Structure-activity relationship analysis of supramolecular antimicrobials

Kaylee Patel (The University of Manchester)
kaylee.patel@hotmail.com

Dr. Cally Haynes (University College London), Dr. Jennifer Hiscock (University of Kent) & Samuel Munday (University of Southampton)

Sponsored by:
optibrium
dotmatics knowledge solutions

1. INTRODUCTION

Supramolecular **S**elf-Associating **A**mphiphiles (SSAs) are a class of novel antimicrobials. The mechanism of action is based on favourable **complexation events** between the **anionic** component and the **phospholipid head groups** within the bacterial cell wall

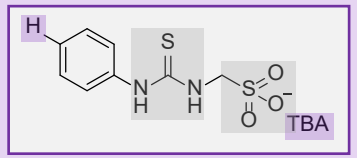

Figure 1: General chemical structure of SSAs. Where *purple* shows the variable groups and *grey* show the important backbone features[1]

2. KEY AIMS & OBJECTIVES

Analyse data to understand trends between **antimicrobial efficacy** and SSA **structural parameters**

Produce a predictive model using **artificial intelligence** and **machine learning** techniques

3. METHODOLOGY

Python programming language

All data analyses were performed on **subsets** of data, split on the basis of their tendency to form **dimers**, **'tapes'** or **'stack'** structures through hydrogen bonding

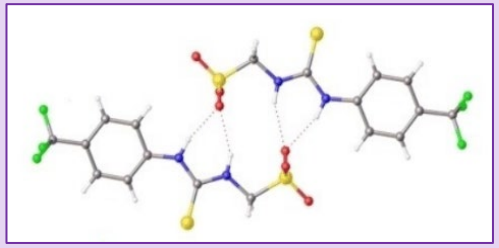

Figure 2: Structural representation of SSA dimers – hydrogen bonding between urea and anionic group shown by grey lines[1]

Linear regression
- Percentage growth of bacteria after 1000 minutes
- Inhibitory concentration for 50% inhibition (IC50)

Key **structural parameters** analysed
- Critical micelle concentration (c.m.c)
- Zeta potential
- Energy minima & maxima of anionic component

4. RESULTS

Structural parameter and respective bacteria % growth was measured with	R-value
C.M.C with E.coli	0.43
C.M.C with MRSA	0.503
Zeta potential with E.coli	0.555
Zeta potential with MRSA	0.558
E min (anionic) with E.coli	-0.015
E min (anionic) with MRSA	-0.14
E max (anionic) with E.coli	-0.71
E max (anionic) with MRSA	-0.315

Table 1: Linear regression output against single parameters. Where *dark green* represents a very good linear fit, *light green* represents good fit, *orange* represents moderate-low fit and *red* represents a poor fit. Other non-linear methods will be explored for poor and moderate-low fits

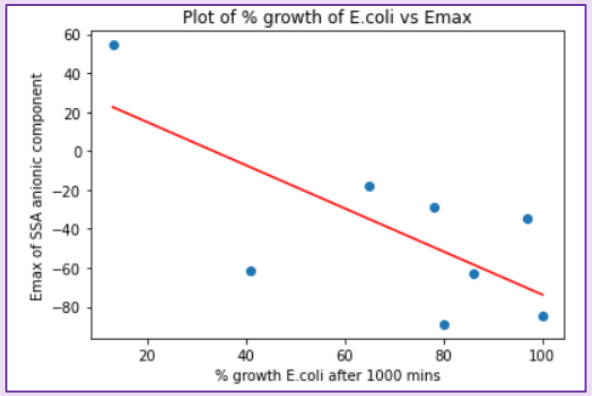

Figure 3: An example plot of energy minima of the anionic component (y-axis) and percentage growth of E.coli after 1000 minutes (x-axis. R=-0.71)

5. FUTURE WORK

Logistic regression
- 1 = antibacterial activity
- 0 = 'no' antibacterial activity

QM/MM Calculations

Random Forest machine learning algorithm

MOPAC computational chemistry package

Long-term aim: To be able to produce novel antimicrobial technology

The predictive model produced will aid the *de novo* design of next-generation SSAs, with high levels of antimicrobial efficacy. At the moment, SSAs with randomly varied R-groups have to be synthesised in the laboratory then tested for efficacy. However, this model can be used to predict R-groups to maximise activity

REFERENCE:
1. N. A. Allen, L. J. White, J. E. Boles, G. T. Williams, D. F. Chu, R. J. Ellaby, H. J. Shepherd, K. K. L. Ng, L. R. Blackholly, J. R. Hiscock, *Chem. Med. Chem.* 2020, **15**, p. 2193-2205

PSDS PHYSICAL SCIENCES DATA-SCIENCE SERVICE
RSC INTEREST GROUP CHEMICAL INFORMATION AND COMPUTER APPLICATIONS GROUP
AI³ Science Network+