Computational Biology and Chemistry xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/compbiolchem

Research article

A combined systems and structural modeling approach repositions antibiotics for *Mycoplasma genitalium*

Denis Kazakiewicz[a,b,c,1], Jonathan R. Karr[d,e,1,**], Karol M. Langner[f,2], Dariusz Plewczynski[a,g,h,*]

[a] Center for Innovative Research, Medical University of Białystok, Białystok 15-089, Poland
[b] Center for Statistics, Universiteit Hasselt, Hasselt BE3500, Belgium
[c] Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt BE3500, Belgium
[d] Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[e] Graduate Program in Biophysics, Stanford University, Stanford, CA 94305, USA
[f] Department of Molecular Physiology & Biological Physics, University of Virginia, Charlottesville, VA 22908, USA
[g] Centre of New Technologies, University of Warsaw, Warsaw 02-097, Poland
[h] The Jackson Laboratory for Genomic Medicine, Farmington, CT 06030, USA

ARTICLE INFO

Article history:
Received 4 March 2015
Received in revised form 5 May 2015
Accepted 24 July 2015
Available online xxx

Keywords:
Systems biology
Metabolic modeling
Homology modeling
Drug repositioning
Mycoplasma genitalium
Thymidylate kinase

ABSTRACT

Bacteria are increasingly resistant to existing antibiotics, which target a narrow range of pathways. New methods are needed to identify targets, including repositioning targets among distantly related species. We developed a novel combination of systems and structural modeling and bioinformatics to reposition known antibiotics and targets to new species. We applied this approach to *Mycoplasma genitalium*, a common cause of urethritis. First, we used quantitative metabolic modeling to identify enzymes whose expression affects the cellular growth rate. Second, we searched the literature for inhibitors of homologs of the most fragile enzymes. Next, we used sequence alignment to assess that the binding site is shared by *M. genitalium*, but not by humans. Lastly, we used molecular docking to verify that the reported inhibitors preferentially interact with *M. genitalium* proteins over their human homologs. Thymidylate kinase was the top predicted target and piperidinylthymines were the top compounds. Further work is needed to experimentally validate piperidinylthymines. In summary, combined systems and structural modeling is a powerful tool for drug repositioning.

1. Introduction

Mycoplasma genitalium is a gram-positive bacterium and an important sexually transmitted pathogen (Weinstein and Stiles, 2012; Cazanave et al., 2012; Citti and Blanchard, 2013; Manhart, 2013; Taylor-Robinson, 2014). It infects 1–3% of all individuals, is the second most common cause of non-gonococcal urethritis in men, and is an increasingly common cause of cervicitis, endometritis, and pelvic inflammatory disease in women (Manhart, 2013). *M. genitalium* infection has also been associated with adverse sequelae including infertility, ectopic pregnancy, preterm birth, and cancer (Manhart, 2013; Zarei et al., 2013).

M. genitalium infection is typically treated with macrolides such as azithromycin, fluoroquinolones such as moxifloxacin, and tetracyclines such as doxycycline (Taylor-Robinson, 2014). However, clinicians have recently reported resistance as high as 40% in some regions (Salado-Rasmussen and Jensen, 2014). This has been attributed to target mutation and gene transfer (Taylor-Robinson, 2014). Macrolide and fluoroquinolone resistance has primarily been attributed to mutations to the central loop of the 23S rRNA V domain (Jensen et al., 2008) and to DNA topoisomerase genes *gyrA* and *parC* (Shimada et al., 2010), respectively. Tetracycline resistance in the closely related species *Mycoplasma hominis* has been shown to be due to acquisition of *tetM*, which dissociates tetracycline from ribosomes (Dégrange et al., 2008; Dönhöfer et al., 2012). However, the mechanism of *M. genitalium* tetracycline resistance has not yet been determined (Cazanave et al., 2012).

* Corresponding author at: Centre of New Technologies, University of Warsaw, Warsaw 02-097, Poland.
** Corresponding author at: Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.
[1] These authors contributed equally to this work.
[2] Current address: Google Inc., Mountain View, CA 94043, USA.

http://dx.doi.org/10.1016/j.compbiolchem.2015.07.007
1476-9271/

New anti-*Mycoplasma* targets and drugs are urgently needed to combat the emerging resistance. Unfortunately, few pharmaceutical companies are engaged in antibacterial development (Cole, 2014; Nambiar et al., 2014; Harbarth et al., 2015) and the development rate of new molecular entities has declined by 20% in recent years (Pammolli et al., 2011; Khanna, 2012). One promising way to develop new anti-*Mycoplasma* drugs is to reposition existing antibiotics which have already been proven effective against other distantly related bacteria. This repositioning strategy takes advantage of prior knowledge and human toxicity studies. New computational methods are needed to identify potential repositioning candidates to increase the productivity of the drug development pipeline (Rosales-Hernández and Correa-Basurto, 2015).

Systems models which explicitly represent individual drug targets have great potential to improve drug target identification (Chan et al., 2010; Duran-Frigola et al., 2013) including repositioning targets from distantly related species. Chavali et al. (2012) have extensively reviewed the existing systems modeling approaches to drug discovery. Previously, Raman et al. (2008) combined qualitative metabolic modeling with structural modeling to identify potential new druggable *Mycobacterium tuberculosis* targets. Shen et al. (2010) expanded this method to include virtual drug screening and applied their methodology to *Escherichia coli* and *Staphylococcus aureus*. Kim et al. (2011) and Folger et al. (2011) have used the same metabolic modeling methodology to suggest new *Vibrio vulnificus* and anti-cancer targets, respectively.

These systems modeling approaches have not been broadly adopted for drug discovery in part because it has not been possible to build comprehensive quantitative systems models. Recently, systems biologists have developed several new integrative modeling methods, which have enabled unprecedented new models including the first whole-cell model of *M. genitalium*, which accounts for the function of every annotated gene product (Karr et al., 2012). These new systems modeling approaches have not yet been used to reposition drugs.

Our goal was to identify potential new anti-*Mycoplasma* therapeutics by repositioning existing drugs from related species. Toward this goal, we developed a novel combination of quantitative systems modeling, structural modeling, and bioinformatics to reposition established antibiotics and targets to new species based on the rationale that effective drug targets are fragile nodes which are unique to bacteria and selectively inhibited by small molecules. First, we used a quantitative systems metabolic model which incorporates kinetic constraints to predict the effect of inhibiting each *M. genitalium* protein. Second, we performed a literature search to identify selective inhibitors of homologs of the most fragile enzymes. Third, we used multiple sequence alignment to preliminarily assess that *M. genitalium* shares the reported inhibitor binding sites, but humans do not. Fourth, we used structural modeling to verify that the reported inhibitors preferentially bind *M. genitalium* proteins over their human homologs. Our approach improves over prior combined systems and structural modeling approaches to drug discovery by focusing on repositioning and by using a quantitative metabolic model which incorporates kinetic constraints. This enabled us to identify the most fragile enzymes and allowed us to focus our subsequent bioinformatic and structural analyses on high quality targets, thereby minimizing the computational cost of virtual drug screening.

2. Materials and methods

We developed a novel combination of systems modeling, bioinformatics, and structural modeling to identify potential new *M. genitalium* targets and drugs by repositioning established antibiotics and targets from distantly related bacteria. Our approach was based on the rationale that effective drug targets are fragile nodes which are not shared with humans and selectively inhibited by small inhibitors (Fig. 1). First, we used systems modeling to identify fragile nodes. Second, we used literature searches to identify fragile nodes whose homologs have known inhibitors. Lastly, we used sequence analysis and structural modeling to identify fragile nodes with conserved inhibitor binding sites.

2.1. Systems metabolic modeling to identify fragile nodes

We used a kinetically constrained FBA model of *M. genitalium* metabolism to identify the metabolic enzymes which have the most control over the cellular growth rate. We chose to use a kinetically-constrained FBA model because kinetic constraints have been shown to improve the accuracy of metabolic flux and growth rate predictions across experimental conditions (Adadi et al., 2012). The model was originally developed as the metabolic sub-model of a more comprehensive whole-cell model (Karr et al., 2012).

The model is a flux balance analysis (FBA) model (Orth et al., 2010) with additional flux constraints equal to the product of the enzyme copy numbers and enzyme turnover rates. The protein copy numbers were predicted from the observed mRNA expression (Weiner et al., 2003), total protein mass (Morowitz et al., 1962), and the N-end rule (Bachmair et al., 1986) as described in Karr et al. (2012). The turnover rates were curated from the BRENDA (Chang et al., 2015) and SABIO-RK (Wittig et al., 2012) databases and several primary research articles (see Karr et al., 2012 Table S3O). The kinetic constraints are similar to those of MOMENT (Adadi et al., 2012), except that the protein expression values are not free variables, but rather are derived from RNA expression measurements and sub-models of translation and protein degradation.

The metabolic model represents 146 of 525 genes, 104 enzymes, 568 metabolites, and 645 reactions including 57, 67, and 102 kinetically constrained enzymes, genes, and reactions, respectively. We used the metabolic sub-model of version 1.1 of the *M. genitalium* whole-cell model (Karr et al., 2015).

For each gene, we used the metabolic model to calculate the expression level at which the predicted growth rate is 50% of that of the wild type strain (EC_{50}; Code S1–2; Table S1). We used this

Fig. 1. Combined systems and structural modeling and bioinformatics approach repositions known antibiotics and targets to new species. We used a kinetically constrained FBA systems model to identify fragile metabolic enzymes from the 525 gene products present in *M. genitalium*. Next, we conducted a literature search for selective inhibitors of homologs of the top scoring enzymes. We then used sequence alignment to verify that the inhibitor binding site is conserved among bacteria and not shared with humans. Lastly, we used molecular docking to verify that the inhibitor binding site is conserved across bacteria and more energetically favorable than its human homolog. Thymidylate kinase best matched all of our drug target criteria. Piperidinylthymines were the top predicted anti-Mycoplasma compounds.

Please cite this article in press as: D. Kazakiewicz, et al., A combined systems and structural modeling approach repositions antibiotics for *Mycoplasma genitalium*, Comput. Biol. Chem. (2015), http://dx.doi.org/10.1016/j.compbiolchem.2015.07.007

analysis to narrow our drug target search to the top scoring genes. We chose to focus on fractional rather than full growth inhibition, which has been used by many previous metabolic modeling drug discovery studies (Raman et al., 2008; Shen et al., 2010; Kim et al., 2011; Suthers et al., 2009), for several reasons. First, we wanted to use a metric which could reduce the number of targets subjected to further bioinformatics and structural analyses. The majority of *M. genitalium*'s metabolic genes are already known to be essential from both experimental (Glass et al., 2006) and computational (Karr et al., 2012) studies. Consequently, filtering out non-essential genes would not significantly narrow the search for effective drug candidates and, furthermore, it could reject good targets whose expression is critical to fast growth. In contrast, a partial inhibition metric could reduce the list of potential drug targets. Second, focusing on partial inhibition enables us to quantitatively rather than qualitatively assess the impact of each metabolic gene on growth. It enables us to assess the degree to which each protein must be inhibited to achieve the same effect on the growth rate. This analysis provides more information than essentiality and can be used to both filter and rank drug targets. Third, we believe that focusing on partial inhibition could lead to more robust antibiotics because this partial inhibition analysis recognizes the fact that growth is a non-linear function of protein activity. Specifically, this partial inhibition analysis identifies targets whose inhibition will significantly decrease the growth rate even at minimal drug levels. In contrast, focusing only on essentiality could suggest targets which must be fully inhibited to significantly reduce the growth rate. Fourth, focusing on partial inhibition leads to targets which can be inhibited by small drug concentrations, minimizing potential off target deleterious effects.

2.2. Literature search for lead inhibitor compounds

We performed literature searches for selective inhibitors of homologs of the top scoring genes using PubMed (National Center for Biotechnology Information, 2015) and Google Scholar (Google, 2015). We used this literature analysis to further narrow our drug target search.

2.3. Sequence analysis of inhibitor binding site conservation

We increased our confidence in our candidate drugs and targets by verifying that the reported inhibitor binding sites are conserved among bacteria and divergent with humans. This was based on the rationale that conserved sites among bacteria suggest that the inhibitor would likely inhibit the *M. genitalium* protein and that conserved binding sites with humans could cause adverse effects on human tissue. We obtained the reported inhibitor binding sites from the previously described literature search. We then performed multiple sequence alignment of the candidate *M. genitalium* drug targets, their bacterial homologs with known inhibitors, and their human homologs. We performed this using T-Coffee version 11.00.8cbe486 (Magis et al., 2014) with the default settings. We inspected the sequence alignment to verify that the reported binding sites are conserved among bacteria and divergent with humans.

2.4. Structural analysis of binding site conservation and energetics

We performed structural modeling to verify that the inhibitor binding sites are conserved between *M. genitalium* and the other species in which the inhibitors were reported. We used BLASTp to identify template structures from the PDB for each *M. genitalium* protein and its bacterial and human homologs. We then used MODELLER (Webb and Sali, 2014) to perform homology modeling to construct structural models of each protein. We assessed the quality of the predicted structures by inspecting the predicted dihedral angle distributions. We then superimposed the structures of the candidate *M. genitalium* targets and their bacterial homologs and inspected the conservation of the reported inhibitor binding sites. We also used AutoDock Vina (Trott and Olson, 2010) to perform molecular docking between the *M. genitalium* proteins and their human homologs and the reported inhibitors to predict the free energies of the protein-inhibitor interactions.

3. Results

3.1. Systems metabolic modeling to identify fragile nodes

We used the quantitative metabolic model to predict the enzyme expression levels at which the growth rate is 50% of that of the wild type strain (EC_{50}). Table 1 lists the predicted EC_{50} values of the three most fragile genes. Table S1 lists all of the predicted EC_{50} values. Thymidylate kinase (*tmk*, MG006) was predicted to be the most fragile gene. The model predicted that 50% inhibition of Tmk is sufficient to decrease the growth rate by 50%, whereas S-adenosylmethionine synthetase (*metK*, MG047), and acetate kinase (*ackA*, MG357) must be inhibited by 83% and 92%, respectively, to decrease the growth rate by the same amount. This suggests that Tmk, followed by MetK and AckA are potentially effective drug targets because even partial inhibition is sufficient to substantially diminish bacterial growth. Furthermore, Fig. S1 illustrates that the growth rate is sensitive to Tmk at all doses and sensitive to AckA at low doses. Based on this analysis, we decided to focus the remainder of our study on these three drug targets.

3.2. Literature search for lead inhibitor compounds

We conducted literature searches for specific inhibitors of homologs of the three most promising targets Tmk, MetK, and AckA. We found that Martínez-Botella et al. (2012, 2013) reported that piperidinylthymines inhibit gram-positive bacterial Tmk. Martínez-Botella et al. (2012) also reported the positions of the Tmk piperidinylthymine binding site. Furthermore, we found that Keating et al. (2012) reported that piperidinylthymines exhibited activity in an in vivo mouse model of *S. aureus* MRSA252 infection and that Martínez-Botella et al. (2013) reported that piperidinylthymines are moderately and weakly active against gram-negative and human Tmk, respectively. This analysis suggests that piperidinylthymines are good potential lead anti-*Mycoplasma* compounds.

We were unable to identify selective inhibitors for MetK. Consequently, we chose to exclude MetK from further bioinformatic and structural analyses. We found that there is already a known natural antibiotic, allicin, which inhibits the acetyl-CoA synthesis pathway which contains AckA (Focke et al., 1990). Allicin is produced by garlic and has broad-spectrum antibacterial activity as well as antifungal, antiparasitic, antiviral activity. Consequently, because there is already a known and safe antibiotic for AckA, we also chose to exclude AckA from further analysis.

Table 1
Predicted protein expression levels at which growth is 50% of that of the wild type strain (EC_{50}) for the most growth-sensitive proteins. See Table S1 for a complete list of predicted EC_{50} values.

Symbol	Name	Locus	EC_{50} (Rel)
Tmk	Thymidylate kinase	MG006	0.5000
MetK	S-adenosylmethionine synthetase	MG047	0.1771
AckA	Acetate kinase	MG357	0.0753

Please cite this article in press as: D. Kazakiewicz, et al., A combined systems and structural modeling approach repositions antibiotics for *Mycoplasma genitalium*, Comput. Biol. Chem. (2015), http://dx.doi.org/10.1016/j.compbiolchem.2015.07.007

3.3. Sequence analysis of inhibitor binding site conservation

We performed multiple sequence alignment of *M. genitalium* (UniProt ID: P47252), *S. aureus* (UniProt ID: Q2G230), and *Homo sapiens* (UniProt ID: P23919) as well as the gram-negative bacterium *E. coli* (UnitProt ID: P0A720) Tmk to preliminarily assess the interaction energies of piperidinylthymines with *M. genitalium* Tmk. As illustrated in Fig. 2, the reported Tmk inhibitor binding sites Arg53, Phe70, Arg74, Ser101, and Gln105 (Martínez-Botella et al., 2012) are all conserved between *M. genitalium* and *S. aureus*, suggesting that piperidinylthymines bind *M. genitalium* Tmk. The sequence alignment also showed that Arg53 and Gln105 are not conserved with human Tmk and that Ser101 is weakly conserved with human Tmk. This suggests that piperidinylthymines preferentially bind *M. genitalium* over human Tmk, increasing our confidence that piperidinylthymines are good potential lead anti-*Mycoplasma* compounds.

3.4. Structural analysis of binding site conservation and energetics

We constructed a 3-dimensional structural model of *M. genitalium* Tmk to more thoroughly verify that the inhibitor binding site is conserved and to evaluate the energetics of its interactions with piperidinylthymines. We used BLASTp to search the PDB database for similar (E-value $< 10^{-10}$ and identity >35%) high-resolution (<1.67 Å) experimentally determined protein structures. This identified four *S. aureus* Tmk structures with the same primary sequence: PDB IDs 4HLC, 4MQB, 4QG7, and 4QGG. Of these four structures, we chose 4QGG because it had the highest quality refinement, as indicated by the fractional difference between the R_{free} and R values. This structure has E-value = 10^{-28} and 36% identity with *M. genitalium* Tmk, 1.62 Å resolution, R = 0.209 and R_{free}-R = 6.2% (Kawatkar et al., 2014). We used MODELLER to construct a 3-dimensional structural model of *M. genitalium* Tmk based on this template and the *M. genitalium* Tmk amino acid sequence (Structure S1). The RMSD value between the predicted and template structures was 0.23 Å. We further assessed the quality of the predicted structure by inspecting its dihedral angle distribution (Fig. S2). This analysis showed that the majority of the predicted dihedral angles fall within the most energetically favorable regions, suggesting that the predicted structure is physically realistic.

We superimposed the *M. genitalium* and *S. aureus* Tmk structures (Fig. S3a) to rigorously evaluate the conservation of the inhibitor binding site. This showed that the inhibitor binding pocket (Fig. S3b), including the reported binding residues Arg53, Phe70, Arg74, Ser101, and Gln105 (Martínez-Botella et al., 2012), and the global structure, including the α-helices and β-strands, are well-conserved between *S. aureus* and *M. genitalium* Tmk. Only one loop (top-left of Fig. S3a) differs significantly between the proteins, and this non-conserved loop likely does not interfere with piperidinylthymine binding because it is distant from the piperidinylthymine binding site and does not affect the global structure as indicated by the high overall structural similarity. This analysis gave us further confidence that piperidinylthymines inhibit *M. genitalium* Tmk.

We used AutoDock Vina version 1.1.2 to perform molecular docking of piperidinylthymines with the predicted *M. genitalium* Tmk structure and with the observed *S. aureus* MRSA252 (PDB ID: 4HLD) and *H. sapiens* (PDB ID: 2XX3) Tmk structures (Structures S2-22). We performed docking with flexibility for all residues located within 4 Å of the reported *S. aureus* binding site. *S. aureus* binding site-residue distances were evaluated using 4HLD. *M. genitalium* and *H. sapiens* binding site-residue distances were evaluated by first using PyMOL (Schrödinger, LLC, 2015) to align their structures to 4HLD. All analyses were based on the predicted interaction energy of the first mode of each protein-compound pair.

Table 2
Predicted interaction energies of seven piperidinylthymine compounds with *M. genitalium*, *S. aureus* and *H. sapiens* Tmk.

	Predicted interaction energy (kcal mol^{-1})		
PDB Ligand ID	*M. genitalium*	*S. aureus*	*H. sapiens*
32E	−12.2	−10.8	−10.4
32K	−14.8	−12.4	−13.0
0YB	−12.1	−11.3	−10.6
31Z	−12.2	−10.9	−11.1
T05	−12.2	−11.1	−11.2
16T	−11.8	−11.1	−10.9
13Y	−11.7	−10.3	−10.9

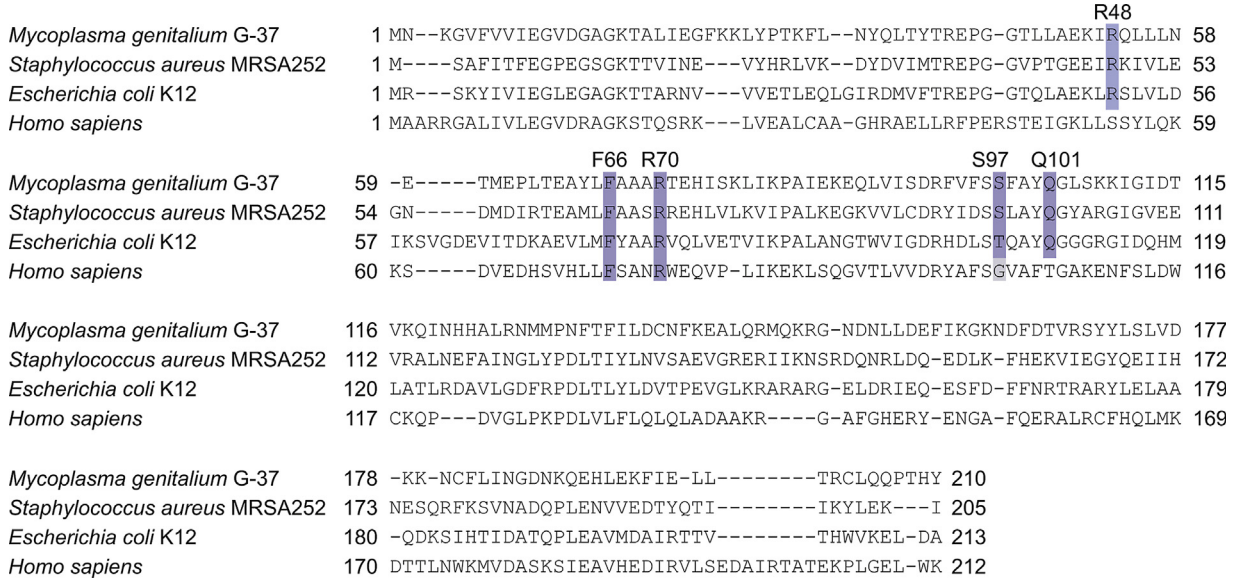

Fig. 2. Multiple sequence alignment of *M. genitalium*,*S. aureus*, *E. coli*, and *H. sapiens* Tmk shows that the reported inhibitor binding site (highlighted residues numbering according to *S. aureus* Tmk; Martinez-Botella et al., 2012) is conserved among bacteria and divergent from *H. sapiens*.

Please cite this article in press as: D. Kazakiewicz, et al., A combined systems and structural modeling approach repositions antibiotics for *Mycoplasma genitalium*, Comput. Biol. Chem. (2015), http://dx.doi.org/10.1016/j.compbiolchem.2015.07.007

We examined the seven piperidinylthymine compounds with experimentally determined structures contained within the PDB ligand database (Shin and Cho, 2005) which also have low reported *S. aureus* 50% inhibitory concentrations (IC_{50}). We curated IC_{50} values from Martínez-Botella et al. (2012), Keating et al. (2012), Martínez-Botella et al. (2013), and Kawatkar et al. (2014). Table 2 lists the predicted interaction energies of each of the seven compounds with *M. genitalium*, *S. aureus*, and *H. sapiens* Tmk. This analysis shows that all seven compounds interact more energetically favorably by 0.8–1.8 kcal mol^{-1} (1.3–3.0 RT) with *M. genitalium* than with human Tmk, suggesting that at least at subsaturating concentrations, piperidinylthymines could selectively disrupt bacterial metabolism. Furthermore, the analysis shows that 32E is likely the most *M. genitalium*-selective compounds because it has the largest free energy difference between *M. genitalium* and human Tmk. Furthermore, Kawatkar et al. (2014) have shown that 32E inhibits *S. aureus* Tmk almost as strongly as 0YB and is equally selective for *S. aureus* over human Tmk as 0YB.

4. Discussion

We have developed a novel approach to drug repositioning that combines systems and structural modeling and bioinformatics. Our approach integrates multiple complementary sources of biological and chemical information to successively increase our confidence in the predicted targets and compounds. First, we used a kinetically constrained FBA metabolic model to identify the most fragile metabolic enzymes whose partial inhibition is sufficient to significantly decrease the cellular growth rate. We then performed literature searches for known selective inhibitors of the most fragile enzymes Tmk, MetK, and AckA. This identified piperidinylthymines as potential lead anti-*Mycoplasma* compounds. Third, we analyzed the *M. genitalium* Tmk sequence to verify that *M. genitalium* Tmk contains the reported piperidinylthymine binding site. Next, we used homology modeling to construct a structural model of *M. genitalium* Tmk. Finally, we used molecular docking to predict the interaction energies of several piperidinylthymine compounds with *M. genitalium* and human Tmk.

Taken together, we presented several computational analyses, which suggest that piperidinylthymines are good lead anti-*Mycoplasma* compounds: we predicted that *M. genitalium* Tmk inhibition strongly decreases the cellular growth rate, that *M. genitalium* Tmk contains the same reported piperidinylthymine binding site as *S. aureus*, and that piperidinylthymines preferentially bind *M. genitalium* over human Tmk.

Tmk is an essential metabolic enzyme, which phosphorylates dTMP, producing dTDP Cui et al. (2013), an important step in dTTP synthesis for DNA replication. Several reports have suggested that Tmk is a promising antibacterial target for multiple microorganisms including *M. tuberculosis* (Vanheusden et al., 2002), *Pseudomonas aeruginosa* (Choi et al., 2012), and *Bacillus anthracis* (Byun et al., 2008). In addition, in vivo mouse studies suggest that piperidinylthymines could be effective human drugs (Keating et al., 2012).

Going forward, in vivo *M. genitalium* infection studies must be conducted to confirm that piperidinylthymines are safe and effective anti-*Mycoplasma* antibiotics. Chemical process engineering will also be needed to make piperidinylthymine synthesis cost-effective and, in turn, affordable for patients.

Our combined systems and structural modeling approach quickly predicts the efficacy of previously described targets and compounds in new species without expensive, time-consuming wet lab experiments. This contrasts with most prior in silico drug discovery approaches, which have focused on identifying entirely new targets and drugs. Our approach is particularly useful for repositioning established antibiotics to new species. In this study, we focused on determining whether piperidinylthymines which have already been described as affective anti-*Staphylococcus* antibiotics, can effectively inhibit *M. genitalium*, which contains a homolog of the piperidinylthymine target, but which has undergone dramatic genome reduction and target mutation since divergence from its common ancestor with *S. aureus*.

Our approach improves over earlier combined systems and structural modeling approaches to drug discovery proposed by Raman et al. (2008), Shen et al. (2010), and Kim et al. (2011) by using a kinetically constrained metabolic model. This enabled us to quantitatively assess the impact of partially inhibiting each metabolic enzyme and identify the most fragile enzymes. In turn, this allowed us to focus our structural modeling on a small set of high quality metabolic targets, reducing the computational cost of homology modeling and molecular docking. This approach is broadly applicable to other microorganisms.

We chose to use a kinetically constrained FBA model which combines enzyme kinetic measurements with protein expression predictions because kinetic constraints been shown to improve the accuracy of metabolic flux and growth rate predictions across experimental conditions (Adadi et al., 2012), and because we believe these predictions would be further improved with additional RNA and protein expression data and models. The kinetically constrained FBA model is similar to MOMENT (Adadi et al., 2012), except that the protein expression values are not free variables, but rather are determined by RNA expression data and translation and protein degradation sub-models. Our approach could also be used with any systems model including other types of constraint-based models such as E-flux (Colijn et al., 2009), GIMME (Becker and Palsson, 2008), iFBA (Covert et al., 2008), iMAT (Zur et al., 2010), MOMENT (Adadi et al., 2012), PROM (Chandrasekaran and Price, 2010), and rFBA (Covert et al., 2004), as well as ordinary differential equation, partial differential equation, logical, and hybrid models. Machado and Herrgård have extensively reviewed the numerous available types of constraint-based models (Machado and Herrgård, 2014). Going forward, we believe our approach would be improved by the use of detailed whole-cell models which represent every gene and cellular function and molecular species (Karr et al., 2012).

Although our approach has several advantages over prior methods, it also has some limitations. It depends on a FBA-based metabolic model, which cannot fully describe cellular metabolism. FBA does not predict unique flux distributions, cannot predict metabolite concentrations, and cannot account for dynamics in the cellular composition. Our approach requires previously described inhibitors from other organisms, precluding its application to de novo target and lead compound discovery. This limitation could be overcome through the use of large-scale virtual screening as in Shen et al. (2010). Our approach requires previously observed protein structures for homologous proteins. In addition, our approach cannot accurately predict drug safety or efficacy. Consequently, it must be coupled with in vitro and in vivo experimental validation. Furthermore, our systems and structural models are not yet completely accurate. For example, structural modeling predicts that 32 K would inhibit *S. aureus* Tmk more strongly than 0YB. However, the IC_{50} and MIC values of 32 K are higher than that of 32 E (Martínez-Botella et al., 2012; Kawatkar et al., 2014). Lastly, our approach could be improved by considering the cost of chemical synthesis.

Going forward, we plan to strengthen our pipeline with additional complementary absorption, distribution, metabolism, excretion, and toxicity (ADMET) analyses. Statistical quantitative structure-activity relationship (QSAR) models, proteome-wide molecular docking, and chemoinformatics are needed to predict potential adverse off-target non-homologous binding. Merlot (2010), Xie et al. (2011), Merino et al. (2010), and Schmidt et al.

Please cite this article in press as: D. Kazakiewicz, et al., A combined systems and structural modeling approach repositions antibiotics for *Mycoplasma genitalium*, Comput. Biol. Chem. (2015), http://dx.doi.org/10.1016/j.compbiolchem.2015.07.007

(2014) have extensively reviewed the existing methods for predicting off-target binding. We also plan to focus on improving the quality and expanding the scope of systems models, which are one of the key bottlenecks to all in silico approaches to drug discovery. In addition, we plan to standardize our integrative systems modeling approach to enable pharmaceutical researchers to take more advantage of the latest developments in systems modeling.

5. Conclusion

In summary, in combination with structural modeling and bioinformatics, systems modeling can be a powerful drug discovery tool. As other researchers have shown, this combined strategy can be used to discover new targets and lead compounds. In addition, as we have demonstrated, it can also be used to reposition established drugs to new species. In this study, we extrapolated the effect of piperidinylthymines from *S. aureus* to the distantly related *M. genitalium*. We envision that this repositioning approach could also be used to repurpose anti-cancer treatments between cancers. Looking forward, we believe that broader systems models, including whole-cell models, combined with more functionally predictive structural models will facilitate de novo target and drug discovery.

Acknowledgments

We thank Markus Covert for thoughtful feedback, W. Andrzej Sokalski and Wladek Minor for enabling this collaboration, Paweł Kafarski for helpful discussions on drug design, and Kira Dashkina for help designing Fig. 1.

This work was supported by grants from the Polish National Science Centre (UMO-2013/09/B/NZ2/00121 and 2014/15/B/ST6/05082) and the Wrocław Centre for Networking and Supercomputing (255). DP was supported by grant from the European Cooperation in Science and Technology (BM1405 and BM1408 actions). DP and DK were supported by the Medical University of Białystok and the European Social Fund. JRK was supported by a National Science Foundation Graduate Fellowship and a James S. McDonnell Foundation Postdoctoral Fellowship Award in Studying Complex Systems. KLM was supported by the University of Virginia Health System.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.compbiolchem.2015.07.007.

References

Adadi, R., Volkmer, B., Milo, R., Heinemann, M., Shlomi, T., 2012. Prediction of microbial growth rate versus biomass yield by a metabolic network with kinetic parameters. PLoS Comput. Biol. 8 (7), e1002575.
Bachmair, A., Finley, D., Varshavsky, A., 1986. In vivo half-life of a protein is a function of its amino-terminal residue. Science 234 (4773), 179–186.
Becker, S.A., Palsson, B.O., 2008. Context-specific metabolic networks are consistent with experiments. PLoS Comput. Biol. 4 (5), e1000082.
Byun, Y., Vogel, S.R., Phipps, A.J., Carnrot, C., Eriksson, S., et al., 2008. Synthesis and biological evaluation of inhibitors of thymidine monophosphate kinase from Bacillus anthracis. Nucleosides Nucleotides Nucleic Acids 27 (3), 244–260.
Cazanave, C., Manhart, L.E., Bébéar, C., 2012. *Mycoplasma genitalium*, an emerging sexually transmitted pathogen. Med. Mal. Infect. 42 (9), 381–392.
Chan, J.N., Nislow, C., Emili, A., 2010. Recent advances and method development for drug target identification. Trends Pharmacol. Sci. 31 (2), 82–88.
Chandrasekaran, S., Price, N.D., 2010. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U.S.A 107 (41), 17845–17850.
Chang, A., Schomburg, I., Placzek, S., Jeske, L., Ulbrich, M., et al., 2015. BRENDA in 2015: exciting developments in its 25th year of existence. Nucleic Acids Res. D439–D446.
Chavali, A.K., Blazier, A.S., Tlaxca, J.L., Jensen, P.A., Pearson, R.D., et al., 2012. Metabolic network analysis predicts efficacy of FDA-approved drugs targeting the causative agent of a neglected tropical disease. BMC Syst. Biol. 6, 27.
Choi, J.Y., Plummer, M.S., Starr, J., Desbonnet, C.R., Soutter, H., et al., 2012. Structure guided development of novel thymidine mimetics targeting *Pseudomonas aeruginosa* thymidylate kinase: from hit to lead generation. J. Med. Chem. 55 (2), 852–870.
Citti, C., Blanchard, A., 2013. Mycoplasmas and their host: emerging and reemerging minimal pathogens. Trends Microbiol. 21 (4), 196–203.
Cole, S.T., 2014. Who will develop new antibacterial agents? Philos. Trans. R. Soc. Lond. B Biol. Sci. 369 (1645), 20130430.
Colijn, C., Brandes, A., Zucker, J., Lun, D.S., Weiner, B., et al., 2009. Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. PLoS Comput. Biol. 5 (8), e1000489.
Covert, M.W., Knight, E.M., Reed, J.L., Herrgard, M.J., Palsson, B.O., 2004. Integrating high-throughput and computational data elucidates bacterial networks. Nature 429 (6987), 92–96.
Covert, M.W., Xiao, N., Chen, T.J., Karr, J.R., 2008. Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. Bioinformatics 24 (18), 2044–2050.
Cui, Q., Shin, W.S., Luo, Y., Tian, J., Cui, H., et al., 2013. Thymidylate kinase: an old topic brings new perspectives. Curr. Med. Chem. 20 (10), 1286–1305.
Dégrange, S., Renaudin, H., Charron, A., Bébéar, C., Bébéar, C.M., 2008. Tetracycline resistance in *Ureaplasma* spp. and *Mycoplasma hominis*: prevalence in Bordeaux, France, from 1999 to 2002 and description of two tet(M)-positive isolates of *M. hominis* susceptible to tetracyclines. Antimicrob. Agents Chemother. 52 (2), 742–744.
Dönhöfer, A., Franckenberg, S., Wickles, S., Berninghausen, O., Beckmann, R., et al., 2012. Structural basis for TetM-mediated tetracycline resistance. Proc. Natl. Acad. Sci. U.S.A. 109 (42), 16900–16905.
Duran-Frigola, M., Mosca, R., Aloy, P., 2013. Structural systems pharmacology: the role of 3D structures in next-generation drug development. Chem. Biol. 20 (5), 674–684.
Focke, M., Feld, A., Lichtenthaler, K., 1990. Allicin, a naturally occurring antibiotic from garlic, specifically inhibits acetyl-CoA synthetase. FEBS Lett. 261 (1), 106–108.
Folger, O., Jerby, L., Frezza, C., Gottlieb, E., Ruppin, E., et al., 2011. Predicting selective drug targets in cancer through metabolic networks. Mol. Syst. Biol. 7, 501.
Glass, J.I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M.R., et al., 2006. Essential genes of a minimal bacterium. Proc. Natl. Acad. Sci. U.S.A. 103 (2), 425–430.
Google, 2015. Google Scholar. (accessed 23.02.15).
Harbarth, S., Theuretzbacher, U., Hackett, J., 2015. Antibiotic research and development: business as usual? J. Antimicrob. Chemother. pii, dkv020.
Jensen, J.S., Bradshaw, C.S., Tabrizi, S.N., Fairley, C.K., Hamasuna, R., 2008. Azithromycin treatment failure in *Mycoplasma genitalium* -positive patients with nongonococcal urethritis is associated with induced macrolide resistance. Clin. Infect. Dis. 47 (12), 1546–1553.
Karr, J.R., Sanghvi, J.C., Macklin, D.N., Gutschow, M.V., Jacobs, J.M., et al., 2012. A whole-cell computational model predicts phenotype from genotype. Cell 150 (2), 389–401.
Karr, J.R., Sanghvi, J.C., Macklin, D.N., Jacobs, J.M., Covert, M.W., 2015. *Mycoplasma genitalium* whole-cell model version 1.1. figshare.1321518 (accessed 03.03.15).
Kawatkar, S.P., Keating, T.A., Olivier, N.B., Breen, J.N., Green, O.M., et al., 2014. Antibacterial inhibitors of gram-positive thymidylate kinase: structure-activity relationships and chiral preference of a new hydrophobic binding region. J. Med. Chem. 57 (11), 4584–4597.
Keating, T.A., Newman, J.V., Olivier, N.B., Otterson, L.G., Andrews, B., et al., 2012. In vivo validation of thymidylate kinase (TMK) with a rationally designed, selective antibacterial compound. ACS Chem. Biol. 7 (11), 1866–1872.
Khanna, I., 2012. Drug discovery in pharmaceutical industry: productivity challenges and trends. Drug Discov. Today 17 (19–20), 1088–1102.
Kim, H.U., Kim, S.Y., Jeong, H., Kim, T.Y., Kim, J.J., et al., 2011. Integrative genome-scale metabolic analysis of *Vibrio vulnificus* for drug targeting and discovery. Mol. Syst. Biol. 7, 460.
Machado, D., Herrgård, M., 2014. Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. PLoS Comput. Biol. 10 (4), e1003580.
Magis, C., Taly, J.F., Bussotti, G., Chang, J.M., Di Tommaso, P., et al., 2014. T-coffee: tree-based consistency objective function for alignment evaluation. Methods Mol. Biol. 1079, 117–129.
Manhart, L.E., 2013. *Mycoplasma genitalium*: an emergent sexually transmitted disease? Infect. Dis. Clin. North Am. 27 (4), 779–792.
Martínez-Botella, G., Breen, J.N., Duffy, J.E., Dumas, J., Geng, B., et al., 2012. Discovery of selective and potent inhibitors of Gram-positive bacterial thymidylate kinase (TMK). J. Med. Chem. 55 (22), 10010–10021.
Martínez-Botella, G., Loch, J.T., Green, O.M., Kawatkar, S.P., Olivier, N.B., et al., 2013. Sulfonylpiperidines as novel, antibacterial inhibitors of Gram-positive thymidylate kinase (TMK). Bioorg. Med. Chem. Lett. 23 (1), 169–173.
Merino, A., Bronowska, A.K., Jackson, D.B., Cahill, D.J., 2010. Drug profiling: knowing where it hits. Drug Discov. Today 15 (17-18), 749–756.
Merlot, C., 2010. Computational toxicology–a tool for early safety evaluation. Drug Discov. Today 15 (1–2), 16–22.
Morowitz, H.J., Tourtellotte, M.E., Guild, W.R., Castro, E., Woese, C., 1962. The chemical composition and submicroscopic morphology of *Mycoplasma gallisepticum*, avian PPLO 5969. J. Mol. Biol. 4, 93–103.

Nambiar, S., Laessig, K., Toerner, J., Farley, J., Cox, E., 2014. Antibacterial drug development: challenges, recent developments, and future considerations. Clin. Pharmacol. Ther. 96 (2), 147–149.
National Center for Biotechnology Information, 2015. PubMed. (accessed 23.02.15).
Orth, J.D., Thiele, I., Palsson, B.O., 2010. What is flux balance analysis. Nat. Biotechnol. 28 (3), 245–248.
Pammolli, F., Magazzini, L., Riccaboni, M., 2011. The productivity crisis in pharmaceutical R&D. Nat. Rev. Drug Discov. 10 (6), 428–438.
Raman, K., Yeturu, K., Chandra, N., 2008. TargetTB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. BMC Syst. Biol. 2, 109.
Rosales-Hernández, M.C., Correa-Basurto, J., 2015. The importance of employing computational resources for the automation of drug discovery. Expert Opin. Drug Discov. 15.
Salado-Rasmussen, K., Jensen, J.S., 2014. *Mycoplasma genitalium* testing pattern and macrolide resistance: a Danish nationwide retrospective survey. Clin. Infect. Dis. 59 (1), 24–30.
Schmidt, F., Matter, H., Hessler, G., Czich, A., 2014. Predictive in silico off-target profiling in drug discovery. Future Med. Chem. 6 (3), 295–317.
Schrödinger, LLC, 2015. The PyMOL Molecular Graphics System. (accessed 23.02.15).
Shen, Y., Liu, J., Estiu, G., Isin, B., Ahn, Y.Y., et al., 2010. Blueprint for antimicrobial hit discovery targeting metabolic networks. Proc. Natl. Acad. Sci. U.S.A. 107 (3), 1082–1087.
Shimada, Y., Deguchi, T., Nakane, K., Masue, T., Yasuda, M., et al., 2010. Emergence of clinical strains of *Mycoplasma genitalium* harbouring alterations in ParC associated with fluoroquinolone resistance. Int. J. Antimicrob. Agents 36 (3), 255–258.
Shin, J.M., Cho, D.H., 2005. PDB-ligand: a ligand database based on PDB for the automated and customized classification of ligand-binding structures. Nucleic Acids Res. 33 (Database issue), D238–D241.
Suthers, P.F., Dasika, M.S., Kumar, V.S., Denisov, G., Glass, J.I., et al., 2009. A genome-scale metabolic reconstruction of *Mycoplasma genitalium*, iPS189. PLoS Comput. Biol. 5 (2), e1000285.
Taylor-Robinson, D., 2014. Diagnosis and antimicrobial treatment of *Mycoplasma genitalium* infection: sobering thoughts. Expert Rev. Anti Infect. Ther. 12 (6), 715–722.
Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31 (2), 455–461.
Vanheusden, V., Munier-Lehmann, H., Pochet, S., Herdewijn, P., Van Calenbergh, S., 2002. Synthesis and evaluation of thymidine-5′-O-monophosphate analogues as inhibitors of *Mycobacterium tuberculosis* thymidylate kinase. Bioorg. Med. Chem. Lett. 12 (19), 2695–2698.
Webb, B., Sali, A., 2014. Protein structure modeling with MODELLER. Methods Mol. Biol. 1137, 1–15.
Weiner 3rd, J., Zimmerman, C.U., Göhlmann, H.W., Herrmann, R., 2003. Transcription profiles of the bacterium *Mycoplasma pneumoniae* grown at different temperatures. Nucleic Acids Res. 31 (21), 6306–6320.
Weinstein, S.A., Stiles, B.G., 2012. Recent perspectives in the diagnosis and evidence-based treatment of *Mycoplasma genitalium*. Expert Rev. Anti Infect. Ther 10 (4), 487–499.
Wittig, U., Kania, R., Golebiewski, M., Rey, M., Shi, L., et al., 2012. SABIO-RK–database for biochemical reaction kinetics. Nucleic Acids Res. 40 (Database issue), D790–796.
Xie, L., Xie, L., Bourne, P.E., 2011. Structure-based systems biology for analyzing off-target binding. Curr. Opin. Struct. Biol. 21 (2), 189–199.
Zarei, O., Rezania, S., Mousavi, A., 2013. *Mycoplasma genitalium* and cancer: a brief review. Asian Pac. J. Cancer Prev. 14 (6), 3425–3428.
Zur, H., Ruppin, E., Shlomi, T., 2010. iMAT: an integrative metabolic analysis tool. Bioinformatics 26 (24), 3140–3142.

Please cite this article in press as: D. Kazakiewicz, et al., A combined systems and structural modeling approach repositions antibiotics for *Mycoplasma genitalium*, Comput. Biol. Chem. (2015), http://dx.doi.org/10.1016/j.compbiolchem.2015.07.007