Fauman EB, Rai BK, Huang ES (2011) Structure-based drugga-bility assessment—identifying suitable targets for small molecule therapeutics. Curr Opin Chem Biol 15:463–468

84. Hale MR, Hill P, Lahiri S, Miller MD, Ross P, Alm R, Gao N, Kutschke A, Johnstone M, Prince B, Thresher J (2013) Exploring the UDP pocket of LpxC through amino acid analogs. Bioorg Med Chem Lett 23:2362–2367

85. Barb AW, McClerren AL, Snehelatha K, Reynolds CM, Zhou P, Raetz CR (2007) Inhibition of lipid A biosynthesis as the primary mechanism of CHIR-090 antibiotic activity in *Escherichia coli*. Biochemistry 46:3793–3802

86. Lee CJ, Liang X, Chen X, Zeng D, Joo SH, Chung HS, Barb AW, Swanson SM, Nicholas RA, Li Y, Toone EJ (2011) Species-specific and inhibitor-dependent conformations of LpxC: implications for antibiotic design. Chem Biol 18:38–47

87. Fuhrer T, Heer D, Begemann B, Zamboni N (2011) High-through-put, accurate mass metabolome profiling of cellular extracts by flow injection–time-of-flight mass spectrometry. Anal Chem 83:7074–7080

88. Patel K, Kumar A, Durani S (2007) Analysis of the structural consensus of the zinc coordination centers of metallopro-tein structures. Biochim Biophys Acta Proteins Proteomics 1774:1247–1253