https://doi.org/10.46344/JBINO.2024.v13i06.05

IN-SILICO IDENTIFICATION OF POTENTIAL DRUG TARGETS IN CLOSTRIDIUM DIFFICILE R20291: MOLECULAR DOCKING OF A CANDIDATE ENZYME SORTASE

Muhammad Imran Qadir*, FarihaYameen,Ayesha Manzoor

Institute of Molecular biology and Biotechnology, BZU, Multan, Pakistan

*E.mail:mrimranqadir@hotmail.com

ABSTRACT

Severe diseases include pseudomembranous colitis, mild sporadic diarrhea, and gastrointestinal ailments are primarily caused by Clostridium difficile. Many medications were created, however they were all marginally resistant to the novel Clostridium difficile strains. We must thus seek out more innovative medications in order to protect ourselves from these natural foes of toxicity. Therefore, the goal of the study is to identify potential therapeutic targets in the pathogen using the four-step in silico approach. These include searching for similarities between the pathogen and the host, using gene databases, and conducting a metabolic functional research with several genomes, the Genome database, and the Genetic Encyclopedia. Based on the aforementioned research, 19 such targets that are not homologous to host proteins and they are associated with four distinct pathogen pathways: the D-alanine metabolism pathway, the phosphotransferase system, the two component system, and the peptidoglycan biosynthesis pathway. Peptides in the list of potential target proteins are the most effective donors in the aforementioned pathways. The peptidoglycan of the gram-positive bacteria's cell wall and some proteins are covalently bonded by the sortase enzyme. Sortase is essential to the pathophysiology of Clostridium difficile, and their inhibitors have been explored as possible new treatments. Out of the inhibitors found in the small molecule database, eight have demonstrated a docking score ranging from -7.9 to -10. Pro6K has the strongest binding strength of all the compounds examined; it consumes only 3 kcal/mol and exhibits a higher binding affinity with the target than the naturally occurring sortase substrate and recognized inhibitors. New drug development will benefit from the chemicals that have the potential to treat disorders linked to C. difficile.

Keywords: QuercitrinClostridium difficile. Drug target: Sortase. Molecular docking. Compound:

Introduction:
Mammal guts are infected with the spore-forming, Gram-positive, anaerobic bacteria C. difficile. Infections known as CDIs are typically brought on by drugs that upset the delicate balance of the gut's microbial flora. This helps the bacteria develop to its full potential, which in turn causes a range of gastrointestinal disorders in people, collectively referred to as C. difficile-associated illnesses (also known as CDAD). CDI is more frequently observed in healthcare settings, including nursing homes, hospitals, and similar establishments. The symptoms of the CDAD-carrying patient range widely, from moderate diarrhea to potentially fatal pseudomembranous colitis. In addition to the problems with mobility and death, CDAD places a significant financial strain on the global health system. The fact that there has been a notable rise in CDAD cases over the last ten years and that the advent of the highly virulent strain C has been noted is concerning. This is the highly contagious and rapidly spreading PCR ribotype 027, which is extremely hazardous and mostly seen in North America and Europe. Numerous studies have demonstrated that, in comparison to non-epidemic strains (ribotype 027 strain CDI 196 and ribotype 012 strain 630), this more virulent strain is the source of our more severe health issues, increased death rate, and diarrhea. Many medications, including vacnomycin and metronidazole, are used to treat CDAD. Whereas vacnomycin is used for more serious infections, metronidazole is often used for moderate infections. Sadly, strains of Clostridium difficile have demonstrated their resistance to the so-called antibiotics. Therefore, we must focus more of our efforts on developing novel antibiotics to combat these strains. Finding new therapeutic targets is by far the most important first step in the design of new medications. By the way, the current work requests that new potential therapeutic targets be investigated using an in silico technique in *C. difficile* R20291, a hypervirulentribotype 027.

Researchers frequently offer several silico methods to find possible pharmacological targets in diverse harmful microorganisms. As a result, these techniques are those that identify the genes responsible for pathogen virulence, certain important genes, factors influencing host-pathogen interaction, and distinctive metabolic pathways specific to each pathogen, among other things. These methods have proven effective in discovering novel therapeutic targets for Helicobacter pylori and Mycobactrium TB, among other infections. Selecting which ones to focus on should be determined by unique standards such as specificity and essentiality. A pathogen's ability to grow and survive depends on its ability to select a target, which it uses exclusively to evade the host proteins' reaction. Consequently, the study analyzed the entire genome sequences of C. difficile R20291 and humans in order to identify the genes that the virus preferentially carries. It has been demonstrated that by looking through databases of metabolic pathway information and essential genes (DEGs), it is possible to determine the essentiality and potential metabolic functional relationship of these proteins. Thus, the structure of the

putative target Sortase(3D) may then be constructed by the study while possible inhibitors can be found through molecular docking using the small molecule databases.

MATERIAL AND METHODS

The NCBI FTP site provided all of the C. difficile R20291 protein sequences, which have also been somewhat humanized. 3,507 protein sequences in the genome of C. difficile are just ready to provide further details about this cunning bacterium. This study's molecular docking of proteins and ligands was made extremely understandable and human-touched thanks to the assistance of Software Discovery studio. Because protein sortase was custom-made for Discovery Studio, it had a more intimate feel. In this case, the ligand Quercitin was tagged in order to enable detection of its distinct location and subsequent modification. After then, work was done on enzyme modification in the Auto Dock software, and we do the same with the ligands here. In layman's words, the Vina software was utilized to determine whether or not Protein would interact appropriately with the medication. It was also used to assess Protein's compatibility with its drug. Ultimately, decisions were taken to use the protein and medication results that were compatible. The sophisticated PyMol program was used to display the protein and ligand structures in an extremely clear and comprehensive manner. Ultimately, the protein with its ligand properties—which were determined with the use of MedChem software—came to be the main emphasis. Such as "Hey, I've downloaded the PDB file of the protein from the RCSB and the SDF file of the ligand from PubChem."

METHOD

Humans with that type of enzyme will be thoroughly investigated because the scientists used homology modeling to create a three-dimensional structure for the enzyme sortase of Clostridium difficile R20291. The following is the humanized version of the given sentence: Sortase's amino acid sequence was discovered on NCBI, and a PDB file was acquired in order to locate an appropriate template. Because laptop enzyme does not function on cells, sortase enzyme was chosen by humans as a homology model for Clostridium difficile. The protein from Discovery Studio 4. 1 was opened from its downloaded file in PDB format, allowing a team of scientists to study it more thoroughly as they would any other living thing. At this stage, the ligand that was previously there was cut to make the protein structure simpler, and the protein structure was then saved as a pdb file. Then, a new aperture was made in this location to identify the various ligand structures and the components that were linked to them. Now that different elements and hydrogen had been added to the ligand, it was possible to preserve the valencies of every atom on the ligand and save the new ligand as a PDB file. As a result, the RMSD and structural similarities may be computed, and Discovery Studio makes it easier to identify structural similarities.

PREPARATION OF TARGET PROTEIN STRUCTURE

Molecular modeling is similar to narrating the universe's narrative again, but with

AutoDock tools. First, we read the molecule and open the protein in the pdb file form. Here, we can choose any atom to bond with the protein. Next, we add hydrogen and remove the water molecules. Ultimately, the protein undergoes modification, after which it is squared and stored as pdbqt files.

PREDICTION OF LIGAND BINDING SITES

To make my task much easier, I downloaded the SDF file of ligand from PubChem. Different binding sites that were present on the ligand were substituted for different components at different places in this ligand. Therefore, we can develop novel medications to treat CDI more effectively by first altering this ligand. I believe that the ligand completes its activity at Discovery Studio. Now that we've completed this task, the pdb stored ligand is opened in AutoDock to adjust its amount of torsions and saved as a pdbqt file.

GRID RECEPTOR GENERATION

Prior to docks, the Receptor Grid file created using Glide's "Receptor Grid Generation" option. Receptor atom van der Walls radii were scaled by one, akin to how scientists streamlining their work for greater comprehensibility. The partial genomic charge of 0 and the methodical Harvard University research team might lead to a significant discovery. This year, the lengthy, non-polar portions of the protein were given a +25. We have determined the grid box size for Glide, a vehicle that needs a pre-assigned grid box for docking trials, to be 10*10*10A with coordinates of X=19. This eliminates the need for more studies on this front. Z=44. 379, Y=6. 999, and Z=254. The ligand was able to freely spin inside the binding pocket and undergo several conformational changes because the Grid box was made possible by the force field located in the core of the active site region.

RESULT

Affinity of different modes in Kcal/mol are given in table 1.Interaction of protein with ligand 1 is given in figure 1 &Interaction of protein with ligand 2 is given in figure 2. Interaction of protein with ligand 3 is given in figure 3.

Table 1: Affinity of different modes in Kcal/mol

Mode	Ligand 1	Ligand 2	Original Q	Ligand 3	Original Rutin
1	-7.8	-7.8	-7.7	-8.0	-7.4
2	-7.5	-6.9	-6.9	-7.5	-7.4
3	-7.4	-6.4	-6.8	-7.3	-7.3
4	-7.4	-6.2	-6.8	-7.1	-7.3
5	-7.2	-6.1	-6.7	-7.0	-6.9
6	-7.1	-6.1	-6.6	-7.0	-6.9
7	-6.9	-6.0	-6.5	-7.0	-6.8

8	-6.9	-5.9	-6.5	-6.8	-6.7
9	-6.5	-5.8	-6.4	-6.6	-6.6

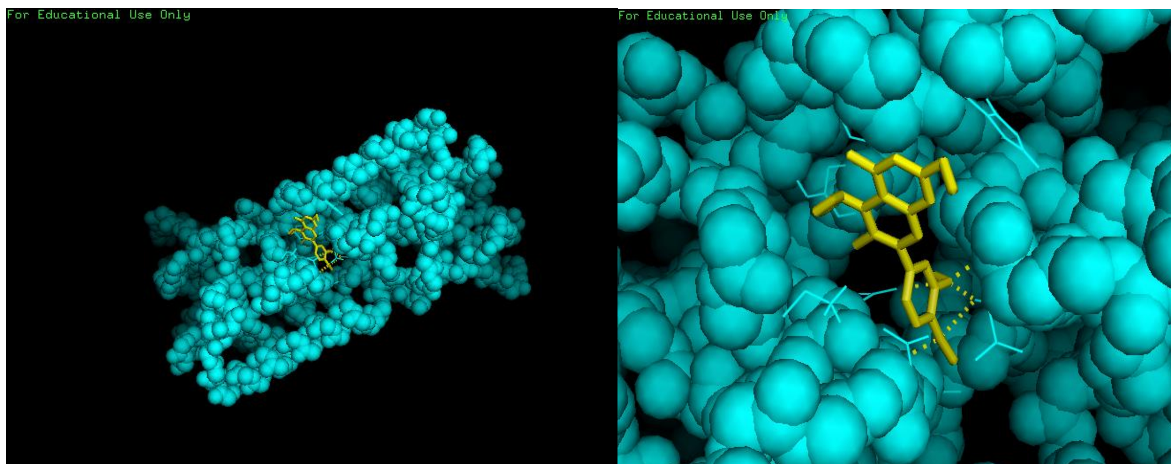


Figure 1: Interaction of protein with ligand 1

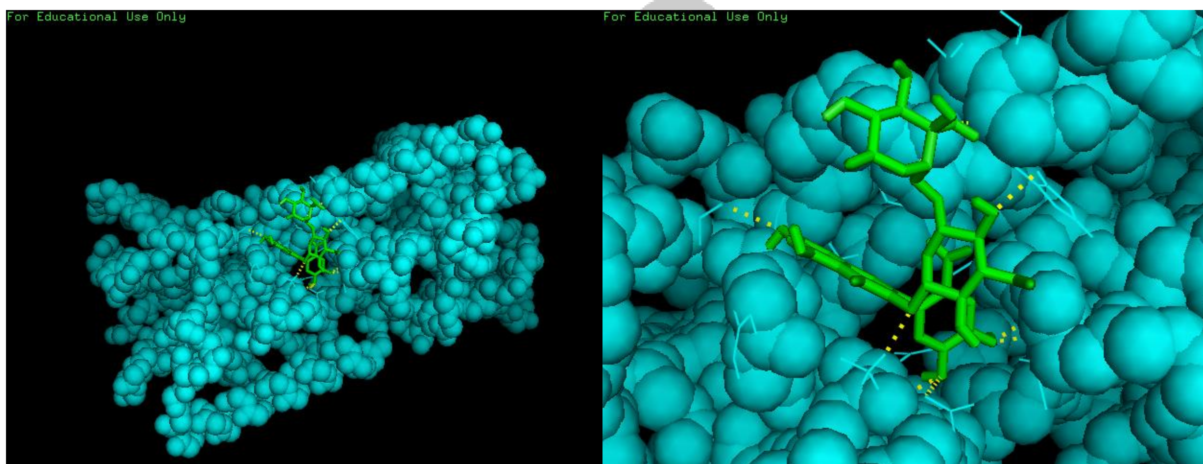
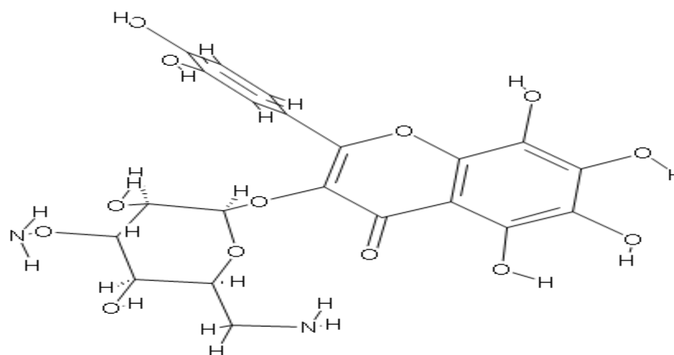


Figure 2: Interaction of protein with ligand 2

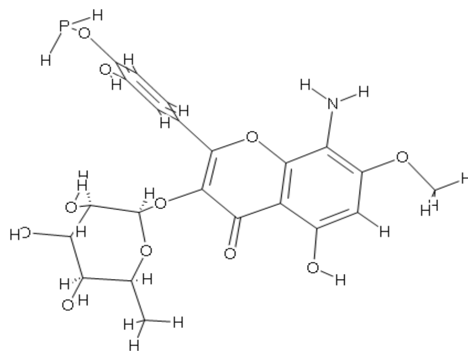
MODIFIED QUERCITRIN 1

In the discovery studio, the ligand underwent its initial modifications, personifying it as a character that required refinement. The modification sites of the ligands H40 and H41 were taken, and N was added to both of them. As a result, we arrived at N40 and N41, respectively. The updated results are therefore displayed in the table.



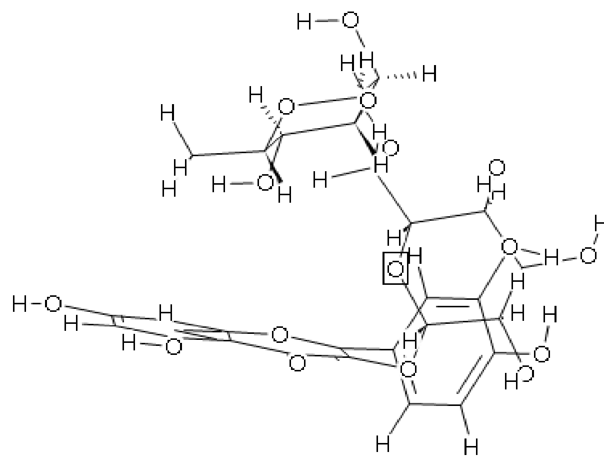
MODIFIED QUERCITRIN 2

The ligand is, after all, the cornerstone of the entire discovery studio. The ligands were changed at the H50 sites to become C50, N44, and P52, respectively, and C and H44 at N and H52 at P. A table containing the updated results is then displayed.



Modified Rutin

The ligand that was altered was the one that was initially encountered in the discovery studio. C60, N67, and O48 were obtained, respectively, as a result of the ligand site modifications at H60, where C and H67 were added, and O and H48 were changed to their corresponding numbers. As a result, the table shows the updated findings.



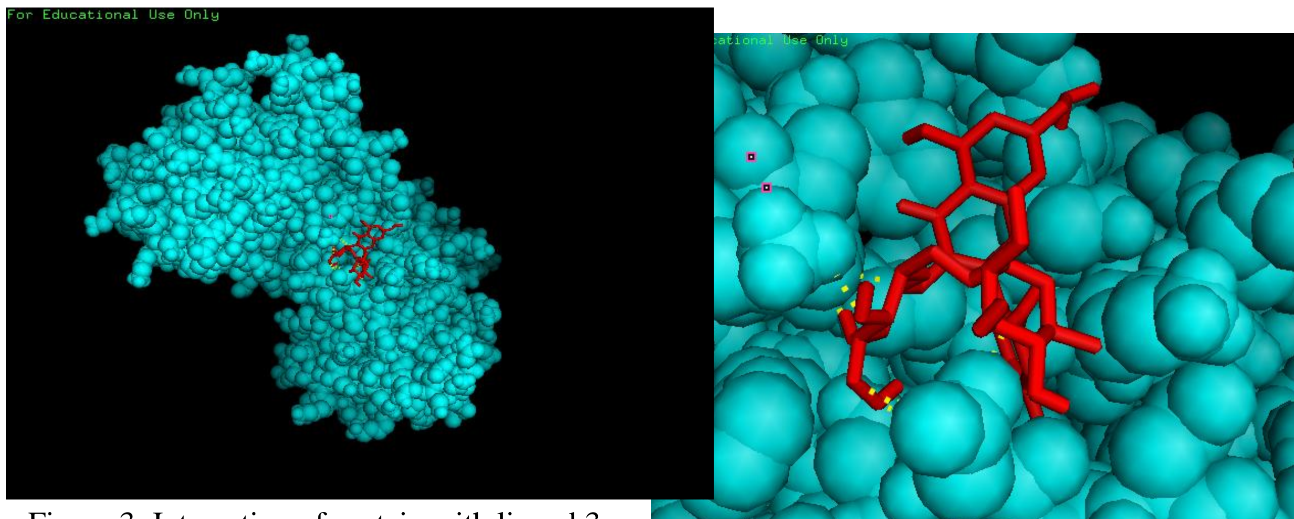


Figure 3: Interaction of protein with ligand 3

ADMET PROPERTIES

Drug discovery uses a set of test categories—absorption, distribution, metabolism, excretion, and toxicity—that are employed concurrently to provide background information on how a pharmaceutical drug interacts with the body as a whole.

Table 2: ADMET properties

Structure Name	Ligand 1	Ligand 2	Original Querictrin	Ligand 3	Rutin Original
MlogP	-3.379	-1.855	1.357	-3.998	-3.223
S+log	-10.949	-0.138	0.035	-1.396	-0.870
S+logD	-2.078	-0.142	-0.066	-1.431	-0.918
Rule of 5	3.000	3.000	2.000	3.000	3.000
Rule of 5 code	Hb,Mw,No	Hb,Mw,No	Hb,Mw,No	Hb,Mw,No	Hb,Mw,No
Mwt	510.414	509.409	448.386	655.570	610.529
M-No	15.000	12.000	11.000	18.000	16.000
T-PSA	271.780	194.300	190.280	304.680	269.430
HBDH	12.000	7.000	7.000	12.000	10.000

Table 3: Drug Scoring

Ligand Name	rmsd	Rank (score)	Score
QUERCITRIN 1 (Modified)	None	1	-136
QUERCITRIN 2(Modified)	None	1	-140
Rutin (Modified)	None	1	-117

Discussion

Figuring out the docking idea The process of identifying qualities related to protein-ligand interactions, such as hydrophobicity, binding energy, electron distribution, and donor acceptor characteristics of hydrogen bonds, is that which is being described here. An essential way to circumvent the costly and time-consuming procedure of conducting screening tests is using computational approaches to foundation-based drug development.

Since sortases are necessary for bacterial survival and pathogenicity, they are thought to be a possible target for novel antibiotics. Consequently, at least one corporation has expressed minimal commercial interest in this. The kinetics of colonization are illustrated with an analysis of the growth rate of the UK colony. Mimivirus causative agent that can be safely treated in mammals such as hamsters. Finally, in the R20291 model, based on the time point at which the clinical end point maximum was reached (20 hours), death and morbidity were determined in 100% of the animals. The image in humans is comparable to that in adults who have a high virulence, which is linked to the illness's fatality. The end point's

high difference may suggest that clindamycin is being utilized in the hamster model, as shown by R20291. More toxins and larger bacterial counts were among the diseases discovered 36 hours after infection and later, which were linked to the increasing damage this infection caused. The expression of the signaling molecule "difficile" is likely what controls the C. quorum-sensing pathway of E. coli. C. Toxin binary CDT is Clec1900, which is further blended to produce the toxin. In a rabbit ileal loop model, CDT likewise causes oedema to occur; however, neither the target protein's 3D atomic structure nor that of a hamster were present in either sample. Subsequently, I used Accelrys Discovery Studio Runtime 4.0 as instructed in my user manual. 1 Next, in order to screen through auto dock vina, the ligand and protein records need to be converted to pdbqt files. For this reason, adjustments are made to both the ligand and the protein using a program called auto dock. The module's Auto doc Vina docking platform is operated by the possibility of docketing.

Conclusion

One key idea in the docking process is the binding relationship between the ligand and the protein. The last few decades have seen a significant rise in the frequency and severity of C. difficile diarrheal illness, which has prompted increased efforts worldwide to create novel anti-C. difficile medications. these days, a risk to getting access to effective medications for diarrhea caused by C. *difficile* Finding the most suitable and ideal targets for treatments is the primary core aim of the drug discovery procedure. Therefore, they have completed the process of joining quercitrin and rutin by adjusting the Sortases protein in order to achieve the goal of shielding humans and other species from the state of being infected by C. difficile. This indicates that the protein complies well with Clostridium difficile species, as evidenced by their superior performance over the original drug. For this reason, I have included the ADMET features of these ligands below, which likewise demonstrated effective outcomes.

REFERANCES

1) Allsop AE (1998) Bacterial genome sequencing and drug discovery. CurrOpinBiotechnol 9(6):637–642

2)Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W,Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402

3)Anishetty S, Pulimi M, Pennathur G (2005) Potential drug targets in Mycobacterium tuberculosis through metabolic pathway analysis.ComputBiolChem 29(5):368–378 Barreteau H, Kovac A, Boniface A, Sova M, Gobec S, Blanot D (2008) Cytoplasmic steps of peptidoglycan biosynthesis. FEMS Microbiol Rev 32(2):168–207

4)Bartlett JG (2008) Historical perspectives on studies of Clostridium difficile and C. difficile infection. Clin Infect Dis 46(Suppl 1):S4–S11

5)Chhabra G, Sharma P, Anant A, Deshmukh S, Kaushik H, Gopal K,Srivastava N, Sharma N, Garg LC (2010) Identification and modeling of a drug target for Clostridium perfringens SM101. Bioinformation 4(7):278–289

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

7)Dutta A, Singh SK, Ghosh P, Mukherjee R, Mitter S, Bandyopadhyay D (2006) In silico identification of potential therapeutic targets in the human pathogen Helicobacter pylori. In SilicoBiol 6(1–2):43–47

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

9)El Zoeiby A, Sanschagrin F, Levesque RC (2003) Structure and function of the Mur enzymes: development of novel inhibitors. MolMicrobiol 47(1):1–12.

10)Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, Pieper U, Stuart AC, Marti-Renom MA, Madhusudhan MS, Yerkovich B, SaliA(2003) Tools for comparative protein structure modeling and analysis. Nucleic Acids Res 31(13):3375–3380.

11)Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47(7):1739–1749.

12)Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobicenclosure for protein–ligand complexes. J Med Chem 49(21):6177–6196.

13)Galperin MY, Koonin EV (1999) Searching for drug targets inmicrobial genomes. CurrOpinBiotechnol 10(6):571–578.

14)Gerding DN (2005) Metronidazole for Clostridium difficile-associated disease: is it okay for Mom? Clin Infect Dis 40(11):1598–1600.

15)Gerding DN, Muto CA, Owens RC Jr (2008) Treatment of Clostridium difficile infection. Clin Infect Dis 46(Suppl 1): S32–S42

jbino