

REVIEW ARTICLE

A model to predict anti-tuberculosis activity: value proposition for marine microorganisms

Miaomiao Liu^{1,2,3}, Tanja Grkovic¹, Lixin Zhang², Xueting Liu² and Ronald J Quinn¹

The development of new antibiotics effective against all strains of tuberculosis (TB) is needed. To evaluate the potential of marine microbe-derived natural products as anti-TB leads, we analyzed and compared the physico-chemical properties of 39 current TB drugs and candidates against 60 confirmed mycobacteria-active natural products. We showed that anti-TB natural products sourced from marine microbes have a large overlap with TB drug-like space. A model to predict potential anti-TB drugs is proposed.

The Journal of Antibiotics (2016) 69, 594–599; doi:10.1038/ja.2016.87; published online 13 July 2016

INTRODUCTION

Tuberculosis (TB) is responsible for the death of millions of people every year and continues to claim more lives than any other single infectious agent.^{1,2} The high incidence of HIV—TB co-infection, together with the emergence of multi-drug resistant and extensively-drug resistant forms, has resulted in TB being one of the most significant threats to global health. Major problems limiting the development and effectiveness of new drugs to combat TB are the profound innate resistance of *Mycobacterium tuberculosis* to host defense mechanisms, as well as its intrinsic tolerance to chemotherapeutic reagents.³ Mycobacteria are unusually successful in surviving the presence of toxic compounds because they produce effective permeability barriers, comprising of the outer membrane and the mycolate-containing cell wall on the cell surface.⁴ As a result, among hundreds of antibiotics commercially available, anti-TB treatment is still limited to drugs that were developed >50 years ago, such as ethionamide (1960), ethambutol (1961), capreomycin (1963) and rifampicin (1963). Insufficient drug options, long treatment regimens and patient non-compliance have led to worldwide emergence of strains resistant to almost all available drugs.⁵

In the search for novel drug candidates, including anti-TB drugs, natural products are evolutionary selected and pre-validated by nature, displaying unique chemical diversity and a corresponding diversity of biological activities.^{6,7} Among the 11 currently used nature-derived TB drugs, 7 of them were either isolated from microbes or semi-synthesized from microbial natural products (Figure 1).

Traditionally terrestrial microorganisms were explored as a source of biologically active natural products, however, natural products sourced from the marine environment are becoming increasingly important as a source of structurally novel and biologically active

compounds.^{8–13} In order to examine the potential of marine microbe-sourced natural products as effective TB agents, we conducted a comprehensive literature search covering the period up to December 2014. In total, 60 marine microbial compounds have been reported to show anti-*M. tuberculosis* activity in *in vitro* models. These compounds coupled with 39 current TB drugs and candidates were used to establish a predictive model for anti-TB activity. In this period, 336 marine natural products that have not been screened for anti-mycobacterial activity were isolated from various marine microbes.

Three data sets (Table 1), TB drugs and candidates (39 compounds), mycobacteria-active natural products from marine microbes (60 compounds) and *M. tuberculosis*-untested natural products from marine microbes (336 compounds), were evaluated in this study.

Oral bioavailability will very likely increase compliance, lack of which is a major problem especially with the long duration of therapy currently required for available therapies. The lipophilicity is clearly not required to cross the mycolic acid-rich cell wall, given the good activity of several RNA-active aminoglycosides.

COMPARISON OF THE CHEMGPS-NP CHEMICAL SPACE BETWEEN MARINE MICROBIAL NATURAL PRODUCTS AND TB DRUGS

Experimentally validated data in data sets 1 and 2 (Table 1) were used to examine the physico-chemical profiles of 39 TB drugs and candidates, and 60 mycobacteria-active natural products. Chemical global positioning system-natural product (ChemGPS-NP) is a principal component analysis-based global chemical positioning system tuned for exploration of biologically relevant chemical space, that is, those areas of chemical space most likely to enclose biologically active compounds.¹⁴ In ChemGPS-NP aspects of size, shape, lipophilicity, polarity, polarizability, flexibility, rigidity and hydrogen bond

¹Eskitis Institute for Drug Discovery, Griffith University, Brisbane, QLD, Australia; ²Chinese Academy of Sciences, Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China and ³University of Chinese Academy of Sciences, Beijing, China
Correspondence: Professor RJ Quinn, Eskitis Institute for Drug Discovery, Griffith University, Nathan Campus, Brisbane 4111, Queensland, Australia.
E-mail: r.quinn@griffith.edu.au

or Professor X Liu, F816, Institute of Microbiology, Chinese Academy of Sciences, NO.1 Beichen West Road, Chaoyang District, Beijing 100101, China.
E-mail: liuxt2010@gmail.com

This article is dedicated to the fond memory of the late Professor Lester Mitscher, a great scholar, teacher and Emeritus Editor of this Journal.

Received 2 April 2016; revised 15 June 2016; accepted 15 June 2016; published online 13 July 2016

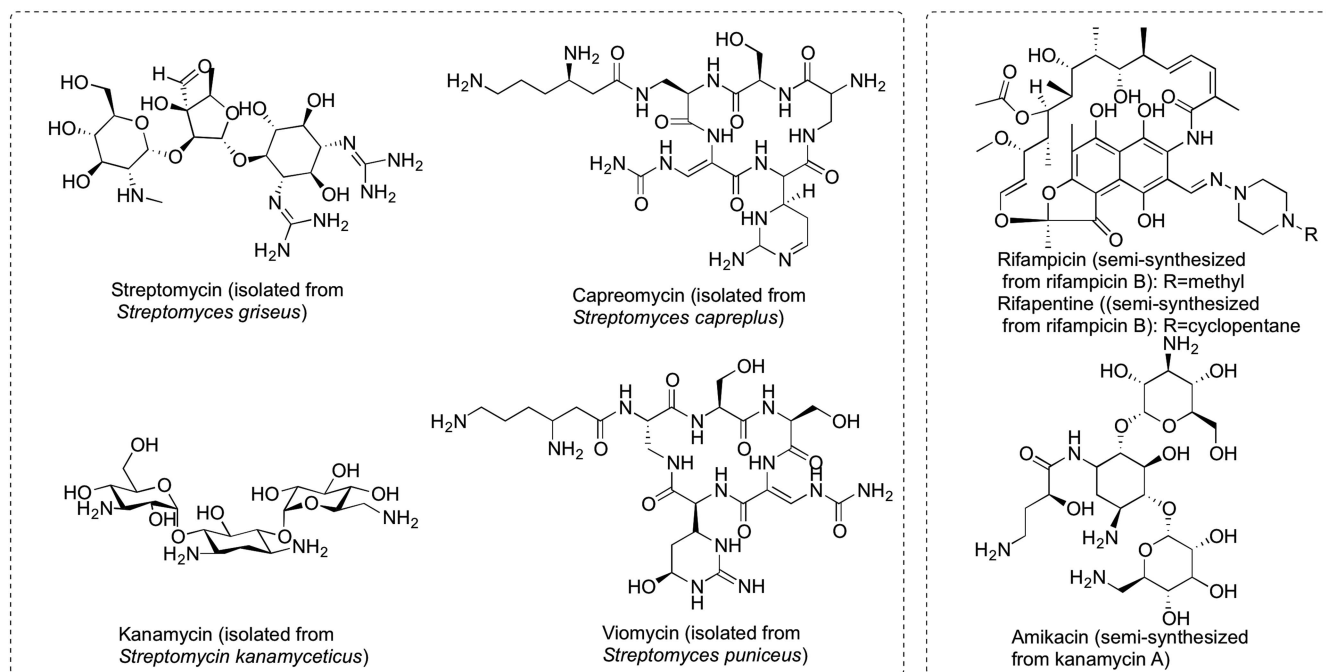


Figure 1 Seven TB drugs sourced or based on the NP scaffold isolated from a microbe. TB, tuberculosis.

Table 1 Information of three data sets used in this review

Data set	Data set name	Anti-M. tuberculosis activity	Number of compounds	Compound types	Compound sources
1	TB drugs and candidates	Active	39	Natural products and synthetic	Various
2	Marine microbial NPs	Active	60	Natural products	Marine microbes
3	Marine microbial NPs	Unknown	336	Natural products	Marine microbes

Abbreviations: NPs, natural products; TB, tuberculosis.

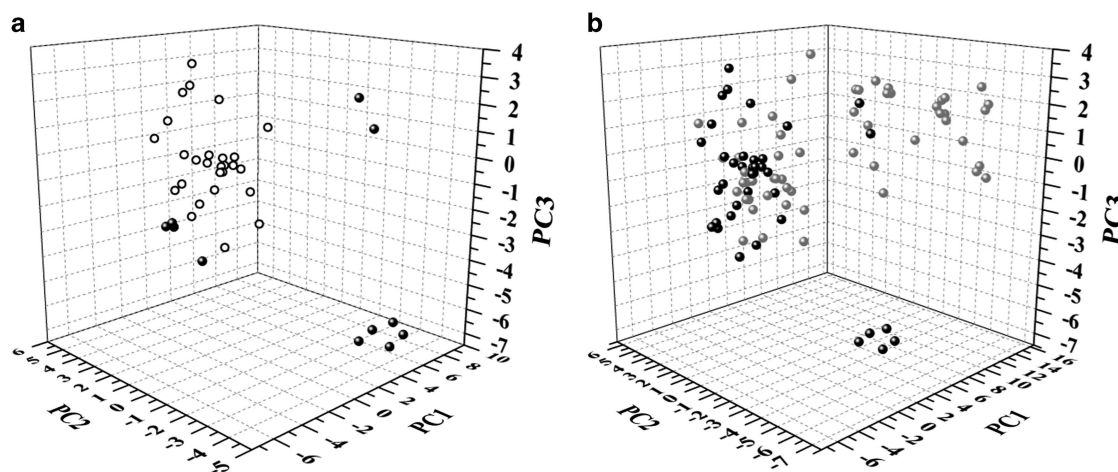


Figure 2 Score plot of mycobacteria-active natural products sourced from marine microbes and TB drugs. (a) PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity) for sets of 28 synthetic TB drugs and candidates (black empty circle) and 11 nature-derived TB drugs and candidates (black filled circle) and (b) PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity) for sets of 39 TB drugs and candidates (black) and 60 mycobacteria-active natural products from marine microbes (gray). TB, tuberculosis.

capacity are compared.¹⁵ The ChemGPS-NP space map coordinates are *t*-scores from principal component analysis using a carefully selected subset of 35 descriptors,¹⁶ which are then analyzed by eight respective principal components that can be mapped onto a consistent eight-dimensional map. The four most significant principle

components (PCs) explain 77% of the variance and can be interpreted as follows: PC1 represents size, shape and polarizability, PC2 corresponds to aromatic- and conjugation-related properties, PC3 describes lipophilicity, polarity and H-bond capacity, and PC4 expresses flexibility and rigidity.¹⁷ Any compound with a known

Table 2 The scores specification of each region

Regions	PC1	PC2	PC3
Total score range	-4.24 ~ 14.14	-5.72 ~ 4.89	-2.84 ~ 3.54
Region 1	-4.24 ~ 2	-5.72 ~ -2	-2.84 ~ -1
Region 2	2 ~ 8	-5.72 ~ -2	-2.84 ~ -1
Region 3	8 ~ 14.14	-5.72 ~ -2	-2.84 ~ -1
Region 4	-4.24 ~ 2	-5.72 ~ -2	-1 ~ 1
Region 5	2 ~ 8	-5.72 ~ -2	-1 ~ 1
Region 6	8 ~ 14.14	-5.72 ~ -2	-1 ~ 1
Region 7	-4.24 ~ 2	-5.72 ~ -2	1 ~ 3.54
Region 8	2 ~ 8	-5.72 ~ -2	1 ~ 3.54
Region 9	8 ~ 14.14	-5.72 ~ -2	1 ~ 3.54
Region 10	-4.24 ~ 2	-2 ~ 1	-2.84 ~ -1
Region 11	2 ~ 8	-2 ~ 1	-2.84 ~ -1
Region 12	8 ~ 14.14	-2 ~ 1	-2.84 ~ -1
Region 13	-4.24 ~ 2	-2 ~ 1	-1 ~ 1
Region 14	2 ~ 8	-2 ~ 1	-1 ~ 1
Region 15	8 ~ 14.14	-2 ~ 1	-1 ~ 1
Region 16	-4.24 ~ 2	-2 ~ 1	1 ~ 3.54
Region 17	2 ~ 8	-2 ~ 1	1 ~ 3.54
Region 18	8 ~ 14.14	-2 ~ 1	1 ~ 3.54
Region 19	-4.24 ~ 2	1 ~ 4.89	-2.84 ~ -1
Region 20	2 ~ 8	1 ~ 4.89	-2.84 ~ -1
Region 21	8 ~ 14.14	1 ~ 4.89	-2.84 ~ -1
Region 22	-4.24 ~ 2	1 ~ 4.89	-1 ~ 1
Region 23	2 ~ 8	1 ~ 4.89	-1 ~ 1
Region 24	8 ~ 14.14	1 ~ 4.89	-1 ~ 1
Region 25	-4.24 ~ 2	1 ~ 4.89	1 ~ 3.54
Region 26	2 ~ 8	1 ~ 4.89	1 ~ 3.54
Region 27	8 ~ 14.14	1 ~ 4.89	1 ~ 3.54

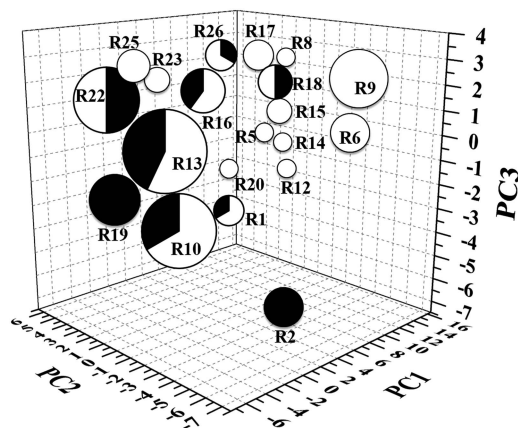


Figure 3 Score plot of mycobacteria-active natural products from marine microbes and TB drugs and candidates in regions. Total numbers of compounds in each region is indicated by sizes of pie charts, percentages of *M. tuberculosis* NPs from marine microbes and TB drugs in each region are showed in white and black, respectively. PC1 (molecular size) versus PC2 (molecular aromaticity) versus PC3 (molecular lipophilicity). TB, tuberculosis.

chemical structure can be positioned onto this map using interpolation in terms of principal component analysis score prediction. From the results, the properties of the compounds can be compared and easily interpreted together with trends and clusters.

The 39 TB drugs and candidates can be divided into two groups according to their biological origin: 11 nature-derived compounds and 28 synthetic compounds. Interestingly, these two groups showed

significant differences in the ChemGPS analysis (Figure 2a). Nature-derived TB drugs occupy a very broad range of physico-chemical space while most of the synthetic TB drugs distributed in the positive PC2 direction (high aromaticity). Furthermore, the nature-derived drugs and candidates can be divided into three sub-groups in the PC1/PC2/PC3 plot. The first subgroup consists of the injectable drugs, kanamycin, amikacin, streptomycin, capreomycin and viomycin, which showed high molecular weight, the least aromaticity and the lowest lipophilicity. The second subgroup is representative of low molecular weight but high aromaticity, including cycloserine, p-aminosalicylic acid, pyrazinamide and isoniazid. The last two rifamycin compounds, rifapentine and rifampicin, belong to the third subgroup, possessing the highest molecular weight and high lipophilicity.

In this model, the mycobacteria-active natural products from marine microbes are largely overlapping with known drugs, at least in the first three dimensions (Figure 2b), strongly arguing that natural products from marine microbes have the potential to serve as an important source for TB drugs. However, 13 mycobacteria-active natural products in the upper right corner, which have higher molecular size, relative low aromaticity and are rather non-polar, do not coincide with any of the known drugs or candidates. These compounds may have a new mode of action for TB inhibition.

The area containing the TB drugs, drug candidates and mycobacteria-active natural products was subdivided into 27 regions in the ChemGPS-NP according to the values of the first three most significant PCs (explaining 71% of the variance). Each of the regions were analyzed in terms of occupancy with regard to both chemical properties and biological activities of the compounds. The scores specification of each region is listed in Table 2.

The basic interpretation of the 27 regions of ChemGPS-NP is as follows: the size of compounds increases in the positive direction of PC1 (that is, small molecules: regions 1, 4, 7, 10, 13, 16, 19, 22 and 25; medium size molecules: regions 2, 5, 8, 11, 14, 17, 20, 23 and 26; large molecules: regions 3, 6, 9, 12, 15, 18, 21, 24 and 27); compounds are increasingly aromatic in the positive direction of PC2 (that is, less aromatic molecules: regions 1–9; moderately aromatic molecules: regions 10–18; more aromatic molecules: regions 19–27); lipophilic compounds are situated in the positive direction of PC3 while predominantly polar compounds are located in the negative PC3 direction (that is, polar molecules: regions 1, 2, 3, 10, 11, 12, 19, 20 and 21; mediate-polar molecules: regions 4, 5, 6, 13, 14, 15, 22, 23 and 24; non-polar molecules: regions 7, 8, 9, 16, 17, 18, 25, 26 and 27).

All of the marine mycobacteria-active microbial natural products were compared with the corresponding TB drugs in the same region (Figure 3). Region 13, representative of low molecular weight, moderate polarity and aromaticity, holds the most active natural products (12) and TB drugs (9) that belong to two different classes distinguished by the mechanisms of DNA inhibition and protein synthesis inhibition.

The 27 regions can be classified into three classes according to the proportion of TB drugs and mycobacteria-active natural products in each region. Firstly, the injectable and nature-derived TB drugs comprised of streptomycin, kanamycin, amikacin, capreomycin and viomycin clustered together in region 2 and no representative mycobacteria-active natural products occupied this region of physico-chemical space. It is important to note that all of these drugs are protein synthesis inhibitors. Another region that lacked representation of mycobacteria-active natural products from marine microbes is region 19, which contains five cell wall inhibitors (isoniazid, pyrazinamide, thioacetazone, ethionamide and prothionamide) and aminosalicylic acid. The low density of natural products in these regions indicates a promising field for new natural products with similar mode of

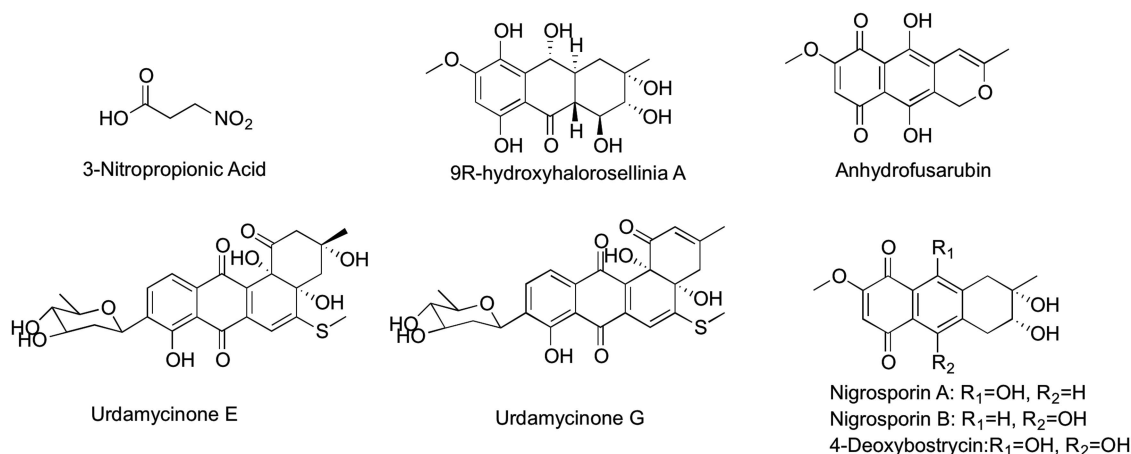


Figure 4 Structures of eight mycobacteria-active natural products in region 10.

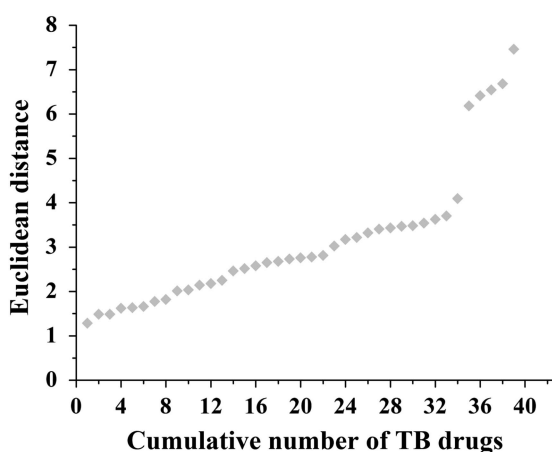


Figure 5 Distribution of ED to the nearest natural product neighbor for the TB drugs and candidates. The cumulative number of drugs is plotted against the ED to the closest natural product neighbor. TB, tuberculosis.

mechanism with current drugs. Secondly, there were eight regions that had only mycobacteria-active natural products, region 9 in particular, the second biggest region with nine natural products of which four exhibited anti-*M. tuberculosis* MIC <1 μm. These active natural products have unique properties in chemical space, as well as excellent biological activities. Finally, seven regions contained TB drugs, as well as mycobacteria-active natural products, and according to the theory that compounds with similar activity profile and chemical properties often show a similar mode of action, it is possible to predict the putative mode of action of these natural products. For example, region 10 contained TB drugs with small sizes and cell wall inhibition activities, as well as eight marine microbe-derived mycobacteria-active natural products (Figure 4). These natural products may be potential mycobacteria cell wall inhibitors. Similarly, natural products in region 22 and 26 could be predicted as promising cell wall inhibitors and ATP synthesis inhibitors, respectively. Moreover, there were twelve mycobacteria-active natural products and nine anti-TB drugs in region 13, including six DNA inhibitors and three protein inhibitors. Based on the same hypothesis, the mode of mechanisms of these compounds would be predicted as inhibitors of DNA or proteins of mycobacteria. This hypothesis was strongly supported by the fact that nine compounds in region 13 with known mode of mechanisms are all protein inhibitors, consistent with our prediction.

NATURAL PRODUCT NEAR NEIGHBORS OF APPROVED DRUGS

Calculation of Euclidean distances based on ChemGPS scores has been found to be a useful tool to identify natural products leads for drug discovery.¹⁵ The EDs were calculated between points $P = (p_1, p_2, \dots, p_n)$ and $Q = (q_1, q_2, \dots, q_n)$ in the Euclidean n -dimensional space, as defined by the following expression:

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

Thereby all mycobacteria-active natural products were assigned with 39 EDs by the use of all eight coordinates calculated by ChemGPS, one ED to each drug. In Figure 5, the 39 drugs are plotted against the ED to their closest natural product neighbor. Interestingly, all drugs have a natural product neighbor closer than ED=8, and around 85% of the drugs have a natural product neighbor closer than ED=4. This forms a strong argument that natural products from marine microbes have the potential to serve as an important source of TB drugs.

Some natural products showed short EDs with more than one drug neighbor (Figure 6). Nanomycin βA and αA, isolated from a marine-derived *Streptomyces* sp., have been reported to inhibit mycoplasma, fungi and Gram-positive bacteria, showing inhibitory activity against *M. tuberculosis* H37Rv with an MIC value of 8.0 μg ml⁻¹. Nanomycins showed a close relationship with nine anti-TB drugs, including one cell wall inhibitor, five DNA inhibitors and two protein synthesis inhibitors, with ED value <2. We suggest that the natural products with short EDs to known drugs should be paid particular attention in developing new anti-TB drugs candidates because of their high similarity with known drugs in physico-chemical space.

The above information can be used to develop a strategy to select compounds when conducting anti-*M. tuberculosis* screening (Figure 7). First, mapping of candidate compounds into the *M. tuberculosis*-relevant chemical space defined by TB drugs and *M. tuberculosis*-active natural products allows for a quick and efficient filter for potential *M. tuberculosis*-activity. The location of compounds in this model can also be used to predict potential mode of action (MoA). Second, compounds with EDs <2 to known TB drugs provides a second filter. Screening of the prioritized compounds in chemical space may directly lead to the identification of biologically relevant structures for TB drug discovery.

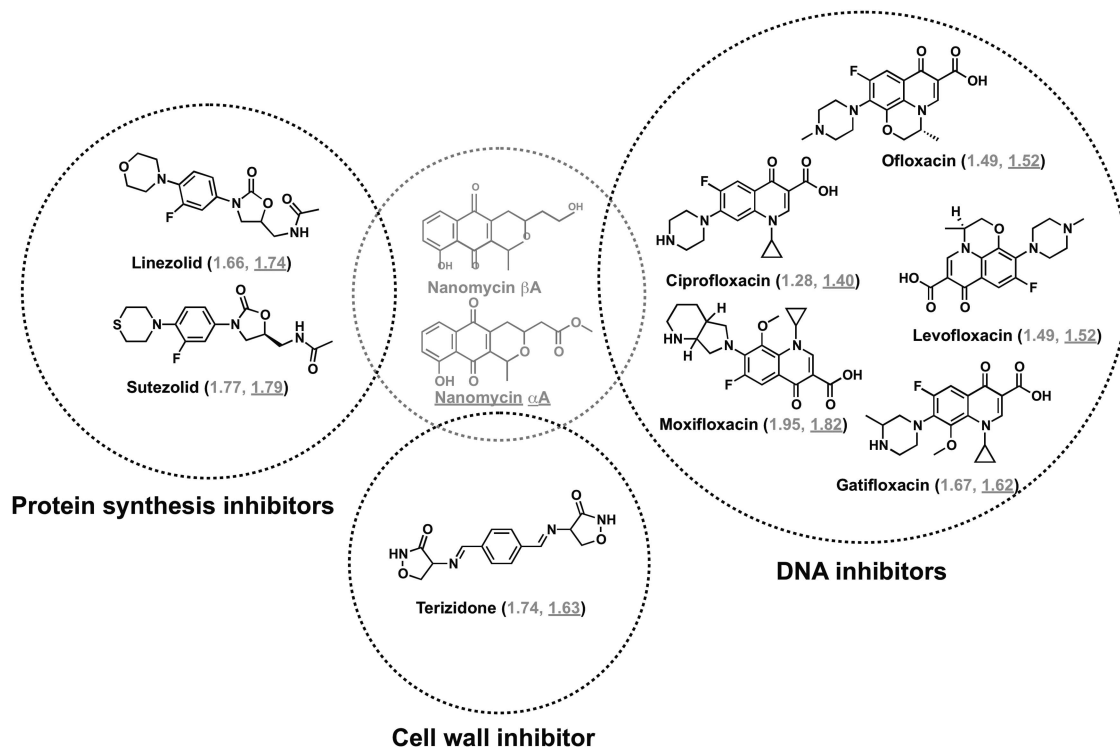


Figure 6 Structures of marine microbe-derived mycobacteria-active natural products nanaomycin α A and nanaomycin β A and anti-TB drugs with EDs shorter than 2. Anti-TB drugs with protein synthesis inhibitor activity are in orange, DNA inhibitors and cell wall inhibitors are in black. EDs between drugs and nanaomycin α A and nanaomycin β A are given in gray without and with underline, respectively. TB, tuberculosis.

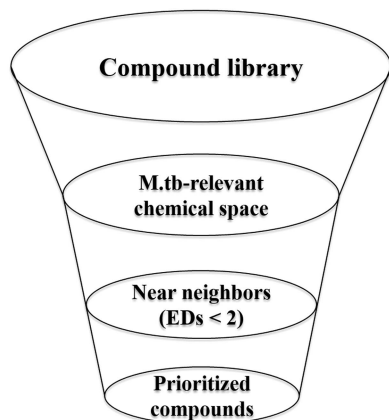


Figure 7 Main factors to design diverse combinatorial libraries or to select diverse compounds to conduct an anti-*M. tuberculosis* screening collection.

APPLYING THE MODEL

Data set 3 (Table 1) was analyzed using the model. 327 of 336 compounds were found to fall into the 27-region chemical space (Figure 8). Similar with *M. tuberculosis*-active natural products, region 13 contained the most untested compounds (90). Surprisingly, three regions 4, 11 and 24, which have no represented TB drugs and *M. tuberculosis* active compounds, occupied with 5, 1 and 2 untested compounds, respectively. These compounds may show unique structure properties.

Filtering these compounds using EDs calculation to identify near neighbors with known TB drugs produced 82 prioritized structures with $ED < 2$ to at least one drug. Structures of the top 20 nearest EDs of untested natural products/drugs are given in Figure 9. Untested

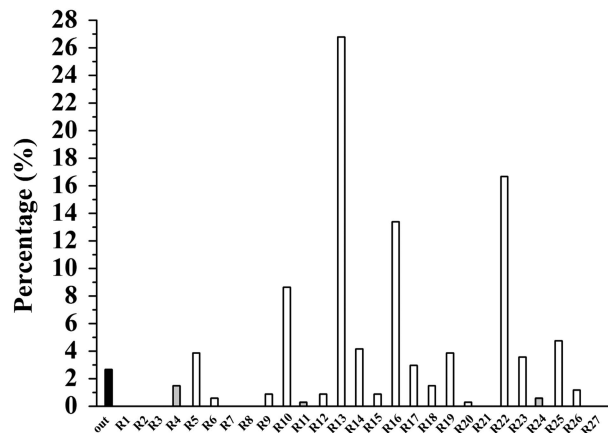


Figure 8 Histogram of the percentage of untested NPs from marine microbes in each regions. Compounds not falling into the 27 regions were plotted in black, compounds in regions 4, 11 and 24 without represented TB drugs and *M. tuberculosis*-active NPs were plotted in gray. TB, tuberculosis.

compounds showing short EDs with more than one drug neighbor may be worth screening.

CONCLUSION

TB threatens people's lives around the world and the appearance of drug resistance has increased the need to identify novel anti-TB drugs. The complex etiology of the disease involving oral bioavailability, lung-alveolar macrophages, granulomas and mycobacterial cell permeability, requires complex modelling. We used the chemical space navigation tool ChemGPS-NP to compare 60 mycobacteria-active

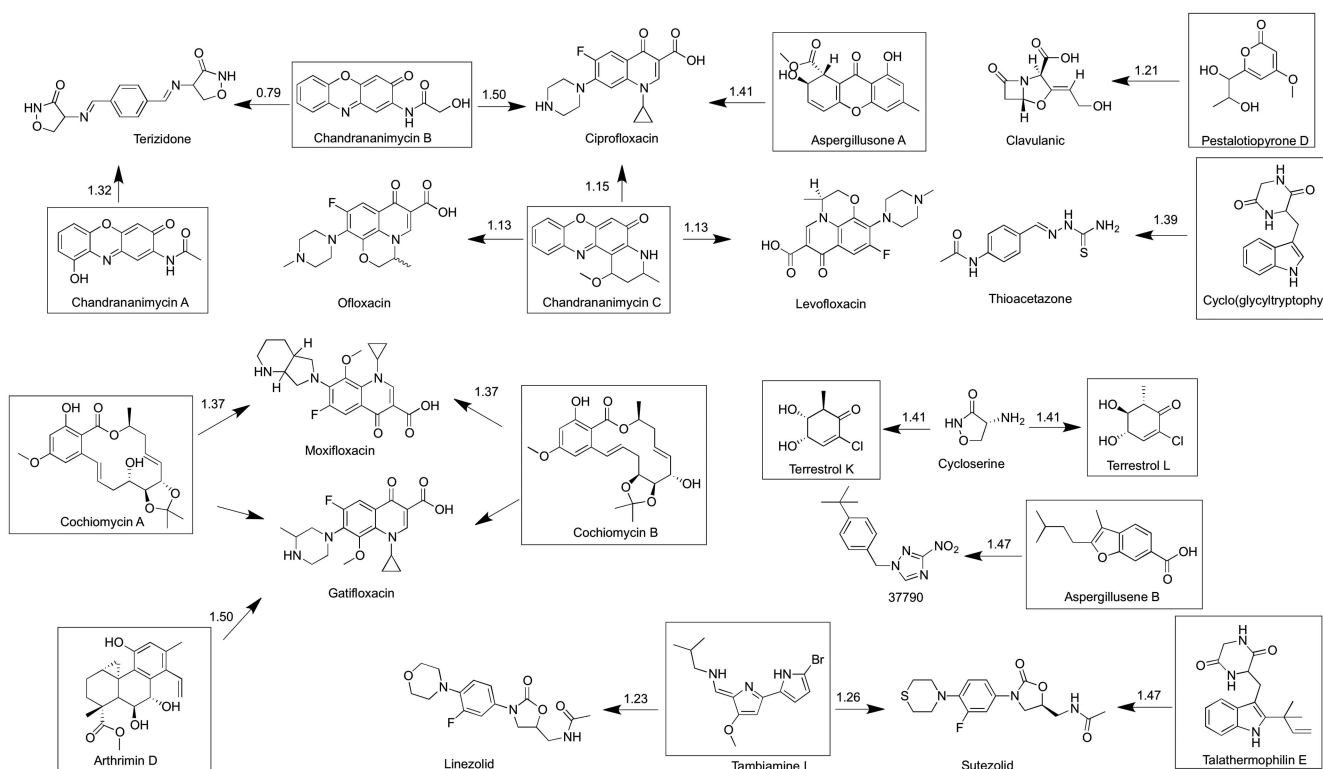


Figure 9 Structures of top 20 shortest EDs of untested natural products/drugs pairs. Untested natural products from marine microbes are in red and TB drugs and candidates are in blue and EDs are given in black numbers. TB, tuberculosis.

natural products and 39 TB drugs and drug candidates with respect to physico-chemical properties and their occupation of chemical space. In physico-chemical space, both sets largely overlapped and defined a region of Chem-GPS-NP space

Further analysis by ChemGPS-NP defined 27 regions within TB active space. Secondly, EDs to known TB drugs may be a better predictive tool. We concluded that if a compound has a ED of <2 to any of the current TB drugs, then the compound has a much higher chance of itself being active.


CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This research was supported by an Australian Research Council (ARC) Linkage Project LP120100485 and the Bill & Melinda Gates Foundation Grand Challenges Explorations Grant Phase II OPP1035218 GCE. M. Liu thanks Griffith University for a GUIPRS scholarship.

- Newman, D. J., Cragg, G. M. & Snader, K. M. Natural products as sources of new drugs over the period 1981-2002. *J. Nat. Prod.* **66**, 1022-1037 (2003).
- Cragg, G. M. & Newman, D. J. Natural products: a continuing source of novel drug leads. *Biochim. Biophys. Acta* **1830**, 3670-3695 (2013).
- Rubiolo, J. A., Alonso, E. & Cagide, E. Marine compounds as a starting point to drugs. *Seafood Freshwater Toxins* **18**, 1141-1178 (2014).
- Damare, S., Singh, P. & Raghukumar, S. Biotechnology of marine fungi. *Prog. Mol. Subcell. Biol.* **53**, 277-297 (2012).
- Liu, X. T. *et al.* Exploring anti-TB leads from natural products library originated from marine microbes and medicinal plants. *Antonie van Leeuwenhoek* **102**, 447-461 (2012).
- Skropeta, D. & Wei, L. Recent advances in deep-sea natural products. *Nat. Prod. Rep.* **31**, 999-1025 (2014).
- Xiong, Z. Q., Wang, J. F., Hao, Y. Y. & Wang, Y. Recent advances in the discovery and development of marine microbial natural products. *Mar. Drugs* **11**, 700-717 (2013).
- Zhao, C. Y., Zhu, T. H. & Zhu, W. M. New marine natural products of microbial origin from 2010 to 2013. *Chinese J. Org. Chem.* **33**, 1195-1234 (2013).
- Larsson, J., Gottfries, J., Bohlin, L. & Backlund, A. Expanding the ChemGPS chemical space with natural products. *J. Nat. Prod.* **68**, 985-991 (2005).
- Rosen, J., Gottfries, J., Muresan, S., Backlund, A. & Oprea, T. I. Novel chemical space exploration via natural products. *J. Med. Chem.* **52**, 1953-1962 (2009).
- Larsson, J., Gottfries, J., Muresan, S. & Backlund, A. ChemGPS-NP: tuned for navigation in biologically relevant chemical space. *J. Nat. Prod.* **70**, 789-794 (2007).
- Rosen, J. *et al.* ChemGPS-NPWeb: chemical space navigation online. *J. Comput. Aid. Mol. Des.* **23**, 253-259 (2009).

 This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

- WHO Global Tuberculosis Report (2015). 1-204 http://www.who.int/tb/publications/global_report/en/. Accessed 12 June 2016.
- Katsuno, K. *et al.* Hit and lead criteria in drug discovery for infectious diseases of the developing world. *Nat. Rev. Drug Discov.* **14**, 751-758 (2015).
- Nguyen, L. & Pieters, J. Mycobacterial subversion of chemotherapeutic reagents and host defense tactics: challenges in tuberculosis drug development. *Annu. Rev. Pharmacol. Toxicol.* **49**, 427-453 (2009).
- Nikaido, H. Preventing drug access to targets: cell surface permeability barriers and active efflux in bacteria. *Semin. Cell Dev. Biol.* **12**, 215-223 (2001).
- Mesfin, Y. M., Hailemariam, D., Biadgign, S. & Kibret, K. T. Association between HIV/AIDS and multi-drug resistance tuberculosis: a systematic review and meta-analysis. *PLOS ONE* **9**, 89-97 (2014).

© The Author(s) 2016