# REQUIREMENTS AND STRATEGIES FOR WINNING THE BATTLE AGAINST ANTIBIOTIC RESISTANCE BY ANTISENSE TECHNOLOGY

BY

SEYFULLAH ENES KOTIL

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biophysics and Computational Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Eric Jakobsson, Chair
Professor Steven Blanke
Professor Saurabh Sinha
Associate Professor Lee DeVille
Assistant Professor Thomas Kuhlman

**Abstract**

In chapter 1 we have investigated the different requirements and conditions for efficiency and specificity of antisense molecules. For specific therapy, an antibacterial RNA must be able to distinguish between its designed targets and its off-targets. This distinction is reflected in the binding energy calculations. The major component of efficiency and specificity is uncovered to be the nature of the off-targets. We have made a new thermodynamic based model to explain in-vivo antisense binding. We have shown that it fits previously unexplained experimental data perfectly. The second chapter deals with how to preserve effective therapy in evolving population. The effectiveness of redesigning on resistance is conditioned on rescuing the hybridization affinity. The hybridization affinity can be rescued if the mutations for acquiring resistance were on the target sequence. However there can be mutations elsewhere in the genome that would confer resistance. We have investigated possible therapy strategies to direct the bacteria to take mutations that are on the target. Having multiple entry mechanisms for RNA therapy seems to be the key to directing bacteria towards a sustainable therapy. Third chapter deals with the following: Using antisense therapy to block progression of antibiotic resistance for trimethoprim. Converge bacteria to desired mutations. Using antisense molecules to induce loss of trimethoprim resistance mutations.

In the name of Allah, the Most Gracious and the Most Merciful

Peace be upon his Messenger Muhammad

This thesis is dedicated to my parents, grandparents, children and my wife

## Acknowledgement

I can never acknowledge enough the sacrifices, the guidance, the help and the good wishes made for me. I deeply appreciate the efforts mentioned below and for the ones I have forgotten. This thesis is clearly a collective product of many people.

I thank my parent with outmost respect. I thank my mother, Remziye, for she has provided me her love and prayers. Her care through my life is beyond any manifestation; even in my thesis years, her multiple visits all the way from Turkey to support is beyond any words. I thank my father, Temel, for being my role model; he showed me how to be a man through tough and good times with unmatched resilience and optimism. With care and rigor he clearly molded me to be a scientist and I cannot thank him enough for that. I thank my father-in-law, Suleyman Erkan, for his support. I thank my mother-in-law, Gulay Erkan, for her unending support by visiting us multiple times to support of us when we needed it the most.

With highest respect, I thank my grandparents for their support and prayers; Adem, and Fadime Kotil; Lutfu, Emine and Gulsum Kurt. I would like to thank Dr. Erdogan Gurmen, whom I met during my thesis work. I respectfully thank you for teaching me how to understand the insight of things, and many more that cannot be feasibly mentioned.

With my deepest sincerity I thank my siblings, Bilal Kotil, Kubra Kotil and Meryem Kotil. I thank Bilal Kotil for the great conversations we had and his help on editing my thesis. My younger sister Kubra, I thank you especially for the years you lived with me during my PhD years. My youngest sister, Meryem, I actually apologize to you, because I could not be by your side while you grew up. I hope I can make it up to you. I thank you for bringing so much joy to my life. I also thank my cousin Mehmet Can Menekse whom always cheered me up when ever I needed. He has alleviated a great deal of my stress during my thesis wor

My most special thanks are to my wife. You have given me unending, unconditional support that I desperately needed. I deeply feel the sacrifices you have made. If it wasn't for you, this thesis could not been finished. I have given a promise to you to finish my thesis, which has driven me through thick and thin to finish this thesis. I have been blessed with two wonderful children, Seyyid Omer Kotil and Ayse Kotil; you have given me the strength to keep going. You clearly were the remedy after a long day of work.

I would like to thank my friends, Zeeshan, Mete, Osman, Ersin, Adem, Lokman, whom always showed their support. It is Impossible to mention all of them; I apologize to ones I have forgotten to write here.

When I came University of Illinois, I had great velocity but no direction; I had the intellect but not the finesse; I had the tools but not the plan. I thank, Eric Jakobsson, my advisor for he has not just provided me what I was missing, but he also elevated the good qualities that I had. He has been and he will be a great example of being a great scientist and a great person. I would like to thank our collaborator Stanley Maloy, whom I wish I had even more opportunities to work with, whom has inspired me with his every single conversation. I would like to thank my committee members: Steven Blanke, Saurabh sinha, Lee DeVille, Thomas Kuhlman, for their very valuable comments and direction that they have provided me.

## Table of Contents

# CHAPTER 1: In vivo thermodynamic modeling of RNA hybridization reveals the importance of including off-target effects. Doing so uncovers the design principles for efficient and precise antisense technology for applications from killing pathogenic bacteria to metabolic engineering.

## 1.1 Abstract

Potential applications made available by methods using in vivo nucleic acid hybridizations are numerous and very important. Major ones include PCR, FISH, microarrays, killing pathogenic bacteria and metabolic engineering. In this study we focus on bacteria killing. Two main potential advantages of antisense are design easiness and precision. Design is possible by a combination of direct target-antisense sequence complementarity combined with data-based in silico models accounting for the effect of cellular context on the accessibility of the target sequence. This paper is aimed at understanding, and contributing to the progress of, the in-silico models. A vital application of this understanding is being able to enhance the precision of antisense. We define "precision" as the ability to hit intended targets (on-targets) but not off-targets. Precision is vital for killing pathogens while not killing commensal bacteria for antibacterials and knocking down selected metabolic genes and not others for metabolic engineering.

Central to these applications is to understand in-vivo hybridization. The earliest computations only included hybridization energy between the antisense and its on-target. However there are many more off-targets that can bind the antisense to some degree. In the model presented here, we represent the binding of antisense to all on- and off-targets, as the off-targets "compete" for the antisense with respect to their respective binding energies. With this systems approach, we show agreement with published experimental results that are otherwise

unexplainable. On the basis of the model, both precision and efficiency were explored. The analysis shows that antisense therapy can be both precise and efficient. The key to designing efficient and precise antisense is knowing not only the energetics of on-target binding, but the energy distributions of off-targets for the designed length of antisense. In general, shorter designs are more precise and longer designs are more efficient. However efficiencies and precision is very widely distributed for sequences, on-target binding energies or lengths.

**1.2 Introduction**

Nucleic acid Hybridization based applications are vast and important. Some are PCR (1-2), FISH (3), microarrays (4), Metabolic engineering (5) and antibacterials (6-10). First three of these methods are well established and but the latter two are developing technologies with great promise. Metabolic engineering applications have used de-novo sRNAs to alter gene expression, aimed to improve yield of a product (5). sRNAs for metabolic engineering were made by changing target binding sequence of the wild type sRNA with a complementary sequence of the target gene.

PNAs (7), morpholinos (8), RNA-guided nucleases (9, 10) are variations of antisense technology used as antibacterials. PNAs and morpholinos are nucleic acid variant with peptide backbone. They are generally designed to bind to vicinity of translation start site (11) to occlude ribosome binding. They have been used to selectively kill pathogens (12). Antisense technology in vitro shown to reverse, amikacin resistance, chloromycetin resistance and multidrug efflux in Escherichia coli. RNA-guided nucleases are small fragments of RNAs delivered by phages to be used by CRISPR system to degrade their targets.

Studies addressing antibacterial antisense design are limited. (11) showed the importance targeting the vicinity of TSS, on the contrary (13) showed in in-vitro conditions targeting inside coding regions can also achieve repression. (12) Showed 10 nucleotide long antisense can kill salmonella strain and spare e. coli strain. Studies addressing designing of sRNAs for metabolic engineering and sRNA efficiency are mostly experimental as well. (14) is a methods paper written for how to design a sRNA for metabolic engineering. In this paper there are 2 major design principles, have a 25 nucleotide long binding site and make the binding free energy to be larger than -25 kcal/mol. These principles are extracted from experiments and experience, there were no thermodynamic basis for these selections.

There is a big disagreement between the needed free energies and the repression that it gives. For example, necessary free energy needed to have 99% repression of a target by an antisense is only -11 kcal/mol. -25 kcal/mol would leave 1/2^100 percent free target, which is substantially small (repression would be (1-1/(2^100)) substantially higher) than needed or reported. This disagreement between experiment's moderate outcomes for very large free energies also reported by (15). Hao, Yue, et al et al, showed the in-silico computed free energies fits the data after approximately halving the energies. Comprehensive repression experimentation on RyhB-sodB mutant library were done in (16). They have fitted a heuristic model to experimental outcome versus the in-silico energies. In these studies energetics were used to understand hybridization. However they are only based on-target energy (free energy of antisense binding to its intended target). Theoretical studies are made for similar analogs, like microRNA (17), and sRNA. For miRNA (17) the concentrations effect on hybridization was investigated. They are as well depended on on-targets.

In cells there are many off-targets. It is known that off-targets decreases the efficiency of on-target hybridization (18), this effect is termed as off-target effect. Unfortunately, computational studies that deal with in-vivo binding do not include the effects of off-targets. Studies that include the off-targets are not based on thermodynamics, they are based on bioinformatics (18,19). However bioinformatics basis were limited to computing number of matched targets (off-targets) in (12). (19) investigated off-target effect by bioinformatics and showed correlations with efficiency.

Overall, there is no comprehensible thermodynamics based model that can explain in-vivo nucleic acid hybridization. We have made a model by using in-silico binding energies. In our model, the antisense binding in-vivo is explained by a systems approach. We modelled on and off-targets with first principles and thermodynamics.

Hybridization can be thought as a reaction, thus it would have a ΔG (free energy) associated. Logarithm of equilibrium constant multiplied by a Boltzmann factor is the ΔG (20).

$$e^{-\frac{\Delta G}{B}} = K \; eq.1$$

Where the K is the equilibrium constant, and the B is the Boltzmann factor with temperature. A free energy completely characterizes an equilibrium state of a reaction. There are many web based and standalone programs that can compute the free energy of RNA hybridization (21, 22). Thus these programs provide link between the equilibrium concentrations of the reaction to hybridizing sequences.

Single stranded nucleic acid will fold and hybridize on itself. The secondary structure can occlude the binding site (13). Therefore, for two sequences to hybridize, first the sites of hybridization must be available. Thus, ΔG calculations between hybridizing sequences depend

on three separate energies: Opening energy of the first sequence, Opening energy of the second sequence, hybridization energy between two open sequences. Since the antisense can be thought as short enough to not have secondary structure. Thus the total free energy can be computed as the hybridization energy of the open sequences minus the opening energy of the target mRNA (17).

$$\Delta G_{binding} = \Delta G_{hyb} - \Delta G_{mRNA}\ (eq.1)$$

The binding energies for on and off-targets in inside a cell are computed by TargetRNA2 (23). System of ODEs were constructed to compute the free mRNA levels after the antisense and mRNA production, degradation and hybridization equilibrated. The systems of ODEs are given in methods. Two options are available to compute the steady state solutions, equate the derivatives to 0 and solve the system of non-linear equations, or solve the ODEs until a steady state solution is obtained. We have used the second method for easier numerical stability. The energies of -30 to -40 kcal/mol are very large and tend to incorporate errors. We have rescaled the system to obtain slower solutions.

Based on the presented model, precision (not hitting off-targets) and efficiency (hitting the on-target) of antisense are investigated for their mechanisms and design principles. Efficiency of antisense were reported for various kinds of nucleic acid based therapies (16). Precision is investigated with experiments (24).

In summary, efficiency increases with length and on the contrary precision decreases with length. We put forward the general trends by energetics calculations that include the off-targets. Central to all discussion, we show how precision and efficiency is linked to the off-target effects and how it can be treated mathematically. By first developing the mathematical

framework to study antisense binding to on and off-targets. Furthermore characterizing bounds of precision and efficiency by considering random samples from the genome for different lengths of antisense design (10mer, 15mer, 20mer, 25mer, and 30mer). We show the possibility of classifying antisense as efficient only by their on-target energy. Also, prediction of precision by only measuring on-target repression for increasing antisense concentration. Thus experimentalist can design and check their antisenses only by experimenting on the on-target. Otherwise it would be hard to measure number of off-targets for precision.

## 1.3 Methods

### 1.3.1 Computing binding energies for antisense target in whole genome

For every given length 200 sequences from random genes were extracted and their complements were taken as antisense candidates.. Each antisense was used as an input to the TargetRNA 2 program. Ecoli MG1665 was used in all of the analysis. For each sequence all of the binding partner's energies were recorded. Parameters that were used are as following:

**********

PROGRAM PARAMETERS

| | |
|---|---|
| TargetRNA version: | 2.01 |
| Replicon: | Escherichia coli str. K-12 substr. MG1655 |
| NTs before start codon: | 80 |
| NTs after start codon: | 20 |
| Seed length: | 5 |
| sRNA conservation and accessibility: | false |
| sRNA window size: | 13 |
| mRNA structural accessibility: | true |

| | |
|---|---|
| Interaction region: | 30 |
| Filter size: | 2000 |
| Single target: | - |
| File of candidate targets: | - |
| p-value: | 0.9 |

**********

The seed length was increased if the search could not finalize. The important tail of the distribution was defined as the 200 most negative targets

### 1.3.2 Calculating equilibrium concentration of the hybridization levels

The model we used in this paper closely resembles the model in (26).

$$\frac{dA}{dt} = \alpha_A - \beta_A A - A\left(\sum_{i=1}^{n} k_i\, T_i\right) \text{ sys. of eq. } 1$$

$$\frac{dT_i}{dt} = \gamma_{T_i} - \beta_{T_i} T_i - A\, k_i\, T_i$$

Where A is the antisense concentration. $\alpha_A$, $\beta_A$ are the antisense entry rate and antisense degradation rate respectively. $\gamma_{T_i}$, $\beta_{T_i}$ are the synthesis and degradation rate of the target *i*. Notice that there are many targets. $k_i s$ are the binding efficiencies. $k_i$ are linked to ΔGs by the equation 6.

$$e^{-\frac{\Delta G_i}{B}} = k_i \; eq. 2$$

The equilibrium concentrations for the hybridization are calculated by numerically solving the sys. eq. 1 until the equilibrium is reached.

## 1.4 Results

### 1.4.1 Background and the tail

The TargetRNA 2 program was used to compute binding partners of a given antisense, and their distributions versus energies were computed. In Figure 1 the energy distribution of targets associated with a 20 nucleotide long antisense (AAAACAGCAGGCGCGCGGTA) candidate is given. Targeted genome was E. coli. Target marked with the arrow is the intended on-target. Besides the on-target, other targets are distributed around 3 kcal/mol with a very large variance.

Our goal is to calculate the inhibition of the intended target in conditions where other targets are competing to bind antisense. Only the targets having most negative binding energies could effectively compete for antisense and effect the binding to the on-target. These members are the left tail of the distributions in the Figure 1.1. Starting from the most negative targets, we have computed the sufficient number of targets to consider for accurate simulations. Simulations with each subset were compared to fully detailed computations. These computations were made for different lengths of antisense, 10, 15, 20, 25, and 30. In all of these computations the intracellular concentration of antisense was set at a high concentration, 100 copies per cell. In Figure 1.2, the mean relative errors for subset of targets are plotted. As few as ~20 targets is sufficient to have less than 10% mean relative error for all lengths. Biggest jump in accuracy happens when 20 targets are considered. This means that first 20 targets shows the most of the effect.

In Figure 1.3, quantiles of the energy distributions of first 20 targets for different lengths are given. Mean energies are increasing with length. This means that for longer the antisense

stronger the backgrounds. Besides the analysis of the tails we have used total of 100 targets in our calculations since the error becomes even more accurate, virtually becomes negligible.

### 1.4.2 Validation of the model with published data

To our knowledge there is no comprehensive knockdown studies performed with exogenous antisense. We used data from a close alternative, sRNA knockdown, in particular repression of sodB by RhyB mutant library as presented in Peterman 2014 et al (16).

In this study the RhyB (sRNA) sequences were mutated to assess the efficiencies. To illustrate the importance of the off-targets in computation we have computed fold repression two ways; without off-targets, with off-targets. For each antisense sequence we have computed the binding energies to their on-target and off-targets by TargetRNA2 program (23). In all simulations mRNA level is fixed at 10 copies per cell (25) and sRNA as 5 times more than the mRNA level, 50, (26). In both Figure 4 A and B, The turquoise dots' x axis is the computed repression level and the y axis is the experimental repression. If a model predicts the data perfectly then all of the data points should land on y=x line, which is drawn as the red line. In Figure 1.4 A, calculations made only considering the on-target binding energy, placed data points to right of the red line. This shows the predictions overestimates repression. Root mean squared deviation is computed to be 8.26. In Figure 1.4 B we show the same repression calculations for the full model including off-targets. The agreement with the data is remarkable considering that these results were attained only by first principles and no fitting was made. RMSD for the full model were calculated to be 1.715. It is important to note that besides the 4 data points out of the 52 points our predictions are quite well. These points are predicted to have very high repression compared to experimental outcome.

One observation made by (15) was the effective in-vivo binding energies are half as much as it is computed in-silico. Our model, including off-targets, predicts a similar outcome, the outside observer measuring only the on-target inhibition would see binding energies significantly less than expected. In Figure 1.5 A we have plotted the energies computed by only observing the on-target inhibition for RyhB-sodB system. In reasonable concentrations (upto 100 copies of antisense) the observed energy estimate is ~ -7 kcal/mol while the real value is -11.5 kcal/mol. In Figure 1.5 B titration curves of inhibition that were used to compute the real and the observed energies are given.

### 1.4.3 Efficiency

Designing efficient antisense is of great importance. To understand the factors that affect the efficiency we have taken 200 gene targets for 5 different lengths of antisense for random genes and positions (between 80 nucleotides upstream and 20 nucleotides downstream of the transcription start site) of E. coli. The lengths we considered were 10, 15, 20, 25, and 30. In all of these calculations and any subsequent simulation in this study is made by the same parameters described in validating the model against experimental data, for more details refer to methods section. In Figure 1.6, quantiles of inhibition levels for increasing lengths are plotted. Efficiencies (the y axis in Figure 1.6) are calculated as in (16). Clearly the thermodynamic efficiencies increase for longer antisense. Also, all of sub

sequent analysis will be made with the same antisenses.

Inside cells, targets of the antisense simultaneously compete to bind the antisense. In Figure 1.7 we have plotted the mean free number of antisense molecules inside a cell for increasing lengths of mRNA. For sequences longer than 15 sequences more than 90 percent of

the antisense is bound. This means that there is a fierce competition. To effectively bind the intended target, on-target needs to favorably bind antisense relative to competitors (off-targets).

We have shown the change of the strength of the background for increasing length in Figure 1.3, however the on-target's change and its relative position is given in In Figure 1.8. The change of on-target binding energy and the change of best off-target (as a representative of the background's strength) with respect to length is plotted. The growth rate of the on-target binding energy, with respect to the background, grows faster. This means the on-target moves up in rank of targets for increased length. It is clear that the efficiency is depended on the energy difference of the on-target with respect to the background.

In Figure 1.9, the difference of energy between the on-target and the best off-target are plotted for all sequences versus the repression level. This energy gap, closeness of on-target energy to best off-target Delta G represents the efficiency very well.

For efficient antisense, since the requirement is to have larger binding energy than the background, then designing antisenses that have larger energies then the 10 percent quantile of the best off-target would suffice. Table 1 shows the classification power of choosing a binding energy threshold of -12.74, -15.55, -16.63, -17.435, -19,975 for lengths 10, 15, 20, 25, 30. In this classification an antisense is efficient if it represses its on-target by at least 3 fold. For all length the accuracy is at least 90 percent. Efficiency also depends on the concentration of the antisense. In Figure 1.10, the mean inhibition levels versus the increasing concentration are plotted. Efficiency increases sub-exponentially with increased concentration

### 1.4.4 Off-target binding

Another important feature of an antisense is precision, inhibiting its on-target while not inhibiting off-targets. This feature is very important in applications in metabolic engineering. In Figure 1.10, distributions of number of inhibited off-targets for each length is illustrated. In this figure a target is considered as inhibited after 3 fold repression (87.5 percent). Distributions of number of off-targets inhibited moves towards more inhibition with length. Our efficiency analysis showed that the on-target binding increases with length, this would mean that in total the off-targets must bind less antisense. However in Figure 1.10 total number of inhibited off-targets increase. Thus it means that the antisense binds more preferentially to more limited number of targets as length of antisense increases. In Figure 1.11, percentages of total bound antisense for first 5 off-targets are shown for each length. Antisense binding is shared more equally for shorter sequences compared to longer sequences. Thus shorter antisense dilutes the antisense among off-targets.

In Figure 1.12, we have shown that each energy gap between subsequent off-targets increases with length. Also for each lengths subsequent energies becomes smaller. These larger energy gaps promotes preferential binding to stronger off-targets. Our earlier analysis showed that more the on-target is away from the off-target more preferentially it binds the antisense. The same principle applies here. As first off-target is away from the second off-target as it gets more preferential binding. The same principle extends to subsequent off-targets.

As Efficiency, precision also depends on the concentration of the antisense. In Figure 1.13, the mean numbers of inhibited off-targets versus the increasing concentration are plotted. Number of inhibited off-targets increases with concentration.

### 1.4.5 Efficiency vs. number of inhibited off-targets

Figure 1.14 have the plots indicating the number of efficient antisense for given number of inhibited off-targets (0 being the best). Antisense are considered to be efficient if they repress their on-target more than 3 folds. The subpanels are plotted for increasing antisense concentrations. For concentrations 30 or less, longer sequences are efficient and precise (no off-targets are hit). As the concentration increases longer sequences inhibits off-targets faster. For example for 70 copies of intracellular the 30mers give 10 precise and efficient sequences while 15mer gives 20 precise and efficient sequences. After 90 copies of intracellular antisense concentrations there are almost no efficient and precise antisenses.

### 1.4.6 Experimental remarks

Experimentally measuring efficiency of an antisense is very convenient compared to measuring precision, effected number of off-targets. Because an experimenter must measure all possible off-target levels, or perform a microarray if the antisense decreases its target with degradation. We wanted to compute the precision with easily obtainable measurements. Considering only on-target efficiencies at increasing concentrations, we have been able to train a classifier for classifying whether any off-target is hit while inside antisense concentration was 50 molecules per cell. A tree classification was trained with 5 fold cross validation. In table 2, results of the classifier are given, this tree predicts with 85% accuracy. In Figure 1.15 the tree for the classifier is given. In Figure 1.16, all titration curves are plotted for the two predictors (f(22)-f(10)), and f(28)-f(22)), where f(x) is repression level at x concentration of intracellular antisense. Almost most of the true positives are in a quadrant defined by f(22)-f(10)<-2.3362, f(28)-f(22)>=-0.857236. As seen in the scatter plot inhibition levels at concentrations, 10, 22, 28 molecules of antisense per cell are sufficient for this classifier. This classifier is useful in sense

that the parameters needed to have an efficient classifier is only 3. More elaborate classifiers such as bagged trees were predicting the precision even higher, 90.8% percent accuracy. Although these classifiers give higher accuracy, because they have many parameters, they can be relatively more prone to over-fitting.

In previous works, in vivo binding energies have been computed with relative decrease/increase in protein product amount compared to control. As in Figure 1.5 the estimates computed by the protein amount will be considerably small in reasonable concentration. Either antisense concentrations should be taken to high extremes or addition to protein levels, free antisense amount must be measured.

**1.5 Conclusion**

The previous models of in-vivo hybridizations included only the considerations about the on-target. The off-targets were only studied in terms of bioinformatics. Here, we have investigated the importance of the off-targets on efficiency and specificity. We have made a comprehensive energetics model. This new model explains previously unanswered questions. These unanswered questions are: sRNA hybridization experiments do not fit the expectations, the B in eq 1 seems to be larger then what it is naturally (15). Our model fitted perfectly the data in Peterman et al, Figure 1.4. Also, the same RhyB and SodB were investigated and the repression efficiencies were linked to the on-target energies by rescaling (practically halving energies) the energies (15). Physically this means that the repression is less sensitive to binding energies than expected. Our model provides a mechanistic explanation to this observation. The target that is measured (sodB) is embedded in other binding partners (the background). When the on-target energy changes it moves along the off-targets it competes to bind the antisense. To outside observer the energies seems to be scaled to be less than expected.

The off-targets' effect on efficiency and precision was also examined. First and foremost this analysis shows that, it is possible to have efficient and precise antisenses. Furthermore, to have an effective and precise therapy the nature of off-targets are as important as the on-target. Generally, efficiency increases with length and precision decreases with length. Decrease of precision with length is consistent earlier findings in (27), where they have shown the magnitude of seed binding is correlated with increased off-target effects.

For all the lengths one can find good and bad designs. One of the most striking outcomes of these simulations is that the antisense inside cells are bound to something. The question is, do they bind preferentially or are they diluted among targets? . This saturation was seen for lengths over 10 nucleotides (not including 10). The rule of the thumb is, if a target's binding energy is substantially (about 1-2 kcal) away from its subsequent target then it preferentially binds antisense. This principle is consistent for efficiency and precision. For efficiency, the energy difference between the on-target and the most negative tail of the off-targets is a good predictor. This condition results in preferential binding of the on-target. An interesting observation for efficiency was made, since the efficiency is depended on the tail of the off-targets and tails become stronger for longer desings, the bending energy requirement for antisenses binding to on-targets increase with length. For example, -13 kcal/mol binding energy for 10mer would be more than enough but, for 30mer it would be very small. For precision, the interplay of energy differences between subsequent off-target binding energies defines precision. It was shown that precision decreases for increased mean energy differences between off-targets. The dilution of antisense increases with decreasing length, decreasing mean energy differences.

It is important to note that for sequences longer then 15 (including) precision is achieved by dilution among targets, only for 10 it may be with not binding. For precision inside a single

cell, for example, hitting one metabolic gene and sparing others, the precision by dilution might be enough. However for precision among different cells for specific antibacterials, it may not be. In this case the pool of off-targets would increase to several folds compared to a single cell. As number of sequences to be spared increase the precision will decrease, perhaps shorter sequences like 10 nucleotides might be more beneficial to use. Our analysis shows that there is a length regime were precision can be achieved by not binding. Even if the in-silico computations might be off for determining this length, we know that it is short, we suspect close to 10. Thus for specific antibacterials approaches shorter sequences are more suitable, although finding efficient ones would be hard.

Efficiency and precision were found to be tightly linked to intracellular antisense concentrations. Efficiencies increase with antisense concentrations. However precision also decreases with antisense concentrations. It was discovered that finding efficient and precision antisense for low concentrations were far more likely. This shows the problem that a therapy might run in to if concentrations are not controlled. Precision might decrease substantially.

The best approach to finding efficient and precise antisense is to pick lengths between 10 and 20 nucleotides. For a given gene compute, in-silico, all the targets for sliding windows at vicinity of TSS. Compute respective efficiencies and precision in-silico. Rank for efficiency and precision and take the best 10. Perform repression experiments for increasing antisense concentration. Analyze the titration curves with the classifier provided here for both efficiency and precision. Refine the energy calculations and the classifiers with the experimental outcome and repeat.

It is important to note that composition off-targets are context dependent. The dynamic transcriptome defines the off-targets, however in our analysis we have included all of the possible targets.

### 1.6 Figures

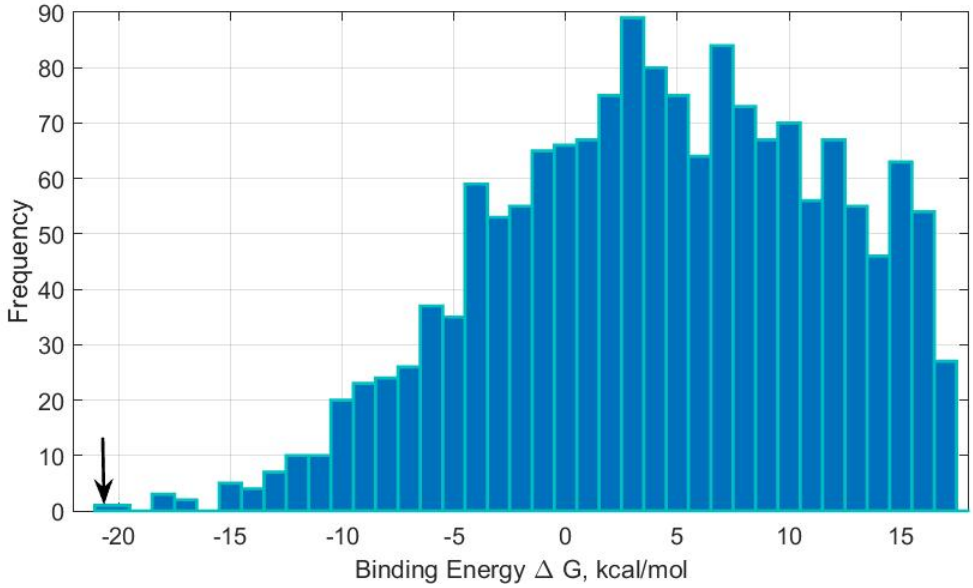


**Figure 1.1.** Energy distribution of an antisense sequence (AAAACAGCAGGCGCGCGGTA).

Distribution of binding energies for all targets to AAAACAGCAGGCGCGCGGTA sequenced antisense. The black arrow marks the on-target which the antisense were designed to bind. The binding energies were computed with TargetRNA2 program (23).

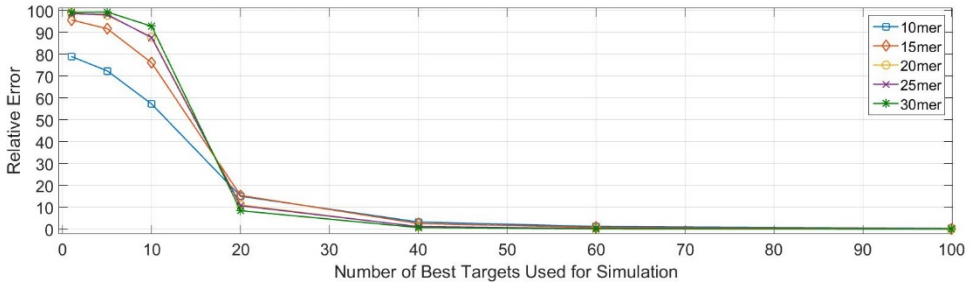


**Figure 1.2.** Considering 20 best targets is sufficient for 10 percent relative error.

Starting from the most negative targets, subsets of best targets were used to compute binding fate for antisense molecules. In these calculations the concentration of antisense is set to be 100 molecules per cell. This high amount is used to have a worse case example. Mean of errors for each computed target were plotted. The accuracies experience a sharp jump at 20 members. The most important effectors are the first 20 members.

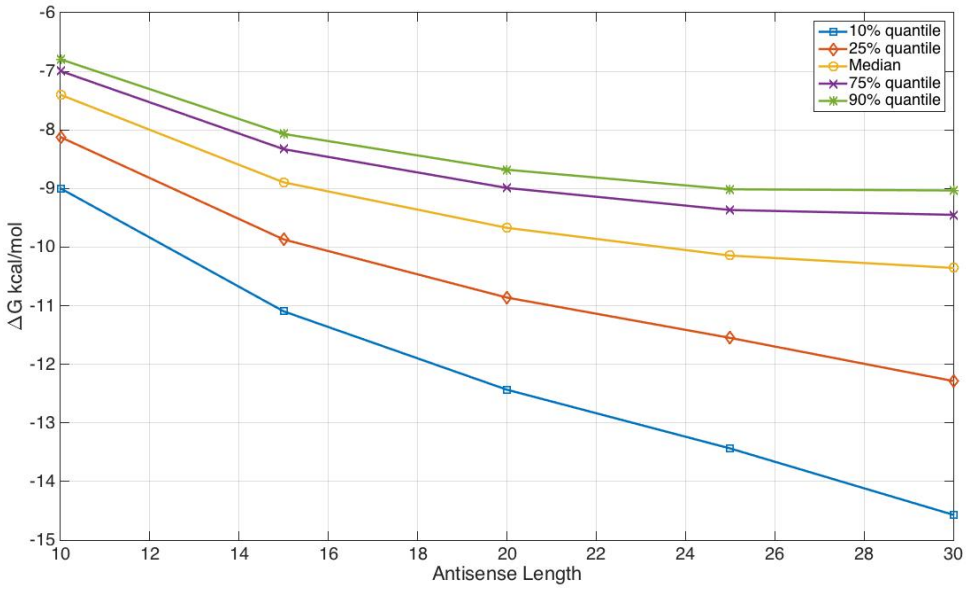


**Figure 1.3.** Longer lengths of antisense have stronger backgrounds.

Energy distributions of the tails for the different lengths of antisenses were analyzed. 10, 25, median, 75, 90 quantiles were marked with square, diamond, circle, cross, star respectivley. Energies of the tails increase with length. This means that, longer the antisense stronger the background in competing with respect to shorter designs.

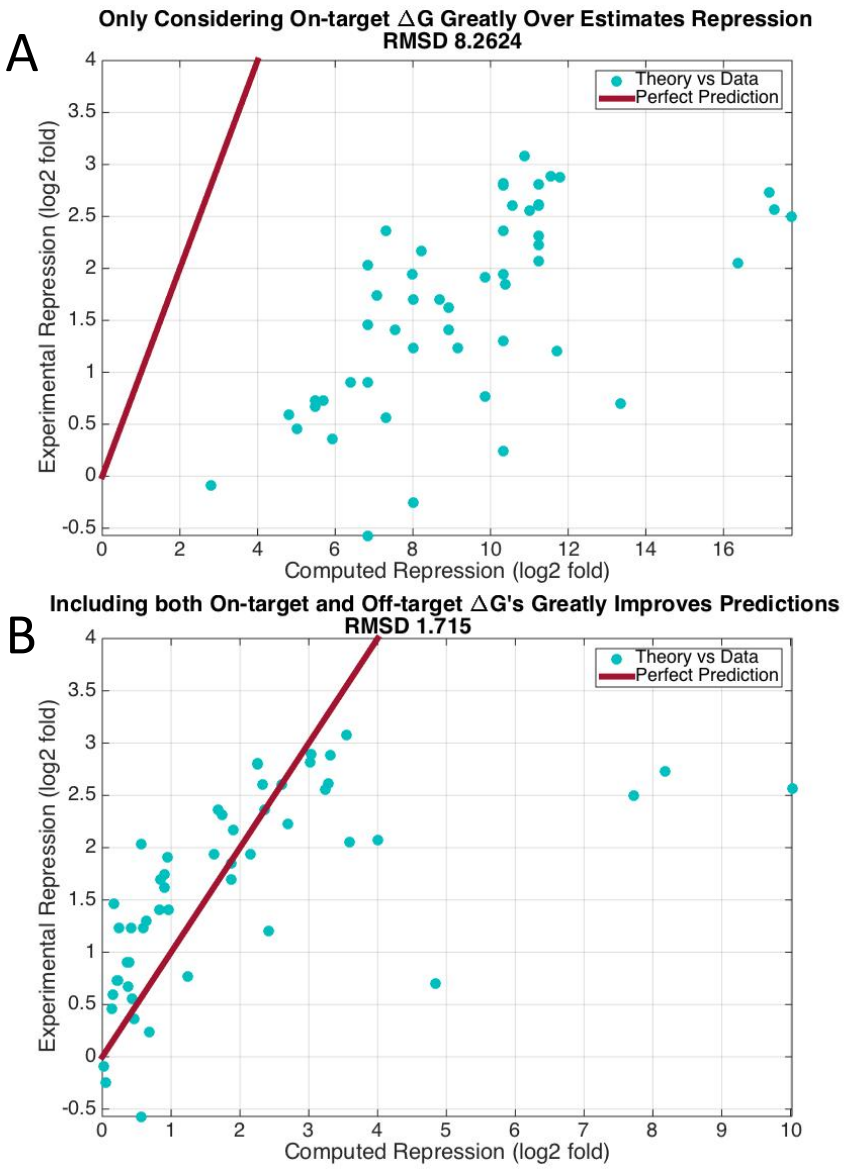


**Figure 1.4.** Computed repression levels for sodB-rhyB system with or without off-targets.

Binding of rhyB to sodB in vivo is modeled from first principles. Computations were made on sequences from mutational study of various rhyB sequences. Binding partners and energies for each rhyB sequence were calculated by TargetRNA2 software. A) Only the binding energy for the on-target (sodB) is used to compute the repression. x axis is the theoretical repression, the y axis is experimental data and the red line indicates the perfect agreement with experiment and data. Using only the on-target energies greatly overestimates the repression amount. Note the range of the x axis.

**Figure 1.4 cont.** The RMSD (root mean square deviation) is very high. B) both the on-target and the off-target binding energies are taken into account. Theory and experiments agree; the RMSD drops to 1.715 . Note that our theory deviates from experiment when the fractional knockdown is the order of $(1-2^{-4})$, approximately 0.938, and higher. However the theory is quite accurate over most of the range of knockdown effectiveness. The difference between subpanels A and B shows that hybridization of the antisense to off-target sequences is the basis for what has previously been introduced as a "scaling factor" in computing antisense hybridization (15).

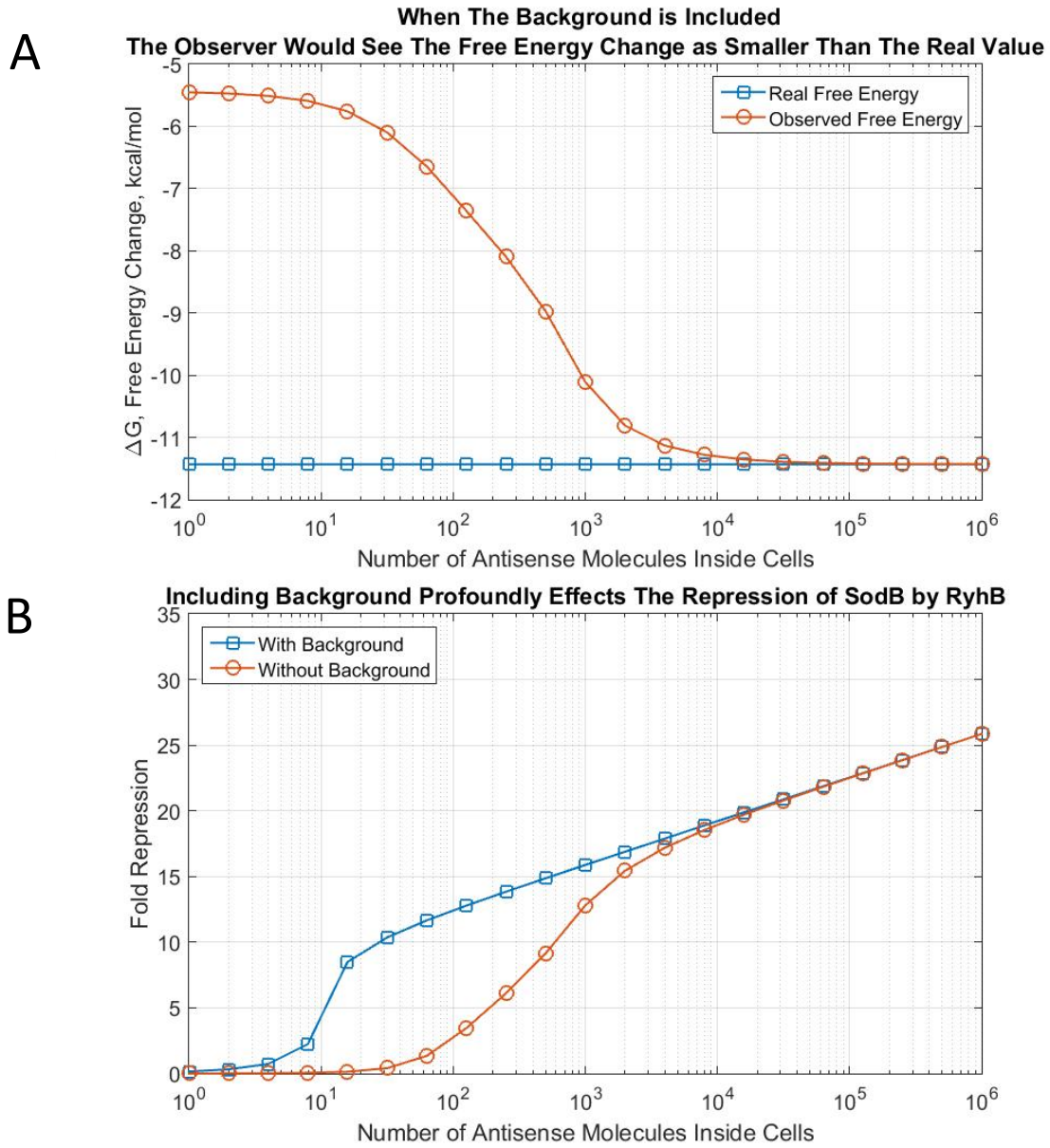


**Figure 1.5.** Hybridization energy changes with respect to observer when off-targets are considered.

Wild type RyhB is used to compute its targets in E. coli by targetRNA2. On-target inhibition were computed with considering the background or without background. A) The real free energy were computed from inhibition titrations without the background. The observed free energies were computed from full computations with the background. This observed free energy is equivalent to experimental outcome. For concentrations of antisense less than 1000 the observed free energy change is substantially different than real value. B) The titration curves for on-target inhibition with background and without

**Figure 1.5 cont.** background are plotted. There is a significant difference of effectiveness up to 1000 antisenses copies per cell.

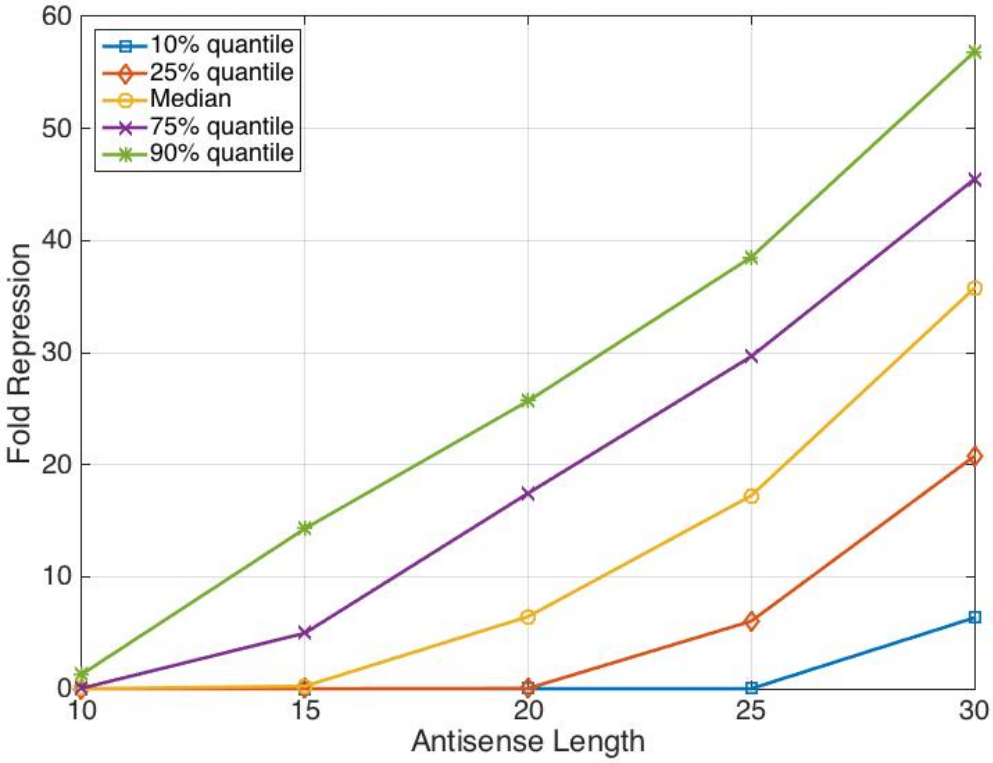


**Figure 1.6.** Fold repression increases with increased length.

200 possible antisenses were designed for E.coli of 5 different lengths. These lengths were 10, 15, 20, 25, 30. For each antisense their on-target repression levels were computed and reported as fold repression (log2(totalmRNA/freemRNA)). For each length 5 different quantiles (10%, 25%, 50%, 75%, 90%) of the distributions of repressions were computed and plotted. Efficiencies increase with increasing length of Antisense.

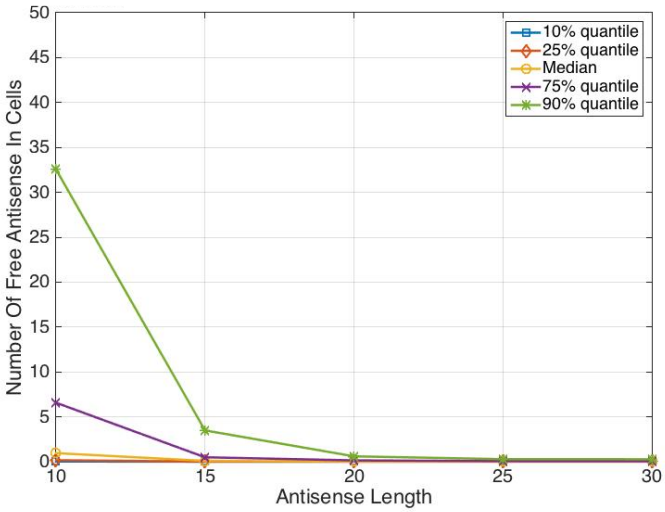


**Figure 1.7.** Antisense longer than 15 nucleotides are in bound state in cells.

In conditions of 50 total antisense inside cells, the free antisense amount were computed for different lengths. As before the conditions of the simulations are given in methods. 10, 25, 50, 75, 90 quantiles were computed and plotted. Antisenses longer than 15 nucleotides are mostly in bound state. Bound to various targets. This means there is substantial competition between the targets, thus an efficient antisense must not be only able to bind to on-target, furthermore preferentially bind to it.

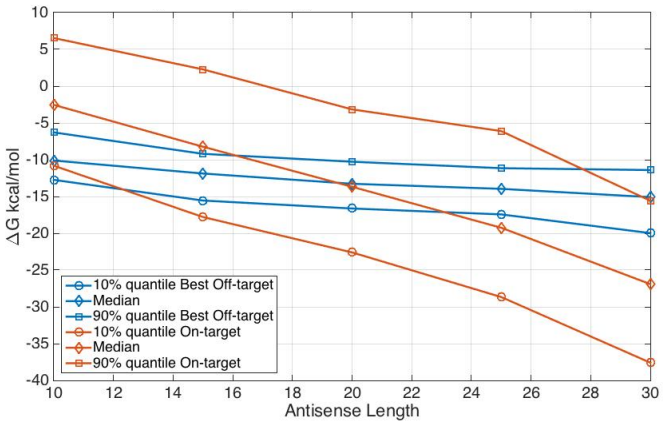


**Figure 1.8.** On-target binding energy grows faster then best off-target.

For all simulations, on-target energy and best off-target energy are compared for all lengths. On-target energies grows (becomes more negative) faster than the best off-target. Median on-targets vary from -10 to -15 kcal/mol while median best off-target changes from -2 to -27 kcal/mol. In short lengths (10, 15) on-median on-target is not more negative than best off-target while in longer sequences (25, 30) median on-target is more negative than the median best off-target. As on-target moves away from the background to be more negative, it binds the antisense more preferably compared to the background. Faster change of the on-target with length is the reason for increased efficiency with longer sequences.

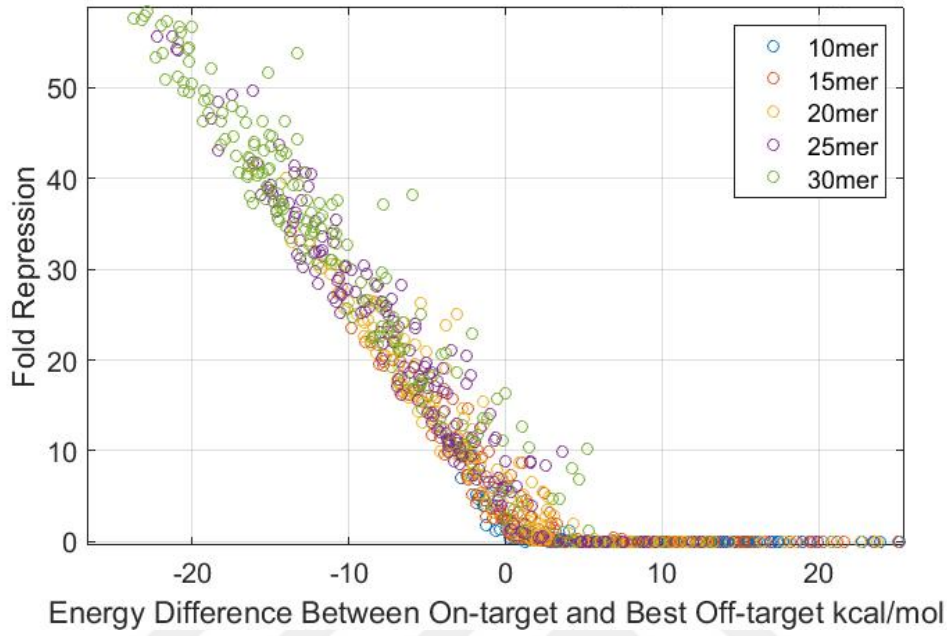


**Figure 1.9.** Fold repression depends on the energy difference between on-target and off-target.

Fold repression versus the energy difference between the on-target and best off-target energy is plotted. This energy difference predicts the efficiency very well. For positive difference, on-target energy is smaller compared to the best off-target, repression is close to zero. As on-target's energy moves away from the best off target then the repression increases. Simulations with different lengths of antisense are plotted with different colors. Although mass of the distributions of increasing length moves towards more repression, the variance is quite high.

| | Predicted as False | Predicted as True | Accuracy |
|---|---|---|---|
| 10 mer False | 183 | 10 | 95 % |
| 10 mer True | 0 | 7 | |
| 15 mer False | 143 | 9 | 90 % |
| 15 mer True | 11 | 37 | |
| 20 mer False | 86 | 8 | 93 % |
| 20 mer False | 6 | 100 | |
| 25 mer False | 43 | 3 | 96 % |
| 25 mer True | 5 | 149 | |
| 30 mer False | 16 | 2 | 96 % |
| 30 mer True | 4 | 178 | |

**Table 1.1.** Accuracy of predicting efficient antisense.

Table of Accuracies for predicting efficiency by mean 10% quantile (mean of second best off-target) of off-targets. For all lengths the accuracy is at least 90 percent.

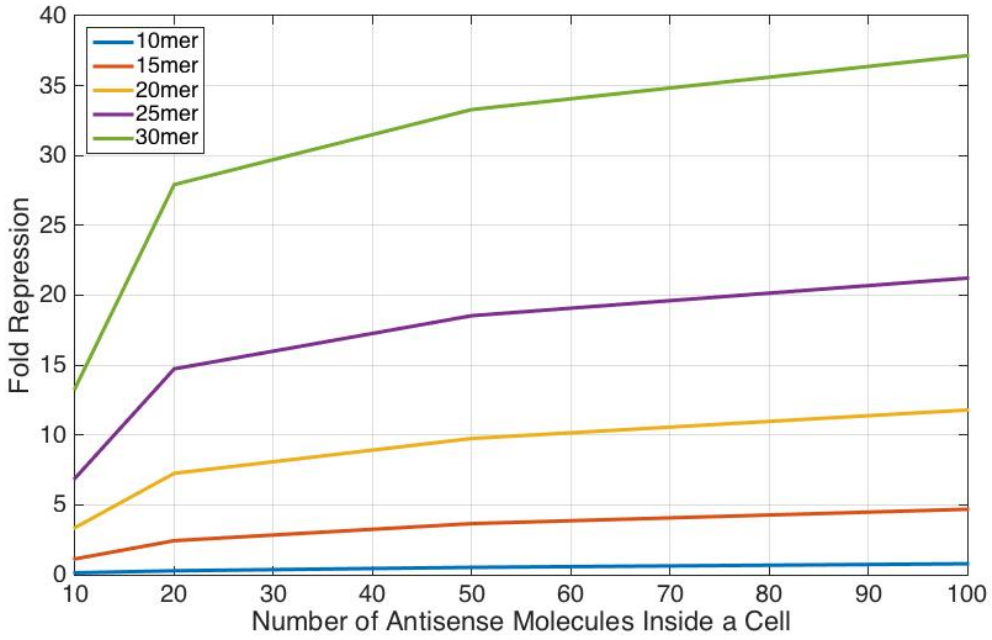


**Figure 1.10.** Repression increases with length.

Mean fold repression for 5 lengths, 10 ,15, 20, 25, 30, for increasing concentration of antisense is plotted. Efficiencies increase monotonically with Concentration. However increase slows down in log scale. Across all concentration longer desings are more effective in repressing the on-target compared to shorter designs.

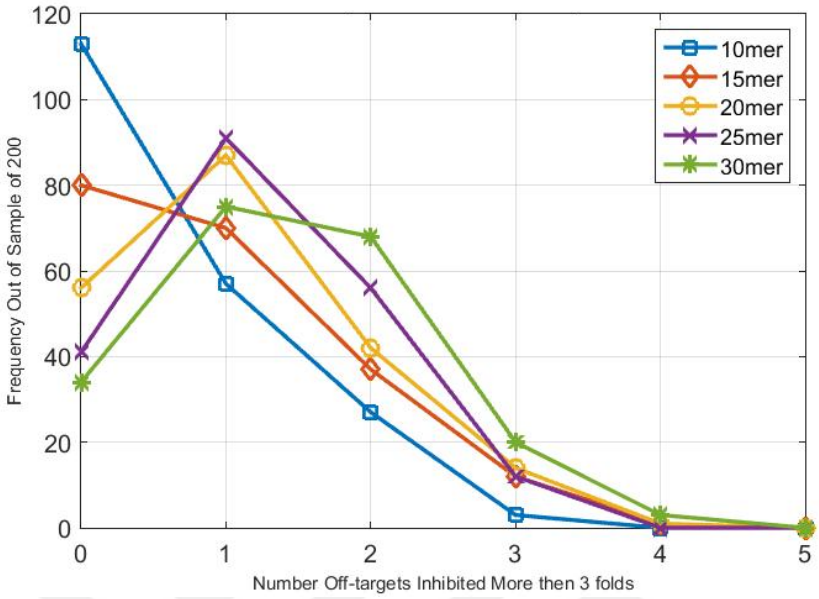


**Figure 1.11.** Off-target hits increases with length.

Distributions of number of inhibited off-targets for 5 lengths, 10, 15, 20, 25, 30 are plotted. Inhibition is taken to be at least 3 fold repression, 87.5%. For 10 mers, most of the trial do not inhibit any off-targets and decreases thereafter. Trend for 15 mers is also same however the mean is shifted to right. 20 mers peak at 1 off-target hit and decreases thereon after. 25 and 30 mers have the same trend as 20 mers however their respective means are more shifted to right.

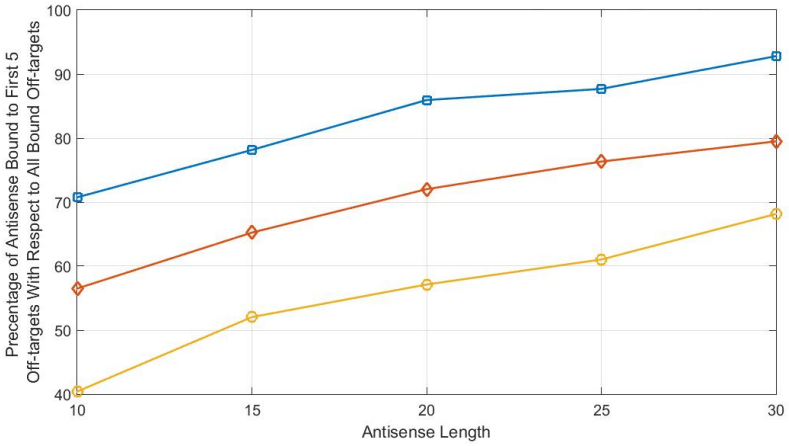


**Figure 1.12.** Antisense binds more preferentially to first 5 targets for longer designs.

Percentage of antisense bound to first 5 off-targets with respect to all bound off-targets were plotted for 5 lengths of design, 10, 15, 20, 25, 30. The conditions of simulations were same as in earlier analysis, more detail can be found at methods. The total inside concentration of antisense is set to be 50. The Median of the distributions vary from 57% to 80%. Thus shorter designs bind less to first five off-targets compared to the longer sequences. Thus shorter antisense gets diluted to many targets compared to longer sequences.

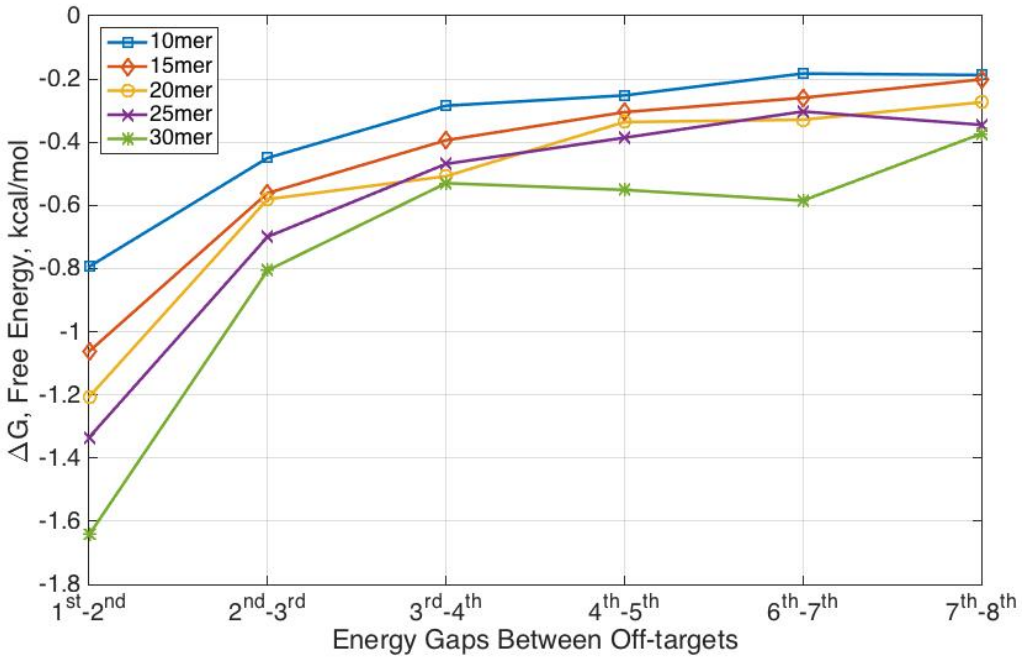


**Figure 1.13.** Energy Gaps between successive off-targets increases with length.

Mean energy gaps between subsequent off-targets were computed. For all lengths the consecutive gaps gets smaller. Largely consistently shorter lengths have smaller gaps between off-targets. As in the case of efficiency, where when the energy gap between the on-target and the best off-target was resulting in preferential binding to on-target, the same principle applies here. Longer energy gaps favors preferential binding.

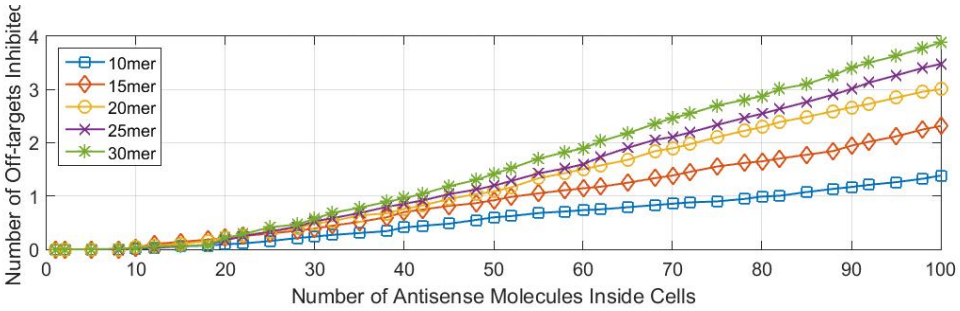


**Figure 1.14.** Inhibited off-targets increase with concentration.

Mean number of inhibited off-targets for each length (10, 15, 20, 25, 30) for increasing antisense concentrations are plotted. The means stay close to zero until 20 copies. Because on-targets for many samples swamps the antisense at low concentrations. After this threshold the means increase almost linearly. Means for longer sequences grows faster. Thus shorter sequences inhibits less targets compared to longer sequences.

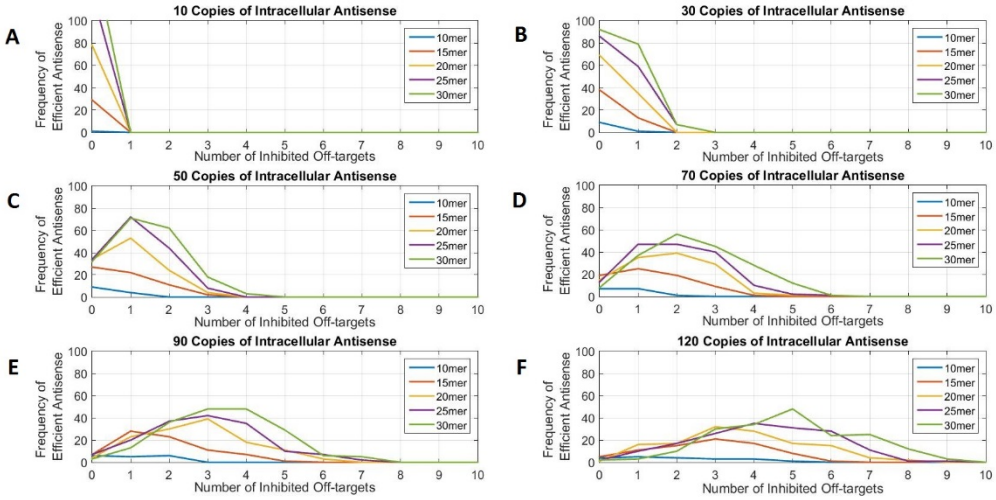


**Figure 1.15.** Distributions of antisense precision for efficient antisense.

For all sequences two results were computed, number off-targets it inhibits and whether if it is efficient (inhibits on-target more than 3 fold). In each panel, distributions of efficient antisense for each level of

**Figure 1.15 cont.** inhibited off-targets are plotted. Each panel has a different intracellular concentration of antisense, 10, 30, 50, 70, 90, 120 for A, B, C, D, E, F respectively.

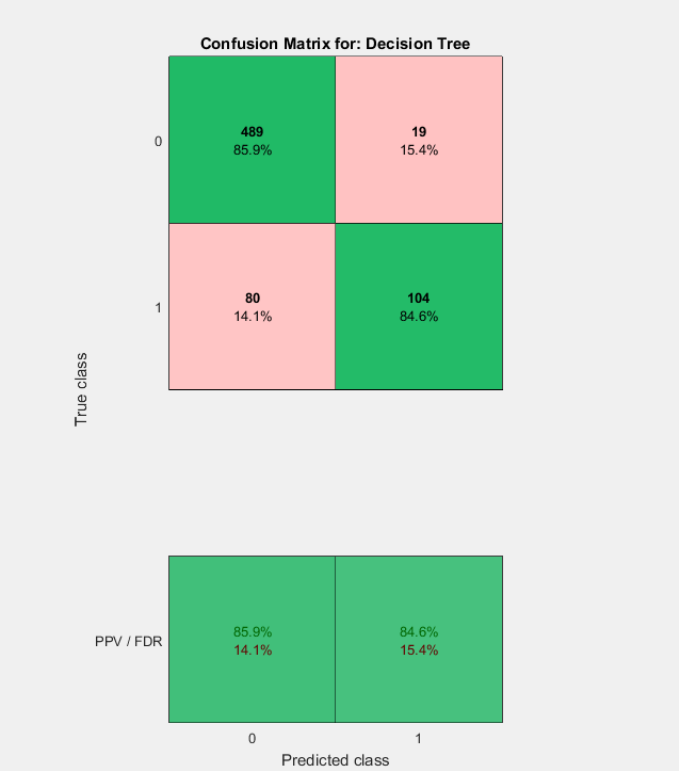


**Table 1.2.** Confusion matrix for classifying precise antisense.

A classifier made for predicting precision from titration curves of on-target inhibition. Precision is defined as not repressing any off-target more than 3 folds. Precise antisenses are labeled as 1 in table. All of the sequences of all lengths are used simultaneously to train and test the data. In this model 5 cross validation was used. After training bagged tree model, the importance of the predictors were computed as frequency of use. According to classifier a precise antisense's titration curve must have the following inequalities as true f(22)-f(10)<-2.3362, f(28)-f(22)>=-0.857236, where f(x) is the repression level at a concentration x.

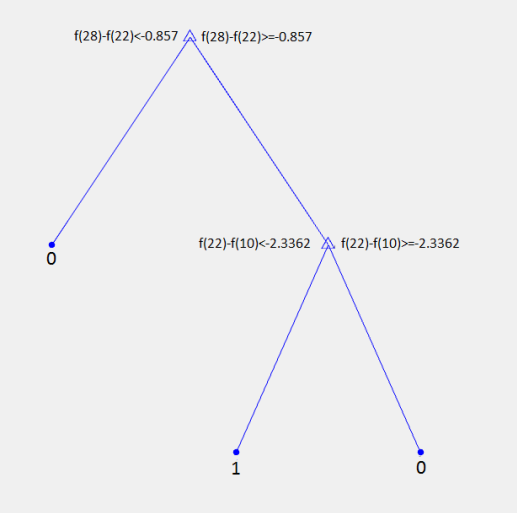


**Figure 1.16.** Classification tree for precision.

Tree that was used to classify the titration curves for predicting precision. Only three repression levels at 10, 22, 28 copies of intracellular antisense.

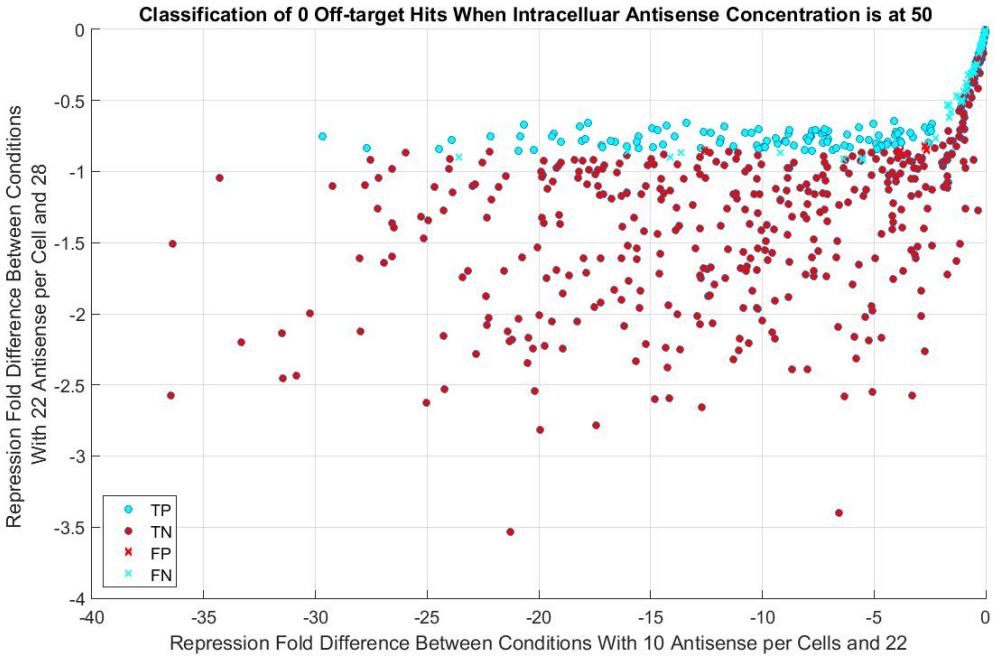


**Figure 1.17.** Use of on-target repression features to classify antisense precision.

When all titration curves are plotted for the two predictors (f(22)-f(10)), and f(28)-f(22)). The True positives are plotted as turquoise dots, the true negatives are red dots, false positives are red crosses and false negatives are turquoise crosses. Most of the true positives are in a quadrant defined by, true f(22)-f(10)<-2.3362, f(28)-f(22)>=-0.857236.

## 1.7 References


1) Saiki, Randall K., et al. "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia." *Science* 230.4732 (1985): 1350-1354.

2) Mullis, Kary B., and Fred A. Faloona. "[21] Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction." *Methods in enzymology* 155 (1987): 335-350.

3) Gall, Joseph G., and Mary Lou Pardue. "Formation and detection of RNA-DNA hybrid molecules in cytological preparations." *Proceedings of the National Academy of Sciences* 63.2 (1969): 378-383.

4) Schena, Mark, et al. "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." *Science* 270.5235 (1995): 467-470.

5) Na, Dokyun, et al. "Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs." *Nature biotechnology* 31.2 (2013): 170-174.

6) Woodford, Neil, and David W. Wareham. "Tackling antibiotic resistance: a dose of common antisense?." *Journal of antimicrobial chemotherapy* 63.2 (2009): 225-229.

7) Good, Liam, and Peter E. Nielsen. "Antisense inhibition of gene expression in bacteria by PNA targeted to mRNA." *Nature biotechnology* 16.4 (1998): 355-358.

8) Greenberg, David E., Kimberly R. Marshall-Batty, Lauren R. Brinster, Kol A. Zarember, Pamela A. Shaw, Brett L. Mellbye, Patrick L. Iversen, Steven M. Holland, and Bruce L. Geller. "Antisense phosphorodiamidate morpholino oligomers targeted to an essential gene inhibit Burkholderia cepacia complex." *Journal of Infectious Diseases* 201, no. 12 (2010): 1822-1830.

9) Citorik, Robert J., Mark Mimee, and Timothy K. Lu. "Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases." *Nature biotechnology* 32, no. 11 (2014): 1141-1145.

10) Silver, Lynn L. "Antibacterials for any target." *Nature biotechnology* 32, no. 11 (2014): 1102-1104.

11) Dryselius, Rikard, et al. "The translation start codon region is sensitive to antisense PNA inhibition in Escherichia coli." *Oligonucleotides* 13.6 (2003): 427-433.

12) Mondhe, Madhav, et al. "Species-selective killing of bacteria by antimicrobial peptide-PNAs." *PloS one* 9.2 (2014): e89082.

13) Shao, Yu, et al. "Rational design and rapid screening of antisense oligonucleotides for prokaryotic gene modulation." *Nucleic acids research* 34.19 (2006): 5660-5669.

14) Yoo, Seung Min, Dokyun Na, and Sang Yup Lee. "Design and use of synthetic regulatory small RNAs to control gene expression in Escherichia coli." *Nature protocols* 8.9 (2013): 1694-1707.

15) Hao, Yue, et al. "Quantifying the sequence–function relation in gene silencing by bacterial small RNAs." *Proceedings of the National Academy of Sciences* 108.30 (2011): 12473-12478.

16) Peterman, Neil, Anat Lavi-Itzkovitz, and Erel Levine. "Large-scale mapping of sequence-function relations in small regulatory RNAs reveals plasticity and modularity." *Nucleic acids research* 42.19 (2014): 12177-12188.

17) Ragan, Chikako, Michael Zuker, and Mark A. Ragan. "Quantitative prediction of miRNA-mRNA interaction based on equilibrium concentrations." *PLoS Comput Biol* 7.2 (2011): e1001090.

18) Worley-Morse, Thomas O., and Claudia K. Gunsch. "A computational analysis of antisense off-targets in prokaryotic organisms." *Genomics* 105, no. 2 (2015): 123-130.

19) Qiu, Shibin, Coen M. Adema, and Terran Lane. "A computational study of off-target effects of RNA interference." *Nucleic acids research* 33, no. 6 (2005): 1834-1847.

20) Dill, Ken, and Sarina Bromberg. *Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience*. Garland Science, 2010

21) Hofacker, Ivo L. "Vienna RNA secondary structure server." *Nucleic acids research* 31, no. 13 (2003): 3429-3431.

22) Zuker, Michael. "Mfold web server for nucleic acid folding and hybridization prediction." *Nucleic acids research* 31, no. 13 (2003): 3406-3415.

23) Kery, Mary Beth, et al. "TargetRNA2: identifying targets of small regulatory RNAs in bacteria." *Nucleic acids research* 42.W1 (2014): W124-W129.

24) Jackson, Aimee L., Steven R. Bartz, Janell Schelter, Sumire V. Kobayashi, Julja Burchard, Mao Mao, Bin Li, Guy Cavet, and Peter S. Linsley. "Expression profiling reveals off-target gene regulation by RNAi." *Nature biotechnology* 21, no. 6 (2003): 635-637.

25) Mackie, George A. "RNase E: at the interface of bacterial RNA processing and decay." *Nature Reviews Microbiology* 11.1 (2013): 45-57.

26) Levine, Erel, Zhongge Zhang, Thomas Kuhlman, and Terence Hwa. "Quantitative characteristics of gene regulation by small RNA." (2007): e229.

27) Ui-Tei, Kumiko, et al. "Thermodynamic stability and Watson–Crick base pairing in the seed duplex are major determinants of the efficiency of the siRNA-based off-target effect." *Nucleic acids research* 36.22 (2008): 7100-7109.

## CHAPTER 2: Using antisense therapy to keep up with bacterial resistance: design principles for therapeutic sustainability

### 2.1 Abstract

Antisense therapy against microbial pathogens has become potentially feasible due to development of delivery vehicles. However, design of antisense therapy requires a strategy to combat anticipated evolved resistance on the part of the pathogen. In this paper we report a simulation model that describes the interaction between patterns of antisense therapy delivery and a population of bacteria evolving resistance to the therapy. Potentially bacterial resistance against antisense therapy can be achieved by two ways; mutating the complementary target sequence (a mutation reducing the effectiveness of a specific antisense), or by enhancing nonspecific resistance mechanisms, such as blocking entry of antisense delivery particles. Of the two, a mutation on the target sequence is the resistance mechanism that can most readily be countered, by modifying the specific antisense sequence. The postulated therapy requires "directing" bacteria to favor attempting to resist therapy by mutating the specific target(s), rather than mutating to resist nonspecific mechanisms. The simulation results suggest: 1) It is possible that a strategy utilizing such direction can prolong the "stalemate" (delay effective development of resistance) against pathogenic bacteria. . It is vital to counter the nonspecific defense system more completely (providing fewer pathways for escape, for example by deploying more than one type of delivery vehicle) than the specific target. To emphasize the latter point, we show simulations in which increasing the relative advantage to the bacteria of mutating the specific knockdown of virulence genes, and decreasing the relative advantage to the bacteria of mutating the nonspecific bacterial resistance pathways, lengthens the time to therapy failure. This

approach appears potentially applicable to cancer cells as they evolve resistance to therapeutic agents.

## 2.2 Introduction

Antibiotic resistance is a major and growing health problem in the United States and around the world, due significantly to overuse of antibiotics in both humans and agricultural animals. [1]. However resistance would arise, albeit more slowly, even with only prudent and appropriate use of antibiotics. It has been estimated that the cost of antibiotic resistance to the US economy is between 20 billion and 35 billion dollars [2]. Hospitals are particularly fertile producers of antibiotic resistant strains, for example MRSA, that also spill out into the larger communities [3]. To exacerbate all of these issues, discovery of new antibiotics is decreasing. Only 12 new antibiotics were approved in the United States between 2000 and 2011. [4]. Keeping up with antibiotic resistance is still a losing battle. [5] Antibiotic research is studied by experiments for identifying molecular mechanisms of resistance [6]. Complementarily, mathematical models have also shown to be effective for predicting and advancement of antibiotic resistance determinants [7,8].

Small molecule antibiotics target essential gene products in bacteria. A molecule has to be in precise shape that could bind to the target of interest and moreover inhibit its activity; the success probability is simply very low. Randomly finding effective antibiotic in this vast space of requirements is very low; thus the discovery rate of antibiotics is very low. An alternative is to rationally design molecules, which requires some level of computational predictability of the effectiveness of designed molecules bound to a variety of targets. Nucleic acids, either free standing or covalently linked to peptides, are an attractive class of rationally designed antibiotics, because complementary sequences in the target organism can be reliably predicted to bind well.

Phosphorodiamidate morpholino oligomers (PMOs) are nucleic acid analogs designed to deliver antisense for knockdown of genes in eukaryotic cells. [9] In addition, morpholinos are further modified by conjugating peptides to form peptide conjugated Phosphorodiamidate morpholino oligomers (PPMOs), which was found to improve cell penetration, targeting and were. used against bacteria [10,11,12,13,14]. PNAs (peptide nucleic acids) are close relatives of morpholinos with a different backbone. They have been used to decrease the resistance of MRSA by hitting the MecA gene [15]. In another approach, RNA-guided nucleases (phagemids) are shown to be selective at killing virulent bacteria and improve survival of infected animals [16,17].

Since designing new antisense sequences that are complementary to essential genes is relatively easy and the sequencing technology is more affordable than ever, in theory repeated redesign of the antisense is an attractive candidate for a sustainable strategy to antibiotic resistance. However this is not the full story. Bacteria can acquire resistance to a potential antisense therapy by three ways: denying entry, mutating the target, and removing the drug from inside of the cell. Furthermore these three resistance mechanisms can be grouped in to two classes, mutations on the specific target sequence and mutations on general mechanisms of resistance, non-specific mutations. The component that can be redesigned most easily is the nucleic acid sequence of the antisense. Mutations at places other then the target puts the antisense therapy in the same position as the conventional antibiotics; that is, requiring a difficult molecular redesign. Since acquiring resistance to the delivery vehicle would have greater survival value for the bacteria than mutating the sequence complementary to the antisense, that would be, the natural course of resistance acquirement . In fact, mutations have been isolated which render PPMOs and PNAs ineffective [18,19]. These mutations happen on an inner

membrane transporter, sbmA, which has been shown to be important for entry of PPMOs and PNAs. The strategy for alleviating this resistance involves changing the peptide portion of the PPMOs.

We modeled the bacterial resistance to antisense therapy by including both escape routes, mutations on target sequence (specific mutation), mutations that impair entry mechanisms (non-specific mutations). Failure effect of entry mechanisms on the system is modeled by decrease in effective antibiotic concentration in cells, whereas specific mutations were chosen by two ways; 1) randomly selecting, 2) parameterization by a polynomial. This set up was used to construct fitness landscapes of antisense resistance. With the landscape we constructed we ask the following questions, what are the parameters for an optimized therapy such that antisense therapy will last longer.

We show that if the antisense therapy is implemented without a strategy to overcome non-specific resistance, then the bacteria will mutate other components than the target itself, which would stifle the potential of nucleic acid therapy. We propose below such a strategy, which involves making mutations to the specific target most beneficial to the bacteria in the short term, thereby inducing the bacteria to preferentially acquire this type of mutation. This strategy will use multiple independent entry mechanisms to "direct" the mutations towards target sequence. This will not necessarily eliminate resistance acquirement altogether, but will "direct" it to a path where we could at least keep up. This new paradigm is analogous to letting bacterial resistance chase us. In meantime the entry mechanisms will be spared. When the target sequence mutates to a certain threshold we will redesign to have an effective therapy again. This cycle will continue until therapy cannot be sustained, without redesigning the entry mechanisms. As the therapy designer we purposefully alter the evolutionary landscape for designed and directed

directing.

Other authors have constructed fitness landscapes of infectious agents. [20, 21, 22]. Nichol et al [22] used fitness landscapes to design a different approach to thwarting microbial resistance to therapy utilizing conventional antibiotics. They designed a sequence of administrations of conventional antibiotics optimized to prevent the resistance development. In the discussion and summary of this paper we will consider how their approach and ours are complementary.

## 2.3 Methods

### 2.3.1 The Model

In Figure 2.1 kinetic diagram of the mutations of each sub-type to its descendants are shown. As noted earlier, bacteria can get two kinds of mutations, specific (on the antisense target sequence) and non-specific (enabling the bacteria to disable one type of delivery vehicle). We have labeled the subtypes of our model with the mutation number of both kinds, e.g. while (1,0) is a label of one mutation in specific and no non-specific mutations, the (2,3) labeled species has two mutations as specific and 3 mutations as non-specific.

Each mutant in (Figure 2.1 ) has an associated growth rate. The kinetic diagram can be viewed as a surface (Figure 2.2) where its heights are growth rates. Each simulation starts at the origin (0,0) of this surface and through time travels on it. The origin is the wild-type bacteria. Under therapy its growth rate is taken to be 10 percent of maximum growth rate, ($0.1\ h^{-1}$ , doubling time of 415 minutes). Once started each trajectory will choose its way according to the fitness landscape. Choosing specific mutations will travel in *j* direction and choosing non-specific mutations will be in *i* direction. The rule of our simulation includes redesigning the antisense

once the population has regained 70% (0.7 $h^{-1}$ , doubling time of 60 minutes) of maximum growth rate (1 $h^{-1}$, doubling time of 40 minutes). The redesign results in a move from (*i,j*) to (*i*,0). First consider the issue when the bacteria acquire specific mutations that compromise the effectiveness of the antisense. Qualitatively, each trajectory will start at the origin (0,0) and then become (0,1) then (0,2) and so on until bacteria has evolved passed the threshold for redesign. The redesign of the antisense sequences will make bacteria susceptible again. This cycle of failure and redesign ideally can repeat indefinitely, because antisense redesign of the antisense sequences is always readily feasible, no matter how the bacteria have evolved. The issue is very different for non-specific mutations that shield the bacteria from being exposed to the cargo of a particular delivery vehicle. In this case, redesign of the delivery vehicle may be possible, but it is likely neither rapid nor obvious. Rather other delivery vehicles must be deployed.

The equations that are linked to the kinetic diagram in Figure 2.1 are;

$$\frac{\mathrm{d}B_{(i,j)}}{\mathrm{d}t} = R_{(i,j)} \cdot B_{(i,j)} \cdot \varphi + mutations\ (eq, 1)$$

$$mutations = all\ the\ terms\ that\ satisfy\ the\ conditions$$

$\left(m\ R_{(i-1,j)}B_{(i-1,j)}\varphi,\ if\ 0 < i, otherwise\ 0\right)$, Specific mutation from ancestor

$\left(m\ R_{(i,j-1)}B_{(i,j-1)}\varphi, if\ 0 < j, otherwise\ 0\right)$, Non-specific mutation from ancestor

$(-m\ R_{(i,j)}B_{(i,j)}\varphi, if\ \ i < 10, otherwise\ 0)$, Specific mutation to descendant

$(-m\ R_{(i,j)}B_{(i,j)}\varphi, if\ \ j < 10, otherwise\ 0)$, Non-specific mutation to descendant

$$\varphi = \left(\gamma \cdot K - \sum_{j=1}^{n}\sum_{i=1}^{n} R_{(i,j)} \cdot B_{(i,j)} \cdot \varphi\right), (eq\ 2)$$

Where the $R_{(i,j)}$'s are growth rates and $B_{(i,j)}$ is population size. $\varphi$ is the nutrient, $\gamma$ is the conversion factor between concentrations and number of cells, K is the carrying capacity, $m$ is the mutation rate and $M_{ij}$ is the mutation matrix. One simplifying assumption should be noted: We consider the bacteria is only capable of taking a particular mutation against a particular antisense, e.g. there is only one first specific mutant, second specific mutant, etc., and there is one particular mutation that disturbs a entry mechanism. In reality there can be different first specific mutants and different first non-specific mutants and so on. A more completely detailed model would consider a web of mutations rather than a linear sequence of mutations. However we believe the simplified model we present captures the essence of the competition of evolving pathogen vs. evolving therapy. In the future more advanced versions of the model to be used to guide therapy will take these differences into account.

$$R_{(i,j)} = \frac{r_{max}}{1 + \left(\rho_j \cdot f(N)_i\right)}, (eq\ 3)$$

$$\rho_j = \frac{A}{IC50_j}$$

The effect of antibiotics on bacterial growth is modeled as a bacteriostatic agent. The overall growth rate under antibiotic treatment is modeled by a first order Hill function, as in eq 3, where $IC50_j$ is the therapeutic concentration that halves growth rate and $A$ is the antisense concentration. To simplify notation $\rho_j$ is introduced as a normalized antisense concentration with $IC50_j$. $r_{max}$ is the maximum growth rate of the WT and $f(N)$ is a function that reflects the effects of the mutations that inactivate entry mechanisms.

In the specific case of the single entry mechanism, if the entry mechanism is operational the $f(N)$ will be 1 Furthermore it is reasonable to assume that pathogen resistance to that single

mechanism will restore growth rate to almost the native untreated level. In this case we represent $f(N)$ to be ~0 in eq 3. Thus

$$R_{0,1} = \frac{r_{max}}{\left(1 + (\rho_j)(0)\right)} = r_{max}\ eq\ 4$$

The same formulation of the growth rates can be extended to double, triple, quadruple and quintuple entry mechanisms. For multiple entry mechanisms the growth rate equation changes to

$$R_{i,j} = \frac{r_{max}}{1 + \rho_j \left(\frac{\#\ total\ of\ entry\ mechanisms - i}{\#\ total\ of\ entry\ mechanisms}\right)} (eq\ 5)$$

If all entry mechanisms are operational then the parenthesis in eq 5 becomes 1. As each nonspecific mutation takes place,one of the independent entry mechanisms becomes completely ineffective. For example if a therapy strategy has 5 entry mechanism the paranthesis in eq 5 becomes 1, 4/5, 3/5, 2/5, 1/5, 0 for 5,4,3,2,1,0 active entry mechanisms respectively. If all of the entry mechanism are mutated then the paranthetical term becomes ~0 and the denominator becomes 1. Thus the total growth rate equals to maximum rate $r_{max}$ $h^{-1}$ .

To simulate the aforementioned system we must have numerical values for the $R_{i,j}$ *'s*. As mentioned earlier $R_{i,j}$*'s* are the growth rates associated for each mutant under therapeutic influence. For example the $R_{i,0}$*'s* (first column in Figure 2.1) are the mutants only having entry mechanism (non-specific mutations) mutations. $R_{0,j}$ *'s* (first row in Figure 2.1) are the mutants only having mutations that disrupts the binding of the antisense (specific mutations). $R_{i,j}$*s* when i,j>0 are the mutants having both mutations. To define all the growth parameters we have to define $R_{0,j}$*s* (Sub-types where all entry mechanisms are operational). In setting up the

computations, we permitted up to 10 specific mutations based on a semi-intuitive expectation that the bacteria would escape the therapy well before 10 mutations, so there would be no artificial boundary effects. In practice (as described in Results) the simulations always ended with five or fewer mutations before the bacteria completely escaped the therapy. $R_{0,0}$ is the growth rate for the WT (wild type) under antisense influence. For $R_{00}$ we assume a maximum growth rate (no antisense) for the WT to be 1 $h^{-1}$ (equivalent to 42 minutes doubling time). Then we assume the growth rate of WT under therapeutic influence, $R_{0,0}$, to be 0.1 $h^{-1}$ (420 minute doubling time) which is 10 percent of the $r_{max}$. The remaining ten mutants are given a wide range of values in different specific simulations, in order to capture the overall behavior of the system. Specifically we have used two methods to set the 10 parameters.

In our first method, in order to keep our simulations generally applicable, we do not assume any particular growth rate increase for any number of specific mutations,. Rather the function of growth rate inhibition vs. specific mutation is computed with random numbers for every simulation run, as follows: In some runs, a single mutation will be sufficient to completely escape from the specific antisense, while in other runs several mutations may be required. For each run we select 10 random numbers in the range 0.1 to 1 (0.1 $h^{-1}$ is the growth rate of WT under influence and 1 $h^{-1}$ is the maximum growth rate attainable by bacteria), and order them from smallest to largest, $R_{0,j}s$ (specific mutants) . Upon the first mutation, the growth rate increases to the fraction represented by the first number. Each subsequent mutation moves the growth rate to the next number.

Unbiased exploration of the parameter space is used to ensure general applicability. To conveniently describe a particular complex mutation profile, we reduce the parameter space to a single parameter, as follows: We consider polynomials which satisfy two conditions, x=0 (x

being number of mutations) R (growth rate) is 0.1 under effective therapy, and R is 1 for x=10. We construct two polynomials which are mirror images of each other with respect to R=0.9 x+0.1. The order of these polynomials is controlled by a parameter $\omega$. The parameter $\omega$ modifies the form of the curve that describes the inhibition of bacterial growth as a function of specific mutations. If $\omega$ is positive and large, only one or a small number of mutations is sufficient to enable the bacteria to escape the treatment. If $\omega$ is negative and large a larger number of mutations is necessary to enable escape from the treatment. The exact form by which $\omega$ exerts its effect is given in the equation below.

$$R_{i,0} = 1 - 0.9\left(\frac{(10-x)}{10}\right)^{\omega}, for\ i = 0{:}10, \omega \geq 1$$

$$R_{i,0} = 0.1 + 0.9\left(\frac{x}{10}\right)^{-\omega}, for\ i = 0{:}10, \omega < -1 \quad \text{eq. 6}$$

Since we have appointed $R_{0,j}s$, and since the parenthetical term in eq 5 is 1, after rearranging we can derive the $\rho_j$.

$$\rho_j = \frac{r_{max} - R_{0,j}}{R_{0,j}} \text{ eq 7}$$

Thus growth rates for mutants having entry mechanism can be computed after we substitute $\rho_j$ to eq 5.

$$R_{i,j} = \frac{r_{max}}{1 + \frac{r_{max} - R_{0,j}}{R_{0,j}}\left(\frac{\#\ total\ of\ entry\ mechanisms - i}{\#\ total\ of\ entry\ mechanisms}\right)} (eq\ 8)$$

The computational experiment is conducted in series of mixed liquid cultures, where a new culture is inoculated from the previous saturated culture. Goal of the simulation is to

calculate therapy failure time, which we defined as the time at which the entry mechanisms are impaired enough that redesigning the sequence will not sustain the therapy.

In the simulated experiments, the starting point is a bacterial population of 1/20000 of the carrying capacity of the vesicle, which is taken to be 10 optical density (OD). The size of the vesicle is 100 ml, giving $8x10^{11}$ bacteria as the carrying capacity. The initial population is thus $4x10^{6}$ cells. In order to do the entire simulation in the growth phase, there is redilution whenever a simulated flask reaches half of carrying capacity, or $4x10^{11}$ cells. This scale of the volume is selected to have a one-to-one comparison with in-vitro evolution experiments that are performed with conventional antibiotics [23]. All of the following simulations are made on this scale unless stated otherwise, for example where we have also included scales of $10^{14}$ (100 L medium), $10^{16}$ (10000 L medium).

The time scales for the two major processes of the simulation (growth and mutation) differ by a factor of approximately $10^{10}$; mutation is a much more rare event than faithful replication of a cell. Specifically we assume a mutation rate of $5 * 10^{-10}$ per generation per nucleotide [24]. This difference in scales led to different mathematical representation of the two processes. Growth is treated continuously and deterministically, while mutations are introduced stochastically (probabilistically). At each simulation step, first the growth portion of the system is simulated deterministically by a fourth order Runge-Kutta method [25], then the cells that are divided are calculated and the mutants are computed analogous to tau leaping methods [26,27], by a Poisson random number with the rate of mutation. This is valid since the entire nutrient that is used makes up a cell, thus the mass is conserved between the total nutrient and the total number of cells. The time step is taken to be 60 minutes. In addition to mutations probability is also introduced when simulating dilution. In this phase of the simulated experiment, .9999 of the

culture is discarded. Each subpopulation is represented by the appropriate fraction of the entire population superimposed on a Poisson distribution scaled according to the expected number of individual cells. For a dominant subpopulation the variance/mean ratio of the distribution is very small. However for very small subpopulations, the corresponding ratio is quite large, so there is a chance of a significant fractional change in this population. Simulations were terminated after 3000 days, if loss of efficacy of the therapy was not achieved earlier.

### 2.3.2 Representation of trajectories

Simulation trajectories are illustrated with Muller plots, an alternative way of showing an evolutionary trajectory. [28] In Muller plots the y axis is the fractional population, the x axis is the time, and each mutant has its own color. At each time the fractional population of a mutant is represented by the fractional coverage of the graph by its color. For example suppose the total population is half WT and its mutant. Let's assume WT is indicated with black and the mutant by white; then the half would be black and other half would be white. Another advantage is easiness of illustrating the ancestor of a mutant. When a mutant arises, it originates from the color of its ancestor. An example trajectory is given in Figure 2.3, for a very simple system. The population starts with wild type (WT)and there are linear mutations e.g. from WT to first, from first to second and so on. The illustrations of classic representation and the Muller plots are provided in different panels of the same figure.

## 2.4 Results

### 2.4.1 Dependency of sustainable therapy on non-specific mutations

After defining all the growth rates, the general goal is to compute the therapy failure time, which we define as the time when the cumulative growth is restored to 90 percent of the

original maximum growth rate ( 46.7 minute doubling time as opposed to 42 minutes for wild type under therapeutic influence). In mathematical terms $\frac{1}{\sum_{j=0}^{M}\sum_{i=0}^{N} B_{i,j}}\sum_{j=0}^{M}\sum_{i=0}^{N} R_{i,j}B_{i,j} \geq 0.9$. Morover, we want to investigate the change of therapy sustainablity among single, double, triple, quadruple, quintuple, independent entry mechanisms

In order to keep our simulation generally applicable, we appointed the 10 specific mutants unbiasedly by random number generation (for details see methods). This ensures that our results do not depend on any specific mutant profile. As before, simulations start with WT bacteria and through time acquire mutations. Once the growth rate recovers to 75% of the native growth (that is, 56 minutes doubling time) times the native growth rate, the specific antisense is redesigned to suppress the effects of specific mutations. The results of the above procedure are plotted in Figure 2.4, in which the variability of the runs between each other is represented by a histogram for the results of multiple runs at the same condition. In this Figure, times to therapy failure for different independent entry mechanisms were simulated. For single and double independent entry mechanisms, the time to failure seems to be comparable and low; with three independent mechanisms the time to failure is significantly extended.. 5 independent mechanisms gave the best result. More independent mechanisms would undoubtedly have been better. In the particular simulation shown in Figure 2.4, if the pathogen did not achieve complete resistance after 1000 days, the simulation was ended.

### 2.4.2 Dependency of sustainable therapy on specific mutations

As evident from Figure 2.4, although increasing number of entry mechanisms increases the time to failure, there is still significant variation from run to run that is caused by the variation in randomly selected growth rates of specific mutations. To understand the effect of

specific mutation series, we consider simplifying our search through the parameter space by characterizing the series of the mutations by the single parameter $\omega$ , as defined in equation 6. The higher the value of $\omega$, the more is the advantage to the pathogen for each mutation against the specific antisense. By this single parameter we can plot the effect of increasing advantage (to bacteria) of specific mutations for therapy failure.

To show how growth rates depend on $\omega$ we have plotted in Figure 2.5A growth rates (for only specific mutants) for $\omega$ equaling to -11,-5,1, 5, 11, 25 in equation 7.. In this figure the y axis is the growth rate and the x axis corresponds to number mutations. The growth rates that are plotted in Figure 2.5A correspond to the first row of kinetic diagram in Figure 2.1, the surface edge represented in black in Figure 2.3, and the $R_{0,j}$ s in eq 5. In Figure 2.5 B, the growth rates for a corresponding number of entry mechanisms were plotted. These correspond to the first column of kinetic diagram in Figure 2.1, the surface edge represented in red in Figure 2.2 and $R_{i,0}$ s in eq 5. With knowing the $R_{0,i}$ $s$ and $R_{j,0}$ $s$ all of the other growth can be computed with eq 5. The corresponding fitness landscapes are given in, Figure 2.5 C. The computational experiments are carried out directly on these fitness landscapes. For all simulations they start on the origin (0,0)(W,T) of the surfaces.

### 2.4.3 Trajectories

Examples of 4 trajectories are illustrated in figures 6 to 9. To assess the result of increasing/decreasing both of the components ($\omega$ and entry mechanism) we have looked at following combinations. In Figure 2.7, $\omega$=1 with 3 entry mechanisms. In Figure 2.8, $\omega$=1 with 5 entry mechanisms. In Figure 2.9 $\omega$=5 with 3 entry mechanisms and lastly in Figure 2.10, $\omega$=5 with 5 entry mechanisms. Resulting trajectories are presented with Muller plots. An illustrative

example for how Muller plots should be interperated is given in Figure 2.1. In a Muller plot, every sub-species is represented with a different colour. In our system, maximally we need 66 colors; for each of 5 potential entry mechanism failures plus fully operational state has 10 specific mutations plus WT. As illustrated in Figure 2.6, we used a different color for entry mechanism failure state and their shades for specific mutations.

In Figure 2.7, the system is simulated with $\omega$=1 and with 3 independent entry mechanisms. At start of each experiment the system only contains WT (0,0). The initial treatment increases the doubling time from 42 minutes to 420 minutes. Through time WT is replaced by other mutants. When the population doubling time falls to a threshold of 60 minutes, all of the specific mutants ($B_{i,j}$ for all *j)* are nullified by redesigned antisense. For simplicity in the mathematics the population is assumed to be homogenous in the target area for the redesigned antisense, which may be a different region of the genome than the preceding antisense. For each Muller plot there is corresponding plot in supplementary showing the same trajectory but with classic representation (supp. 1-4). Part B shows the doubling time of bacteria. At start of each experiment the doubling time starts from per 400 minutes. As it increases by mutations, it reaches the treshold (per 60 minutes). When treshold is reached, then we redesign the antisense sequence. As noted noted earlier this event nullifies all of the specific mutants. In the meantime, nonspecific mutants (mutants reducing effectiveness of delivery vehicles) continue to accumulate and degrade the performance of the therapy. This cycle continues until the redesign cannot rescue the therapy, because none of the delivery vehicles work anymore. At this point the therapy is doomed, until new delivery vehicles are found. Part C shows dynamics of bacterial populations with different numbers of entry mechanism failures. These trajectories are plotted in log 10 scale for y-axis. The log scale was chosen because this representation

enables us to observe all of the scales at once, made necessary by the fact that the range of size of subpopulations varies by up to $10^{13}$. In progression of time, specific mutants take over the population in the order of (0,1), (0,2), (0,3), (0,4). Each mutant arises from the previous mutant. In the meantime, subpopulations having developed resistance to a single entry mechanism arise on three occasions (15.3 days, 24.6 days, 29.8 days). On first two occasions they eventually die off. At third occasion, the mutant gives arise to a mutant having two of its three entry mechanisms mutated. Eventually third entry mechanism also fails and the therapy halts. In Figure 2.7b, The doubling time starts at 420 minutes per generation. The doubling time decreases (growth rate increases) as specific mutants takeover. The threshold of 60 minutes is reached at 41.62 days. A cargo is redesigned at this time. However the doubling rate does not recover, because entry mechanism have already failed. Soon after the therapy fails. In Figure 2.7c, the entry mechanism failures are seen much more easily. The purpose of the log scale is to monitor all scales at the same time. The first entry mechanism mutant stays low but flactuates. This mutant gives rise to second entry mechanism mutant which takes over at 40 days and paves the way for the final and third mutant. Eventually the therapy fails at 41.8 days.

In Figure 2.8, trajectory of evolution is simulated with parameters $\omega$=1 and 5 independent entry mechanisms. 8A, shows the Muller plot. Until the first redesign specific mutants takes over the population in the order of (0,1), (0,2), (0,3), (0,4), (0,5), (0,6), (0,7). Each mutant arises from the previous mutant. Single entry mechanism mutants arises at (39 days and 53 days). On both occasions they eventually die off. Qualitatively the same cycle reoccurs six more times. After $7^{th}$ redesign second entry mechanism fails, eventually third, fourth and fifth entry mechanism fail at 460 days. It is particularly interesting that the single entry mechanism mutant is not detectable before double entry mechanism mutant arises.. This phenomenan of skipping an intermediate

mutant has been called stochastic tunneling in the context of cancer genetics [29]. In our simulations, as in other systems, stochastic tunneling happens because the number of cells in the simulation is higher order compared to the reciprocal of the mutation rate. At 463.46 days, the last redesign is made, but the system is beyond recovery by cargo redesign. The therapy eventually fails at 463.66 days. In Figure 2.8b, change of doubling time through course of the simulation is plotted. The doubling time starts at 420 minutes per generation. The doubling time decreases (growth rate increases) as specific mutants take over. The threshold of 60 minutes is reached at 65 days. Cargo is redesigned at this time and doubling rate is recovered to per 420 minutes. This cycle repeats 6 more times. In Figure 2.8c, the entry mechanism mutants are plotted with log scale. This representation enables us to spot the stochastic tunneling that happened in this trajectory. The first entry mechanism mutant stays low but fluctuates. Many times the second entry mechanism mutates but eventually dies. At 410 days the second entry mechanism mutant arises from the low single mutant and increases to take over the population. Eventually all of the entry mechanisms fails.

In Figure 2.9, trajectory of evolution is simulated with parameters $\omega$=5 and 3 independent entry mechanisms. Figure 2.9A, shows the Muller plot. Until the first redesign, specific mutants take over the population in the order of (0,1), (0,2). Just before the redesign, (0,3) grows to substantial amount. Each mutant arises from the previous mutant, e.g (0,1) arises from (0,0), etc. Qualitatively change of population composition are similar between redesign cycles until the $9^{th}$ redesign (105 days). Single entry mechanism mutants arises sporadically until the $9^{th}$ redesign and dies off. Just before the $9^{th}$ redesign, one of the entry mechanisms fails and it is conserved after the redesign. After this time, the population becomes dominated with mutants having one failed entry mechanism. For the second entry mechanism failure same occurence happens at the

$10^{th}$ redesign. Soon after third mechanism also fails and the therapy fails at 113.2 days. In Figure 2.9b, doubling time versus time is plotted. As usual doubling time starts at 420 minutes per generation. The doubling time decreases (growth rate increases) as mutants takeover. The threshold of 60 minutes is reached at 10 days. Cargo is redesigned at this time and doubling rate is recovered to per 420 minutes. This cycle repeats 10 more times, after which redesign cannot rescue the therapy. In Figure 2.9c, the entry mechanism mutants are plotted with log scale. The first entry mechanism arises sporadically until 110 days, at this time it dominates the population. At 113.2 days the third entry mechanism mutant arises and the therapy fails.

In Figure 2.10, trajectory of evolution is simulated with parameters $\omega$=5 and 5 independent entry mechanisms. Figure 2.10A, shows the Muller plot. Until the first redesign, composition of the populations changes as in Figure 2.9A. These sequence of events continues until 990 days. After this time the population loses one entry mechanism. Now the specific mutations accumulate on the population lacking one entry mechanism. Thus redesign can partially restore the population. However this time the starting population has only four out of five entry mechanisms. As before, the second entry mechanism mutates and dominates the population. During this time redesign can still partially restore the doubling rate. Eventually third mechanism fails and very rapidly the fourth and even more rapidly the fifth entry mechanism fails. In Figure 2.10b, doubling time versus time is plotted, the time range is taken to be greater than 900 days. The dynamics of doubling rates before 900 days resembles dynamics shown between 900 days and 970 days. As usual doubling time starts at 420 minutes per generation. The doubling time decreases (growth rate increases) as mutants takeover. The threshold of 60 minutes is reached per 10 days. Cargo is redesigned at this time and doubling rate is recovered to per 420 minutes. This cycle repeats until 1000 days; after this time redesigning the cargo

becomes partially effective. At this time population has lost its first entry mechanism. Redesign cycles are similar until failure of the second entry mechanism. After second entry mechanism fails redesigning becomes even less effective and eventually the fifth entry mechanism fails and therapy halts. In Figure 2.10c, the entry mechanism mutants are plotted with log scale. The first entry mechanism arises sporadically until 1000 days, at which time it dominates the population. At 1100 days the second entry mechanism mutant arises and dominates the population. Soon after the fourth and yet faster the fifth entry mechanism fails.

### 2.4.4 Mean time to therapy failure

The presented trajectories were a small sample of the considered parameter range. We have varied $\omega$ from -11 to 25 by increments of 2 and varied number of entry mechanisms from 1 to 5. Results of mean time to failure are illustrated in Figure 2.11. It is evident that mean time to therapy failure increases with number of independent entry mechanisms, for any fixed number of $\omega$. For single and double entry mechanisms the mean day of failure is virtually very low. The biggest jump happens from 2 independent entry mechanism to 3. After 3 independent mechanism the trend is consistent and the mean failure time increases and increases. In the direction of $\omega$ the mean failure time also increases. This suggest that, an effective therapy should have at least 3 independent mechanisms and specific mutations should give high advantages even with single mutation.

### 2.4.5 Effect of bacterial populations size on therapy failure

All of the preceding analysis was made with population sizes that would fit 100 ml rich medium. This was intentional since in-vitro evolution experiments are done with this configuration and they provide a direct comparison. Modeling the antibiotic resistance in its full

realism is very complex. Two of the complexities would be, in-vivo modeling of immune system, and population scale modeling with transmission probabilities, demographics etc. To show the effect of population size we made simulations in bigger volumes, 100 L and 10000 L.

Results of the mean failure time for 100 L medium is given in Figure 2.12. For any fixed W, more independent mechanisms provides more sustainability. In the direction of increasing W therapy failure time also increases. Results of the mean failure time for 10000 L medium is given in Figure 2.13. Again, for any fixed W, more independent mechanisms provides more sustainability and also W increases therapy failure time. Trends stay the same for all sizes. However failure time decreases with increased size. This is expected since more bacteria mean more mutations. Here we make the simplifying assumption that effective population size is directly proportional to actual population size[30].

## 2.5 Conclusion

In principle antisense therapy holds the potential to provide sustainable therapy against the changing bacterial population. Sustainability is due to easiness of redesigning a new, effectively binding antisense. However disrupting the binding of antisense to its target is not the only way that bacteria can achieve resistance. Mutations that impair the entry mechanism of the antisense, or mutations that reduce the efficiency of the antisense, can also render a therapy ineffective. For example, it is shown that mutations in protein sbmA impair the entry of PNAs and PPMOs. These components are not as readily redesigned as the antisense. Therefore the natural course of resistance is by impairing the entry. Upon resistance to current entry mechanisms we would be left with a requirement to find new entry mechanisms. This task is in order of difficulty of finding new small molecule antibiotics. Thus if antisense therapy is used

clinically without considering the whole picture, then the great potential of the antisense therapy cannot be realized.

If the natural course of resistance is by impairing entry mechanisms, then it must be relatively more beneficial compared to disrupting the hybridization by mutations at the target. This relation can be reversed by two ways, either by making impairment of entry less advantageous or making disruption of hybridization more advantageous. We have shown that both methods of reversing this relation increases the time to failure of the therapy, Figure 2.10.

Two important suggestions arise from this analysis. First is the inevitable need of multiple entry mechanisms. Secondly, and counter-intuitively, hitting more then one target at a time can speed time of failure, since increasing targets decreases the relative advantage to the pathogen that can be achieved by a single antisense resisting mutation, compared to the advantage of mutations that defeat entry mechanisms. In the parameter space that we investigated simultaneous deployment of more than 3 independent mechanisms is necessary. It is shown that an important challenge to the scientific community is to make various entry systems.

At this stage there are 2 known unique entry mechanisms for antisense. First one is the sbmA protein for PPMOs and PNAs. The second one is via viruses (17, 18). However it is important to note that it may be possible to have great number of different entry mechanisms with viruses. Engineering viruses that recognize different surface markers would be regarded as independent entry mechanisms.

One important consideration should be made about the time scales. The time scales of our results should not be thought as the real clinical time that a therapy would fail. Same phenomenon is seen in in-vitro evolution experiments for conventional antibiotics like

trimetroprim (19). Substantial resistance develops in 20 days; however in real life antibiotics do not acquire widespread resistance in 20 days. This also is the precise reason of simulating in-vitro like conditions since there is a direct comparison. Thus the results of this paper should be compared with in-vitro studies. With this comparison, carefully designed antisense therapy (lasts at least 3000 days) is at least 100 times more sustainable than conventional antibiotics. Furthermore in our simulations we have considered a fast doubling time of 40 minutes.

Although strategizing as noted above makes a significant improvement of the therapy, it eventually fails. Bacterial evolution is a like a running stream: blocking it is not an effective way, since it will pass it in another way. The strategy we outline is to make it run in circles. The failure for a therapy is stochastic; the better strategies give smaller chances of escape. In the current strategy, we have constructed an endless cycle for mutations against antisense; by redesigning the resistance can be reduced to that of the native state. However there is problem, bacteria can get out of this cycle by mutating against the entry mechanism. Our strategy is targeted at minimizing this, but it can't be completely eliminated in the long run. Therefore, a complete cycle, on a longer time scale, must also be constructed that includes redesign of entry mechanisms. However, one wants time to failure to be as long as possible. Here the strategy suggested by Nichols et al [22] may be usefully applied to antisense therapy. Just like conventional antibiotics, various entry mechanisms mutations might have complex fitness landscapes. Directing of evolution to keep entry mechanisms functional would be possible. In this application the fitness landscape for the antisense delivery vehicles would be entered into the Nichols et al formulation, rather than the fitness landscape for conventional antibiotics.

A very good analogy to our approach is a simple electrical system. This system includes wires, expensive equipment and a fuse. Fuse is placed in this system for it to fail first. But

replacing the fuse is easy and cheap. In our system, we design our system to cast the antisense sequence as a fuse. We achieve this by increasing the entry mechanisms and increase the advantages gained by specific mutations, so that they would be preferred.

We considered entry mechanisms as a non-specific way of impairing antisense therapy. There can be other non-specific mechanisms that can impair antisense therapy; for example degradation mechanisms. The same “directing ”game can be played here as well. Two antisenses targeting a single gene are harder to escape against than a single antisense, since a mutation would only render half of the drug ineffective. It is possible that this tool can be used to make the difference that is needed. In the degradation case, hitting the degradation mechanism with multiple antisenses would achieve the same result. In both cases, entry and degradation, another key is to use the same antisense with different entry mechanisms; this will ensure that mutating one entry system will not give high enough alleviation of selection pressure.

Our analysis potentially translates to eukaryotic cells, especially in cancer. Cancer cells gain mutations as they develop [32]. If a potential nucleic acid based therapy uses entry vehicles they will suffer from drug resistance. In these cases as well, we hypothesize that directing evolution to make changes in the target mRNA will be essential for sustained therapy. In one sense having a sustainable therapy for cancer should be easier than having a sustainable therapy for pathogens, since cancers do not spread as readily from person to person as do microbial infections. Thus sustained therapy need continue only as long as an individual human being life span. On the other hand, we understand that cancer genomes are more complex than microbial genomes.  On balance we believe that our analysis method may usefully be extended to cancer.

## 2.6 Figures

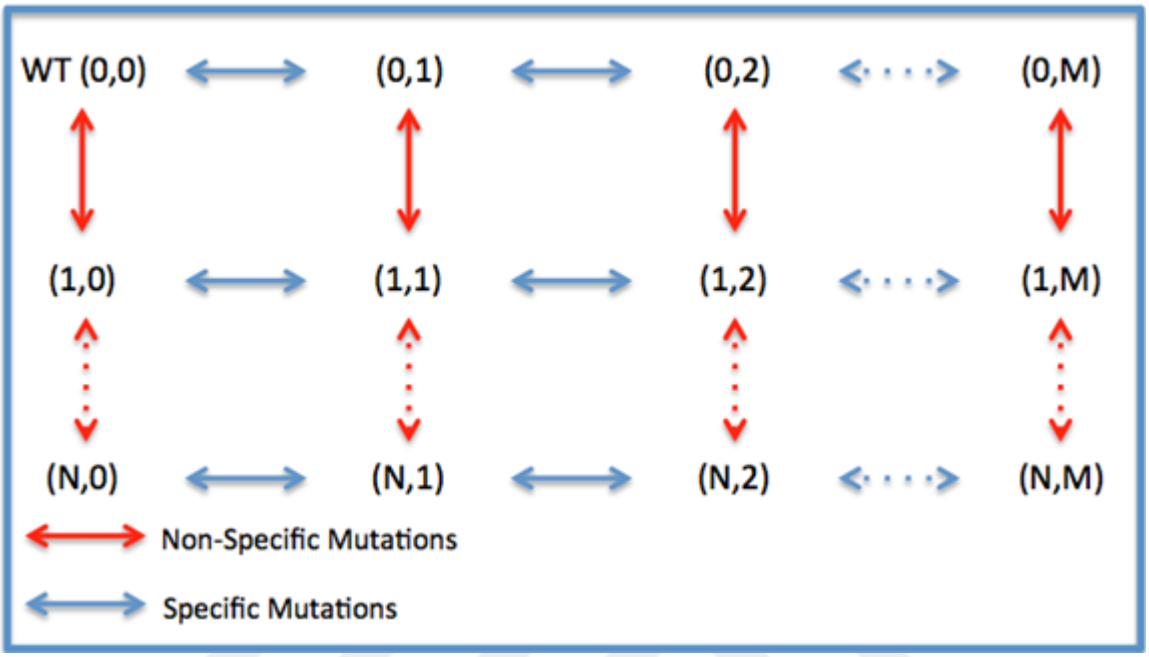


**Figure 2.1.** Kinetic Diagram of Subspecies mutations.

The Kinetic diagram represents the pathways by which resistance mutations arise in the target population in the specific virulence genes (horizontal direction) and the nonspecific resistance mechanisms (vertical direction), for example mutations that inhibit entry of the therapeutic delivery vehicle or that increase the activity of antisense degradation. In the simulations, the maximum value of "n" was set at 10, but there is no theoretical maximum. The direction of horizontal change is from left to right for pathogen mutations, and from right to left for redesign of specific antisense agents. The direction of vertical change is from top to bottom for loss of effectiveness of delivery mechanisms and from bottom to top for redesign of delivery vehicles. Our results indicate that a sustainable therapeutic strategy must "direct" the bacteria to evolve preferentially in the horizontal direction. Specific strategies for this are discussed in the body of the text. The core of these strategies is that rapid redesign of antisense can readily overcome resistance in the horizontal direction, whereas it is much more difficult to redesign entry vehicles. Entry vehicle redesign will not be considered explicitly in the simulations in this paper.

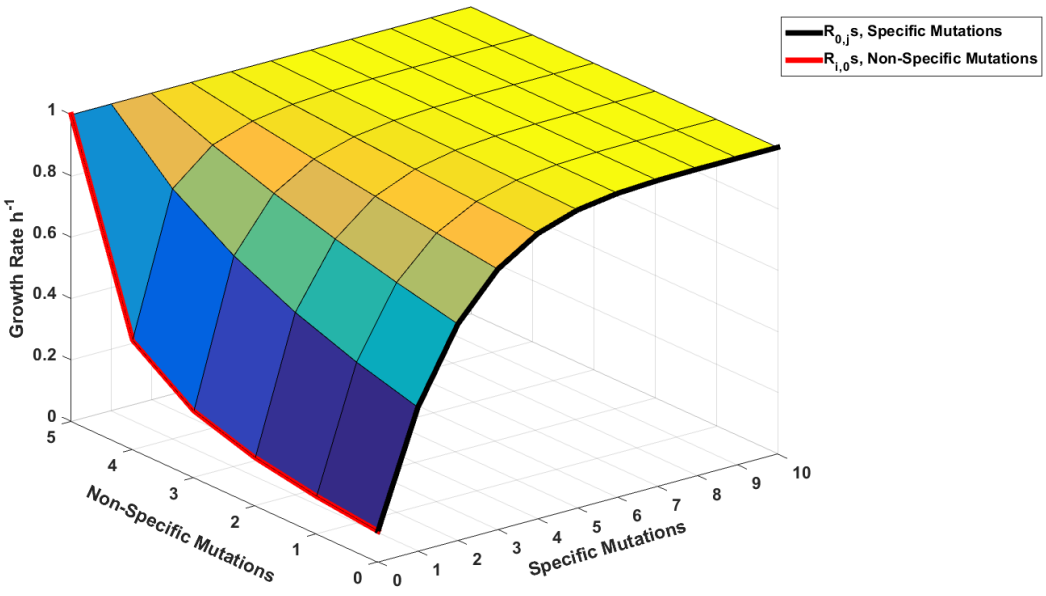


**Figure 2.2.** Example Surface of Growth Rates.

Surface of growth rate shows all the growth rates associated with kinetic diagram in Figure 2. The free parameters of the system is the $R_{0,j}$s (specific mutants), edge marked by black a line. All other growth rates are derived by equation 6. This surface was made by the following parameters: 5 independent entry mechanisms are used, specific mutations are modeled by the equation 7 with $\omega$=5. Color coding indicates height of surface.

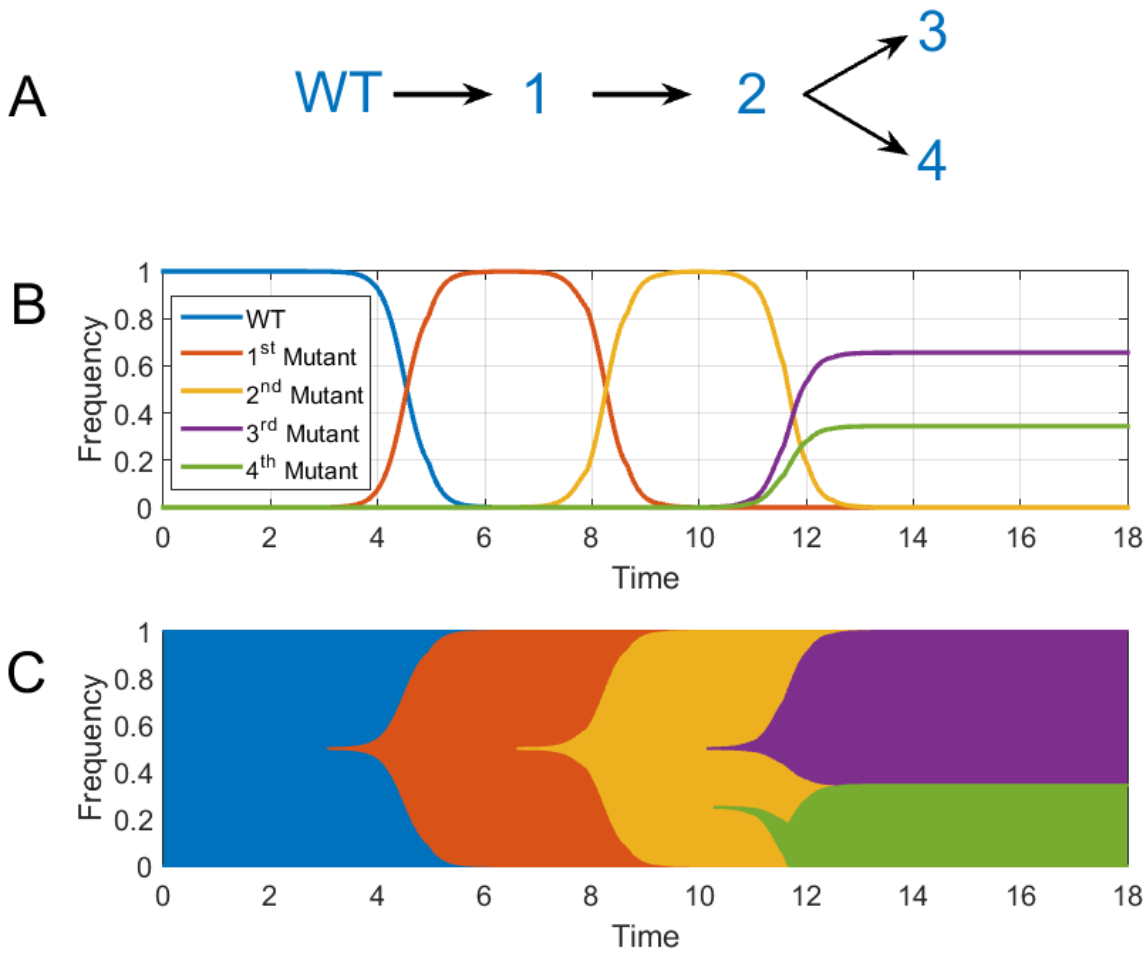


**Figure 2.3.** An Example of a Muller Plot.

A) Simulated system had successive mutations from the WT, as represented by the kinetic diagram; first mutant arises from WT, second mutant arises from first and third and fourth branches off of second mutant B) Classic representation of population dynamics of mutants are presented. Mutations proceed in the order, 1st, then 2nd , and then 3rd, 4th almost simultaneously. First Mutant arises at 3 days, second mutant arises at 6.5 days, and finally third and fourth mutant arises at 10 days. The time units are days in this simulations. The same trajectory is also plotted as a Muller plot in C) The blue indicates that all of the population is WT at start of the trajectory. At 3 days first mutant arises from the WT and grows substantially until 6.5 days. At 6.5 days first mutant gives rise to second mutant and it dominates the population by 9 days. the third and fourth mutant arises from second mutant around 10 days. They

**Figure 2.3 cont.** dominate and partition the population at 13 days. Muller plots advantage is ease of visualizing which mutant arises from which predecessors. However, it lacks the quantitative information accessibility that the classic view (B) provides. In this paper we have used Muller plots for illustration where it is more important to underline the course of evolution and its origins. On other occasions we plotted the classic representation because it can be transformed to log scale, which enables the viewer to observe what happens in all scales.

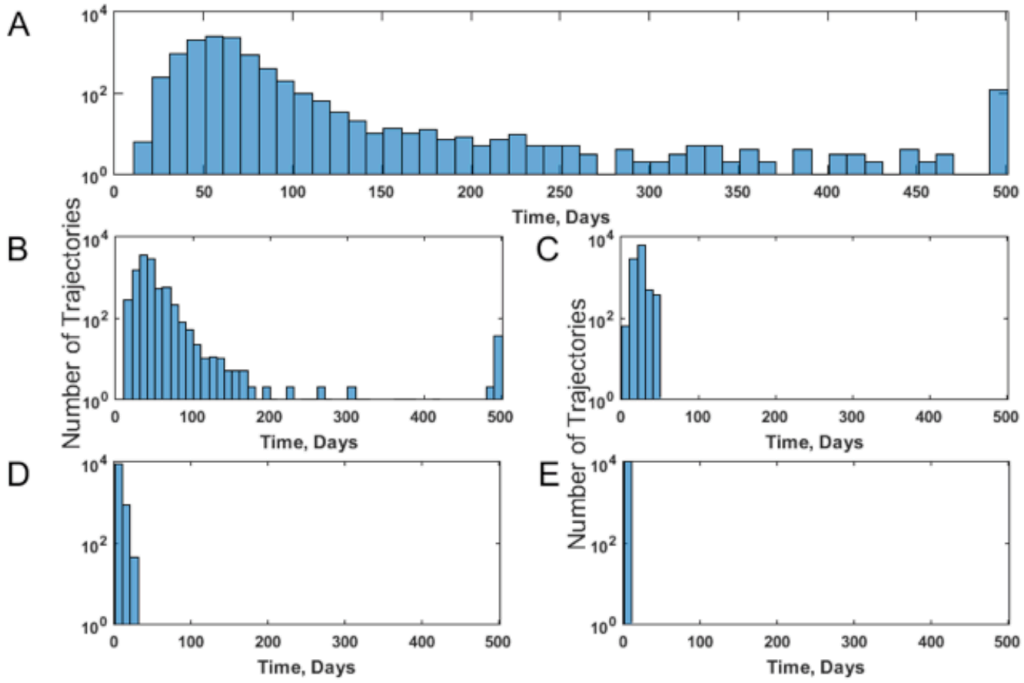


**Figure 2.4.** Therapy Failure Time for Randomly Constructed Specific Mutants.

Results of simulations using Equations 1-3 to explore effect of independent entry mechanisms for nucleic acid delivery. Histograms are number of computer experiments that resulted in indicated time to failure. Parameters for effects of particular mutations on bacterial viability (as measured by growth rates) were chosen randomly, in order to span a wide range of possibilities. Bacterial generation time was 20 minutes. Simulated system was a monoculture with carrying capacity of $10^{13}$ bacteria. Total of 10000 computer experiments were done for 1000 days, or time to failure, whichever came first. A) Represents simulations with 5 independent entry mechanisms. B) Represents simulations with 4 independent entry mechanisms. C) Represents simulations with 3 independent entry mechanisms. D) Represents simulations with 2 independent entry mechanisms. E) Represents simulations with 1 independent entry mechanism. Failure time increases with multiple independent entry mechanisms; the specific mutation growth rates were sampled randomly.

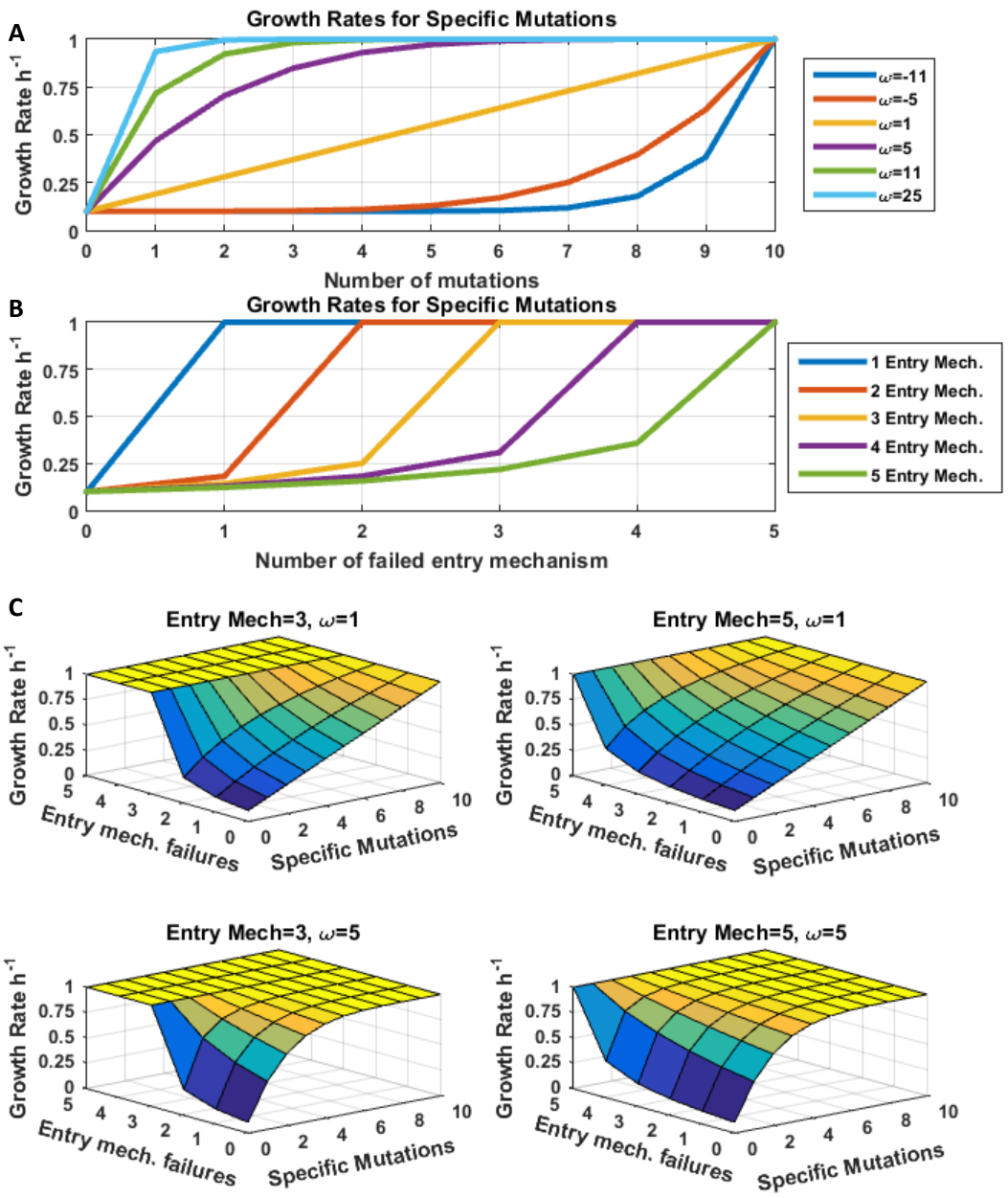


**Figure 2.5.** Illustration of Fitness landscapes for Different Parameters.

A) Functional dependence of degree of escape of pathogen vs number of mutations in specific target region as a function of the parameter w in equation 7. The higher the ω the more advantage the specific mutations provide. B) Functional dependence of growth rate of pathogen vs number of mutations that inactivate one of the delivery mechanisms for the antisense therapy. Curves represent 1, 2, 3, 4, or 5 delivery mechanisms. C) Sample fitness landscapes are made by combining curves from a. and b. by eq 5.

**Figure 2.5 cont.** These surfaces define possible mutational pathways of evolutionary escape of the pathogens from the antisense therapy.

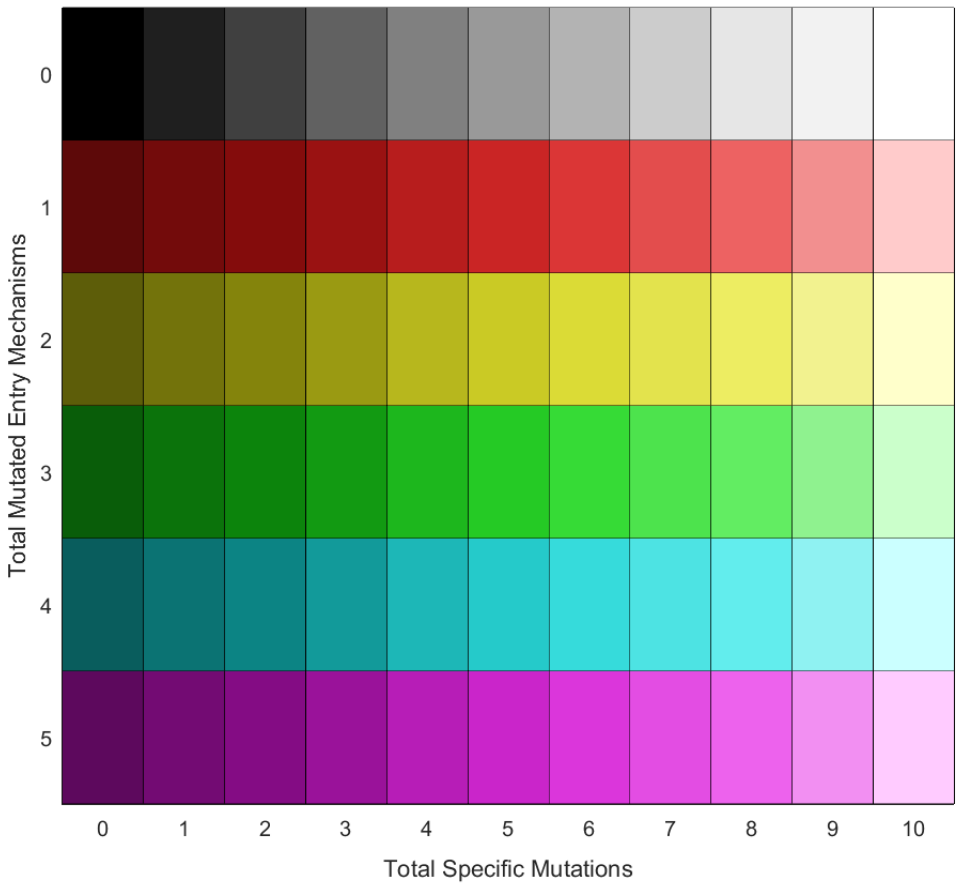


**Figure 2.6.** Color coding for the Muller plots.

We have chosen different color for number of mutated entry mechanisms and their shades for their specific mutants. Coordinates of a colored box is the corresponding mutant. For example the (5,10) is 5 entry mechanism failures with 10 specific mutatations, and its color is the brightest pink, lower right corner.

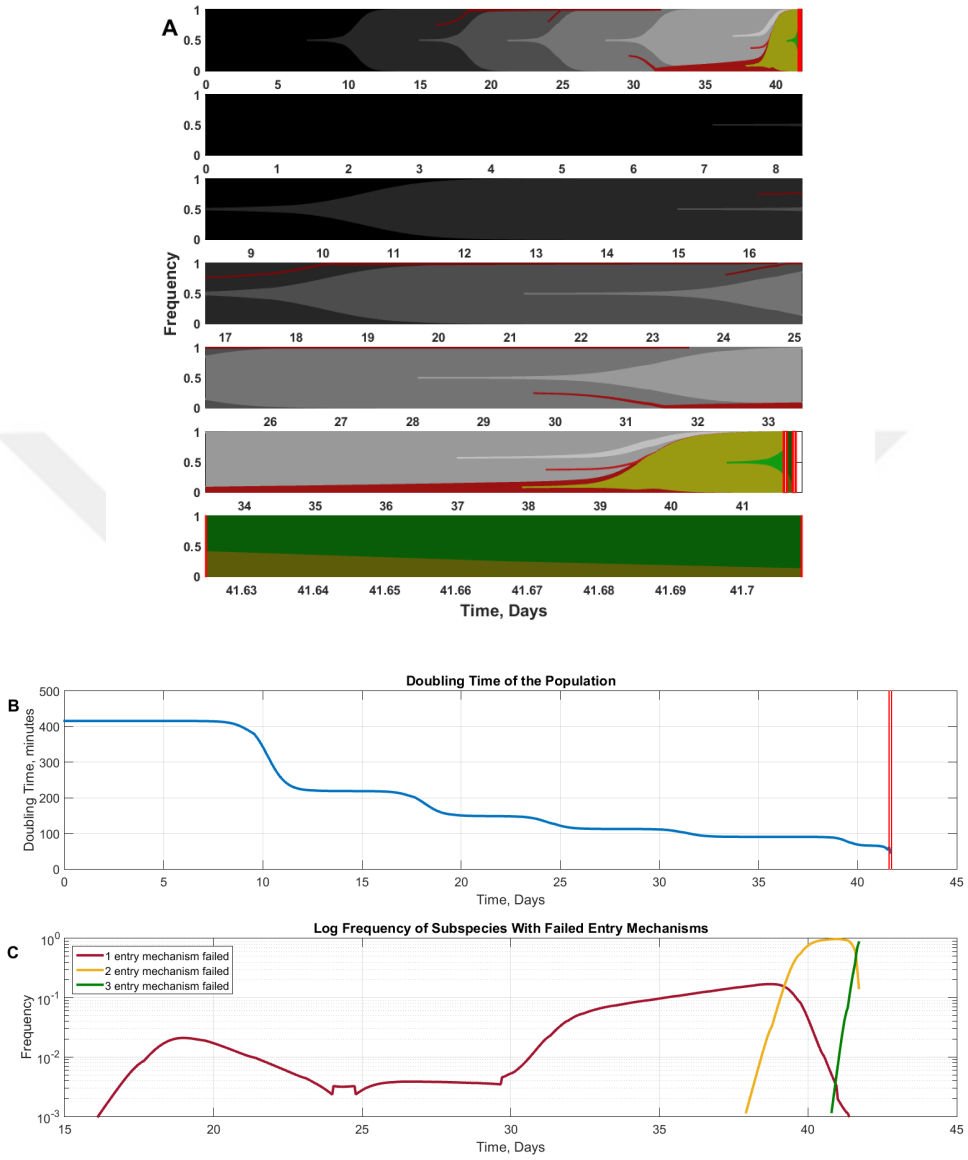


**Figure 2.7.** Sample Trajectory With Three Entry Mechanism and ω=1.

A) The top panel shows the full trajectory. Last panel is the last redesign cycle (41.62 days to 41.8 days), and the panels in between are shorter parts of the full trajectory for more detailed inspection. Therapy lasts for 46 days. The vertical red lines mark the time of redesign. In this trajectory there is only one redesign at 41.62 days. Until failure we see that the first specific mutant (0,1) takes over the population in 11 days. Then the second specific mutant (0,2) takes over at 20 days. Then the third specific mutant (0,3) takes over at 26 days. Before the third specific mutant takesover, first entry mechanism fails on two

**Figure 2.7 cont.** occasions at 16th and 24$^{th}$ days respectively, however they die off by competetion. Right after third specific mutant (0,3) dominates, the non-specific mutant (1,3) arises from (0,3). In the mean time the fourth (0,4) specific mutant also increases. At 33 days (1,3) and (0,4) constitutes the most of the populations. The mutant (1,3) gives rise to second entry mechanism mutant (2,3) around 37 days. The (2,3) having very high growth rate takes over the entire population very rapidly and gives rise to (3,3). (3,3) increases slightly but the overall doubling rate exceeds the threshold for redesign. Redesign happens, there are two subpopulations at the start of the redesign (3,0) and (2,0), in very rapid time (3,0) increases for the full escape. The therapy can no longer be rescued by redesigning the cargo. All of the three entry mechanism have failed. B) Doubling time versus time is plotted. The doubling time starts at 420 minutes per generation. The doubling time decreases step-wise (growth rate increases) as specific mutants takeover. The threshold of 60 minutes is reached at 41.62 days. A cargo redesign is made at this time. Redesign is indicated with red vertical line. However the doubling rate does not recover since all off the entry mechanisms have failed and the therapy fails. C) Entry mechanism mutants are plotted with log10 scale. The purpose log scale is to monitor all scales at the same time. The first entry mechanism mutant stays low but flactuates. This mutant gives rise to second entry mechanism mutant which takes over at 40 days and paves the way for the final and third mutant. Third mutant arises at 41 days and dominates the population at 41.8 days. At this time the therapy halts.

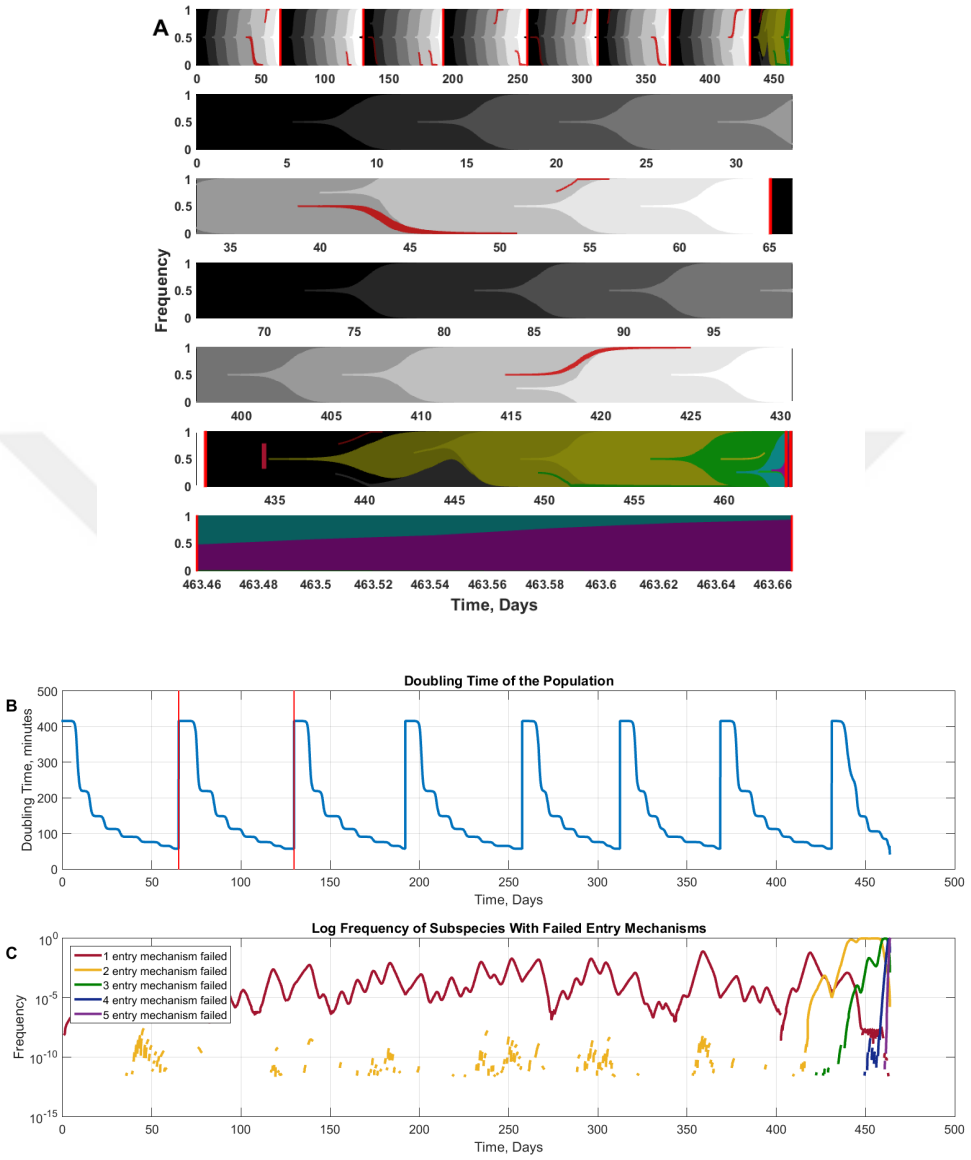


**Figure 2.8.** Sample Trajectory With Five Entry Mechanism and ω=1.

A) The top panel shows the full trajectory, last panel is the last redesign cycle (463.46 days to 463.66 days), and the panels in between are shorter parts of the full trajectory for more detailed inspection. The vertical red lines mark the time of redesign. Therapy lasts for 463 days. In this trajectory are cargo is redesign 8 times. Until first redesign we see (0,1), (0,2), (0,3), (0,4), (0,5), (0,6), (0,7) dominates the populations at 11 days, 16 days, 26 days, 35 days, 45 days, 56 days, 62 days respectively. On two occasions, 39 days and 53 days, one of the five entry mechanisms is mutated but the mutant is shortlived.

**Figure 2.8 cont.** At 65 days the cargo is redesigned and overwhelming majority of the population returns to WT. The first cycle of redesign is also showed in second and third panel. After the redesign, same sequence of events happens until the next redesign at 125 days. Qualitatively this redesign cycle is repeated for 5 more times, total of 7 times. After the seventh redesign the subsequent events are significantly different. This portion of the trajectory is also shown in $6^{th}$ panel. First mutant that arises is the mutant having two of its entry mechanism mutated (0,2), this is not expected since (0,2) must arise from (0,1). This phenomonan is called stochastic tunneling, a mutant arises from an ancestor that was not significantly large as a population. The tunneling is indicated by the vertical bold red line at 434 day (26). While (0,2) takes higher portion of the population, at 438 day (0,1) and (1,0) mutants arises, however they are shortlived. The mutant (0,2) gives rise to (1,2) at 443 days, (1,2) gives rise to (2,2) at 448 days, (2,2) gives rise to the mutant having three mutated entry mechanisms (2,3) at 456 days and this mutant takes over the population at 460 days. Soon after the (2,2) arose, (1,2) gives rise to (1,3) at 450 days however this third mutant is short lived. (2,3) gives rise to fourth entry mechanism mutant (2,4). Eventually the fifth and final entry mechanism mutates at 464 days. Soon after final mutation there is a redesign at 463.46. This redesign drives back all strains back to either (0,4) or (0,5), this population cannot be rescued any further and the therapy fails at 463.67 days. B) Change of doubling time through time is plotted. The doubling time starts at 420 minutes per generation. The doubling time decreases step-wise (growth rate increases) as specific mutants takeover. The threshold of 60 minutes is reached at 65 days. Redesign is indicated with vertical red line. Red lines for first two redesigns are plotted but the later ones are omitted for clear presentation. Redesign times are easily distinguishable by sudden jumps in doubling time. Cargo is redesigned at this time and doubling rate is recovered to per 420 minutes. This cycle repeats 6 more times.C) Entry mechanism mutants are plotted with log scale. This representation enables us to spot the stochastic tunneling that happened in this trajectory. The first entry mechanism mutant stays low but flactuates. Many times the second entry mechanism mutates but eventually dies. At 410 days the second entry mechanism mutant arises from the low single mutant and increases to take over the population. Soon after the second entry mechanism, the third mechanism fails at and dominates the

**Figure 2.8 cont.** population. After third entry mechanism, fourth and fifth entry mechanism fails very rapidly and therpt is halted at 460 days.

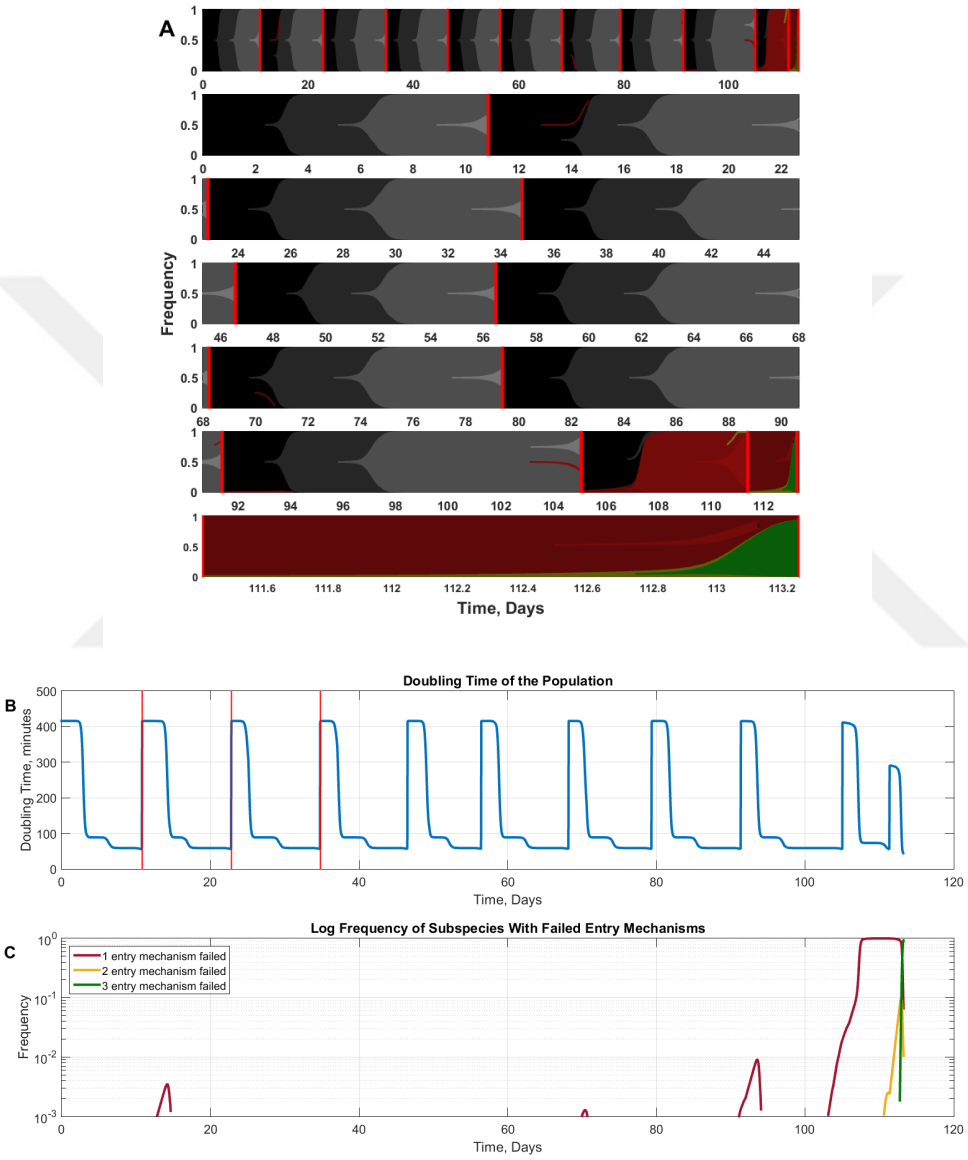


**Figure 2.9.** Sample Trajectory With Three Entry Mechanism and ω=5.

A) The top panel shows the full trajectory, last trajectory is the last redesign cycle (111.4 days to 113.2 days), and the panels in between are shorter parts of the full trajectory for more detailed inspection. The vertical red lines mark the time of redesign. Therapy lasts for 113.2 days. In this trajectory cargo is redesigned 10 times. Until first redesign we see (0,1), (0,2) dominates the populations at 4 days, 7 day

**Figure 2.9 cont.** respectively. At 9 days (0,3) arises from (0,2) and increases until the redesign. First redesign is made at 11 days. Until first redesign entry mechanisms stays intact.The first cycle of redesign is also showed in first panel. After the redesign, same sequence of events happens until the next redesign at 23 days. Qualitatively this redesign cycle is similar to first cycle, however this time single entry mechanism fails but eventually dies. Similar redesigns cycles happen 6 times more. After the redesign at 91 minutes, the trajectory proceed similarly to other cycles. However there is a siginificant difference at the end of this cycle. At 103 days (2,1) mutant arises from (2,0) and does not die off until the redesign. After this redesign at 105 days, 10 percent of total population has a single entry mechanism mutation. This population, (0,2), grows and gives rise to (1,2) at 107 days and (1,2) dominates the population at 107.5 days. (1,2) gives rise to two mutants (2,2) and (1,3) mutants. Both of these mutants survives until the redesign at 111.5 days. Mutants having mutations at all of the entry mechanisms is present at the start. This small population grows and gives rise to (1,3) at 112.7 days. Mutant (1,3) grows and takes over the population beyond recoverablity. Therapy continues until 113.2 days. B) Doubling time versus time is plotted. As usual doubling time starts at 420 minutes per generation. The doubling time decreases stepwise (growth rate increases) as mutants takeover. The threshold of 60 minutes is reached at 10 days. Cargo is redesigned at this time and doubling rate is recovered to per 420 minutes. Redesign is indicated with vertical red line. Red lines for first two redesigns are plotted but the later ones are omitted for clear presentation. Redesign times are easily distinguishable by sudden jumps in doubling time.This cycle repeats 10 more times, at last redesign cannot rescue the therapy. C) The entry mechanism mutants are plotted with log scale. The first entry mechanism arises sporadically until 110 days, at this time it dominates the population. At 113.2 days the third entry mechanism mutant arises and the therapy fails.

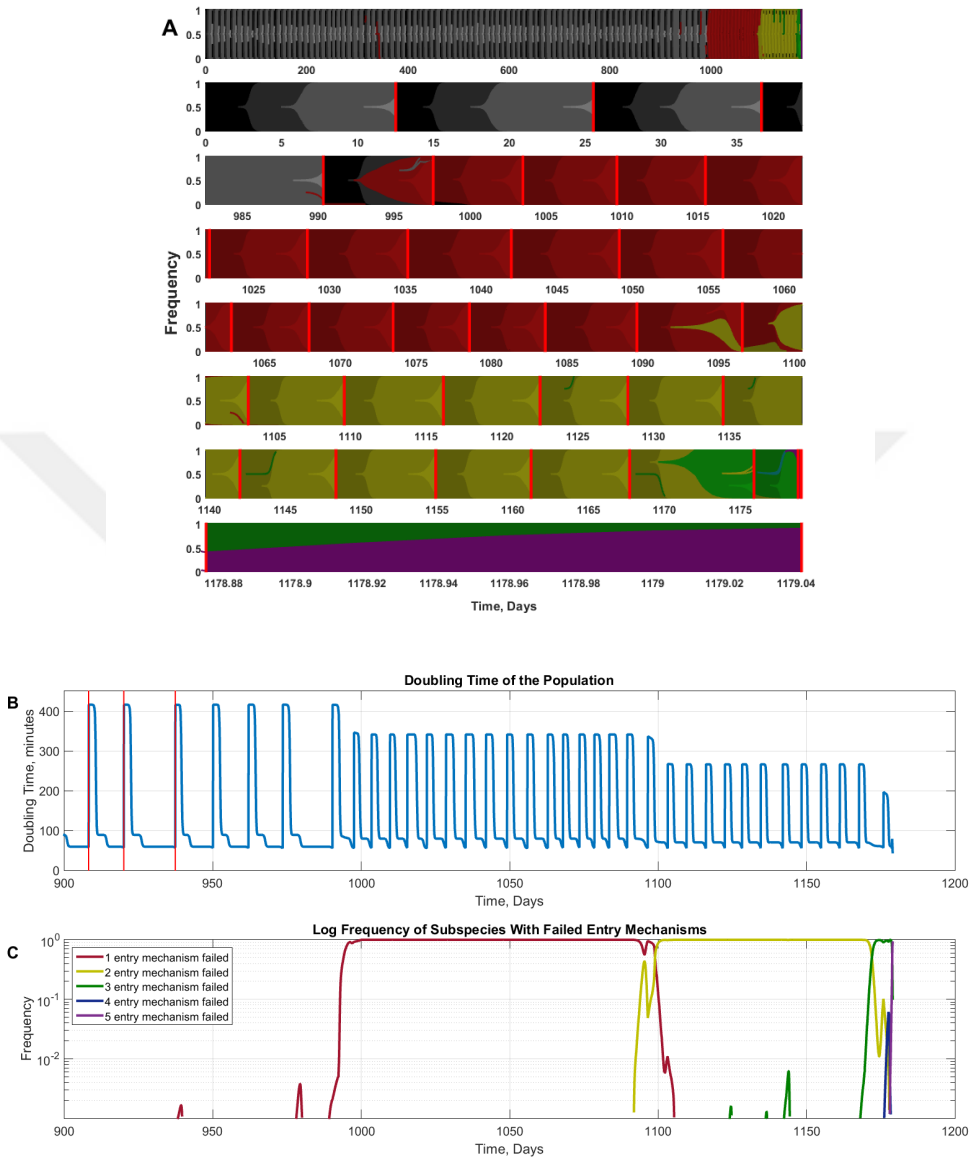


**Figure 2.10.** Sample Trajectory With Five Entry Mechanism and ω=5.

A) The top panel shows the full trajectory, last trajectory is the last redesign cycle (1178.88 days to 1179.04 days), and the panels in between are shorter parts of the full trajectory for more detailed inspection. The vertical red lines mark the time of redesign. Therapy lasts for 1179.04 days. In this trajectory cargo is redesigned about 100 times. Until first redesign we see (0,1), (0,2) dominates the populations at 3.5 days, 7 day respectively. At 10.5 days (0,3) arises from (0,2) and increases until the redesign. First redesign is made at 12.5 days. Until first redesign entry mechanisms stays intact. After the

**Figure 2.10 cont.** redesign, same sequence of events happens, arisel of mutants and redesign, happens qualitatively similar until 991 days. After 991 days, WT, (0,0), gives rise to (0,1) at 992 days, (0,1) gives rise to (1,1) 92.5 days. (1,1) mutant increases rapidly and dominates the population at 996 days. At this point the overwhelming majority of the population has lost one of the entry mechanisms. (1,1) gives rise to (2,1) at 996 days. At 997 days cargo is redesigned and the majority of the population is (0,1). The mutant, (1,1) arises from (0,1) at 999 days, and (2,1) arises from (1,1) at 102 days. The growth of (2,1) continues until the redesign at 1003 days. The same sequence of events that happened in the previous cycle occurs until the redesign at 1090 days. It is important to note the time interval between redesigns, when the system lost a single entry mechanism the interval was around 10 days while the cycles where all entry mechanisms were intact had 13 days. After the redesign at 1090, (1,2) arises from (1,1) but dies off. Later, (1,2) arises from (1,1) at 1098 days and takes over the population by 1100 days. Subsequently, (2,2) arises from (1,2) at 1102 days and the system passes the threshold for redesign at 1104 days. After the redesign at 1104 days, (1,3) arises from (0,3) at 1106 days and subsequently (2,3) arises, then another cycle of redesign is made. This cycle also continues until redesign at 1168 days. After 1168 fourth entry mechanism fails, and after the redesign at 1176 the fifth also is lost. After losing all entry mechanisms, the final redesign happens at 1178.88 days. At this point there are mutants who lost all entry mechanisms and one who lost 4. The mutant who lost all entry mechanisms quickly takes over the population from 1178.88 to 1179.04, as also seen at last panel. The therapy ends at 1179.04. B) Doubling time versus time is plotted for times greater than 900 days. The doubling rates before this time resembles the beggings of the shown trajectory. As usual doubling time starts at 420 minutes per generation. The doubling time decreases (growth rate increases) as mutants takeover. The threshold of 60 minutes is reached per 10 days. Cargo is redesigned at this time and doubling rate is recovered to per 420 minutes. This cycle repeats until 1000 days. At this time population has lost its first entry mechanism. Until losing the second mechanism redesign cycles are similar. However after redesigns doubling rate only recovers to per 340 minutes. After second entry mechanism fails redesigns can only recover the therapy by increasing doubling time to per 280 minutes. Eventually the third, fourth and fifth entry mechanism fails and therapy

**Figure 2.10 cont.** halts. C) The entry mechanism mutants are plotted with log scale. The first entry mechanism arises sporadically until 1000 days, at this time it dominates the population. At 1100 days the second entry mechanism mutant arises and dominates the population. Soon after third mechanism fails 1170 days. Very close to therapy failure, 1179.04 days, fourth and fifth entry mechanism fails.

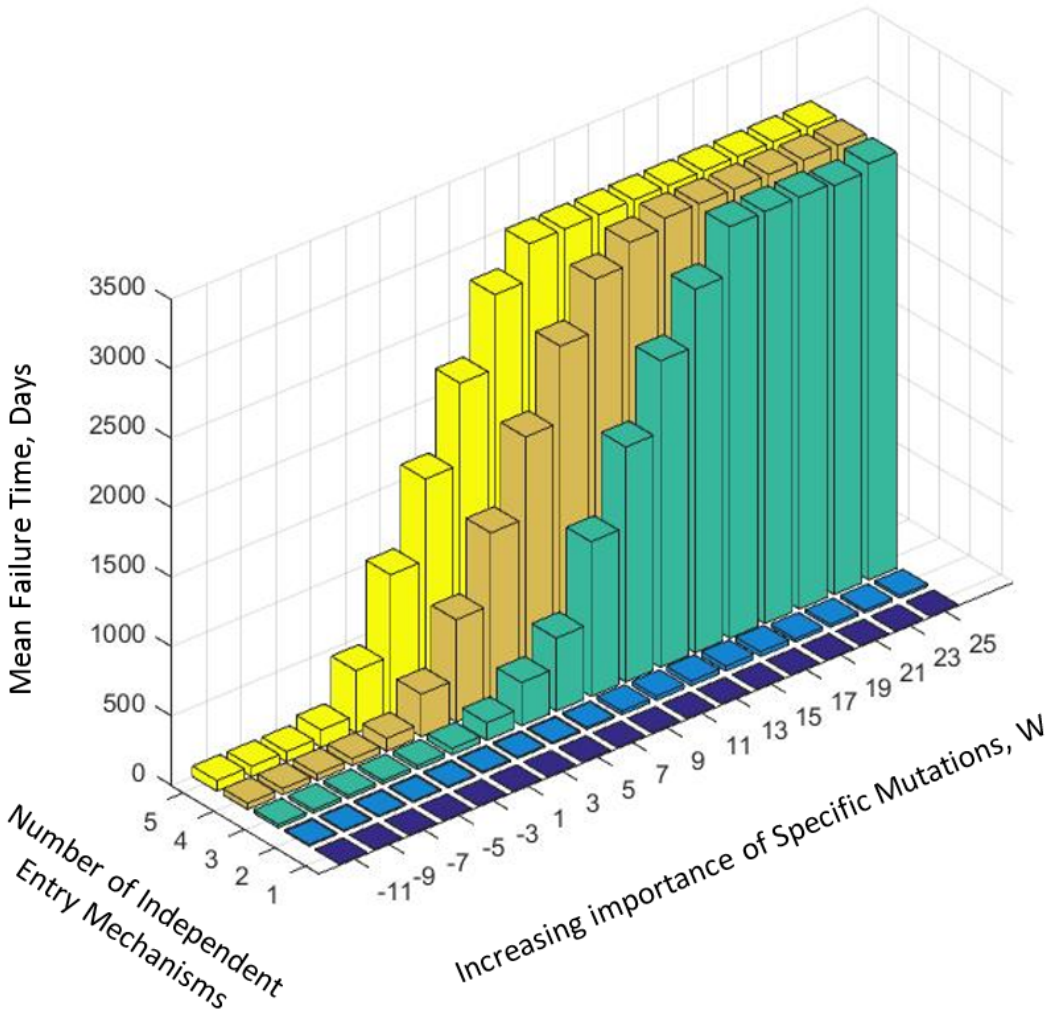


**Figure 2.11.** Mean Therapy Failure Time for Simulations Made in 100 ml Rich Medium.

This figure shows a summary of 1000 trajectories at each of the 90 combinations of number of independent delivery vehicle (5) and number of different values of W (18), (or 12500 total) of which Figures 7-10 show individual examples. This figure is the result of stochastic simulations as described in the text. It is seen that mean time to failure of the therapy is lengthened at higher W and larger number of independent delivery mechanisms. These parameter choices have the effect of making non-specific mutations less advantageous and making specific mutations more advantageous. The simulations were performed as if the experiment were conducted in 100 ml rich medium.

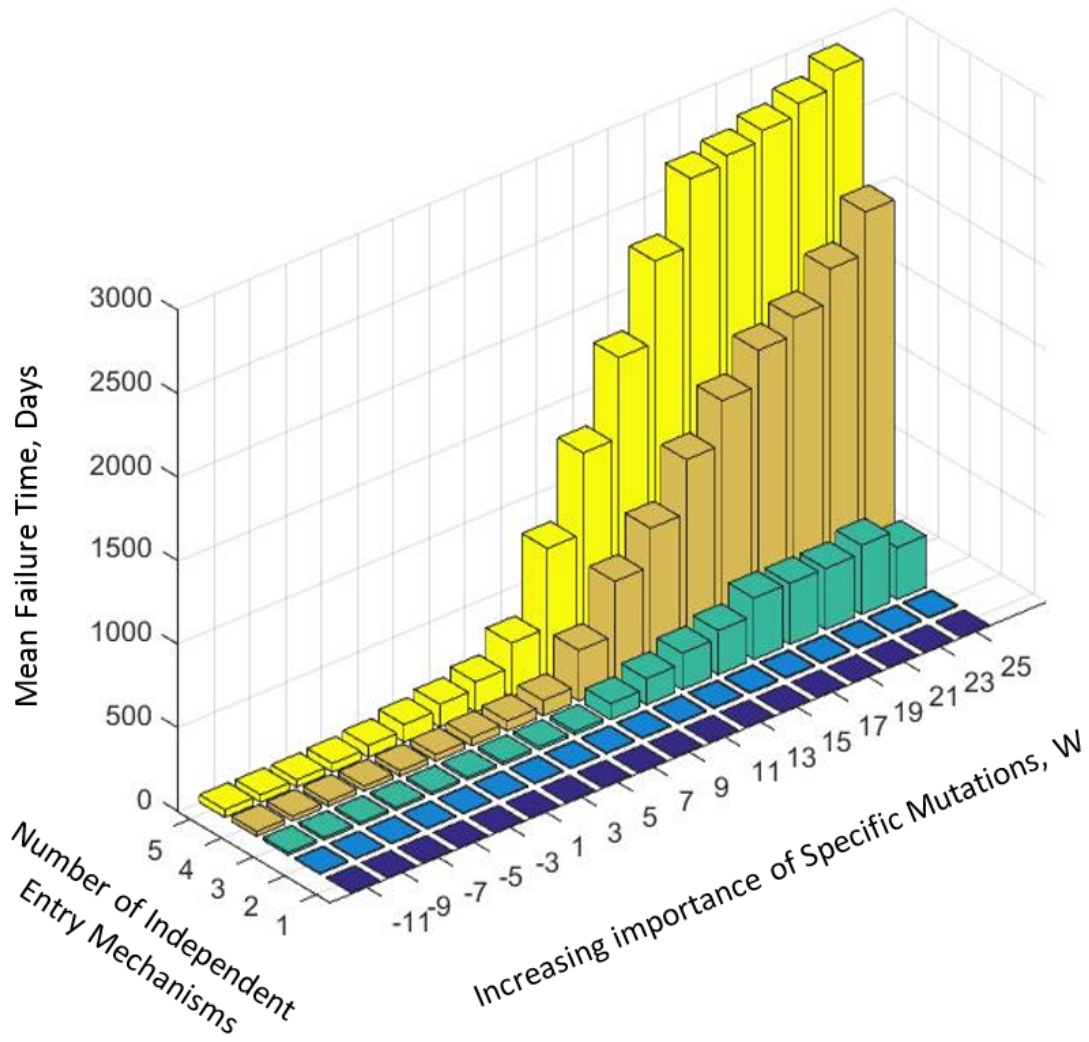


**Figure 2.12.** Mean Therapy Failure Time for Simulations Made in 100 L (100 kg) Rich Medium.

This figure shows a summary of 1000 trajectories at each of the 90 combinations of number of independent delivery vehicle (5) and number of different values of W (18), (or 12500 total). This figure is the result of stochastic simulations as described in the text. It is seen that mean time to failure of the therapy is lengthened at higher W and larger number of independent delivery mechanisms. These parameter choices have the effect of making non-specific mutations less advantageous and making specific mutations more advantageous. The simulations were performed as if the experiment were conducted in 100 liters (100kg) rich medium.

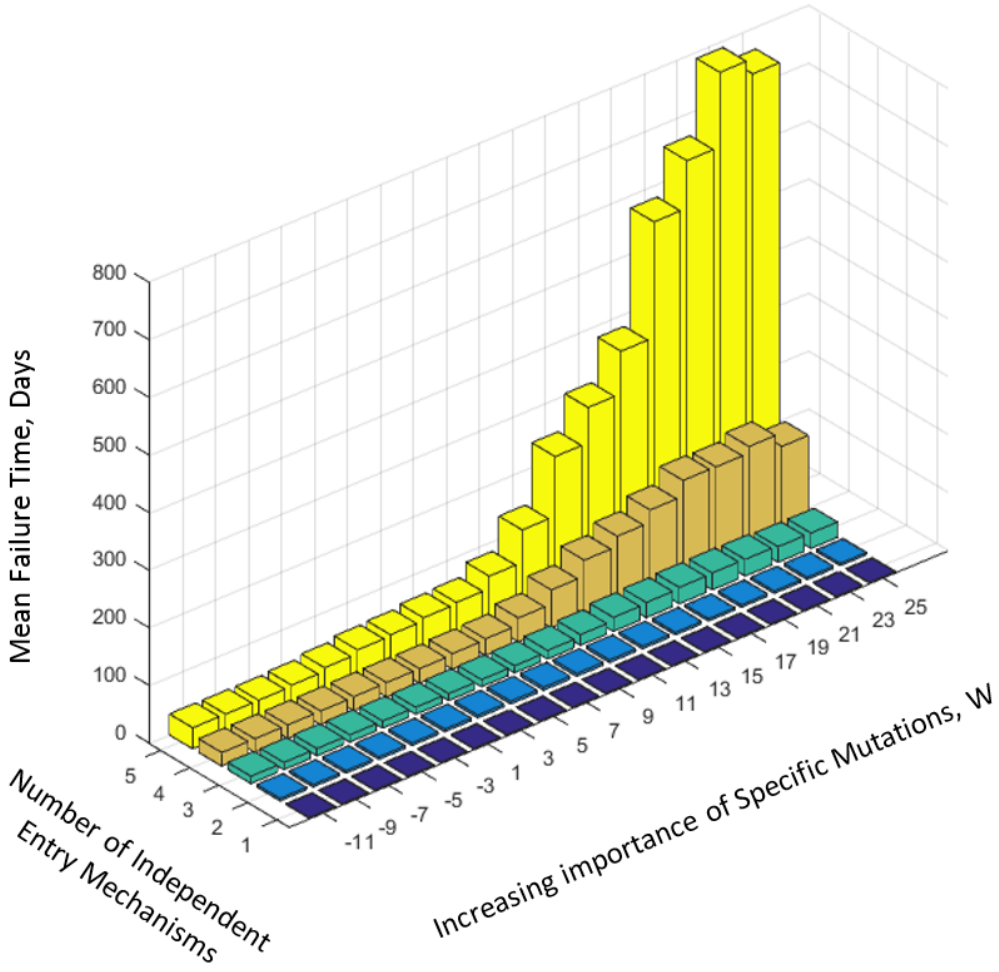


**Figure 2.13.** Mean Therapy Failure Time for Simulations Made in 10000 L (10000 kg) Rich Medium.

This figure shows a summary of 1000 trajectories at each of the 90 combinations of number of independent delivery vehicle (5) and number of different values of W (18), (or 12500 total). This figure is the result of stochastic simulations as described in the text. It is seen that mean time to failure of the therapy is lengthened at higher W and larger number of independent delivery mechanisms. These parameter choices have the effect of making non-specific mutations less advantageous and making specific mutations more advantageous. The simulations were performed as if the experiment were conducted in 10000 L (10000 kg) rich medium.

## 2.7 References


1) Laxminarayan R, Duse A, Wattal C, Zaidi AK, Wertheim HF, Sumpradit N, Vlieghe E, Hara GL, Gould IM, Goossens H, Greko C. Antibiotic resistance—the need for global solutions. The Lancet infectious diseases. 2013 Dec 31;13(12):1057-98.

2) Roberts RR, Hota B, Ahmad I, Scott RD, Foster SD, Abbasi F, Schabowski S, Kampe LM, Ciavarella GG, Supino M, Naples J. Hospital and societal costs of antimicrobial-resistant infections in a Chicago teaching hospital: implications for antibiotic stewardship. Clinical infectious diseases. 2009 Oct 15;49(8):1175-84.

3) Taubes G. The bacteria fight back. Science. 2008 Jul 18;321(5887):356-61.

4) Bassetti M, Merelli M, Temperoni C, Astilean A. New antibiotics for bad bugs: where are we?. Annals of clinical microbiology and antimicrobials. 2013 Aug 28;12(1):1.

5) Brown, E.D. and Wright, G.D., 2016. Antibacterial drug discovery in the resistance era. *Nature*, *529*(7586), pp.336-343.

6) Lin J, Nishino K, Roberts MC, Tolmasky M, Aminov RI, Zhang L, editors. Mechanisms of antibiotic resistance. Frontiers Media SA; 2015 Jun 1.

7) Tello A, Austin B, Telfer T. Selective pressure of antibiotic pollution on bacteria of importance to public health. Environmental Health Perspectives. 2012;120(8):1100-6.

8) Berendonk TU, Manaia CM, Merlin C, Fatta-Kassinos D, Cytryn E, Walsh F, Bürgmann H, Sørum H, Norström M, Pons MN, Kreuzinger N. Tackling antibiotic resistance: the environmental framework. Nature Reviews Microbiology. 2015 May 1;13(5):310-7.

9) Summerton J, WELLER D. Morpholino antisense oligomers: design, preparation, and properties. Antisense and Nucleic Acid Drug Development. 1997 Jun;7(3):187-95.

10) Wesolowski D, Alonso D, Altman S. Combined effect of a peptide–morpholino oligonucleotide conjugate and a cell-penetrating peptide as an antibiotic. Proceedings of the National Academy of Sciences. 2013 May 21;110(21):8686-9.

11) Sawyer AJ, Wesolowski D, Gandotra N, Stojadinovic A, Izadjoo M, Altman S, Kyriakides TR. A peptide-morpholino oligomer conjugate targeting Staphylococcus aureus gyrA mRNA improves healing in an infected mouse cutaneous wound model. International journal of pharmaceutics. 2013 Sep 10;453(2):651-5.

12) Geller BL, Greenberg DE. Peptide-conjugated phosphorodiamidate morpholino oligomers: a new strategy for tackling antibiotic resistance. Therapeutic delivery. 2014 Mar;5(3):243-5.

13) Wesolowski D, Tae HS, Gandotra N, Llopis P, Shen N, Altman S. Basic peptide-morpholino oligomer conjugate that is very effective in killing bacteria by gene-specific and nonspecific modes. Proceedings of the National Academy of Sciences. 2011 Oct 4;108(40):16582-7.

14) Geller BL, Marshall-Batty K, Schnell FJ, McKnight MM, Iversen PL, Greenberg DE. Gene-silencing antisense oligomers inhibit Acinetobacter growth in vitro and in vivo. Journal of Infectious Diseases. 2013 Oct 14:jit460.

15) Meng J, He G, Wang H, Jia M, Ma X, Da F, Wang N, Hou Z, Xue X, Li M, Zhou Y. Reversion of antibiotic resistance by inhibiting mecA in clinical methicillin-resistant Staphylococci by antisense phosphorothioate oligonucleotide. The Journal of antibiotics. 2015 Mar 1;68(3):158-64.

16) Citorik RJ, Mimee M, Lu TK. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nature biotechnology. 2014 Nov 1;32(11):1141-5.

17) Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, Fischetti VA, Marraffini LA. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nature biotechnology. 2014 Nov 1;32(11):1146-50.

18) Ghosal A, Vitali A, Stach JE, Nielsen PE. Role of SbmA in the uptake of peptide nucleic acid (PNA)-peptide conjugates in E. coli. ACS chemical biology. 2012 Nov 20;8(2):360-7.

19) Puckett SE, Reese KA, Mitev GM, Mullen V, Johnson RC, Pomraning KR, Mellbye BL, Tilley LD, Iversen PL, Freitag M, Geller BL. Bacterial resistance to antisense peptide phosphorodiamidate morpholino oligomers. Antimicrobial agents and chemotherapy. 2012 Dec 1;56(12):6147-53.

20) Ferguson AL, Mann JK, Omarjee S, Ndung'u T, Walker BD, Chakraborty AK. Translating HIV sequences into quantitative fitness landscapes predicts viral vulnerabilities for rational immunogen design. Immunity. 2013 Mar 21;38(3):606-17.

21) Mann JK, Barton JP, Ferguson AL, Omarjee S, Walker BD, Chakraborty A, Ndung'u T. The fitness landscape of HIV-1 gag: advanced modeling approaches and validation of model predictions by in vitro testing. PLoS Comput Biol. 2014 Aug 7;10(8):e1003776

22) Nichol D, Jeavons P, Fletcher AG, Bonomo RA, Maini PK, Paul JL, Gatenby RA, Anderson AR, Scott JG. Steering evolution with sequential therapy to prevent the emergence of bacterial antibiotic resistance. PLoS Comput Biol. 2015 Sep 11;11(9):e1004493.

23) Lee H, Popodi E, Tang H, Foster PL. Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. Proceedings of the National Academy of Sciences. 2012 Oct 9;109(41):E2774-83.

24) Toprak E, Veres A, Michel JB, Chait R, Hartl DL, Kishony R. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nature genetics. 2012 Jan 1;44(1):101-5.

25) Stoer J, Bulirsch R. Introduction to numerical analysis. Springer Science & Business Media; 2013 Mar 9.

26) Gillespie DT. Stochastic simulation of chemical kinetics. Annu. Rev. Phys. Chem.. 2007 May 5;58:35-55.

27) Cao Y, Gillespie DT, Petzold LR. Efficient step size selection for the tau-leaping simulation method. The Journal of chemical physics. 2006 Jan 28;124(4):044109.

28) Barrick JE, Lenski RE. Genome dynamics during experimental evolution. Nature Reviews Genetics. 2013 Dec 1;14(12):827-39.

29) Iwasa Y, Michor F, Nowak MA. Stochastic tunnels in evolutionary dynamics. Genetics. 2004 Mar 1;166(3):1571-9.

30) Charlesworth, Brian. "Effective population size and patterns of molecular evolution and variation." *Nature Reviews Genetics* 10, no. 3 (2009): 195-205.)

31) Greaves, Mel, and Carlo C. Maley. "Clonal evolution in cancer." *Nature* 481, no. 7381 (2012): 306-313.

# CHAPTER 3: Antisense therapy can be used to arrest, guide and diminish trimethoprim resistance

## 3.1 Abstract

Discovery of antibiotics might be the most important accomplishment of medicine; however, it is challenged by ever growing antibiotic resistance. In this study, antisense molecules are used to tackle trimethoprim resistance. The mode of action is that antisense molecules will bind to mutant mRNAs and disable their translation. Simulations carried out in this study show that: 1) If antisense molecules are targeted to potential mutations which are not yet acquired, then evolutionary trajectories are blocked from getting those mutation; as a result, blocking evolutionary pathways completely arrests further progression of resistance. 2) If antisense molecules are targeted to incomplete list of potential mutations, then evolutionary trajectories converge to those mutations; Lures bacteria to mutations of interest.3) If antisense molecules are targeted to existing mutations, then resistance mutations are lost completely; during this time antisense efficiency is disrupted many times by target sequence mutation, which shows the requirement to counter by redesigning the antisense molecules. 4) Complete sensitization without any disruption to antisense efficiency is possible if a competing sensitive strain, e.g. wild type pathogen, is introduced alongside with antisense. 5) Once a population is sensitized, and its potential mutations are blocked, then resistance can only be gained by a simultaneous double mutation, which takes at least 1821 years in this setting.

## 3.2 Introduction

First introduced antibiotic was penicillin discovered in 1928 [1] and it was introduced at 1942 [2]. Penicillin resistant strains of Pneumococcus were already isolated in 1942 [3]. Today,

many bacteria are resistant to original penicillin. Antibiotic resistance is also prevalent for any antibiotic. Resistant strains for every antibiotic have been isolated. The total economic burden on US is estimated to be $55 billion a year [4]. Monetary cost is the small price compared to 99,000 [5] lives per year only from health care associated infections.

Battling this war requires three strategies; 1) inventing new antibacterial agents, 2) keep existing antibiotics away from invoking further resistance, 3) sensitize antibiotics that are already encountering high levels of resistance.

Inventing new antibacterial agents is an immense challenge. The majority of antibiotics are discovered based on trial and error tests of different chemicals. This process is costly and slow; according to Tufts center in Boston, cost is $800 million on average to bring a new drug in to market [6,7];only 13 new antibiotics were discovered between 2000 and 2014 [8]. However, there are newly discovered antibiotics like lugdunin [9]. Mode of action of lugdunin and its resistance mutations are not yet catalogued. Unlike lugdunin trimethoprim is an older and well-studied antibiotic. Trimethoprim binds and Inhibits dihyrdrofolate reductase (DHFR) protein [10]. Recently its 7 resistance mutations were cataloged [11]. In this study we have used the cataloged data of trimethoprim resistance.

Efforts on inventing new antibiotics is not enough by any measure but however latter two strategies, keeping resistance from progressing, and sensitizing resistant strains, are much less studied. Up to this day there is no effective measure to prohibit bacteria from invoking further resistance to small molecule antibiotics. Sensitizing resistant bacteria is studied by strategy of collateral sensitivity [12]. In this approach, different small molecule antibiotics are used alternatingly to sensitize a resistant strain. Antisense molecules targeting resistance gene were used to increase antibiotic effectiveness [13]. Antisense molecules are nucleic acid based

molecules that bind to mRNA of a vital gene and block its translation [14]. In principle this is a drug combination, however the evolution of interacting of antisense and conventional antibiotics are not widely studied. There are studies on interactions of conventional antibiotics[15].

Our study uses antisense molecules to tackle both problems: prohibiting antibiotics to invoke further resistance, and sensitizing the genotype of resistant bacteria. Our study is a computer simulation study performed on cataloged mutations of trimethoprim resistance. We show that antisense technology can be used to prohibit the progression of resistance for any strain (including wild type). Moreover antisense technology can revert back resistant mutations; antisense molecules induce loss of resistance mutations.

We show that if antisense molecules are designed for prospective resistance mutations for small molecule antibiotic, then those mutants are rendered inaccessible to evolutionary trajectories. Blocking prospective mutations is very effective since a bacterium has to gain two mutations at the same time: 1) resistance mutation to the antibiotic, 2) mutation to disrupt antisense binding. Furthermore we show that if antisense molecules are designed for present mutations, then resistance mutations are lost.

### 3.3 Methods

#### 3.3.1 Data

Data was obtained from Palmer et.al.. In this publication, 7 mutations at 6 sites for trimethoprim resistance were studied. Total of 89 combinations out of these 7 mutants were constructed. Growth curves at 303 different concentrations of Trimethoprim for each subtype were determined. Three replicates were used at each concentration. With courtesy of authors we have received and used this data.

### 3.3.2 Mutational Pathways

There are 7 mutations in 6 loci; the fifth locus can have 2 mutually exclusive mutations. The codes of mutations are as follows:

100000 = -35C>T

010000 = P21L

001000 = A26T

000100 = L28R

000010 = W30R

000020 = W30G

000001 = I94L

Where '1' and '2' indicates a mutation. Furthermore, '3' or '4' are used for locus, which is resistant to antisense molecule. '3' is used for '1's and 4 is used for '2'.

### 3.3.3 Determining growth rates for subtypes

Each growth curve were fitted with a logistic growth equation, eq 1:

$$B(t) = \frac{K}{1 + \left(\frac{K}{Xo} - 1\right) \mathrm{e}^{-rt}}, eq1$$

Where B is the measured optical density (OD), K is carrying capacity in OD, Xo is initial OD, r is the growth rate. Growth rates, r, were determined for all replicates at all concentrations. r, for a particular subtype at a particular concentration were determined by taking the average of the replicates. After average growth rates were determined at each concentration, smoothing was performed for each subtype to reduce the experimental noise.

### 3.3.4 Minimum inhibitory Concentration (MIC) determination for antibiotics

MIC is defined as the minimum concentration of inhibitory substance that inhibits visual growth. Visual growth in our simulation is taken to be 0.018 OD at 15 hours. Growth rate for the corresponding case is 0.15 $h^{-1}$. MIC is determined by finding the zero of eq 2.

$$f(c) = \left(0.15 - \frac{1}{B_{tot}} \sum_{i \in M} r_i(c)\, B_i\right)^2 , eq2$$

Where M is the set of all subtypes, $r_i(c)$ is the growth rate of the $i^{th}$ subtype at concentration c, Btot is the total bacteria at the calculation time.

### 3.3.5 Modeling Joint effect of antisense molecules and antibiotic concentration

We have modeled the effect of antisense as a probability of occluding mRNA molecules. Effect of multiple antisense are modeled as independent of each other. Furthermore growth rate of the bacteria is assumed to be proportional with respect to maximum free mRNA level. Specifically we computed the growth rates of effected subtypes by eq. 3.

$$R_i(c, A) = r_i(c) \prod_{k \in S} (1 - a_k) , eq. 3$$

Where $a_k$ is the effect of antisense, $r_i(c)$ is growth rate of subtype *i* at antibiotic concentration of c, *S* is an intersection of two sets; set of mutations that are present in subtype *i*, and employed antisense types. $a_k$ 's are determined such that eq2 equals to 0 for fix *k*. All $a_k$ 's were kept same.

### 3.3.6 Simulations

Simulation method is same as in chapter 2. Parameters are adapted to this chapter; growth rates are substituted with $R_i(c)$'s, $\varphi$ is taken to be 3.2 $10^{10}$ cells (equivalent of 100 ml flask with 0.4 OD bacteria), and mutational pathways for all subtypes are determined and substituted to eq.1 in chapter 2.

## 3.4 Results

### 3.4.1 Evolution of trimethoprim resistance at MIC concentration

7 different mutations were successfully isolated from trimethoprim resistant bacteria. There were 6 loci that these 7 mutations could have taken place. In our study the 5$^{th}$ loci has two potential mutations. Viable combinations of 7 mutants were cataloged; total of 89 mutants. For each mutant growth curves at different trimethoprim concentration were measured. Each mutant is named two ways: its genotype, e.g. [0,0,0,1,0,0] and its code e.g. [0,0,0,1,0,0] mutant is regarded as the subtype '7'. This nomenclature is used interchangeably and in figures both are included. For full list of codes and genotypes refer to supplementary table 1.

Growth data of 89 mutants with various trimethoprim concentrations were fitted with logistic growth curve to extract growth rates. Thus for every mutant, growth rates for range of Trimethoprim concentrations were determined. These growth rates enable us to simulate the evolutionary course in full detail, e.g. population dynamics of all mutants under any antibiotic concentration through time.

As a control for antisense therapy we started with simulating the system only with trimethoprim. In a liquid culture there are two independent parameters that can be changed, the volume of the culture (number of bacteria the culture can have), and selection pressure. We have

used 100ml (holds 4 $10^{10}$ bacteria) volume. Simulations are made by inoculating 200 fold diluted bacteria into flask. Then the flasks are left to grow on standard conditions. After reaching stationary phase the bacteria are inoculated again into fresh media.

Antibiotic concentration for each flask is determined by calculating the necessary concentration needed to decrease the growth rate to preset value. This specified "threshold growth rate" is taken to be the selection pressure. We have set the selection pressure threshold to be 0.15 $h^{-1}$. This growth rate brings 1/200 (0.0020 optical density (OD) in this experimental setting) diluted bacteria to 0.018 OD at end of 15 hours of growth. This can be regarded as the growth rate that would be achieved at Minimum inhibitory concentration (MIC). MIC is the lowest concentration of antibiotic needed to prevent visible growth. Also the 0.15 $h^{-1}$ is the lowest threshold that is usable with the data. Lower thresholds introduce high noise.

In Figure 3.1 we have reported the results of simulations performed within 100 ml flask and with selection threshold of 0.15 $h^{-1}$. In Figure 3.1A, minimum inhibitory concentration (MIC) for the trajectories are plotted. In this particular simulation 100 trials were performed. The red line indicates the mean MIC of all trajectories. The mean MIC increases step wise. MIC increases to 1053 microgram/ml at 5 days, then increases to 2041 by 14 days, and furthermore increases to 2450 micrograms/ml by 25 days. The final MIC is 2450 micrograms/ml at 80 days. The underlying genotypes for each simulation are further analyzed in Figure 3.1B. For each trajectory dominant genotype is determined. Percentages of dominant genotypes among 100 simulations were plotted. For example if percentage is 50, it means that particular genotype is dominant in 50 out of 100 simulations. At the start of the experiment all of the simulations has wild type (WT, [0,0,0,0,0,0]) as a dominant genotype. At 5 days the $4^{th}$ loci mutant ([0,0,0,1,0,0]) is dominant in all simulations. At 14 days double mutant of $3^{rd}$ and $4^{th}$ loci

([0,0,1,1,0,0]) becomes dominant in all simulations. At 25 days triple mutant ([1,0,1,1,0,0]) becomes dominant in all simulations. Exemplary trajectory is plotted in Figure 3.2. Frequencies of subtypes are plotted in log scale through time. Log scale is chosen to monitor every scale at once. Different colored curves indicate different subtype's frequency as indicated by the legend of the figure. Population goes through 3 changes; WT to ([0,0,0,1,0,0]) at 5 days; ([0,0,0,1,0,0]) to ([0,0,1,1,0,0]) at 14 days; ([0,0,1,1,0,0]) to ([0,1,1,1,0,0]) at 25 days.

The sequence of events is easily predictable when growth profiles for antibiotic concentrations for each subtype is considered. In Figure 3.3A, growth profiles of WT and its mutant are plotted; the arrow indicates the direction of population change at MIC. Specifically, MIC for WT is 12,27 micrograms/ml. At this concentrations two subtypes ([0,0,0,1,0,0], and [1,0,0,0,0,0]) are the most favored type. As seen Figure 3.2, single mutants arise around same time and they both increase substantially, however as mutant take over MIC increases and eventually the subtype ([0,0,0,1,0,0]) wins. In Figure 3.3B, growth profiles of ([0,0,0,1,0,0]) and its mutant are plotted; the favored double mutant is ([0,0,1,1,0,0]). In Figure 3.3C, growth profiles of ([0,0,1,1,0,0]) and its mutant are plotted; the favored double mutant is ([1,0,1,1,0,0]).

### 3.4.2 Using Antisense to attenuate resistance

Using antisense molecules as an antibiotic has a big advantage. In theory any gene and sequence can be targeted specifically. This opens up 3 major venues: 1) Killing bacteria with antisense by targeting essential and/or virulence genes. 2) Attenuation of resistance by blocking mutants from arising; designing antisense for potential mutations. 3) Sensitizing conventional antibiotic resistance by targeting resistance mutations. We have investigated the first venue in chapter 1 and chapter 2. In this paper we will discover the second and third venue.

From the design perspective, second and third venues are the same. They are same in the sense that the goal is to design antisense for particular mutations. For example, the trimethoprim mutants in our study have single nucleotide changes. Procedure of designing an antisense for these mutations is similar for either blocking or killing. The difference is; if an antisense is used against a mutant that was already present in population, it will kill it or at least decrease its growth rate. If an antisense is used against a mutant that is not already present in population, then the antisense will prevent it from arising; as soon as that mutant arises it will be faced with a selection pressure. We call this blocking a mutant by antisense.

For attenuation we started 3 replicates of simulations for each subtype. A particular simulation starts with one of the 89 subtypes; meanwhile we block any potential mutations with antisense. For example, if we start with single mutant ([1,0,0,0,0,0]) then the following mutations are targeted with antisense: $2^{nd}$, $3^{rd}$, $4^{th}$, both mutations at $5^{th}$, and $6^{th}$ loci. In Figure 3.4, we have plotted the MIC values for subtypes with proper blocking; MICs of all represented trajectories stays fixed, thus resistance progression is halted. Some subtypes were omitted in this figure because they have gained resistance thorough loss of mutations rather then gain of mutations; in nature it is unlikely that a deleterious mutation would be fixed in the first place.

WT is also among the subtypes that have arrested resistance acquisition. Its growth profiles is given in Figure 3.5; due to antisense treatment, all descendants have lower growth rates compared to WT. This particular fact is clearly shown in an example trajectory in Figure 3.6. The trajectory starts with WT and it never diminishes. All of the descendent are selected against and they stay low thorough out the simulation.

The MICs of all subtypes with blocking are plotted in Figure 3.7; subtypes having deleterious mutations gain resistance by loss of mutations, and then their resistance level gets

arrested. For comparison we have included resistance acquisition of all subtypes without blocking in Figure 3.8 MIC for All trajectories increases substantially to 2450 micrograms/ml.

Blocking strategy can also be used to lure bacteria to a mutation of interest. If potential mutations are partially blocked, then the blocked mutations will be selected against, but leftover mutations will readily mutate. Thus with antisense particular descendant can be favored rationally. To demonstrate this we have simulated 100 trials with the following conditions: simulations start with WT, $1^{st}$, $2^{nd}$, $4^{th}$, second mutation of $5^{th}$, and $6^{th}$ loci are blocked, and $3^{rd}$ and first mutation of $5^{th}$ loci are not blocked. Outcome of these simulations are shown in Figure 3.9. In Figure 3.9A, MIC values for trajectories and mean MIC value are plotted. MIC value starts at 12.27 micrograms/ml and ends up in 131 micrograms/ml. In Figure 3.9B, percentage of dominant-subtypes among trajectories are plotted; WT bacteria is replaced with subtypes '3' and '12', then double mutant '14' takes over the population by 9 days. Subtype '14' has the genotype [0,0,1,0,1,0], which are expectedly the mutations that we did not target. Bacteria has converged on a designated mutations, we termed this 'luring bacteria'. Exemplary trajectory is plotted in Figure 3.10. Population starts with a WT and subsequently, single mutants, [0,0,1,0,0,0] and [0,0,0,0,1,0] arise and grow at the same time; ultimately their growth displaces the WT by 5 days. They both give rise to the double mutant [0,0,1,0,1,0], and then eventually that double mutant dominates the population by 8 days.

These sequences of event are easily observed from the growth profiles in Figure 3.11. In Figure 3.11 growth profiles of WT, subtypes '3', '12', '14' are plotted in subfigures A, B, C, D respectively. Black arrows indicates the most probable path as: WT to '3' and '4', '3' to '14', '4' to '14'.

Also, this ability to block mutants opens up an alternative use for antisense. That is, antisenses can be used to block known mutations so unknown mutations takes place. By this method one can discover unknown mutants and have a more comprehensive catalog of mutations. This is again an example of directed evolution to discover different mutants.

### 3.4.3 Sensitizing bacteria

We have dealt with blocking evolution for arresting resistance progression. The last case to consider is to sensitize already resistant bacteria. For our simulations we took a similar approach as we did in blocking; we started with all subtypes, but this time we proposed antisense pressure to all mutations.

Bacteria possess two options to alleviate the selection pressure that antisense exerts: loose trimethoprim resistance mutations, or mutate another nucleotide in the target site. The latter is quite probable in the odd case of trimethoprim resistance; some resistance mutations for trimethoprim have no cost to replication. Natural course events would be as following: some mutations might be lost but eventually trimethoprim resistant subtypes will gain antisense resistance. This means that in expense of antisense resistance we can sensitize trimethoprim resistance, at least to some degree. Moreover this price may be tolerable since redesigning antisense to altered sequences is relatively easy. We devised a simulation study in which the conditions are: simulations will start with one of the subtypes, every subtype will be simulated in 3 replicates, trimethoprim concentration set to 0, antisenses targeted to all mutations and redesigning antisense sequences every 5 days. In Figure 3.12A, MIC and mean MIC values are shown; mean MIC diminishes all the way to 20 micrograms/ml by 33, however fluctuates between this value and 100 micrograms/ml. In Figure 3.12B, dominant subtypes by mutation count among trajectories are plotted. System starts with mix of mutation counts and ends up with

WT or single mutants. Single mutants persist because odd nature of trimethoprim resistance; resistance mutations does not come with a cost. Eventually we suspect that they would die out, since all trajectories had to be single mutant at one point and majority are fully recovered. Exemplary trajectory is plotted in Figure 3.13; it shows that after 7 cycles of redesigning, subtype '63' converts to 'WT'; subtype '63' is the mutant resulted from unrestricted trimethoprim treatment. Growth profiles of sensitization of '63' are plotted in Figure 3.14; they signify the high growth rate of antisense resistant strains. Thus redesigning is a requirement for sensitization.

We have devised a strategy to sensitize bacteria with antisense without need for redesigning. The key is realize that all mutants are impaired with antisense and WT is not, only problem is antisense fails before population converts to WT. One solution would be to introduce 'WT' exogenously. Alongside with antisense molecules we introduced 'WT' in ratio of 1:100000. In Figure 3.15A, we show that all strains remarkably sensitize by 2 days, mean MIC decreases to 12.17 micrograms/ml by 2 days and subsequently increases very slightly. In Figure 3.15B, dominant subtypes by mutation count among trajectories are plotted. Trajectories start out with mix of mutations and very rapidly WT dominates all trajectories, however some trajectories ends up with single mutants. Close inspection of shows these trajectories also started with single mutations. Any subtypes starting with double or more mutations were sensitized fully. The reason underlying this observation is that, with double mutations an antisense resistance mutation can only alleviate part of the selection pressure, thus its growth rate becomes less than WT.

It is important note that, mutants are not being lost in this case; they are out computed with their ancestor. Example trajectory of sensitizing subtype '63' is given in Figure 3.16. At

start of simulation only subtype ’63’ and WT are present, after 1.5 days WT dominates completely. In Figure 3.17, growth profiles of subtype ‘63’ and its descendants with addition of WT are given. The profiles signify how antisense molecules have impaired all mutants and WT is unaffected. The black arrow shows the likely jump from ‘63’ to WT.

By the strategies we have developed we devised an additional way to sensitize single mutants as well. The strategy is to lure the single mutants to be double mutant and then introduce full antisense targeting and a small amount of ‘WT’. In Figure 3.18, we are reporting 50 trials of sensitizing a nasty single mutant ‘7’,([0,0,0,1,0,0]). Specifically the strategy is to let bacteria to gain an additional mutation and convert to ‘10’, ([0,0,0,1,2,0]); by not targeting $3^{rd}$ and second mutation of $5^{th}$ loci. Afterwards sensitization with WT will be carried out. In Figure 3.18A, MIC and mean MIC values are plotted. MIC increases from 1052 micrograms/ml to 1414 micrograms/ml by 25 days. At 40 days sensitizing with WT is carried out; WT is introduced with ratio of 1:100000 and antisenses are targeting all mutations. Eventually, mean MIC decreases to 12.5 micrograms/ml by 42 days. One example trajectory is plotted in Figure 3.19. The population starts with ‘7’, then changes to ‘10’ and finally by introduction of WT, WT dominates the population. In Figure 3.20, growth profiles of subtype ‘7’ and ‘10’ are given. In Figure 3.20A, the growth profiles of subtype ‘7’ and its descendants are plotted. The most probable move is from ‘7’ to ‘10’. In Figure 3.20B, the growth profiles of ‘17’ and its descendants with addition of WT are given. The most likely path is from ‘17’ to ‘WT’.

It is important to note that once a population is dominated fully by WT and if the potential mutations are blocked, then bacteria needs double mutation to gain resistance: mutation to increase Trimethoprim resistance and an addition mutation to render antisense ineffective. The average time of this event can computed by eq 4.

$$\tau = \frac{1}{\max\,(B_{tot})\,m^2} = 1821\ years, eq\ 4.$$

Where maximum of $B_{tot}$ is 2.5415e+09 and m is 5.0000e-09.

### 3.5 Conclusion

In this study we show that antisense therapy can be used to block further resistance to a particular small antibiotic. We have tested our strategy on every possible mutant and none invoked further resistance by gain of mutations. We show that this method is very safe; therapy can only fail by double mutation, which happens every 1821 years in an experimental setting. This strategy can be vital to slow down antibacterial resistance progression. We show additional implications of blocking strategy: converging evolution by restricting paths, and providing a tool for discovering new mutations.

Converging evolutionary path to a certain mutation can be used to lure the bacteria to collaterally sensitive state for another antibiotic. At this instant, sensitized antibiotic can be used jointly with different set of antisense to converge to different collaterally sensitive state. This strategy might delay or inhibit invoking multiple drug resistance.

If all of known mutations are blocked with antisense, then unknown mutations will surface. It is important to note that these mutations will most likely be less rigorous then known mutations. However expanding the catalog of possible mutations might reveal higher order interactions.

Blocking resistance progression is much easier, compared to sensitizing bacteria by inducing loss of resistance mutations. Whenever existing mutations are targeted, bacteria can circumvent the problem by single mutation: loosing the resistance mutation, or mutating other

nucleotide in target sequence. We show that redesigning counters this problem; after 7 redesigns, mean MIC of all subtypes diminishes significantly. This shows the importance of redesign for sensitization procedure.

An easier but more controversial alternative to sensitizing was considered: introducing WT bacteria alongside with antisense therapy. This setup results in full sensitization of double and higher mutants in 2 days, compared to 35 days with redesign strategy. It might seem controversial to swallow pathogenic bacteria; in fact any bacteria that competes with the same resources would be usable. A saying seems fitting here “one nail drives out another”. Single mutations proved to be hardest to get rid of, however we have devised a lure and kill strategy to get rid of them. The strategy is to guide a single mutant become a particular double mutant and then introduce WT and antisense to kill bacteria.

As we have shown antisense technology has potential applications on blocking further progression, guide the population to designated mutations for collateral sensitivity, induce resistance mutation to be lost, favor WT to outcompete mutants.

### 3.6 Figures

| Code | Genotype | Code | Genotype | Code | Genotype |
|---|---|---|---|---|---|
| 1 | '000000' | 31 | '010110' | 61 | '101020' |
| 2 | '000001' | 32 | '010111' | 62 | '101021' |
| 3 | '000010' | 33 | '010120' | 63 | '101100' |
| 4 | '000011' | 34 | '010121' | 64 | '101101' |
| 5 | '000020' | 35 | '011001' | 65 | '101110' |
| 6 | '000021' | 36 | '011010' | 66 | '101111' |
| 7 | '000100' | 37 | '011011' | 67 | '101120' |
| 8 | '000101' | 38 | '011020' | 68 | '101121' |
| 9 | '000110' | 39 | '011021' | 69 | '110000' |
| 10 | '000120' | 40 | '011100' | 70 | '110001' |
| 11 | '000121' | 41 | '011101' | 71 | '110010' |
| 12 | '001000' | 42 | '011110' | 72 | '110011' |
| 13 | '001001' | 43 | '011111' | 73 | '110020' |
| 14 | '001010' | 44 | '011120' | 74 | '110021' |
| 15 | '001011' | 45 | '011121' | 75 | '110100' |
| 16 | '001020' | 46 | '100000' | 76 | '110110' |
| 17 | '001100' | 47 | '100001' | 77 | '110111' |
| 18 | '001101' | 48 | '100010' | 78 | '110120' |
| 19 | '001110' | 49 | '100011' | 79 | '110121' |
| 20 | '001111' | 50 | '100021' | 80 | '111000' |
| 21 | '001120' | 51 | '100100' | 81 | '111001' |
| 22 | '001121' | 52 | '100101' | 82 | '111010' |
| 23 | '010000' | 53 | '100110' | 83 | '111011' |
| 24 | '010001' | 54 | '100111' | 84 | '111021' |
| 25 | '010010' | 55 | '100120' | 85 | '111100' |
| 26 | '010011' | 56 | '100121' | 86 | '111101' |
| 27 | '010020' | 57 | '101000' | 87 | '111110' |
| 28 | '010021' | 58 | '101001' | 88 | '111111' |
| 29 | '010100' | 59 | '101010' | 89 | '111121' |
| 30 | '010101' | 60 | '101011' | | |

**Table 3.1.** Codes of Genotype.

Mutations are indicated by '1' and '2'. '2' is the second mutation at 5$^{th}$ loci. Codes higher then 89 are strains having antisense resistance. They are indicated in figures when needed.

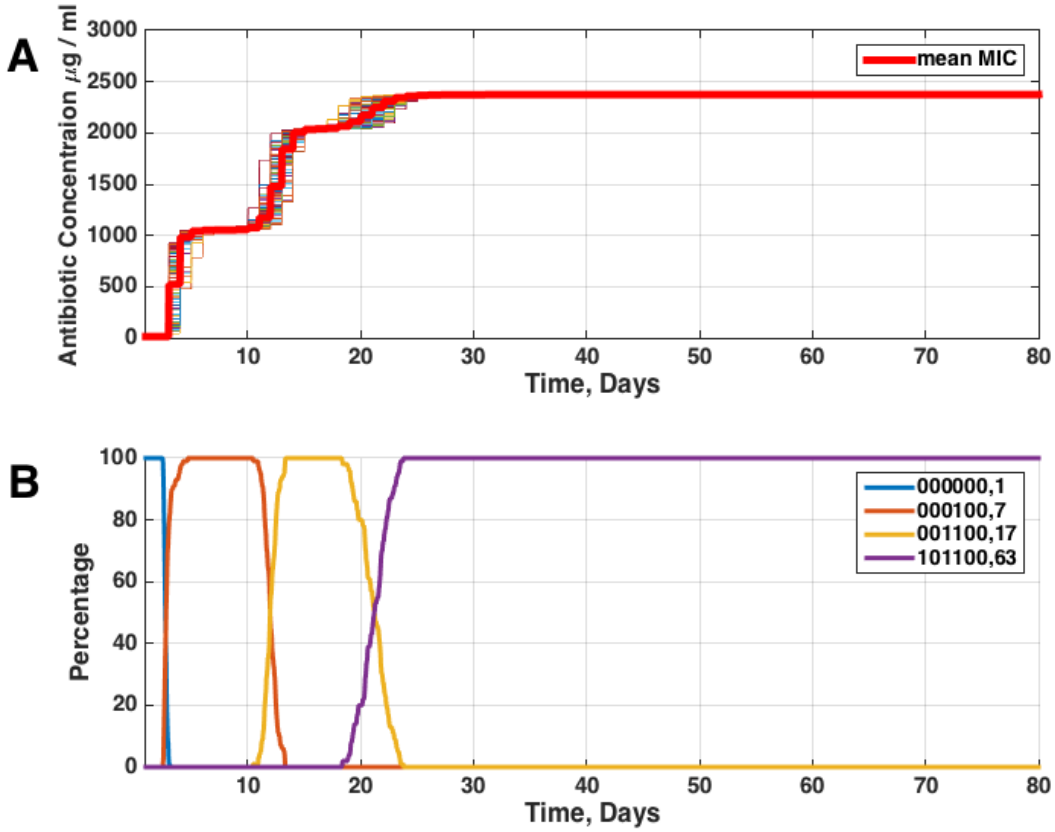


**Figure 3.1.** Antibiotic Resistance Evolution of 100 trials under increasing Trimethoprim concentrations.

100 trials of WT bacteria were evolved for trimethoprim resistance in conditions: 100 ml flask, selection threshold of 0.15 $h^{-1}$. A) Mean MIC values are reported through time. More specifically the evolution of MIC occurs in stepwise fashion; MIC increases from 12.27 micrograms/ml to 1053 micrograms/ml at 5 days; increases to 2041 micrograms/ml at 14 days; increases to 2450 micrograms/ml at 25 days. B) Number of trajectories having a particular subtype is determined and plotted. All of the trajectories start with WT, at 5 days all trajectories are dominated by a single mutant, [0,0,0,1,0,0], at 14 days double mutant, [0,0,1,1,0,0], takeover, and finally triple mutant, [1,0,1,1,0,0], dominates all trajectories by 25 days. The numbers in the legend are the numerical identification numbers for each subtype; they will be used instead when necessary.

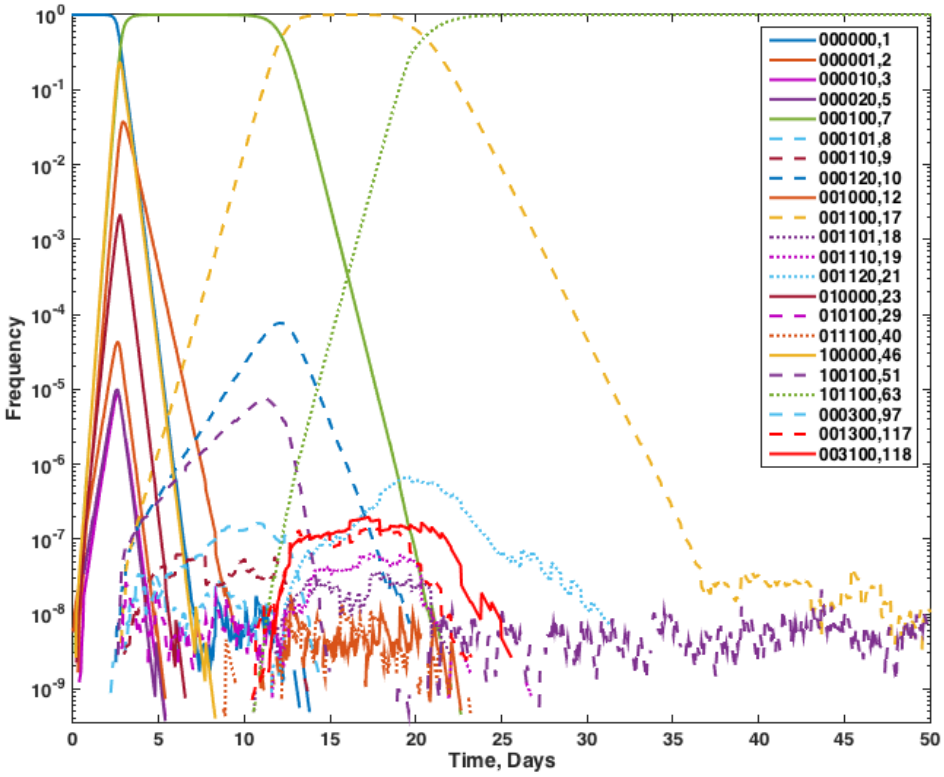


**Figure 3.2.** Sample trajectory of WT bacteria evolved with minimum inhibitor concentration of trimethoprim.

Number of bacteria is normalized to total cell number and frequencies for each genotype is plotted; frequencies are plotted with log scale to show dynamics at each scale. WT bacteria are plotted with solid blue line, whereas single mutants with different colored solid lines, double mutants as dashed lines, and triple mutants as dotted lines. We have intentionally provided non-significant single and double mutants, to show more detailed dynamics. The prominent subtypes are: WT, ‘7’, ‘46’, ‘17’, and ‘63’; as WT decreases subtypes ‘7’ and ‘46’ increases at about the same rate; increase of subtypes ‘7’ and ‘46’ increases the growth rate of the population, as a result trimethoprim concentrations are increased; increase of antibiotic concentration favors subtype ‘7’ and it eventually dominates the population; subtype ‘7’ becomes out competed by its descendant ‘17’ by 14 days; finally subtype ‘63’ dominates the population by 25 days.

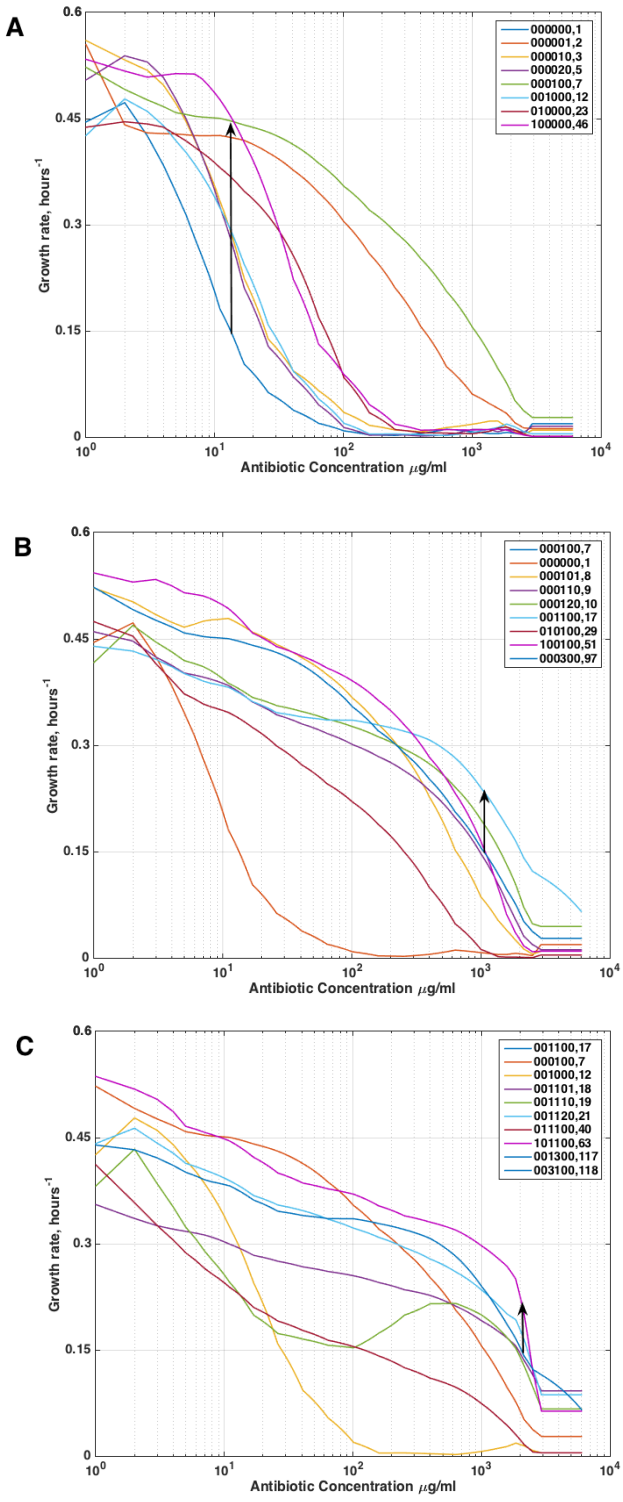


**Figure 3.3.** Growth profiles of subtypes under trimethoprim concentration.

A) Growth profiles of WT bacteria and its 7 single mutant descendants are plotted. WT is plotted by solid blue line. The MIC of WT is easily determinable; concentration at which WT growth rate is 0.15 $h^{-1}$ , 12.27 micrograms/ml. At MIC, subtypes of 7 and 46 have growth rates of 0.4471 $h^{-1}$, and 0.4563 $h^{-1}$ respectively. The black arrow indicates the most probable direction of change for the population; WT to either '7' or '46'. B) Growth profiles of '7' and its descendants are given. '7' is plotted by solid blue line. Most robust subtype at MIC is subtype '17'. The black arrow, from'7 'to '17', indicates most probable change of the population. C) Growth profiles of '17' are plotted. '17' is plotted by solid blue line. Most probable change is from '17' to '63'. In all three images, subtypes having antisense resistance mutations (indicated by 3 in legend) are plotted as solid blue line; Antisense resistant strains have the same growth rate as their ancestor since we do not assume an associated cost.

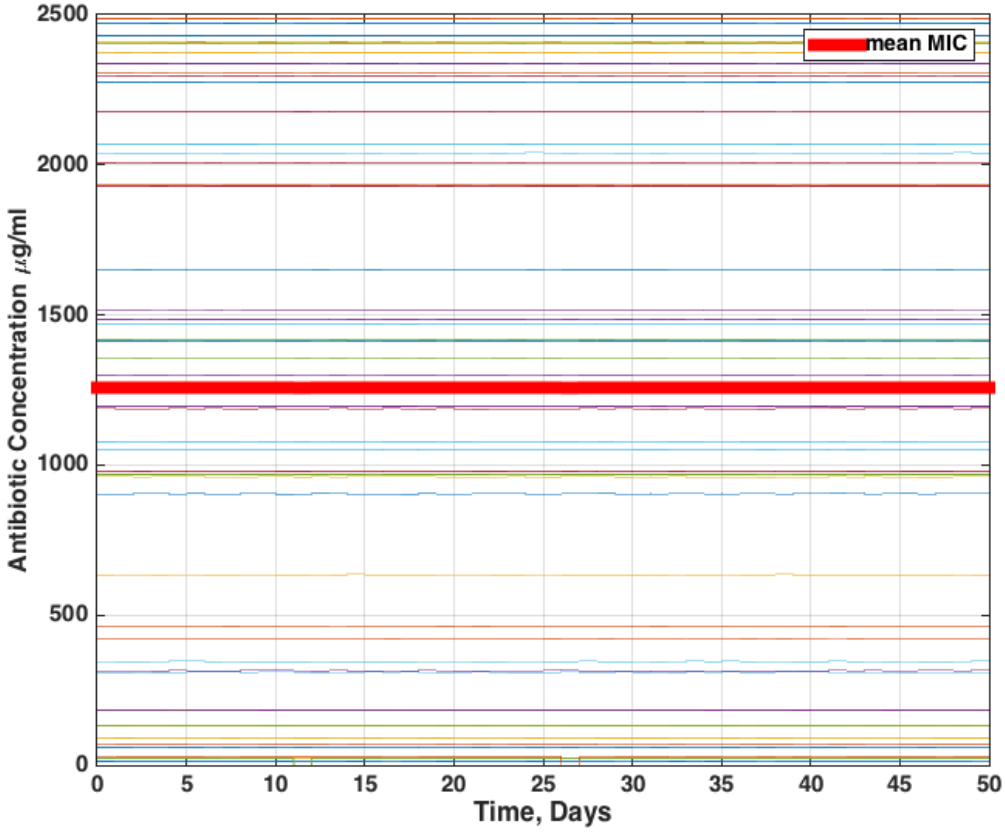


**Figure 3.4.** Blocking potential mutations by antisense halt progression of resistance.

Subtypes not having deleterious mutations were evolved with their potential mutations blocked by antisense. MIC values stayed constant for all of the trajectories. Blocking strategy by antisense has successfully halted further resistance progression.

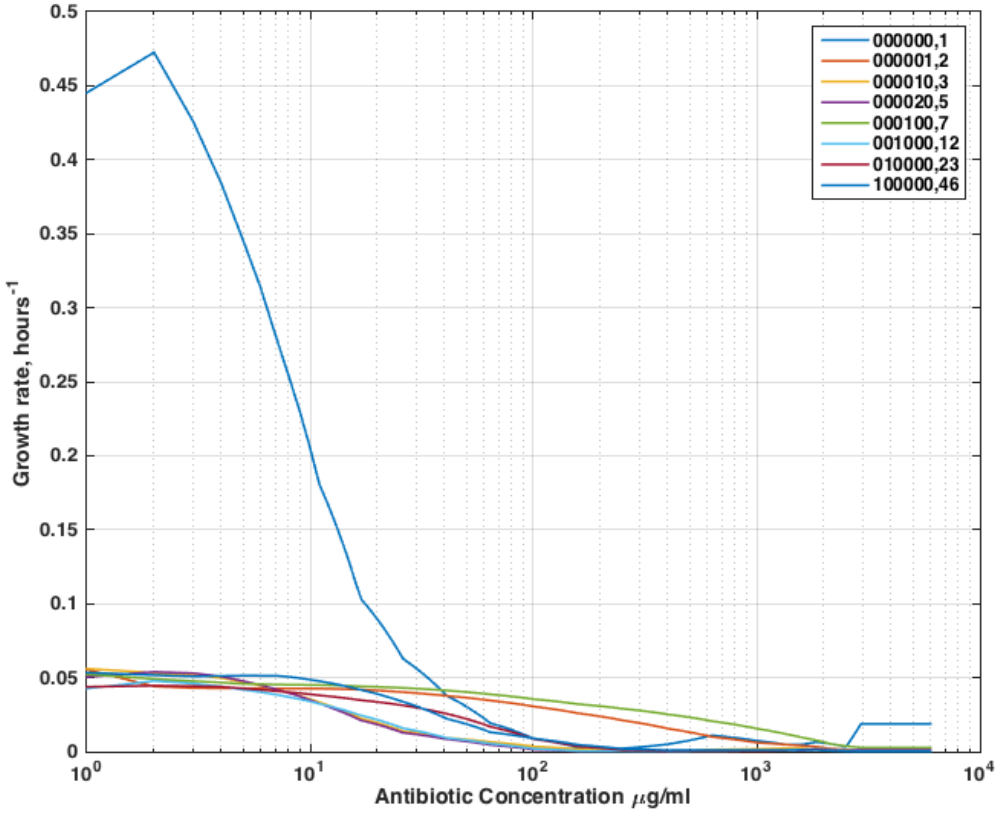


**Figure 3.5.** Growth profiles of WT bacteria with fully blocked mutations.

Growth profile for WT was plotted. In this simulation all of the potential mutations: 1$^{st}$, 2$^{nd}$, 3$^{rd}$,4$^{th}$, both mutations of 5$^{th}$, 6$^{th}$ loci were targeted. This selection pressure decreases descendent significantly and their growth rates become negligible with respect to WT.

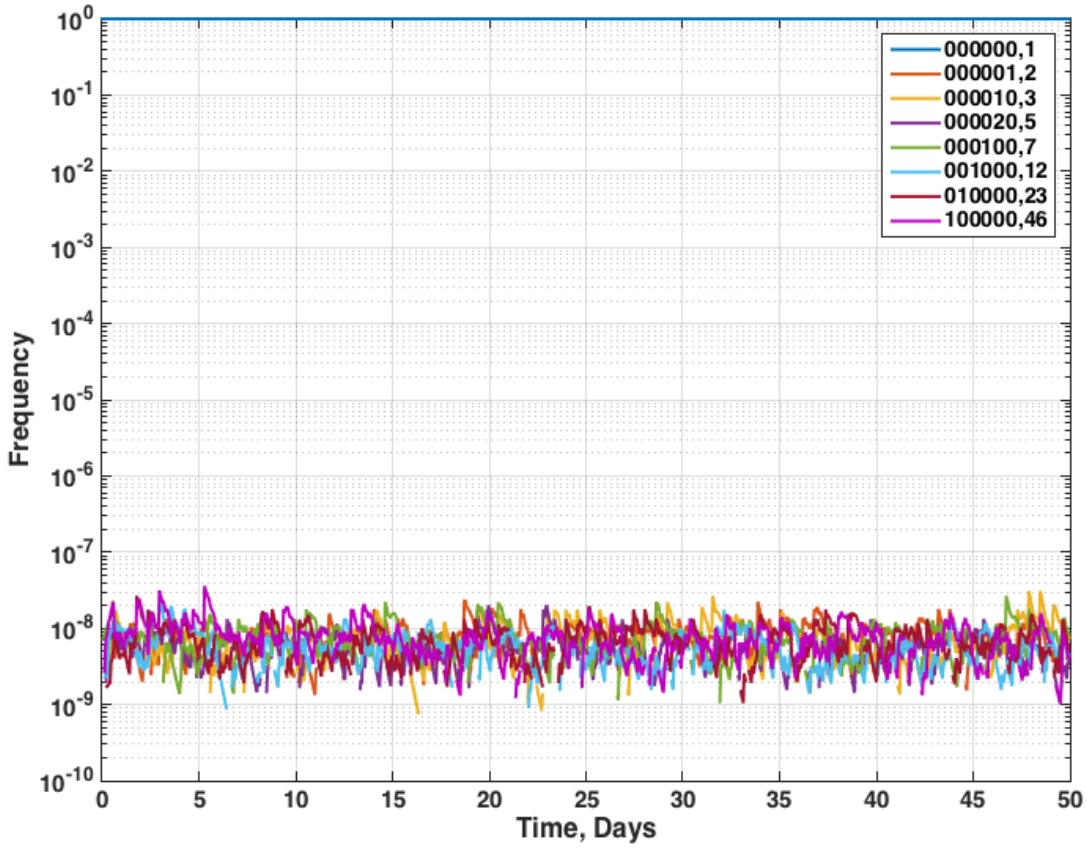


**Figure 3.6.** Trajectory of WT bacteria with fully blocked mutations.

Simulation starts with the WT and it never diminishes. WT's mutants are in the population but since they are selected against they can never grow.

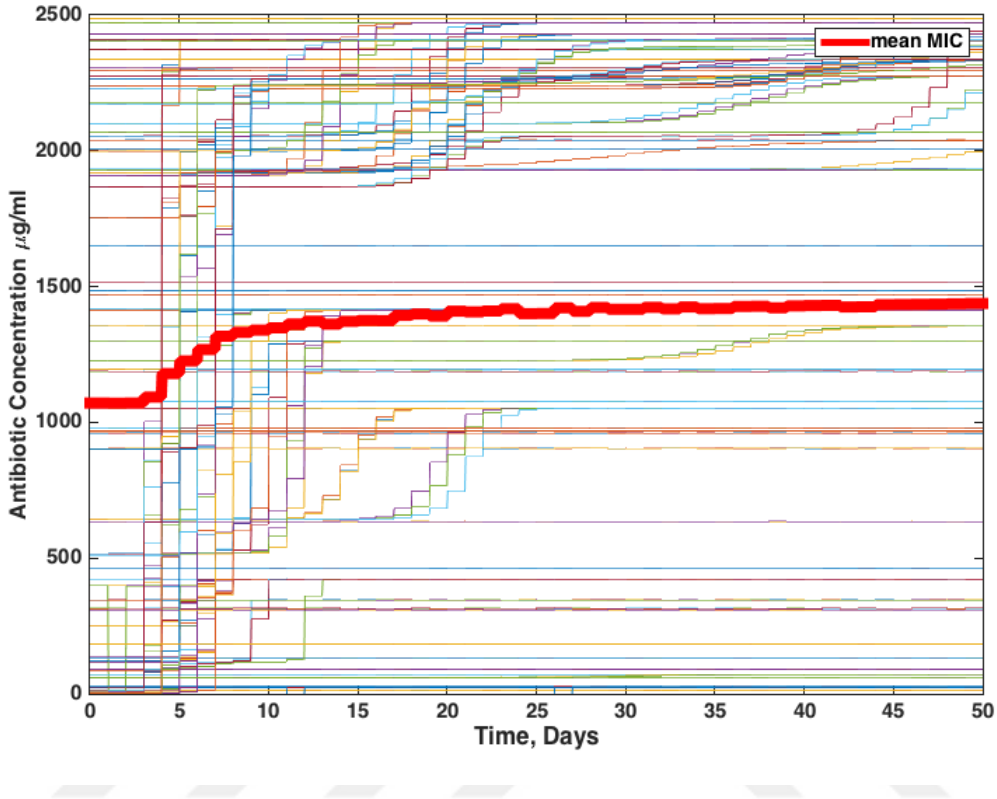


**Figure 3.7.** Blocking potential mutations by antisense for all subtypes.

All subtypes were simulated with antisense blocking. Many of the trajectories stayed at constant MIC levels as shown in Figure 4. However some trajectories gains resistance by loss of mutation, thus they gain some resistance. Mean MIC value increases from 1100 micrograms/ml to 1400 microgram/ml.

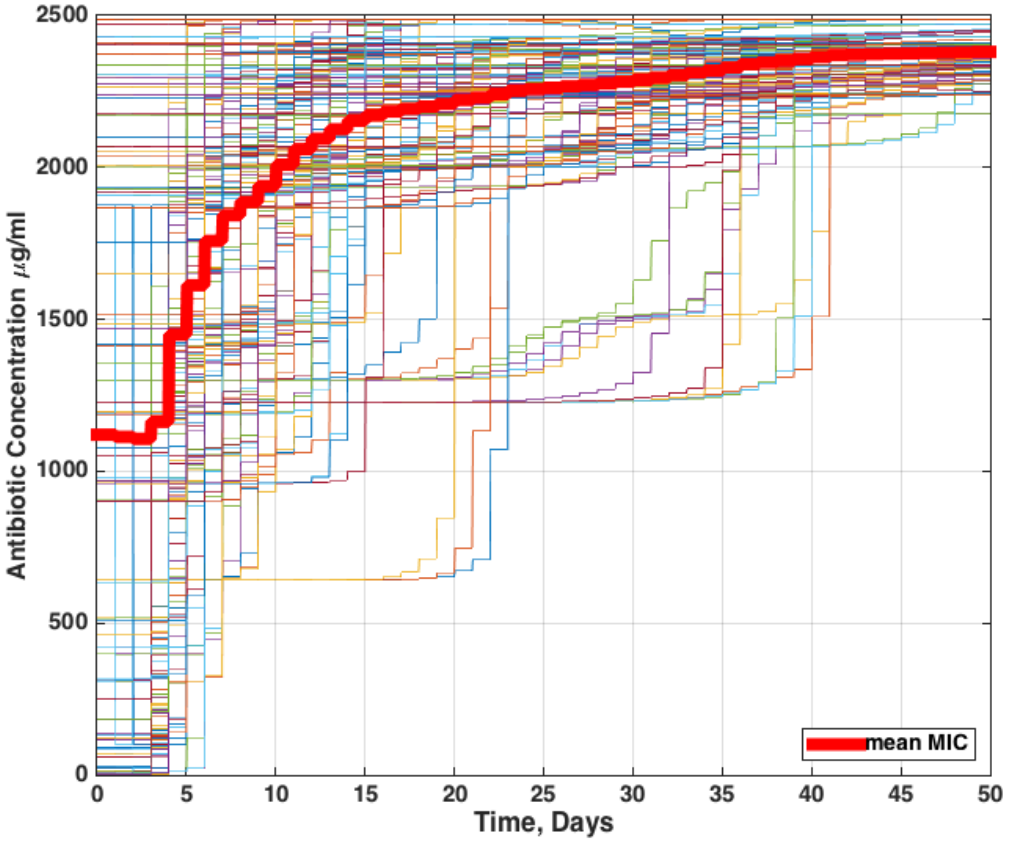


**Figure 3.8.** Resistance acquisition of all subtype without blocking.

All subtypes were evolved just with antibiotic treatment. All of the trajectories increase substantially to reach MIC of 2450 microgram/ml.

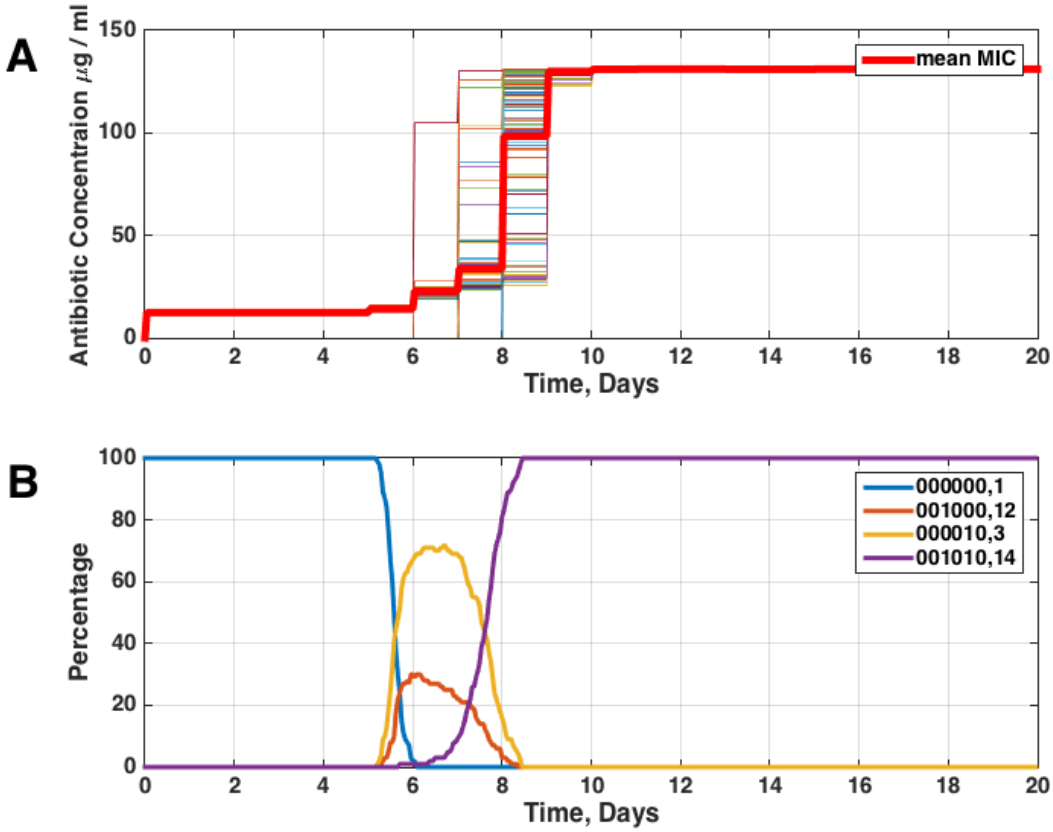


**Figure 3.9.** Blocking some mutations and intentionally not blocking others, lures the bacteria to a defined state.

In these simulations following are the conditions: all trajectories start with WT, 1$^{st}$, 2$^{nd}$, 4$^{th}$ , second mutation of 5$^{th}$ ,and 6$^{th}$ loci are blocked, 3$^{rd}$ and first mutation of 5$^{th}$ loci are not blocked. A) Shows the MIC values for trajectories and the mean MIC value. Mean MIC value starts at 12.27 micrograms/ml and reaches 131 micrograms/ml by 10 days B) All of the 100 trajectories started with WT and probabilistically ended up in either '12' or '3'. Eventually double mutant '14' takes over by 9 days

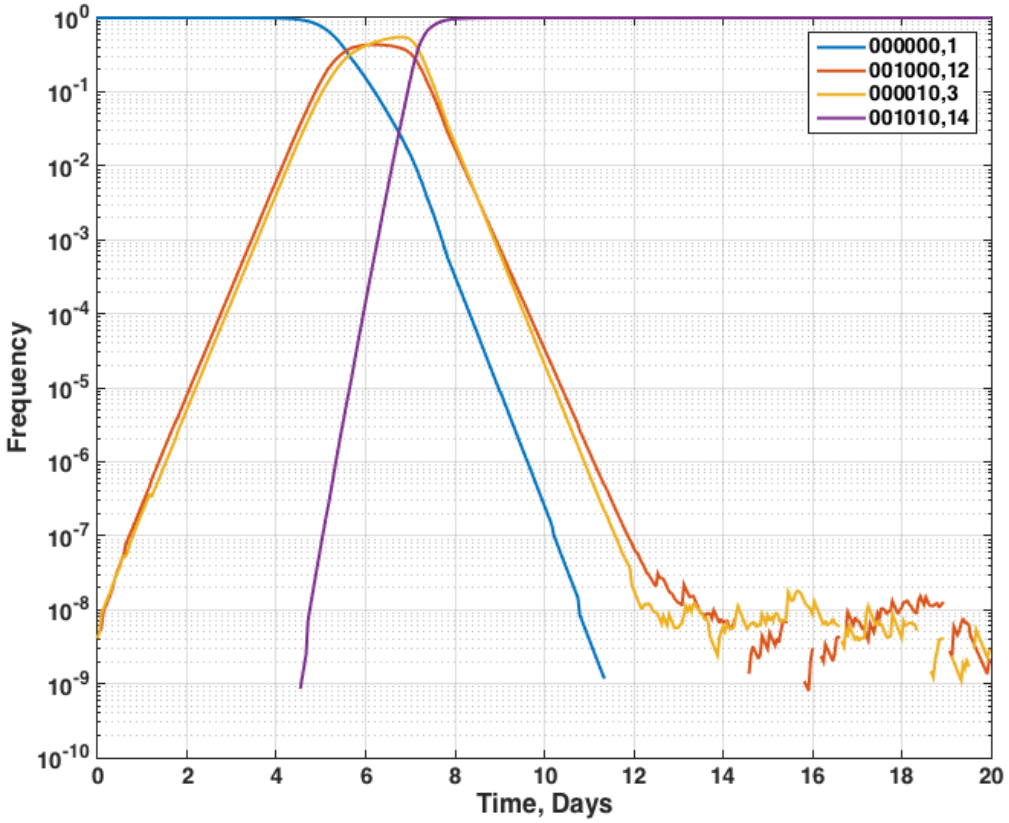


**Figure 3.10.** Trajectory of WT bacteria with partially blocked mutations.

This simulation is made with conditions: starts with WT, 1st, 2nd, 4th , second mutation of 5th ,and 6th loci are blocked, 3rd and first mutation of 5th loci are not blocked. Trajectory starts with WT, and then subtypes '12' and '3' increases to substantial amount to give rise to '14'. Eventually '14' dominates the population.

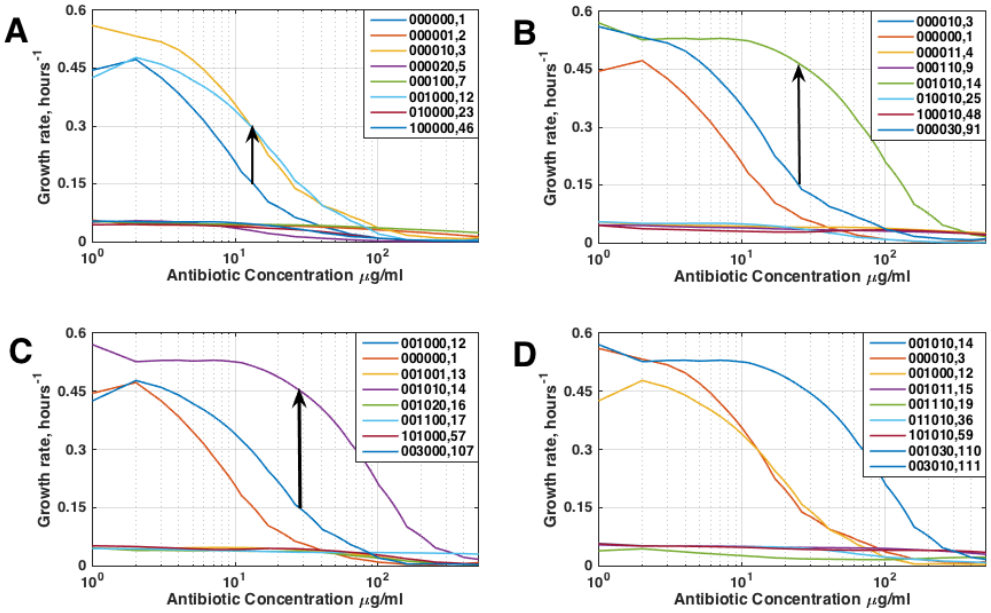


**Figure 3.11.** Growth profiles of simulations with partial blockage.

Growth profiles for: WT, subtypes '3', '12', and '14'. Conditions of the experiment are: 1st, 2nd, 4th , second mutation of 5th ,and 6th loci are blocked, 3rd and first mutation of 5th loci are not blocked. A) Growth profile of WT with its descendants; all of the blocked descendant have significantly lower growth rates, which enable non-blocked subtypes ('3', '12') to be most probable descendant. B) Growth profiles of subtype '3' and its descendants are plotted; most probable descendent is subtype '14', since other mutations are blocked. C) Growth profiles of subtypes '12' with its descendants are plotted; most probable descendent is subtype '14', since other mutations are blocked. D) Growth profiles for subtype '14' and its descendants; none of the descendants have higher growth rate compared to '14', thus resistance halts.

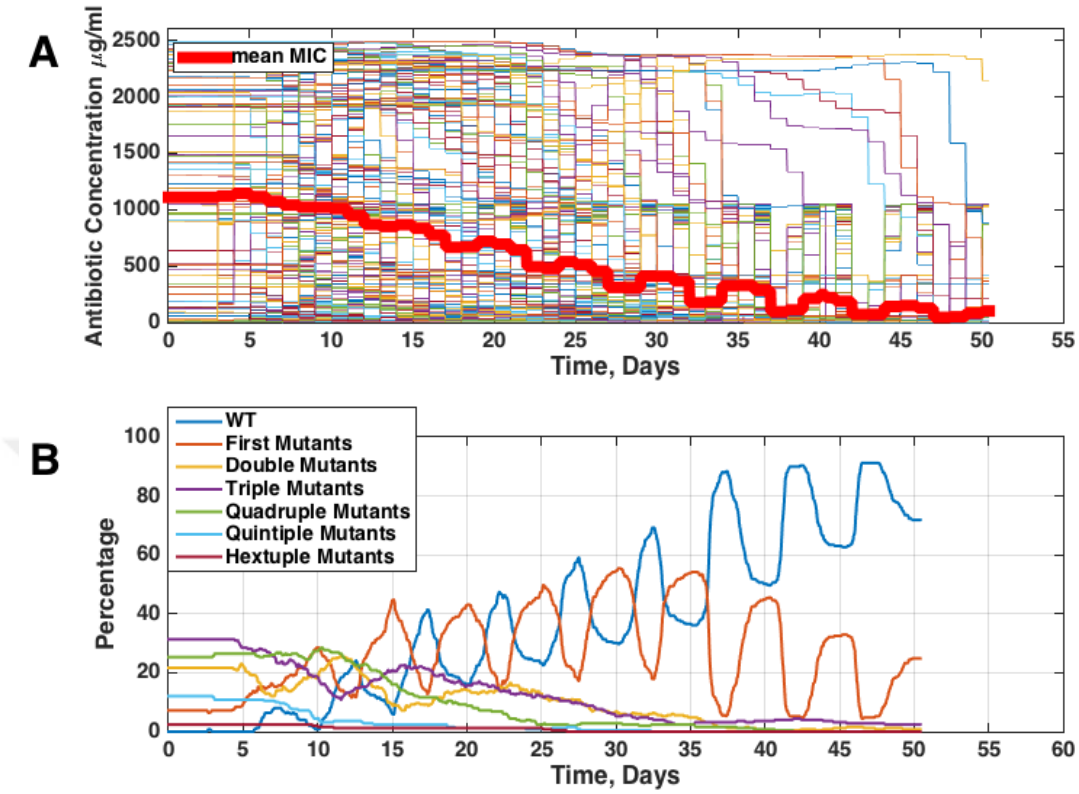


**Figure 3.12.** Antisense directed at all mutations, induces loss of resistance mutations.

All subtypes are simulated with the following conditions: 3 replicates, trimethoprim concentration set at 0, redesign to counter antisense resistance every 5 days. Antisense is proposed to attack all mutations, existing and potential mutations. A) MIC and mean MIC values are plotted; mean MIC decreases substantially to 20 micrograms/ml, however fluctuates between this value and 100 micrograms/ml. B) Shows the percentage of dominant subtypes by mutation count for all trajectories. The system starts with mix of WT, single, double, triple, quadruple, quintuple, hextuple mutants and through time only WT and first mutants are seen in the system. Definitely, number of mutations decrease; however it is harder to get rid of single mutants. We suspect that if this strategy is continued resistance will be lost all together.

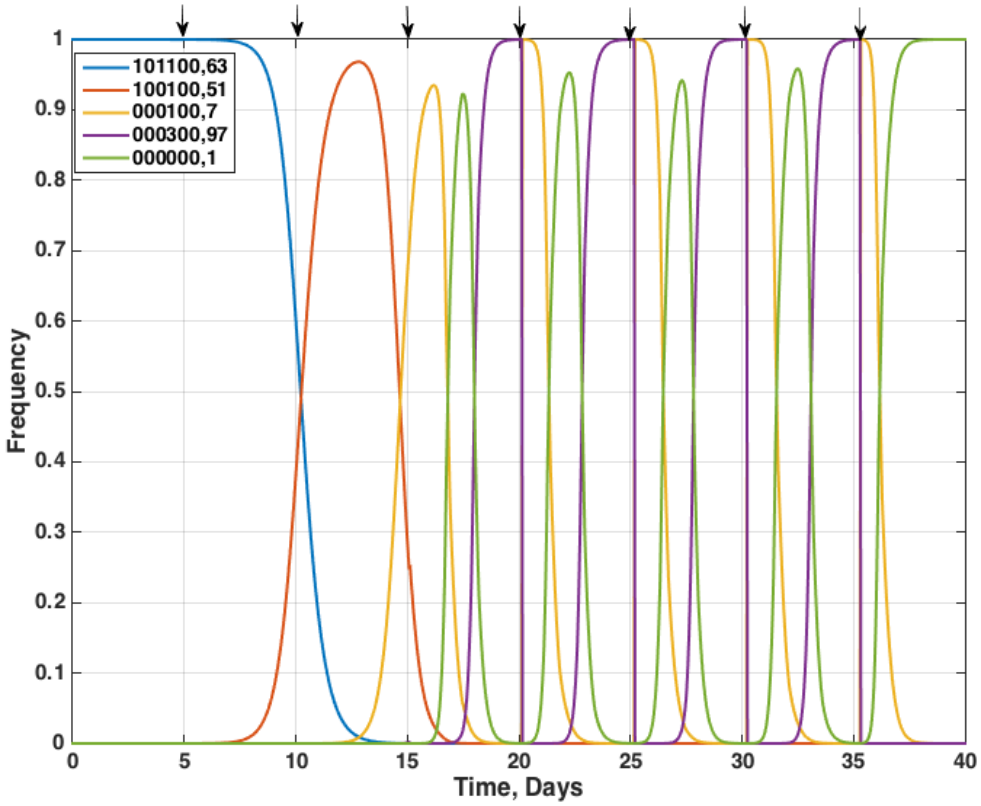


**Figure 3.13.** Sensitization of subtype ‘63’ by antisense molecules.

Triple mutant, ([1,0,1,1,0,0]), or subtype ‘63’ was the result of our simulations for unconstraint trimethoprim resistance evolution. In this simulation we used antisense targeting to sensitize subtype ‘63’. This simulation is made with conditions: starts with subtype ‘63’, all of the mutations are hit by antisense, trimethoprim concentration set at 0, every 5 days redesign is made to counter antisense resistance. Trajectory starts with ‘63’, and then subtype ‘51’ dominates the population until subtype ‘7’. After first appearance of subtype ‘7’, WT temporarily dominates the populations before antisense resistant strain ‘97’, [0,0,0,3,0,0]. Subtype ‘97’ is converted to ‘7’ with redesigning the antisense molecules at 20 days. Until the next redesign similar sequence of events happen; WT temporarily dominates before subtype ‘97’. This cycle reoccurs two more times and at the third cycle ‘WT’ eventually dominates the population permanently. Note that this trajectory is not presented with log scale.

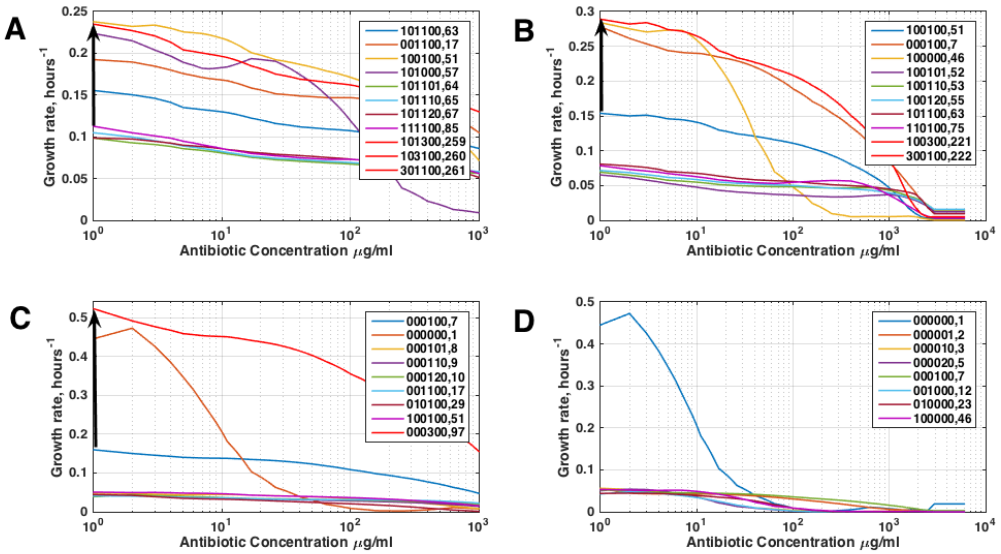


**Figure 3.14.** Growth profiles of major subtypes in sensitization of subtype ’63’.

Growth profiles for the following subtypes are plotted: WT, ‘3’, ’12’, and ’14’. Conditions of the experiment are starts with subtype ‘63’, all of the mutations are hit by antisense, trimethoprim concentration is set at 0, every 5 days redesign is made to counter antisense resistance. A) Growth profile of ‘63’ with its descendants; most probably descendent is ‘51’ however antisense resistant mutants are almost as probable. B) Growth profiles of subtype ‘51’ and its descendants are plotted; most probable descendants are : subtype ‘221’, ‘222’, ‘7’, ’46’. C) Growth profiles of subtypes ‘7’ with its descendants are plotted; most probable descendants are: ‘WT’ and an antisense resistant mutant ‘97’ D) Growth profiles for subtype ‘WT and its descendants; none of the descendants have higher growth rate compared to ‘WT’.

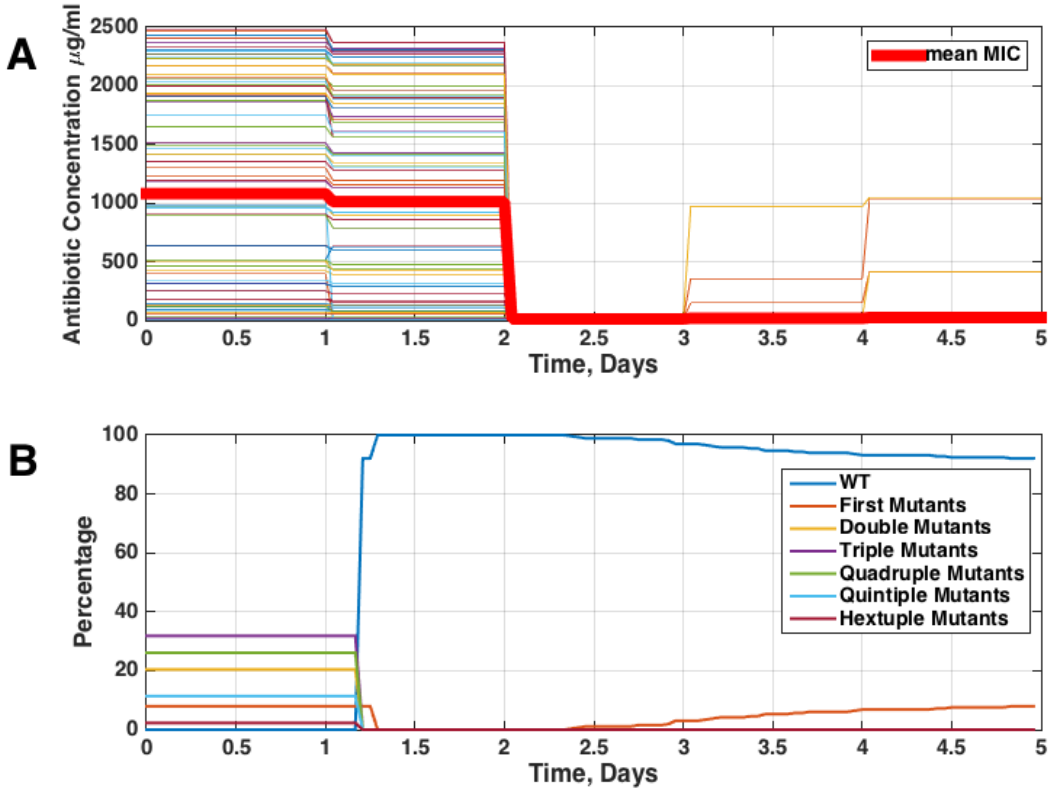


**Figure 3.15.** Introducing small amount of 'WT' with antisense therapy sensitizes populations very fast, even without need of redesigning sequences.

All subtypes are simulated with the following conditions: 3 replicates, trimethoprim concentration set at 0, 1:100000 ratio of 'WT' is introduced at the start of the simulation; antisense is proposed to attack all mutations. A) MIC and mean MIC values are plotted; mean MIC decreases substantially to 18 micrograms/ml by 5 days. B) Shows the percentage of dominant subtypes by mutation count for all trajectories. The system starts with mix of WT, single, double, triple, quadruple, quintuple, hextuple mutants and through time only WT and first mutants are seen in the system. Remarkably only 1.5 day is needed to sensitize any subtype. Single mutants persist because odd nature of trimethoprim resistance; resistance mutations does not come with a cost.

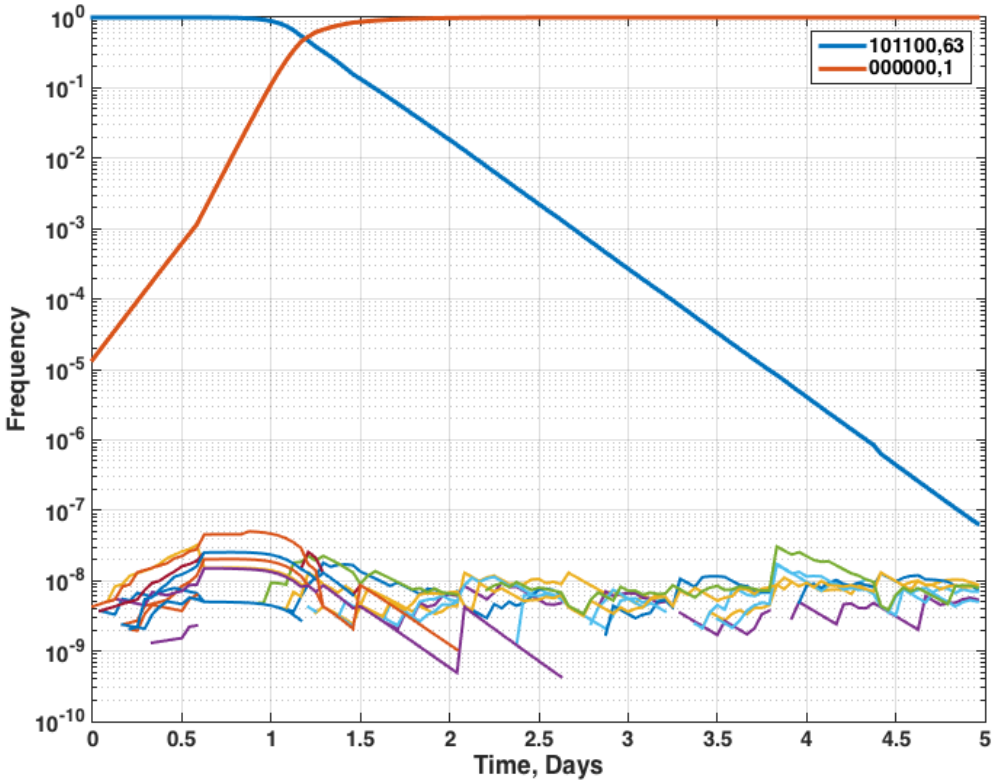


**Figure 3.16.** Sensitization of subtype ‘63’ by antisense molecules and introduced ‘WT’.

Triple mutant, ([1,0,1,1,0,0]), or subtype ‘63’ was the result of our simulations for unconstraint trimethoprim resistance evolution. In this simulation we used antisense targeting and ‘WT’ introduction to sensitize subtype ‘63’. This simulation is made with conditions: starts with subtype ‘63’ trimethoprim concentration set at 0, 1:100000 ratio of ‘WT’ is introduced at the start of the simulation; antisense is proposed to attack all mutations. Trajectory starts with ‘63’ and ‘WT’. ‘WT’ wins over ’63’by 1.5 days. Other descendants of ‘63’ and ‘1’ are not commented on the legend, they stay very low and they never grow.

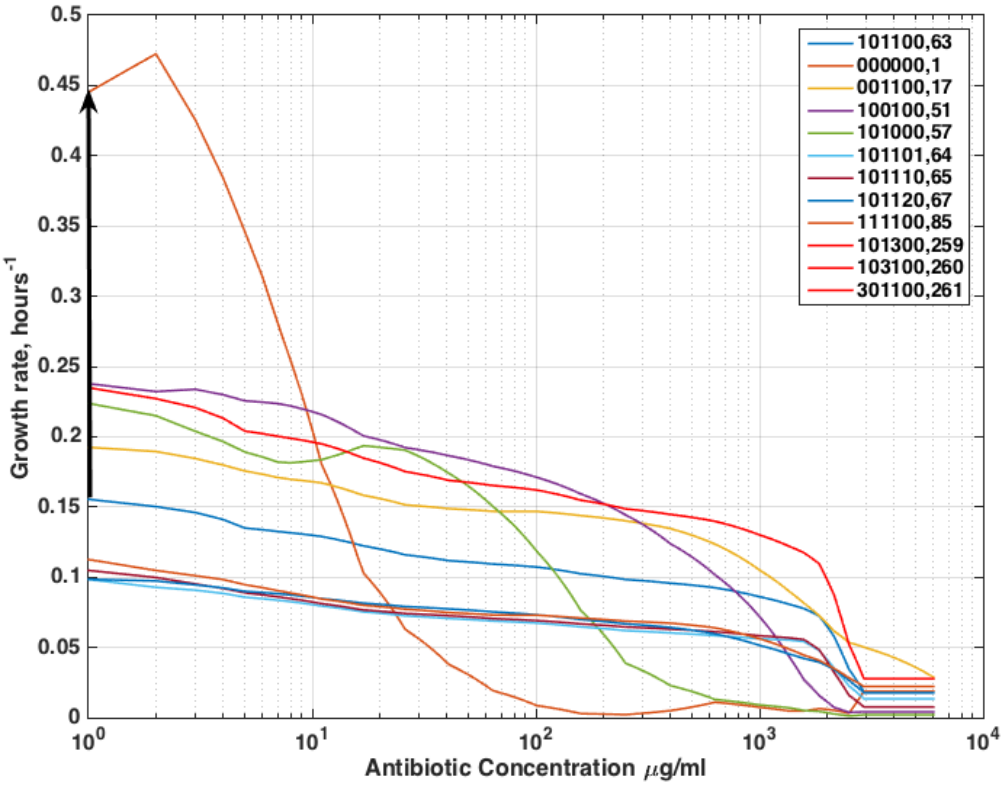


**Figure 3.17.** Growth profile of '63' and its descendent with addition of 'WT'.

Growth profile of '63' and its descendants with addition of 'WT'; most probably outcome is that population will jump from '63' to 'WT'. The rest of the descendants are impaired by antisense molecule.

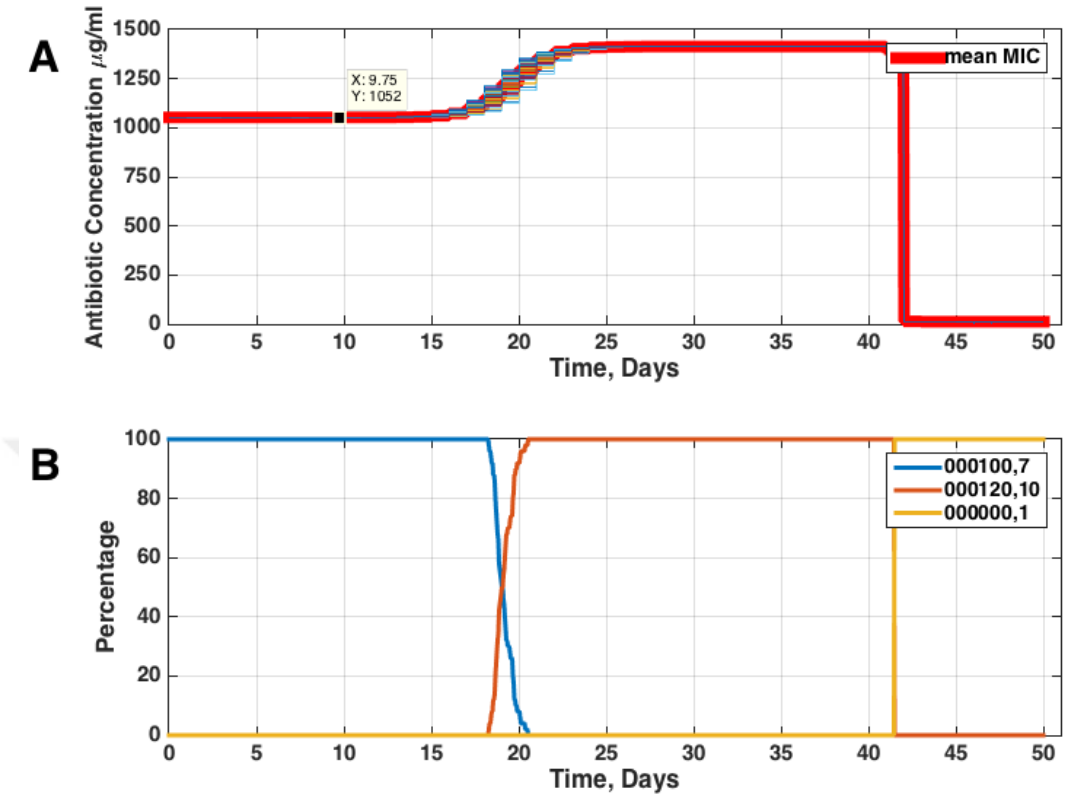


**Figure 3.18.** Luring single mutants to be double mutant, then hitting with antisense and introduced 'WT'.

In these simulations following are the conditions until 40 days: all trajectories start with subtype '7', all mutations besides 4$^{th}$ and second mutation of 5$^{th}$ loci are targeted, Trimethoprim is used at minimum inhibitory concentration. Following conditions are imposed after 40 days: all mutations are targeted, trimethoprim concentration kept at 0, and small 'WT' is introduced at 40 days A) Shows the MIC values for trajectories and the mean MIC value. Mean MIC value; starts at 1052 micrograms/ml and reaches 1414 micrograms/ml by 25 days; After long stasis, mean MIC drops to 12.5 micrograms/ml by 42 days. B) All of the 100 trajectories started with subtype '7' and subsequently converted to '10'. Eventually, at 42 days WT takes over the population very rapidly.

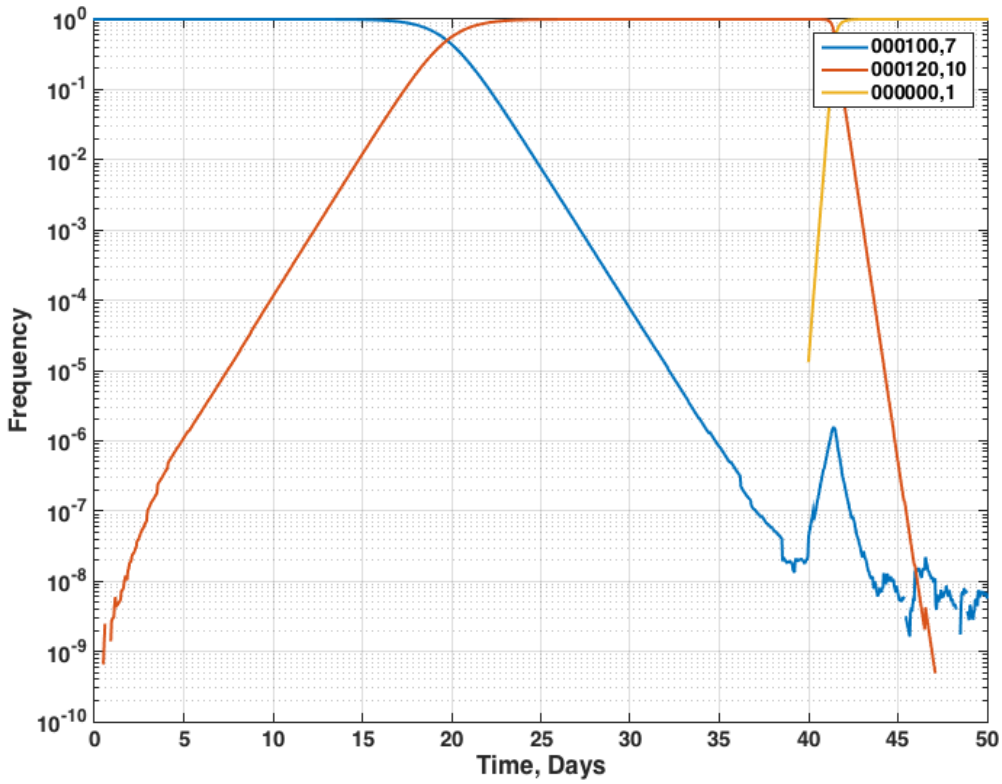


**Figure 3.19.** Example trajectory for lure and kill strategy used on subtype '7' ([0,0,0,1,0,0]).

In these simulations following are the conditions until 40 days: all trajectories start with subtype '7', all mutations besides 4$^{th}$ and second mutation of 5$^{th}$ loci are targeted, Trimethoprim is used at minimum inhibitory concentration. Following conditions are imposed after 40 days: all mutations are targeted, trimethoprim concentration kept at 0, and small 'WT' is introduced at 40 days. Trajectory starts with '7', then subsequently subtype '10' dominates the population. At 40 days, 'WT' is introduced and it eventually dominates the population by 42 days

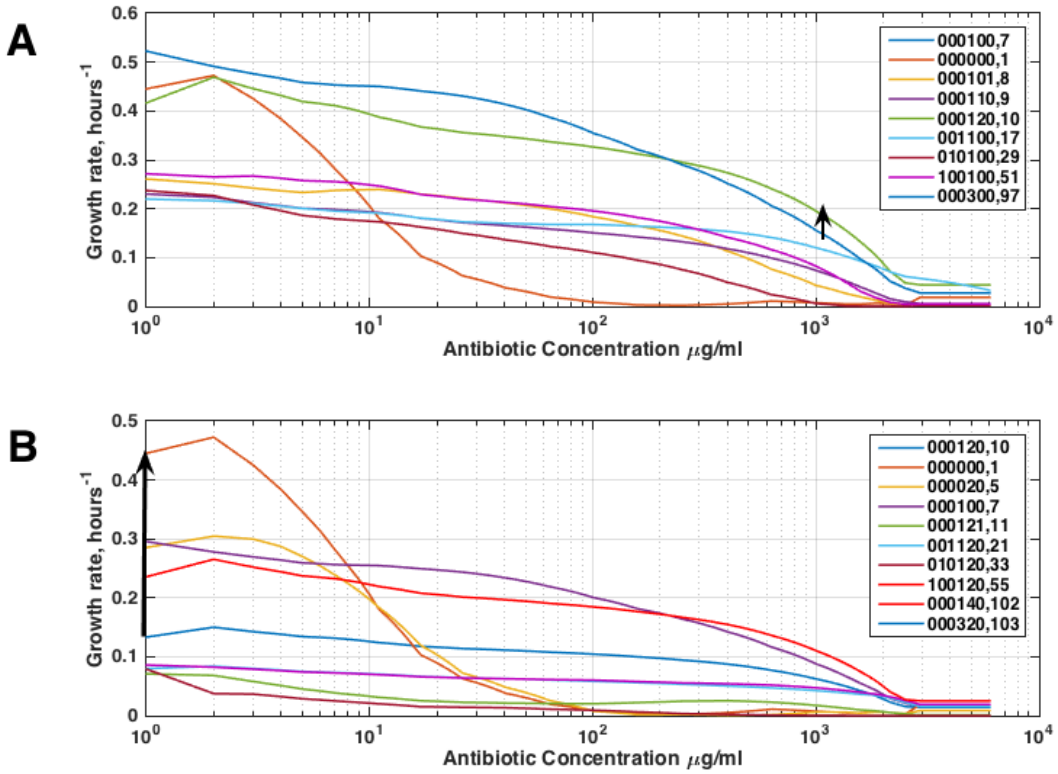


**Figure 3.20.** Growth profiles of major players at luring and killing strategy imposed on subtype '7', ([0,0,0,1,0,0]).

A) Growth profiles of subtype '7' and its descendant are shown, where antisense are targeting all mutations except 4$^{th}$ and second mutation of 5$^{th}$ loci. Most probable move is from '7' to '10'; for unrestricted case most probable move would be from '7' to '17'. B) Growth profiles of '17' and its descendants with addition of WT are shown. Most probable move is from '17'to WT.

### 3.7 References


1) Fleming A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. British journal of experimental pathology. 1929 Jun;10(3):226.

2) Grossman CM. The first use of penicillin in the United States. Annals of internal medicine. 2008 Jul 15;149(2):135-6.

3) Rammelkamp CH, Maxon T. Resistance of Staphylococcus aureus to the Action of Penicillin. Experimental Biology and Medicine. 1942 Dec 1;51(3):386-9.

4) Golkar Z, Bagasra O, Pace DG. Bacteriophage therapy: a potential solution for the antibiotic resistance crisis. The Journal of Infection in Developing Countries. 2014 Feb 13;8(02):129-36.

5) Klevens RM, Edwards JR, Richards Jr CL, Horan TC, Gaynes RP, Pollock DA, Cardo DM. Estimating health care-associated infections and deaths in US hospitals, 2002. Public health reports. 2007 Mar 1:160-6.

6) DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. Journal of health economics. 2003 Mar 31;22(2):151-85.

7) Leeb M. Antibiotics: a shot in the arm. Nature. 2004 Oct 21;431(7011):892-3.

8) Ventola CL. The antibiotic resistance crisis: part 1: causes and threats. Pharmacy and Therapeutics. 2015 Apr;40(4):277.

9) Zipperer A, Konnerth MC, Laux C, Berscheid A, Janek D, Weidenmaier C, Burian M, Schilling NA, Slavetinsky C, Marschal M, Willmann M. Human commensals producing a novel antibiotic impair pathogen colonization. Nature. 2016 Jul 28;535(7613):511-6

10) Dauber-Osguthorpe P, Roberts VA, Osguthorpe DJ, Wolff J, Genest M, Hagler AT. Structure and energetics of ligand binding to proteins: Escherichia coli dihydrofolate reductase-trimethoprim, a drug-receptor system. Proteins: Structure, Function, and Bioinformatics. 1988 Jan 1;4(1):31-47.

11) Palmer AC, Toprak E, Baym M, Kim S, Veres A, Bershtein S, Kishony R. Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. Nature communications. 2015 Jun 10;6.

12) Imamovic L, Sommer MO. Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. Science Translational Medicine. 2013 Sep 25;5(204):204ra132-.

13) Ayhan DH, Tamer YT, Akbar M, Bailey SM, Wong M, Daly SM, Greenberg DE, Toprak E. Sequence-Specific Targeting of Bacterial Resistance Genes Increases Antibiotic Efficacy. PLoS Biol. 2016 Sep

14) Summerton J, WELLER D. Morpholino antisense oligomers: design, preparation, and properties. Antisense and Nucleic Acid Drug Development. 1997 Jun;7(3):187-95.

15) Yeh PJ, Hegreness MJ, Aiden AP, Kishony R. Drug interactions and the evolution of antibiotic resistance. Nature Reviews Microbiology. 2009 Jun 1;7(6):460-6.