Machine learning methods are becoming increasingly used as predictive tools in biology. By learning the relationships between an ‘input’ — such as a drug treatment, a genetic perturbation or a DNA sequence variant — and the ‘output’, such as the resultant molecular or organismal phenotype, these methods can be used to predict outputs when faced with new input data that were not in the original training data set. Despite this predictive value, one major challenge is that the machine-learned relationships often remain enigmatic in the model; this ‘black-box’ nature hinders the elucidation of the actual biological mechanisms determining the output phenotypes.

A new study integrates metabolic networks into a machine-learning strategy to understand biological mechanisms underlying antibiotic lethality. Yang et al. sought to understand mechanisms of susceptibility to three major classes of antibiotics: ampicillin (AMP; a β-lactam that inhibits cell wall biosynthesis), ciprofloxacin (CIP; a fluoroquinolone that inhibits bacterial DNA topoisomerases) and gentamicin (GENT; an aminoglycoside that inhibits bacterial protein synthesis). Although target inhibition is commonly understood to be the primary mechanism of action for most antibiotics, the authors focused on cellular metabolism as a secondary means to alter antibiotic sensitivity.

In their screen, the authors tested Escherichia coli cells across ~24,000 conditions, consisting of supplementation with 206 diverse metabolites and a range of concentrations of the antibiotics. They calculated the half-maximal inhibitory concentration (IC50) for each antibiotic in the different metabolic conditions. Instead of using standard black-box machine-learning analyses to predict output antibiotic lethality based on input metabolic conditions, the authors incorporated a genome-scale computational model of E. coli metabolism and metabolic network simulations of the predicted effects of the supplemented metabolites. Thus, the machine-learning analysis was underpinned with mechanistic metabolic information for the conditions being screened, which the authors call a ‘white-box’ machine learning approach.

As this machine-learning strategy yields pathway mechanisms, the investigators quantified the relative contributions of different metabolic pathways to the lethality profiles of each antibiotic. Consistent with the known stress that each of these antibiotics exerts on central carbon metabolism, the tricarboxylic acid (TCA) cycle scored highly as a determinant of toxicity for each of the antibiotics, thus validating the approach.

Shedding light on less-well-characterized mechanisms of antibiotic sensitivity, the team identified the early stages of purine biosynthesis as increasing sensitivity to AMP and CIP but decreasing sensitivity to GENT, which probably reflects the different cellular process targeted by the drugs. These effects were validated through various means: supplementation with purine biosynthesis metabolites, or genetic mutation or pharmacological inhibition of purine biosynthesis enzymes had the expected differential effects on GENT versus AMP and CIP lethality. Further metabolic simulations resulted in additional mechanistic predictions, which validated experimentally and led to an overall model whereby antibiotic stress triggers adenine limitation, purine biosynthesis, ATP demand, central carbon metabolism and respiration.

It will be interesting to determine whether nucleotide analogues have clinical value in potentiating antibiotic actions, as well as whether this overall network-based white-box strategy can provide useful insights into other aspects of bacterial pathology or for optimizing biotechnology applications. Although the study relied on a genome-scale model of bacterial metabolism, the authors highlight potential applications in human precision medicine, such as in understanding metabolic mechanisms of action of cancer drugs.