Breaking through bacterial cell walls with machine learning

University of Dundee
medical research scotland

David Helekal
School of Life Sciences, University of Dundee

What is the problem?

Antibiotic resistance has recently emerged as a worldwide healthcare crisis. While there are many new drug candidates in development, most of these molecules are only effective against particular types of bacteria and there are virtually no new candidates to tackle a group of bacteria called Gram-negative bacteria, which are the cause of about 80% of all multi-resistant bacterial infections. The cell wall of these bacteria has an extra outer layer, which presents a formidable barrier to antibiotics, which need to get inside the bacteria to reach their target. As well as making it difficult for antibiotics to get into the cell, bacteria also have a way to eject antibiotics that do get inside them by using highly efficient pumps in their cell wall, called efflux pumps (see Figures 1 and 2). It is, therefore, very difficult to develop effective antibiotics which both get into and stay inside the bacteria for long enough for them to kill the bacteria.

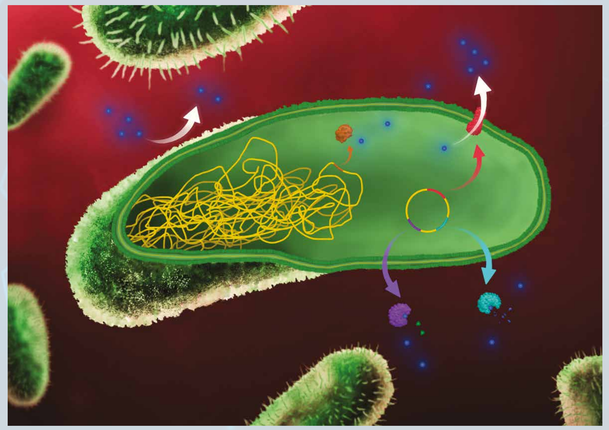

Figure 1 Antibiotics (small dark blue circles) may not be able to get in to Gram-negative bacteria or they may be ejected out of the cell through efflux pumps (shown in red) in the cell wall. Note the layers making up the cell wall. Images created by Cristina Sala Ripoll.

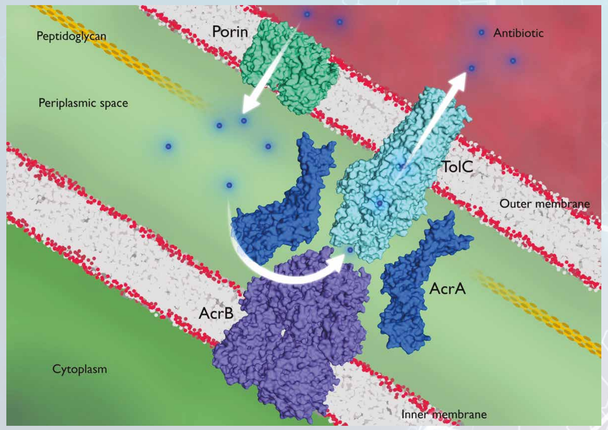

Figure 2 A computer generated image showing the wall of a Gram-negative bacterium. Antibiotics (small blue circles) may get into bacteria via a protein called porin, but they are often ejected from the bacteria by efflux pumps that cross the cell wall (shown in blue). Images created by Cristina Sala Ripoll.

What are we interested in?

We want to understand and predict the features of drugs that enable them to permeate the Gram-negative bacterial cell wall and to evade the action of efflux pumps. We wanted to use computational tools to analyse different molecules in order to design a filter which would tell us which candidate molecules are unlikely to be effective antibiotics against Gram-negative bacteria.

What did we do?

The details regarding inward permeation and outward efflux of molecules are extremely complex and poorly understood, so we decided to make use of machine learning methods in order to identify potential patterns within the molecular structures and chemical properties of antibiotic compounds that make them more likely to reach their targets in Gram-negative bacteria.

There are several constraints to tacking the problem in this way, including the suitability of the features of molecules to enable them to be represented in ways required by machine learning, as well as the limited number of antibiotics for which high-quality information is available to use for training a classifier, which would categorise the molecules on the basis of their features. Because of this, we used a technique called kernel learning.

Kernel functions are mathematical objects which allow us to embed some arbitrary object, in our case a molecular structure, in such a way so that we can compute distances and “angles” between two different molecules. A kernel effectively tells us how ‘similar’ to each other two objects are, i.e. 0 if they are completely different, and the higher the number the more similar they are.

What did we find?

We found that we were able to construct a program which takes a structure description of a molecule in the form of a one-dimensional string of characters (SMILES code), computes various features integrating data from models related to the structure and charge profile of the compound, and then predicts how likely this molecule is to reach its intended target within Gram-negative bacteria.

What does this mean?

We hope that our machine learning system can be further refined and applied to analyse a collection of candidate drugs to identify potential candidate molecules that should be effective against a range of different bacteria. We may also be able to use our machine learning system to derive structural patterns within molecules which determine whether they are capable of getting inside Gram-negative bacteria and evading efflux so that the greater scientific community can use this knowledge in their research.

Who am I?

I am a 3rd year mathematics student at the University of Dundee. Originally, I was enrolled in a mathematical biology course, however later on in my degree I decided to fully focus on applied mathematics. I’ve chosen mathematics because I believe it can help us model, explain and understand complex biological processes which are otherwise impenetrable.

David was supervised during his Vacation Scholarship project, *“Breaking the membrane permeability barrier for antimicrobials in Gram-negative bacteria”* by Dr Ulrich Zachariae.