BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(10)

In Silico analysis of *Escherichia coli* polyphosphate kinase (PPK) as a novel antimicrobial drug target and its high throughput virtual screening against PubChem library

Saurav Bhaskar Saha[1]* & Vivek Verma[2]

[1]Department of Computational Biology and Bioinformatics, JSBB, SHIATS, Allahabad –211007, Uttar Pradesh, India; [2]Post Doctoral Fellow, Clinical Vaccine R&D Center, Chonnam National University Hwasun Hospital, 160 Ilsim-Ri, Hwasunup, Hwasun-County, Jeonnam 519-809, South Korea; Saurav B Saha – Email: SSB-saurav.saha@shiats.edu.in; VV-bioverma@gmail.com; Tel: +91-9621675416; +82-61-379-8465; Fax: +82-61-379-8455; *Corresponding author

Received May 14, 2013; Accepted May 17, 2013; Published June 08, 2013

Abstract:
Multiple drug resistance (MDR) in bacteria is a global health challenge that needs urgent attention. The 2011 outbreak caused by *Escherichia coli* O104:H4 in Europe has exposed the inability of present antibiotic arsenal to tackle the problem of antimicrobial infections. It has further posed a tremendous burden on entire pharmaceutical industry to find novel drugs and/or drug targets. Polyphosphate kinase (PPK) in bacteria plays a crucial role in helping latter to adapt to stringent conditions of low nutritional availability thus making it a good target for antibacterials. In spite of this critical role, to best of our knowledge no in-silico work has been carried out to develop PPK as an antibiotic target. In the present study, virtual screening of PPK was carried out against all the 3D compounds with pharmacological action present in PubChem database. Our screening results were further refined by interaction maps to eliminate the false positive data respectively. From our results, compound number 5281927 (PubChem ID) has been found to have significant affinity towards affinity towards PPK active ATP-binding site indicating its therapeutic relevance.

Keywords: Autodock Vina, *Escherichia coli,* Ligplot, Multidrug resistance, Polyphosphase kinase, Virtual Screening.

Background:
The generation and spread of multiple drug resistant (MDR) bacteria is a global health challenge. Emergence of MDR pathogens have thwarted the action of antimicrobial agents and resulted in substantial increase in infection associated deaths and costs of treatment of these infections. Bacteria were thought to be simple, less complex organisms lacking cell-to-cell communications that highly evolved organism posses. But last few decades have witnessed a drastic change in our understanding of bacteria as latter have the ability for intercellular communication that acts as an adaptation mechanism **[1].** Owing to horizontal transfer of drug resistance amongst microorganisms even a commensal bacterium like *Escherichia coli*, which is considered to be a harmless intestinal inhabitant has become a highly versatile and deadly pathogen **[2].**

This very common bacterium, also called 'laboratory workhorse' **[3]** has proved to be much smarter than many other pathogenic strains. It has acquired resistance through directed mutations and system modulation and poses great challenge for novel drug discovery **[3, 4].** From late 1990s, MDR *E. coli* that produce extended-spectrum β lactamases (ESBLs), such as the CTX-M enzymes, have emerged within the community establishing itself as a cause of several fatal diseases like enteric/diarrheal disease, urinary tract infections (UTIs) and sepsis/meningitis **[3, 5].** The development of drug resistance in this pathogen against commonly used antibiotics has

necessitated the search for new antimicrobial agents and their targets.

The importance of Inorganic polyphosphate (polyP) in bacterial pathogenesis and stationary phase survival has been well documented **[6, 7, & 8].** It is a linear, highly conserved polymer of inorganic phosphate (P_i) linked by phosphoanhydride bonds. PolyP is involved in regulation of rpoS and recA expression at the transcriptional level and affects the expression of many stress-inducible and stationary-phase inducible genes **[9].** Polyphosphate kinase (PPK) is the key enzyme responsible for catalyzing the reversible conversion of the γ-phosphate of ATP to polyP. Reports have shown its involvement in motility, quorum sensing and biofilm formation **[10].** Studies with the ppk-knockout bacteria and PPK specific inhibitors indicate that PPK is an attractive antimicrobial drug target, and that the mechanism of action of PPK inhibitors may be distinct from that of existing antibiotics **[9].**

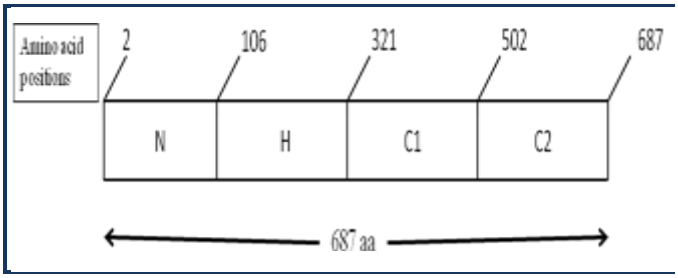

Figure 1: Schematic representation of *E. coli* Polyphosphate kinase domains. The abbreviation used: N, H, C1 and C2 for Amino terminus domain, Head domain and two Carboxy terminus C1 and C2 domain.

Protein sequence of a PPK monomer can be divided into four domains i.e. the N-terminal domain, the head domain (H-domain), the C-terminal domain C1, and domain C2 **(Figure 1).** In a PPK monomer, the N – terminus domain is highly conserved and is comprised of three long antiparallel α-helices. It lies on the upper surface of the C-terminal domains and provides the upper binding interface for adenine ring of ATP. The H-domain, which shows the lowest degree of homology among these domains has a core α /β/ α fold in the middle and forms the protruding 'head' of the PPK monomer. The H-domain interacts with the C1 – domain of the other PPK monomer in the asymmetric unit and is involved in dimerization. Dimerization of PPK is crucial for the synthesis of polyP where H1 domain of one PPK molecule interacts with C1-domain of other PPK molecule as shown in **(Figure 2).** Both the C1 and C2 domains consist of a seven-stranded mixed β – sheet flanked by five α –helices. However, the structural topology and relative orientations of the helices to the β – sheet in these two domains are different. The C1 and C2 domains are highly conserved in the PPK family and some of the residues that are crucial for enzyme catalytic activity are located in these domains **[9].**

But the most important structural feature of PPK is the tunnel that penetrates the center of each PPK monomer. This tunnel is mostly formed by the intersection of the N, C1 and C2 domains. Surprisingly, most of the conserved residues lie in the PPK indicating its functional significance. One end of the tunnel accommodates only one ATP molecule whereas the other end contains highly conserved positively charged residues. These positively charged residues interact with polyP chain during polyP elongation. When ATP binds to the active site, the dimension of the rest of the tunnel can accommodate only one linear polyP chain of about 6–10 residues, indicating that the polyP chain is unlikely to elongate on the autophosphorylated Histidine at position 435 **[9].** It is plausible that ATP enters from one side of the tunnel and the synthesized polyP chain exits from the other side, with the newly synthesized end of the polyP chain remaining at the active site to accept the γ-phosphate group from the next ATP bound to the tunnel. Moreover, studies have also shown that mutations in Arginine at position 564 and 375 abolish PPK enzymatic activity and both of them directly interact with ATP phosphate group **[9].** In this study we have carried out high throughput in-silico screening of a library of 6225 compounds with pharmacological activity to find out potent drug compounds that can act as anti-microbial agents against *E. coli*.

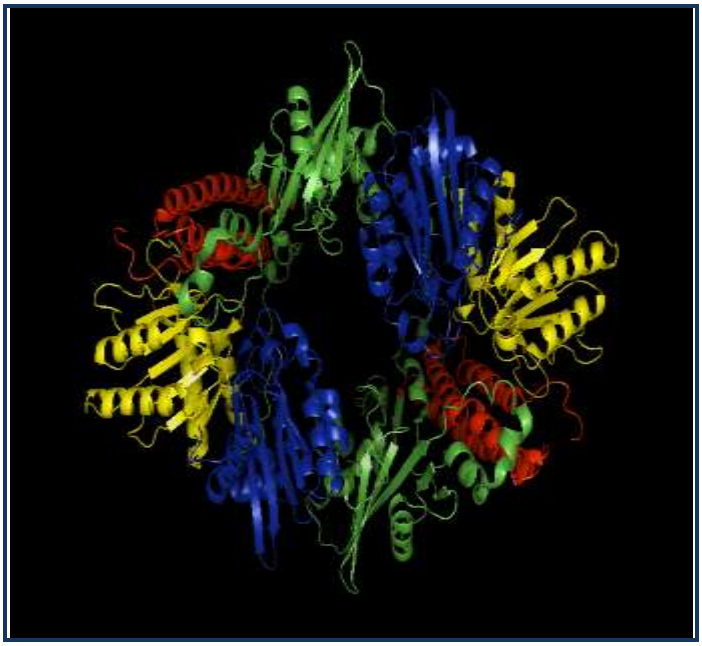

Figure 2: Structure of *E. coli* PPK Dimer (PDB ID: 1XDO) where N terminal domains of both subunits (2-106 residue) are shown in red, the head domains (107-321 residue) are shown in green, the C-terminus domains C1(322-502 residue) are shown in blue and domains C2 (503-687 residue) are shown in yellow.

Methodology:

To identify potential inhibitors of *E. coli* PPK, the crystal structure of PPK was retrieved from Protein Data Bank **[11]** (PDB ID code 1XDO). The compound database used in our virtual screening was PubChem **[12].** PubChem is an open repository for experimental data identifying the biological activities of small molecules. PubChem is a part of the Molecular Libraries and Imaging (MLI) component of the National Institutes of Health (NIH) Roadmap for Medical Research initiative. In this study, the PPK tunnel was the target of interest. Autodock vina was used for docking with a large search space encompassing entire PPK tunnel. Autodock vina is a new open-source program for drug discovery, molecular docking and virtual screening, offering multi-core capability,

high performance and enhanced accuracy **[13].** As Autodock vina required files in pdbqt format, our PDB file and library (SDF format) were converted into pdbqt using openbabel. According to their binding affinity, top 50 compounds were extracted and subjected to ligand interaction analysis using Ligplot and then the interaction maps were checked with the annotation provided on PubChem website for each top 50 compounds. To find optimal compounds, we set our priority to those which interacted with the most conserved domain in the tunnel (Arginine 375, Arginine 564 and Histidine 435) **[9].** The docking and analysis process were automated through in-house perl and shell scripts. The work flow is outlined in **(Figure 3).**

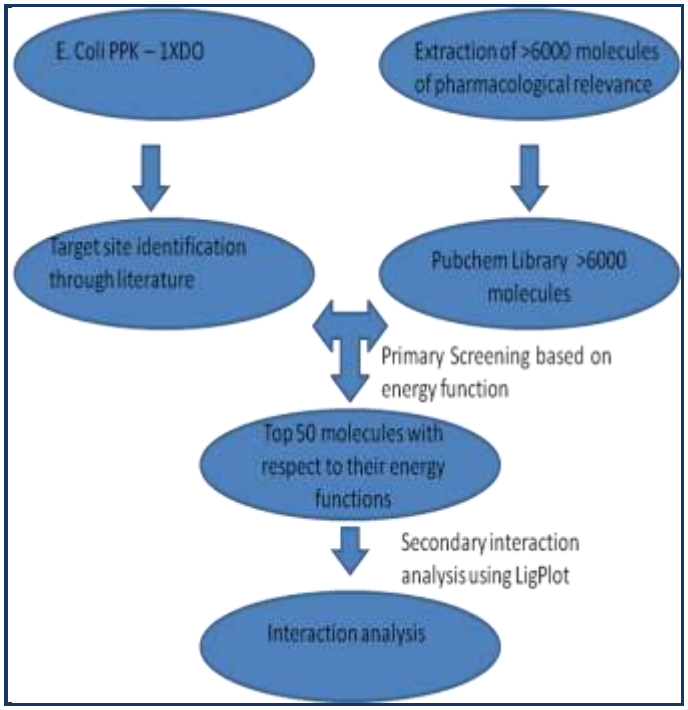

Figure 3: Virtual screening workflow used in the present study.

Results & Discussion:

Our understanding of *E. coli* pathogenesis was primarily based on studies of Laboratory K-12 strains grown in standard media which many a times hindered our understanding about *E. coli* pathogenic strains. PPK being one virulence marker we carried out a BLAST search of *E. coli* PPK protein sequence (Accession number for *E. coli* K-12 strain: BAA16389). Surprisingly a very high degree of similarity with pathogenic strains like *E. coli* O157:H7 EDL933 (Identities = 688/688 (100%), Positives = 688/688 (100%), Gaps = 0/688 (0%)) and *E. coli* O26:H11 str. 11368 (Identities = 687/688 (99%), Positives = 687/688 (99%), Gaps = 0/688 (0%)) was reported indicating PPK protein's canonical role in pathogenesis also.

After virtual screening a ranking list was generated as per energy affinity for all the compounds. Surprisingly, when we checked annotations provided by PubChem, none of the top 5 compounds in our ranking list reported any anti microbial action. So, we carried out further analysis to see interaction maps for all the top 50 compounds. From our analysis we found three compounds; PubChem Identification number (UID): 5281927, 5351232, 11758093 viz., Dynemicin A, Muconomycin A and Etoposide respectively, which demonstrated significant interactions with PPK conserved key domains residues (Histidine 435 as autophosphorylation site and Arginine 564 and 375 are involved in its enzymatic activity) **[9].** The interaction maps are shown in **(Figure 4)**. **Dynemicin A** (CAS# 124412-57-3) is an antineoplastic agent that is hydrogen bonded with PPK at Arginine 375 and 564, and Asparagines 617, 45 and 539. **Muconomycin A** (CAS# 3148-09-2) was found to form hydrogen bonds with Arginine 375, 564, and 53 respectively. The third compound **Etoposide** (CAS# 33419-42-0) is an anticancer drug forming hydrogen bonds with Arginine 53, 375 and 564, Tryptophan 14, and Asparagines 459 and 461 of PPK in the most conserved tunnel region. Further information about these compounds is summarized in **Table 1 (see the supplementary material).**

PPK in bacteria is associated with pathogenicity, motility and drug resistance through quorum sensing, regulation of error-prone replication and biofilm formations **[6, 7, 8, 14, & 15].** It is highly conserved in prokaryotes and nearly absent in eukaryotes which makes it a potential drug target in antibacterial therapeutics. In this study, we used computational screening to identify compound that inhibit PPK. Among top 50 compounds obtained after docking of chemical library, we further screened compounds on the basis of their interactions with PPK conserved domains. We found three potential compounds viz., Dynemicin A, Muconomycin A and Etoposide. It is evident from the analysis of the docked complex that the ligands are located in the center of the active site, and is stabilized by hydrogen bonding interactions. Among these Dynemicin A is a member of the enediyne family of antibiotics that was discovered in 1989 and was isolated as a natural product from a fermentation broth of *Micromonospora chersina* as a violet-colored solid **[16].** Muconomycin A is an antibiotic which was found during a search for new fungicides from natural sources. It is a crystalline compound isolated from *Myrothecium verrucaria* **[17].** While etoposide is a semisynthetic derivative of podophyllotoxin isolated from *Podophyllum* species **[12].**

PPK is a protein that is involved in bacterial stationary phase survival **[6, 7, 8, 14 & 15].** As shown by Zhu Y. et al **[9],** the active site of PPK is located in a tunnel containing a unique ATP-binding pocket essential for polyP synthesis. In our present study, our virtual screening and interaction map data demonstrated promising results. The shortlisted compounds interacted intimately with the PPK molecule in the tunnel region forming hydrogen bonds and thus blocking the site of ATP attachment. Among these compounds, Muconomycin A has been found highly toxic and capable of inducing inflamation **[18]** and Etoposide has been reported to produce significant genotoxic effect to eukaryotes **[19].** However, Dynemicin A remains a promising compound due to its low cytotoxicity **[20]** and high affinity to PPK active site. Thus we envisage that Dynemicin A can be a good candidate for development as antibacterial drug, especially against stationary phase *E. coli* that inherently are more resistant towards available antibiotics. Preliminary experimentation shows promising results and we plan to test the compound in in-vitro conditions in wet lab experiments to understand the pharmacodynamics and pharmacokinetics of the drug.

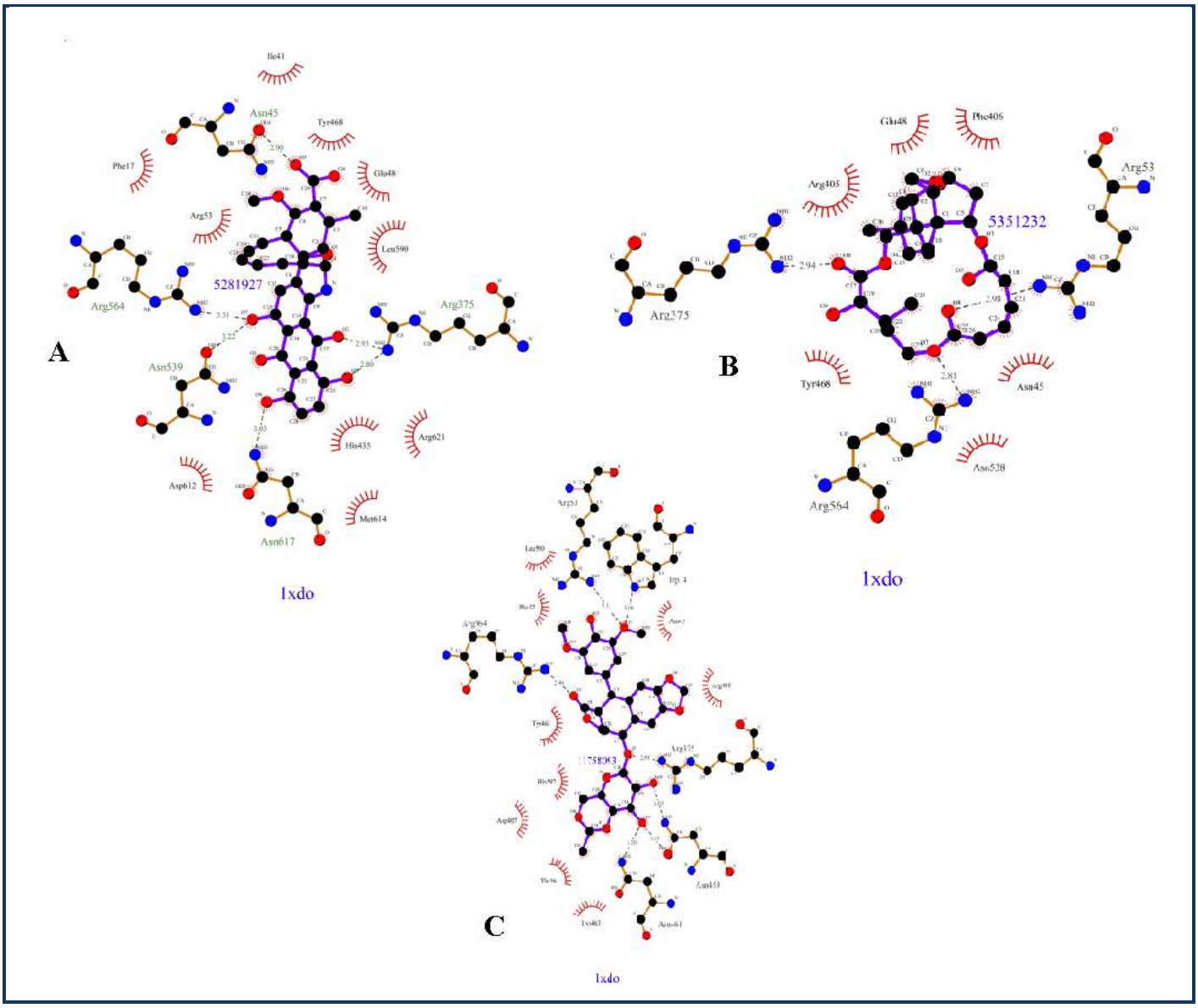

Figure 4: Interaction maps of shortlisted compouds from PubChaem library; A) Dynemicin A; B) Muconomycin A; and C) Etoposide.

Acknowledgement:
The authors acknowledge the support of Department of Computational Biology and Bioinformatics, Sam Higginbottom Institute of Agricultural, Technology and Sciences, Allahabad, U.P., India.

Conflict of interests:
None to declare.

References:

[1] Njoroge J & Sperandio V, *EMBO Mol Med.* 2009 **1**: 201 [PMID: 20049722]
[2] Scheutz F *et al. Euro Surveill.* 2011 **16**: pii: 19889. [PMID: 21699770]
[3] Kaper JB *et al. Nat Rev Microbiol*. 2004 **2:** 123 [PMID: 15040260]
[4] Levy SB, *Pediatr Infect Dis J.* 2000 S120 [PMID: 11052402]
[5] Darwish RM & Aburjai TA, *BMC Complement Altern Med.* 2010 **10**:9 [PMID: 20187978]
[6] Kornberg A *et al. Annu Rev Biochem*. 1999 **68:** 89 [PMID: 10872445]
[7] Brown MR & Kornberg A, *Proc Natl Acad Sci U S A.* 2004 **101:** 16085 [PMID: 15520374]
[8] Rao NN & Kornberg A, *J Bacteriol*. 1996 **178**: 1394 [PMID: 8631717]
[9] Zhu Y *et al. EMBO Rep.* 2005 **6**: 681 [PMID: 15947782]
[10] Rashid MF *et al. J Bacteriol*. 2000 **182:** 225 [PMID: 10613886]
[11] Berman HM *et al. Nucleic Acids Res* 2000 **28:** 235 [PMID: 10592235]
[12] http://pubchem.ncbi.nlm.nih.gov
[13] Trott O & Olson AJ, *J Comput Chem.* 2010 **31:** 455 [PMID: 19499576]
[14] Kwang-Seo K *et al. Proc Natl Acad Sci U S A*. 2002 **99**: 7675 [PMID: 12032342]

[15] Stumpf JD & Foster PL, *Mol Microbiol*. 2005 **57**: 751 [PMID: 16045619]
[16] Tuttle T *et al. J Am Chem Soc.* 2005 **127:** 9469 [PMID: 15984874]
[17] Vittimberga BM *Journal of Organic Chemistry*. 1963 **28**: 1786
[18] Guarino AM *et al. Biotechnology and Bioengineering*. 1968 **10**: 457
[19] Gupta RS *et al. Mutagenesis*. 1987 **2**: 179 [PMID: 3325741]
[20] Konishi M *et al. J Antibiot (Tokyo).* 1991 **44:**1300 [PMID: 1778782]

Edited by P Kangueane
Citation: **Saha & Verma,** Bioinformation 9(10): 518-523 (2013)

BIOINFORMATION

open access

Supplementary material:

Table 1: Common name, Source, Virtual Screening ranking according to their affinity, Autodock vina affinity score, important Hydrogen bonds with PubChem annotation of the shortlisted compounds.

S. No.	PubChem Identification number	Name of the compound	Source	Virtual screening Ranking	Autodock vina affinity result	Important H-Bonding	Action
1	11758093	Etoposide	*Podophyllum* species	29	-10.5	R564, R375	Antineoplastic Agents, Phytogenic
2	5351232	Muconomycin A	*Myrothecium verrucaria*	45	-10.3	R564, R375	Anti-Bacterial Agents, Antibiotics, Antineoplastic and Antineoplastic Agents, Phytogenic
3	5281927	Dynemicin A	*Micromonospora chersina*	44	-10.3	R564, R375	Anti-Bacterial Agents and Antibiotics, Antineoplastic